

Titre: Developing Models to Evaluate Pedestrian-Cyclist Interaction
Title: Satisfaction on Pedestrian Streets

Auteur: Elnaz Mohsenin
Author:

Date: 2023

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Mohsenin, E. (2023). Developing Models to Evaluate Pedestrian-Cyclist Interaction
Citation: Satisfaction on Pedestrian Streets [Mémoire de maîtrise, Polytechnique Montréal].
PolyPublie. <https://publications.polymtl.ca/56993/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/56993/>
PolyPublie URL:

**Directeurs de
recherche:** Francesco Ciari
Advisors:

Programme: Génie civil
Program:

POLYTECHNIQUE MONTRÉAL

Affiliée à l'Université de Montréal

**Developing Models to Evaluate Pedestrian-Cyclist Interaction Satisfaction
on Pedestrian Streets**

ELNAZ MOHSENIN

Département des génies civil, géologique et des mines

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie civil

Octobre 2023

POLYTECHNIQUE MONTRÉAL

Affiliée à l'Université de Montréal

Ce memoir intitulé :

Developing Models to Evaluate Pedestrian-Cyclist Interaction Satisfaction on Pedestrian Streets

présentée par **Elnaz MOHSENIN**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Owen WAYGOOD, président

Francesco CIARI, membre et directeur de recherche

Alejandro QUINTERO, membre

DEDICATION

To my beloved parents who are always my rock

To my friends for their unwavering support and encouragement

To my dear professor Francesco Ciari who gave me the chance to be part of the CPP project

ACKNOWLEDGEMENTS

I would like to express my utmost appreciation to Professor Francesco Ciari, my thesis supervisor, for affording me the invaluable opportunity to participate in the noteworthy project of Cohabitation Pietons et Cyclists in Montreal. His guidance and support have been instrumental in enabling me to complete this research successfully.

I am also deeply grateful to Mitacs, the national research organization, for partially funding my studies, which provided me with the resources to conduct this research.

Furthermore, I extend my sincere gratitude to my dear friends Kaveh Rezvani Dehaghani, Mehran Mehri, Hamed Malekzadeh and Shabnam Abdollahi for their unwavering support throughout my thesis journey. Their motivation and unwavering faith in me have been pivotal in making this accomplishment a reality. I am profoundly grateful for the myriad ways in which they have provided me with support throughout this process, serving as a continuous source of inspiration.

RÉSUMÉ

Les modes de transport actifs (la marche et le vélo) sont essentiels au développement de systèmes de transport accessibles, sécuritaires et durables. Pour promouvoir le transport actif, les villes aménagent des rues piétonnes et améliorent les espaces de marche et de vélo tout en priorisant les préoccupations de sécurité. De nombreuses études ont été menées pour évaluer la satisfaction des personnes à l'égard de divers lieux et espaces, et il est largement reconnu que la satisfaction est influencée par un large éventail de facteurs. Les études sur la satisfaction intègrent de nombreuses variables, ce qui peut rendre le modèle trop complexe. De plus, les méthodes d'analyse simples peuvent ne pas fournir suffisamment d'informations lorsqu'il s'agit de problèmes tels que l'identification des facteurs qui affectent la satisfaction des utilisateurs. Pendant ce temps, les méthodes d'apprentissage automatique (Machine Learning/ML) sont généralement plus aptes à gérer et à étudier la relation entre plusieurs variables, en particulier dans les situations où des interactions complexes entre les facteurs peuvent exister.

Lors de la pandémie de Covid-19 à l'été 2021, la Ville de Montréal et ses arrondissements ont mis en place 12 projets de rues piétonnes. Le projet pilote « Cohabitation Cyclistes et Piétons (CCP) » a été lancé sur l'avenue du Mont-Royal et la rue Wellington. Dans le cadre de ce projet, la Ville de Mont-Royal souhaitait créer de multiples lieux de promenade et de vélo. Non seulement la Ville de Mont-Royal permet aux gens de rester actifs, mais elle réduit également le risque de propagation de la maladie. La présente recherche visait à évaluer le niveau de satisfaction des piétons et des cyclistes avec leur coexistence mutuelle par les données obtenues à partir de l'enquête CCP. S'appuyant sur la littérature existante, dans la présente étude, des techniques d'apprentissage automatique comprenant à la fois des méthodes de régression et de classification ont été proposées et appliquées pour identifier les facteurs clés influençant les niveaux de satisfaction des utilisateurs.

Cette étude s'est déroulée en deux étapes principales. Dans un premier temps, les données ont été préparées pour analyse via un processus de nettoyage sélectionné en fonction du type de données. Puis, avant d'appliquer les modèles de régression, des tests analytiques ont été effectués. Par la suite, le premier modèle de régression (régression Lasso) a été utilisé pour déterminer les relations et les dépendances entre les variables et pour identifier les caractéristiques les plus influentes ayant un impact sur la satisfaction des utilisateurs. Le deuxième modèle de régression appelé Multinomial Logistic Regression (MNL) a été utilisé pour analyser les caractéristiques

sélectionnées à partir de la régression Lasso et pour évaluer leur contribution à la détermination des niveaux de satisfaction des utilisateurs. Dans un deuxième temps, deux modèles de classification, à savoir Extreme Gradient Boosting (XGB) et Random Forest (RF), ont été mis en œuvre pour évaluer l'importance de toutes les caractéristiques du modèle. Ces modèles ont également évalué la contribution des utilisateurs avec différents niveaux de satisfaction sur la sélection de ces facteurs. De plus, les performances de précision des méthodes de classification et de régression utilisées dans cette recherche ont été évaluées et comparées.

Il a été constaté que, dans les modèles de régression et de classification, l'accent mis sur la sécurité et le comportement mutuel des cyclistes et des piétons sont apparus comme les facteurs les plus influents pour déterminer le niveau de satisfaction des usagers dans les rues piétonnes. De plus, les résultats ont montré que les modèles de classification avaient une meilleure précision que les modèles de régression pour prédire la satisfaction des utilisateurs dans cette étude. De plus, lorsqu'il s'agit de comparer la précision de la prédiction entre XGB et RF, le modèle XGB a démontré une précision légèrement supérieure dans cette recherche, suggérant ses performances supérieures par rapport à la RF.

En guise de recommandation, les résultats de cette recherche pourraient être utilisés pour établir des lignes directrices pour la conception et la gestion des rues piétonnes en termes d'amélioration de la satisfaction des usagers. La présente étude comporte également certaines limites qui doivent être prises en compte, notamment en ce qui concerne la non-prise en compte de certains facteurs d'analyse socio-économique et spatiale dans l'analyse qui sont discutés dans l'étude. À titre de suggestion pour les études ultérieures, puisque l'analyse vidéo a été effectuée dans le cadre de ce projet, l'intégration de ces analyses et des résultats de cette recherche pourrait être utile pour se concentrer davantage sur les problèmes liés à la sécurité entre les piétons et les cyclistes.

ABSTRACT

Active transportation modes (walking and cycling) are vital for developing transportation systems that are accessible, safe, and sustainable. To promote active transportation, cities are developing pedestrianized streets and improving walking and cycling environments while prioritizing safety concerns. Numerous studies have been conducted to assess people's satisfaction with various locations, and it is widely recognized that satisfaction is influenced by a broad range of factors. Studies on satisfaction incorporate many variables, which may make the model overly complex. Furthermore, simple analysis methods may not provide enough insight when it comes to problems such as identifying the factors that affect user satisfaction. Meanwhile, Machine Learning (ML) methods are generally more adept at handling and investigating the relationship between multiple variables, particularly in situations where complex interactions between factors may exist.

During the Covid-19 pandemic in the summer 2021, the City of Montreal and its boroughs introduced 12 pedestrian street projects. The pilot project called Cohabitation Cyclists et Piétons (CCP) was launched on Mont-Royal Avenue and Wellington Street. As a part of this project, the City of Mont Royal wanted to create multiple places where people could walk and cycle. Not only does the City of Mont Royal allow people to stay active but it also reduces the risk of spreading the disease. The present research aimed to assess the interaction satisfaction of pedestrians and cyclists with their coexistence on pedestrian streets by the data obtained from the CCP survey. Drawing from the existing literature, in the present study, ML techniques including both regression and classification methods were proposed and applied to identify the key factors influencing user satisfaction levels.

This study was done in three main steps. As the first step, the data were prepared for analysis through a cleaning process selected based on the data type. As the second step, before applying regression models, analytical tests were conducted. Afterward, the first regression model (Lasso regression) was employed to determine the relationships and dependencies among the variables and to identify the most influential features impacting users' satisfaction. The second regression model called Multinomial Logistic Regression (MNL) was used to analyze the selected features from Lasso regression, and to assess their contribution to determine users' satisfaction levels. As the third step, two classification models, namely Extreme Gradient Boosting (XGB) and Random Forest (RF), were implemented to assess the importance of all features in the model. These models also evaluated the contribution of users with varying levels of satisfaction on the selection of

these factors. Additionally, the accuracy performance of both classification and regression methods used in this research was evaluated and compared.

It was found that, in both regression and classification models, the emphasis on safety and mutual behavior of cyclists and pedestrians emerged as the most influential factors in determining the satisfaction level with the coexistence of pedestrians and cyclists in pedestrian streets. Additionally, the results showed that classification models had better accuracies than regression models in predicting the users' satisfaction in this study. Furthermore, when it comes to comparing the accuracy of prediction between XGB and RF, the XGB model demonstrated slightly higher accuracy in this research, suggesting its superior performance over RF.

As a recommendation, the findings of this research could be employed in establishing guidelines for designing and managing pedestrian streets in terms of improvement of users' satisfaction. The present study also has some limitations that must be considered particularly in terms of not considering certain socio-economic and spatial analysis factors in the analysis that are discussed within the study. As a suggestion for the further studies, since the video analysis has been done within this project, integrating these analyses and findings of this research could be helpful to focus more on safety-related issues between pedestrians and cyclists.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGEMENTS	iv
RÉSUMÉ.....	v
ABSTRACT.....	vii
TABLE OF CONTENTS	ix
LIST OF TABLES	xii
LIST OF FIGURES.....	xiii
LISTE OF SYMBOLS AND ABBREVIATIONS	xiv
LIST OF APPENDICES	xv
CHAPTER 1. INTRODUCTION.....	1
1.1 General Context and Background	1
1.2 The Pilot Project.....	4
1.3 Objectives.....	6
1.4 Thesis Structure.....	7
CHAPTER 2. LITERATURE REVIEW.....	8
2.1 Benefits of Active Transportation: From Health to Pandemic Response.....	8
2.2 Shift from Car Dominance to Shared Spaces	9
2.3 Research Background.....	10
2.3.1 Technical Terms	11
2.3.2 Pedestrian Satisfaction Indices	12
2.3.3 Pedestrian-cyclist interaction satisfaction indices	15
2.3.4 Pedestrians' satisfaction studies by Machine Learning.....	17
2.4 Importance of the Subject.....	18
2.5 Research gaps	20
CHAPTER 3. METHODOLOGY	23
3.1 Handling missing values and data preparation technique	24
3.1.1 Transforming Categorical Data into Numeric Data	25
3.2 Analytical tests	25
3.2.1 Chi-squared	25
3.2.2 Spearman's correlation coefficient.....	26
3.2.3 Variance Inflation Factor (VIF).....	26
3.3 Supervised Machine Learning methods	27

3.3.1	Overfitting and Underfitting.....	29
3.3.2	Optimization techniques and Hyperparameters.....	29
3.4	Regression Models	30
3.4.1	Feature Selection technique.....	30
3.4.2	Lasso Regression.....	31
3.4.3	Multinomial Logistic (MNL) Regression.....	35
3.5	Classification Models	36
3.5.1	Extreme Gradient Boosting	37
3.5.2	Random Forest	39
3.5.3	Feature Importance computation	41
3.5.4	Performance Evaluation and Visualization	41
CHAPTER 4. CASE STUDY		45
4.1	Survey type, Questionnaire and Categorization	45
4.2	Variables of the study.....	48
CHAPTER 5. RESULTS.....		54
5.1	Descriptive Data Analysis	54
5.2	Statistical tests' result.....	64
5.2.1	Variance Inflation Factor.....	66
5.3	Lasso regression results.....	67
5.3.1	Performance evaluation results for Lasso regression	69
5.4	Multinomial Logistic (MNL) Regression results	70
5.4.1	McFadden's pseudo-R-squared	74
5.4.2	Classification evaluation performance for MNL.....	75
5.5	XGBoost results	77
5.5.1	Classification evaluation performance by Confusion Matrix for XGB.....	80
5.5.2	Visualization of XGBoost results by SHAP.....	82
5.6	Random Forest results	86
5.6.1	Classification evaluation performance by confusion matrix for RF.....	89
5.6.2	Visualization of Random Forest results by SHAP	91
5.7	Comparison of top five features in XGBoost, Random Forest and Lasso regression	95
5.8	Category importance comparison for XGB and RF	97
5.9	Comparison of XGBoost and Random Forest by Confusion Matrix	99
CHAPTER 6. CONCLUSIONS, CONTRIBUTIONS and PERSPECTIVES		102
6.1	Practical implications	105

6.2	Linking Research Gaps with Findings	106
6.3	Contribution	107
6.4	Limitations	108
6.5	Perspectives	109
APPENDICES		111
REFERENCES		154

LIST OF TABLES

Table 3.1: Confusion matrix components	42
Table 4.1: Information derived from the survey employed in this research.....	47
Table 4.2: Variables taken from the survey, possible answers, and suggested categories.	49
Table 5.1: Frequency of people with different satisfaction levels who have chosen their favorite pedestrian street.	56
Table 5.2: Frequency of trip purpose of people with different satisfaction levels	57
Table 5.3: Gender frequency of people with different satisfaction levels.....	58
Table 5.4: Frequency of Being a Regular cyclist and level of satisfaction.	59
Table 5.5: Frequency of Trip Companionship	60
Table 5.6: Frequency of Street Visiting Time	61
Table 5.7: Frequency of Accessibility Transportation Options.....	62
Table 5.8: Frequency of Walking during Covid-19	63
Table 5.9: Frequency of Cycling during Covid-19	64
Table 5.10: Spearman correlation result for highly correlated variables	65
Table 5.11: Moderately to highly correlated predictor variables of the model	67
Table 5.12: Lasso best coefficient for both dependent variables	68
Table 5.13: Lasso regression evaluation metrics for both dependent variables	70
Table 5.14: MN Logit Regression Results using Very Dissatisfied as the reference variable.....	71
Table 5.15: Classification evaluation performance for MNL.....	76
Table 5.16: Feature Importance of each independent variable in XGB	78
Table 5.17: Classification performance for XGB.....	81
Table 5.18: Feature Importance of each independent variable in RF.....	86
Table 5.19: Classification performance for RF	90
Table 5.20: Top-five important features of XGB	95
Table 5.21: Top-five important features of RF	95
Table 5.22: Top-five selected features by Lasso	96
Table 5.23: Comparison of categories importance in XGB and RF.....	97
Table 5.24: F1-score and Accuracy comparison for XGB and RF	101

LIST OF FIGURES

Figure 1.1: Satellite view of 12 pedestrian streets in Montreal	5
Figure 2.1: Research diagram.....	22
Figure 4.1: Location of pedestrianized streets on Montreal map	47
Figure 5.1: Confusion matrix for MNL_Y1	77
Figure 5.2: Confusion matrix for MNL_Y2	77
Figure 5.3: SHAP value for Y1_XGB.....	84
Figure 5.4: SHAP value for Y2_XGB.....	85
Figure 5.5: SHAP value for Y1_RF	93
Figure 5.6: SHAP value for Y2_RF	94
Figure 5.7: Confusion matrix for XGB_Y1	100
Figure 5.8: Confusion matrix for XGB_Y2	100
Figure 5.9: Confusion matrix for RF_Y1	100
Figure 5.10: Confusion matrix for RF_Y2	100
Figure 6.1: XGB Feature importance Y1	150
Figure 6.2: XGB Feature importance Y2	151
Figure 6.3: RF Feature importance Y1.....	152
Figure 6.4: RF Feature importance Y2.....	153

LISTE OF SYMBOLS AND ABBREVIATIONS

CCP	Cohabitation Cyclists et Piétons
CCTV	Closed-Circuit Television
FS	Feature Selection
FN	False-negative
FP	False-positive
MDG	Mean Decrease Gini
ML	Machine Learning
MNL	Multinomial Logistic Regression
RD	Rather Dissatisfied
RF	Random Forest
RS	Rather Satisfied
SHAP	Shapley Additive explanations
TN	True negative
TP	True positive
VD	Very Dissatisfied
VIF	Variance Inflation Factor
VS	Vary Satisfied
VAS	Voies Actives Sécuritaires
XGB	Extreme Gradient Boosting

LIST OF APPENDICES

APPENDIX A SURVEY	111
APPENDIX B SPEARMAN'S RANK COEFFICIENT CORRELATION TEST	116
APPENDIX C CROSS-TABULATION (CONTINGENCY TABLE).....	119
APPENDIX D CHI-SQUARED TEST.....	139
APPENDIX E VIF (VARIANCE INFLATION FACTOR).....	144
APPENDIX F LASSO REGRESSION COEFFICIENT	148
APPENDIX G BEST HYPERPARAMETERS FOR XGB AND RF	149
APPENDIX H XGBOOST FEATURES IMPORTANCE PLOT	150
APPENDIX I RANDOM FOREST FEATURES IMPORTANCE PLOT	152

CHAPTER 1.

INTRODUCTION

This chapter explores the thesis topic, identifying research gaps, and clarifying objectives. It provides an introduction to the main topic of the research, which focuses on the assessment of interaction satisfaction in pedestrian streets. It begins by providing an overview of the topic, followed by a discussion of existing research gaps in the field. The chapter concludes by presenting the research objectives that will guide the study.

1.1 General Context and Background

Walking and cycling are two essential elements of sustainable transportation, contributing to various social and economic benefits for our societies (Nguyen et al., 2021). Considering the positive effects of active mobility on health, safety, livability, social interaction, climate change, economy, environment, and air quality, planners and designers should include active transportation in their policies.

Due to high reliance on private cars, numerous negative externalities have been created. Most notably traffic injuries, fatalities, air, and noise pollution. In recent years many European and American cities have invested on expanding walking and cycling paths to encourage people to walk or bike. As an example, Montreal is expanding its bike and pedestrian paths by creating 327 kilometers of new infrastructure, transforming the city's roads during the summer to better accommodate pedestrians and cyclists (Lorraine Carpenter, 2022) .This includes the creation of an Active Pathways Circuit to enhance active transportation (Teddy Elliot, 2020) .

Changing trends in urban transportation planning over the past few years have led to a shift toward sustainable transportation planning. The coexistence of different types of users within the same infrastructure, such as shared streets and spaces, is not a novel concept. Due to the car-oriented urban areas many problems regarding congestion, pollution and safety emerged, in a way that sustainable transportation planning tried to decrease the use of motor transportation (Litman, 2021). One of the feasible ways to fulfill the sustainable transportation goals is through pedestrianization due to limited space (Soni, 2016). According to the definition of shared spaces, they are streets or intersections in which pedestrians and vehicles can use them in a consistently low-speed environment without any obvious physical segregation between them, providing a sense of place and making it possible to perform multiple functions (Karndacharuk et al., 2013)

Over the years, the idea of shared spaces gained popularity, while the risk of conflicts between vehicles and pedestrians in shared spaces became apparent. Various road user groups have expressed agreement or disagreement with the shared space concept. Some road users (including elderly and disabled people) have expressed discomfort about the concept since they perceive it as less safe (Hamilton-Baillie, 2008; Thomas, 2008). In fact, shared spaces were introduced for the purpose of avoiding the construction of more and more roads devoted to motor vehicles, but throughout the years some minor accidents between motor vehicles and pedestrians have occurred which have demonstrated that safety is a concern for the users (Methorst et al., 2007).

Due to the mentioned reasons, pedestrianization is becoming more prevalent in urban areas. The spread of pedestrian streets necessitated the need for specialized studies on the factors that affect people's satisfaction with the cyclists and pedestrians' interaction in these streets. Thus, further studies on the interaction satisfaction of pedestrians and cyclists need to be done in order to better understand how pedestrians and cyclists might coexist in pedestrian streets. It is therefore important to evaluate satisfaction before permanently or temporarily converting streets close to motor vehicles to pedestrian streets for the purpose of increasing active transportation travel (Soni, 2016) .

As pedestrianization satisfaction assessment might be affected by a wide range of personal and environmental factors, it is often challenging to consider all effective factors simultaneously by conventional methods of analysis due to time and budget constraints. Based on the findings of Sahani et al., pedestrian perceived satisfaction is influenced by several factors, including the physical conditions of walking, environmental elements, and overall satisfaction levels of users with the current situation (Sahani et al., 2020). A study which was performed by Distefano showed that different physical conditions and urban design elements have been positively correlated with the walkability satisfaction (Distefano et al., 2021). In a study which was done by S Lee et al., in South Korea, it has been shown that urban design of the physical environment affects pedestrian satisfaction (S. Lee et al., 2021).

However, due to time and budget constraints, few of these studies considered a broad range of related factors simultaneously. The related factors range from the physical condition of the place to travel behavior and characteristics of travelers with the consideration of weather condition, which will be addressed in this research. Despite the growing popularity of shared spaces, very

few studies have examined the interactions between pedestrians and bicycles (Philippe Brodeur-Ouimet, 2023).

According to studies looking at user safety on shared spaces there is negligible chance of a collision in those spaces (Chong et al., 2010). Few studies have also shown correlations between the number of interactions, the speed of bikes, and the density of users on these paths (Gkekas et al., 2020). There would be greater interactions with cyclists in an environment with a larger pedestrian density, although the cyclists' pace would likely be slower in this scenario. Conversely, while the density of the pedestrians' circulation is low, interaction is fewer but cycling speed is higher (Beitel et al., 2018). Likewise, Australian research used direct observations to examine behaviors linked to bike and pedestrian safety on shared paths. The findings demonstrate that bikers' distraction or excessive speed might lead to safety concerns. Furthermore, the majority of pedestrians who interacted with bikers kept their original positions, whereas cyclists typically avoided them (Hatfield et al., 2016).

Most previous studies examined how active modes of transport relate to urban or environmental factors, but rarely concentrated on the co-existence of bicycles and pedestrians on pedestrianized streets and their interactions. They mostly examined what causes people to choose active transport, regarding physical conditions of walking, environmental elements, and satisfaction level of the walking experience (S. Lee et al., 2021). Few studies have investigated the safety of pedestrians and motor vehicles on shared streets (Karndacharuk et al., 2013) or Pedestrian Level Of Service (PLOS) (Nikiforiadis et al., 2021) or the relationship with the physical environment (Nag et al., 2020) but a broad range of variables have not been considered simultaneously in a setting like pedestrian streets which both pedestrians and cyclists circulate which will be addressed in this study. Previous studies have extensively examined interactions between vehicles and pedestrians, as well as between vehicles and cyclists (Narelle Haworth, 2014), and overlooked pedestrian-cyclist interaction satisfaction on pedestrian streets. A more comprehensive comprehension of the factors influencing such interactions within pedestrianized environments will guarantee the secure and harmonious coexistence of these two user groups.

Acknowledging that satisfaction studies are influenced by a wide range of factors, and including all these factors can lead to complex models, researchers have turned to Machine Learning (ML) methods. ML methods offer a simplified approach to modeling in terms of flexibility and adaptability as well as not relying on specific behavioral assumptions or pre-existing knowledge

to construct models (Liew et al., 2022). Additionally, for understanding factors that impact user satisfaction with a place, simple analysis methods like bi-variate analysis may be insufficient. These methods can only indicate relationships between factors and cannot determine the most influential ones, particularly in classification problems like the present research. In contrast, ML models are algorithmically created based on available data, enabling them to handle and investigate relationships among multiple variables, especially when complex interactions between factors are present (Liew et al., 2022)

1.2 The Pilot Project

Many urban policymakers and activists in the field of the urban environment have taken an interest in pedestrianization on a short- or long-term basis. Among Canadian cities, Montreal is known for implementing a policy of converting certain stretches of roads into pedestrian-only spaces during the summer months. Montreal started an initiative to expand the proportion of public areas designated for pedestrians in 2015 (Philippe Brodeur-Ouimet, 2023). In this regard in order to transform streets into public spaces, the Pedestrian and Shared Streets Implementation Program was designed. This implementation was called "small steps" which involved the communities in some neighborhoods and promoted walking and socializing (Philippe Brodeur-Ouimet, 2023).

During the summer of 2020, during the COVID-19 pandemic, Montreal provided an opportunity for outdoor activities and social interaction while adhering to the necessity of physical distancing. This was achieved by introducing a temporary initiative known as Active Safe Routes (ASR) or Voies Actives Sécuritaires (VAS in French). ASR gave pedestrians and cyclists more space on more than 100 km of streets during pandemic (INRS, 2021).

As a result of pedestrianization in the summer of 2020 during Covid-19, the Economic Development Department (Service du Développement Economique (SDÉ) in French) of Montreal launched a new version in April 2021. EDD announces 13 pedestrian street projects for the summer of 2021 to address the requirements of citizens (mobility, recreation, and social activities) while respecting the imperatives of physical distance. Along the commercial arteries, these seasonal pedestrian malls provide people with multiple outdoor activities. In the summer of 2022, Montreal Mayor Valérie Plante announced that 12 streets would be pedestrian-only in a way that the city would fund the pedestrianization of Montreal's commercial arteries with \$12 million over a three-year period (Lorraine Carpenter, 2022). The new pedestrian street project

includes budget for various facilities and data collection, based on the results of pedestrianization in the summer of 2020. These 12 streets were Mont-Royal (Plateau Mont-Royal), Wellington (Verdun), Ontario (Hochelaga-Maisonneuve), Masson (Rosemont-Petite-Patrie), Duluth (Plateau Mont-Royal), Bernard (Outremont), Castelnau (Villeray), Saint-Catherine west in downtown (Ville-Marie), Crescent in downtown (Ville-Marie), Sainte-Catherine east in Marie Village (Ville-Marie), Saint-Denis in Quartier Latin (Ville-Marie), and Marché Jean Talon (Rosemont Petite-Patrie).

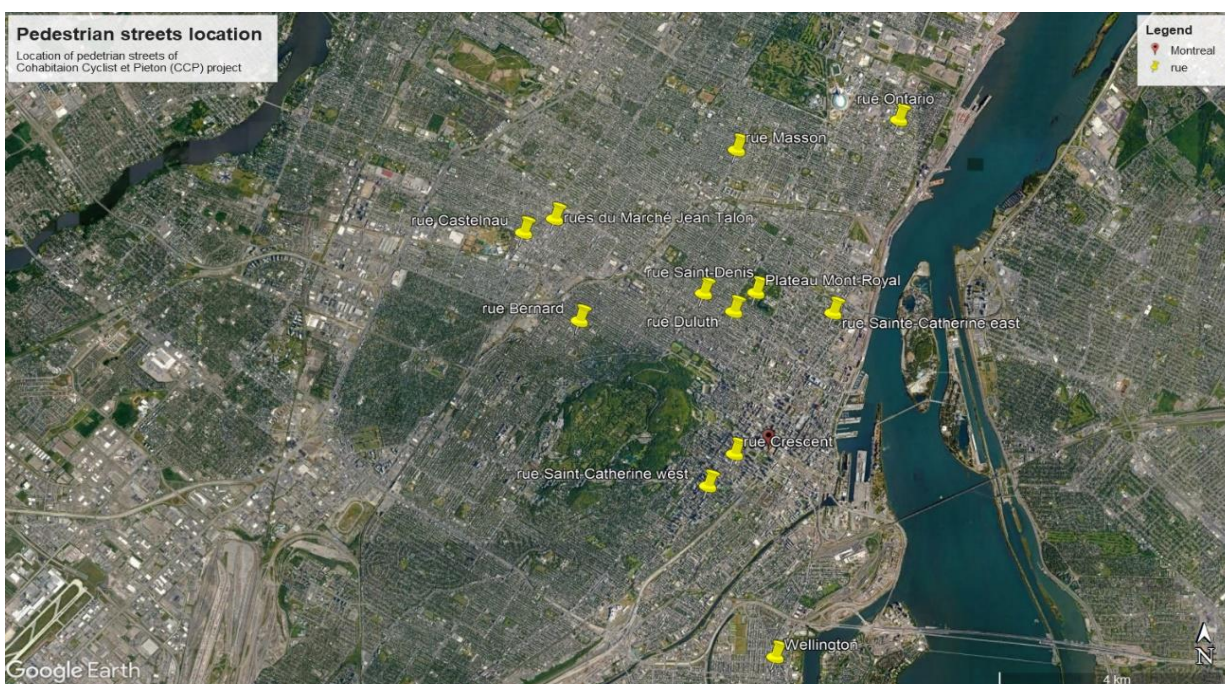


Figure 1.1: Satellite view of 12 pedestrian streets in Montreal

In the initial stage of the pedestrian street in 2020, Mont-Royal city banned bicycles. However, shortly thereafter, the City realized that this approach was not practical, and despite a ban, there may still be instances of noncompliance. As a result, the City decided to have taken up a gamble of coexistence of pedestrians and cyclists but with the priority of pedestrians (Simone Caron, 2023).

During the summer of 2021, a pilot project was implemented to facilitate the coexistence of pedestrians, cyclists, scooters, skateboards, rollerblades, or any other two-wheeled vehicle on two specific pedestrianized streets. The selected sections for this pilot project were Avenue Mont-Royal, located in the Le Plateau-Mont-Royal neighborhood, between Saint-Laurent and Fullum, and Wellington Streets, (Verdun), between Regina and 6th Avenue. Stakeholders, pedestrians,

cyclists, merchants, borough officials, and elected officials of the two mentioned neighborhoods introduced these streets as a pilot project. In these streets, the cyclists, other two-wheeled vehicles, and pedestrians could coexist at the walking speed and yield to pedestrian traffic (Ville de Montreal, 2021). Thereby this will create a "slow zone" (zone lenteur), a term proposed by the neighborhood Plateau-Mont-Royal.

Soon after, concerns arose about access and safety due to the presence of cyclists on pedestrian streets, which were originally designated for pedestrians only. In fact, while considering the coexistence of pedestrians and cyclists in shared spaces, several factors came into play. Pedestrian zones were primarily designed to create safe and accessible environments for walking and socializing. However, cyclists typically have higher speeds compared to pedestrians, which can potentially lead to conflicts and safety issues and pedestrians may feel uneasy sharing the space with cyclists (Philippe Brodeur-Ouimet, 2023). Thus, this was why the evaluation of the safety of pedestrian streets to permit bicycle access became a significant consideration.

1.3 Objectives

To our knowledge, there is a lack of academic studies addressing pedestrian-cyclist interaction satisfaction in previous research on pedestrian streets specially in Montreal. Previous research has not concurrently explored the comprehensive influential factors of satisfaction in pedestrianized streets, especially in our case study. Furthermore, the relative significance of each impactful factor related to satisfaction in pedestrian streets has not been examined in a city like Montreal. Additionally, no previous study has investigated the contribution of each group of respondents in predicting influential factors of interaction satisfaction assessment on pedestrian streets in a city like Montreal. Therefore, this study aims to answer the following questions:

- 1- Which factors have the greatest influence on interaction satisfaction assessment in pedestrian streets?
- 2- What is the relative importance of each factor in interaction satisfaction assessment in pedestrian streets?
- 3- What is the contribution of each class of respondents to the influential factors in the assessment of interaction satisfaction on pedestrian streets?

This study aims to assess pedestrians and cyclists' interaction satisfaction on pedestrianized streets, and to evaluate the importance of different satisfaction factors. In addition, this study tries to address the gap of previous research as it includes a comprehensive spectrum of related factors affecting the pedestrians and cyclists' interaction satisfaction on pedestrian streets. Considering a comprehensive spectrum of related factors provides a better understanding of what influences user satisfaction on pedestrian streets. This research also helps to identify which group of respondents contributes more significantly to the selection of these influential factors.

1.4 Thesis Structure

This research is mainly focused on the pedestrian-cyclist interaction satisfaction on pedestrianized streets in six chapters. The first chapter provides a brief introduction provided here to the pedestrian street followed by some successful projects throughout the world, as well as the problem that prompted this study and its main objectives.

Chapter two discusses the origins and previous studies on shared spaces, pedestrianization, followed by pedestrian satisfaction indices, pedestrian-cyclists interaction indices and methodologies that led to the current study and research gaps. Moreover, this chapter provides a definition of the specific technical terms which used throughout the current study, and previous methodologies for analyzing satisfaction with transportation users.

Chapter three presents two-step methodology including the statistical tests followed by employed Machine Learning methods in this study.

Chapter four consists of a description of the data derived from the survey, the survey, applied variables, and the study area.

Chapter five includes conventional data analyses (contingency tables) of respondents to the survey, followed by the machine learning results from the applied methods introduced in chapter three. In addition, the outcome of the Machine Learning methods along with their accuracy and prediction performance are discussed in more detail in this chapter as well. The results of deployed Machine Learning methods are compared and the more accurate technique for assessing pedestrian satisfaction assessment of the present study are presented.

As a final chapter, the study is summarized with a summary of each part's results as contribution and conclusion, followed by, limitations, and perspectives on pedestrianized streets.

CHAPTER 2. LITERATURE REVIEW

This chapter comprehensively explores the subject, underlining its significance and tracing its origins. It reviews previous studies on the emergence of shared spaces and the importance of active transportation, shedding light on existing knowledge in the field. Furthermore, this chapter delves into pedestrian satisfaction indices, the interactions between bicycles and pedestrians, and previous pedestrian satisfaction studies conducted using Machine Learning. Finally, after reviewing the literature, the chapter discusses the significance of the research and highlights any gaps in current knowledge.

2.1 Benefits of Active Transportation: From Health to Pandemic Response

In recent years many Europeans have been using Active transport (cycling and walking), specifically for commuting to work, associated with higher mental well-being, physical health, and travel satisfaction (Mouratidis et al., 2019). Moreover, active mobility requires less space than motor vehicles as well as reducing congestion and improving economic benefits (Voss, 2018). Non-motorized transportation, such as walking, cycling, their variants, or public transportation can reduce personal vehicle emissions and vehicle ownership and maintenance costs, as well as lower the costs of the city or local government maintenance of roads and parking (Ariffin et al., 2013). However, walking can offer many benefits that go beyond health; according to Ann Forsyth and Michael Southworth's statement that "walkability is the foundation for the sustainable city" (Forsyth et al., 2008) Walking and cycling can enhance livability, equity, social interaction, and economic growth, as well as reducing society's health costs as part of sustainable development (Qiu et al., 2018). Another major advantage of active transportation is that it has a low environmental impact, and it is a pollution-free mode of transportation. In a way that by using active transportation, emissions and pollutants emitted from tailpipes and congestion could be greatly reduced as well. Nowadays many people in developed countries prefer to walk to the nearest destination instead of using a car to avoid traffic congestion, especially during peak hours (Litman, 2017). As Litman in "Evaluate the Value of Walking" in the study for Victoria Transport Policy estimates that non-motorized transport can save 25 percent vehicle-mile reduced annually (Litman, 2017).

Following the outbreak of the pandemic and the lockdown in 2020, people began to become more interested in active transportation in a way that walkable and pedestrian streets have shown greater value throughout the world. Due to global lockdowns, many urban citizens have been

confined to their homes which have completely altered their daily travel patterns. As of early April 2020, approximately half of the world's population, or 3.9 billion people, had been placed under lockdown, which restricted their mobility (Sandford, 2020). It was reported that during the Covid-19 pandemic, physical inactivity among urban dwellers increased when outdoor mobility was restricted (Payne, 2021). Due to the increasing inactivity level of people, the World Health Organisation recommended cycling or walking to improve health (Alasdair Sandford, 2020). During the Covid-19 pandemic, with fears of crowded public transportation and limited indoor socializing, cycling, and walking became the preferred and safe options for short trips (Nikitas et al., 2021). Canadian cities followed this global trend of prioritizing pedestrianization, with examples such as the successful project on Bernard Avenue in Kelowna, a mountain city in British Columbia Canada. Bernard Avenue is a busy pedestrian corridor and an important commercial street that was redesigned in 2017 to make it a pedestrianized street in summer. In response to the pandemic, this street was transformed into a pedestrianized zone during the summer of 2020. Pedestrian traffic counters were installed due to the street's high activity during the summer months. The recorded counts during this period were found to be approximately twice as high as the previous year (Matt Worona, 2020).

2.2 Shift from Car Dominance to Shared Spaces

Since the early 1900s, researchers explored ways to get traditional cities back to their smaller, healthier, walkable streets. Cities in the late 19th century were dominated by walking as a means of transportation. In the meantime, the automobile paved the way for the modernization of cities following the industrial revolution. The negative effects of automobiles began to surface in European cities and later in the United States in the late 1960s (Watson AY, 1988). From the 1950s to 1960s, western countries began creating pedestrian areas as a way to regenerate and prevent downtown commercial areas from further decline (U. Kim et al., 2021). In Germany, Denmark, and, later, North America, pedestrian malls which are known as pedestrian streets were built in the center of cities to increase livability and encourage pedestrian activity (Parkhurst, 2003; Yassin, 2019). After the rapid suburbanization and urban decay of the mid-20th century in Germany, over 200 pedestrian streets were installed in the 1960s and 1970s to encourage shoppers to return to downtowns (Judge, 2013). In 1961, the League of German Cities (Deutscher Städtetag) recommended the pedestrianization of main shopping streets to protect their attractiveness against the pressures of rapidly growing motorization which was a boom in

pedestrian malls that began to prevail in the mid-1960s, with Cologne and Essen as models (ECMT, 1996).

Consequently, the negative effects of automobiles changed the approach to pedestrian-oriented urban space development (Kashani Jou, 2011). Two ways to solve negative transportation externalities are either to develop a strong public transportation system or to increase soft forms of transportation, such as biking and walking, which first initiated the notion of shared streets.

The evolution of shared (street) space evolved as a reaction against car-dominated societies. The coexistence of different categories of users within the same infrastructure is not a new idea; in contrast, traditionally, the streets have been an area of social, cultural, and economic interaction between people (Hamilton-Baillie, 2008a). Planners in the Netherlands attempted to rearrange streets and integrate traffic into social spaces already in the late 1960s (Johnes et al., 1991). Commonly, the term "shared street" is used in English, which originates from the Dutch term "woonerf," which means "street for living" or "living yard" (Biddulph, 2012). They are sometimes referred to as "green streets", "Curb less" and "pedestrian malls" in American cities and Germany's "Play Streets", and Finland's "Yard Streets" are similar concepts as well (DVRPC, 2018). The conceptual and philosophical roots of 'woonerf' date before its actual implementation and introduction moment in the city of Delft in the 1960s by British urban designer Colin Buchanan (Mike Biddulph, 2001). "Woonerfs" or "home zones" which were created in the Netherlands in the 1960s shares the streets with pedestrians, bicyclists, and motor vehicles without clearly separating bicycle lanes and sidewalks from motor vehicles (Mike Biddulph, 2001). The objective of shared streets is to achieve peaceful coexistence between all user types on urban residential streets with pedestrians and cyclists getting priority which integrates all vehicular and pedestrian traffic into one living space and simultaneously enhances safety and quality of life. (Mike Biddulph, 2001). According to studies by Biddulph and Baillie, shared space primarily pertains to urban design and planning disciplines that care about the uses and appearances of public space (Biddulph, 2010, 2012; Hamilton-Baillie, 2008).

2.3 Research Background

Pedestrian streets play a vital role in promoting active modes of transport in urban environments and contribute to shaping a positive image of the city's public space. They serve as a representation of the city and its attractions (Olgun, 2019). Thus, it is essential to examine the interaction satisfaction of pedestrians and cyclists in these places throughout the year to ensure

a continuous and safe flow of pedestrians. Before introducing any case studies, defining the concept and related technical terms within the pedestrian street's context can be beneficial.

2.3.1 Technical Terms

In the literature review, several essential technical terms were pinpointed that are closely related to pedestrian streets and the interaction satisfaction of pedestrians and cyclists. These terms are central to the present research objectives, and they will be frequently employed in the context.

Pedestrianization, pedestrianized/ pedestrian streets: refers to the process of closing or reducing streets to vehicular traffic from neighborhood city streets in order to make them pedestrian-friendly (ECMT, 1996). Based on the study, which was done by Hussein, pedestrianization might be a limit on the number of hours or a permanent limit. Depending on the context and purpose of the projects, pedestrianization in urban spaces can be divided into three types.

- 1- Full-time pedestrian streets: In this design method, vehicles are not allowed to travel inside the streets. It means that full-time pedestrian streets can only be entered by emergency service vehicles, and they will only be accessible to pedestrians.
- 2- Part-time Pedestrian Streets: These pedestrian streets are those that are open only at certain times to vehicular traffic; on-street parking is prohibited, but loading bays are available.
- 3- Traffic Calming Streets: The third method of pedestrianization is traffic calming streets, which aim to diminish the dominance and speed of motor vehicles. In spite of the fact that vehicular access is not restricted, footpaths are widened, and parking spaces are reduced in this area (Hussein, 2018).

User Satisfaction: The assumption is that there is a direct link between the actual service and the customer's perception of it which is considered as satisfaction (Friman et al, 2009). To increase the acceptability of any transport services, the service should be designed and performed in a way that meets customer expectations (Beirão et al., 2007). User satisfaction is a key component of travel satisfaction research, which has been widely studied in domains ranging from marketing to psychology and is one of the crucial examinations of changes in transportation fields and facilities (Fornell et al., 1996). It is often users' experience with the facilities that determine how satisfied commuters are with the service and whether it meets their needs and expectations (Grigoroudis et al., 2010). Among urban transportation agencies, customer satisfaction surveys are the most common method of gathering user satisfaction or

perception. Over the past few years, several countries have begun measuring non-motorized user satisfaction alongside satisfaction with public transport (Rissel et al., 2016) . Moreover, user satisfaction with active transport has gained global popularity as soft mobility becomes prevalent in modern societies (Majumdar et al., 2021). According to a study in Sydney, Australia, those who cycled or walked to work, or study reported higher levels of satisfaction than those who drove cars (Rissel et al., 2016) .

Interaction Between Pedestrians and Cyclists: According to the previous studies, although the number of interactions that had led to accidents between cyclists and pedestrians in shared spaces has been very few, these collisions and near misses cannot be ignored. A study in Finish cities showed that near accidents are 50 times more frequent than collisions and could affect pedestrians' and cyclists' sense of safety (Mesimäki et al., 2021). Another study by O'Hern et al., in Melbourne Australia during 2006-2016 discovered that the possibility of pedestrians' injuries caused by collisions with cyclists are less frequent than the possibility of pedestrians' injuries in collisions involving motor vehicles (O'Hern et al., 2019). Poulos et al., discussed near accidents between pedestrians and cyclists that occurred on the road, shared paths, and bicycle paths as well. The results showed that near misses were associated with pedestrian observation errors and unexpected maneuvers as well as pedestrian misuse of shared or bicycle infrastructure (Poulos et al., 2017).

In the following sections some factors related to the topic of the present research will be discussed.

2.3.2 Pedestrian Satisfaction Indices

Based on previous research worldwide, the study of pedestrian satisfaction has gained popularity among urban designers, planners, and transport engineers in recent years. Despite the growing interest in user satisfaction in urban transport development, there have been relatively few studies that specifically assess the satisfaction factors of pedestrianized environments. Generally, studies on satisfaction of walking have fallen into two categories (Sheng et al., 2021); *the first category* focuses on walkability based on "comfort and safety," which assesses how the physical environment (including pavement, shading, street furniture, and other physical conditions) affects walking (Borst et al., 2009). In these studies, generally, questionnaires are used to measure the influence of these variables on people's willingness to walk, their sense of safety and comfort while walking, and the frequency with which they walk (Borst et al., 2009). A study which was performed by Rahman et al., in Saudi Arabia showed that streets' design and facilities had a

significant impact on the walking satisfaction (Rahman et al., 2019). *The second* category explores the spatial conditions of walking, studying how land use, density, and street connectivity influence walking behavior (Cervero, 1996). In a study which was performed by Oakes in Minneapolis-St. Paul, it was shown that neighborhood design had a tremendous impact on walking behavior and habits of residents (Oakes et al., 2007). On-site observations (such as gate count surveys) and questionnaires are used to analyze the effects of the spatial conditions on pedestrian volume distribution (Chen et al., 2017). Various subjective and objective indices influence the satisfaction of users in pedestrianized streets, as satisfaction assessment is a multi-faceted subject. Some of the subjective and objective attributes identified in various studies are as follows.

Built environment and streetscape characteristics: In a recent systematic review conducted by Fonseca, documents were analyzed to explore the influence of the built environment and streetscape characteristics on pedestrian satisfaction and walking purposes. The study revealed that spatial factors, such as intersection density, residential density, and land use mix, were frequently utilized to evaluate pedestrian satisfaction. In addition, streetscape design and security attributes were much less widely used in pedestrian satisfaction assessments (Fonseca et al., 2022). In a study which was performed in Korea by S. Kim et. al, showed that built environment attributes can improve both pedestrian satisfaction and activities (S. Kim et al., 2014).

Trip-specific information: In a comprehensive study by Lee et al, pedestrian satisfaction will probably differ based on why people walk ,where they walk, and when they choose to walk. In another word, the importance of environmental factors influencing pedestrian satisfaction in a pedestrianized environment may vary depending on the purpose of travel, whether it is for work, school, shopping, or leisure activities (S. Lee et al., 2021).

Aesthetics and attractiveness: Lee et al. have also highlighted the importance of attractiveness and aesthetics as factors that can shape pedestrians' satisfaction levels (H. Lee et al., 2021). Moreover, research performed by Sahani et al. indicated that aesthetics, and maintenance of the sidewalks and pedestrian zones can primarily motivate people to walk (Sahani et al., 2016). Furthermore, pedestrian overall satisfaction with the place can be impacted by cleanliness and attractiveness of the place (Methorst et al., 2010).

Urban design characteristics of streets: A study by Blečić showed urban features and design characteristics of streets are most strongly correlated with a qualitative perception of street quality

and walking satisfaction. According to the result, useful sidewalk width, architectural, urban, and environmental attractions, the density of shops, bars, services, economic activities, vehicle-pedestrian separation, cyclability, opportunities to sit, shelters and shades, and streetlighting were the top predictors of walking satisfaction (Blečić et al., 2016). Another study performed by Zainol et al. showed that street furniture and pedestrian facilities such as lighting and benches influenced pedestrian satisfaction (Zainol et al., 2014). Ann Forsyth proved that some urban design features such as building orientation and setback, block length, building height and street enclosure influenced walking satisfaction (Forsyth, 2015) .

Climatic issues' impact on active transportation: In many countries such as Canada, the indicator of weather should not be underestimated as it has a harsh winter and climate should be a barrier for active transport. A study conducted in Finland revealed that weather conditions had a significant influence on active transportation satisfaction in winter cities. It was found that individuals spend only 4% of their total time outdoors during cold weather (Mäkinen et al., 2006). Another study conducted in East Lansing, Michigan identified natural factors such as snow, ice, treacherous footing, and slippery surfaces are the primary barriers to walking and bicycling during the winter months. Respondents also expressed concerns about natural obstacles in the fall, including leaf litter and debris on sidewalks and bicycle lanes, which could impede walking (Westphal et al., 2012). A study in Canada conducted by Clark et al. suggested that increasing walking as a mode of transport might be difficult to achieve in climates where there are consistently low temperatures or high amounts of precipitation and climate variables such as snow. He also noted that total precipitation has a significant impact on walkability and satisfaction level of walking (Clark et al., 2014). In a survey which was done in Toronto, the most important impacts of City services on walkability and satisfaction level of walking were identified as snow removal on sidewalks (86%), lighting on sidewalks (78%), and general sidewalk maintenance/repairs (78%) (Medical Officer of Health, 2014). As Campos et al. found weather/climatic conditions was highly effective in encouraging people to walk as a mode of transport (Campos et al., 2003). In both Clark et al.'s and Campos et al.'s studies, the weather was associated with coldness since their studies were conducted in cold countries. In Vancouver, Canada, a study found that car-dependent neighborhoods (with longer block lengths, fewer intersections, and greater distances to amenities) became inaccessible in snow, while older adults who lived in walkable neighborhoods walked 25 percent fewer routes in the snow (Clarke et al., 2017). A survey conducted in Vancouver, Canada, by Winters et al. revealed that snowy and icy

conditions significantly discouraged potential cyclists (Winters et al., 2011). Flynn et al. studied bicyclists reported commuting habits in a northern US state, indicating that around 2.5 cm of snow on the ground decreased cycling likelihood by approximately 10%. Enhancing snow removal and preventing ice formation could boost winter cycling rates (Flynn et al., 2012). A Swedish study by Bergström and Magnusson emphasized that improving winter maintenance could potentially increase winter bicycle trips by 18%, with snow clearance being the most critical factor for achieving this outcome (Bergström et al., 2003). A study conducted in Sweden suggests a direct connection between soft mobility (walking and cycling), and climatic barriers, requiring future urban planning and design for winter cities to consider a wider range of weather conditions, especially rain (Chapman et al., 2017).

As seen, the lack of a study that explores a wide range of subjective and objective elements that contribute to pedestrian satisfaction simultaneously can be seen. These elements can include socio-economic, urban design, aesthetic and infrastructure, trip-specific information, mutual behavior of cyclists and pedestrians, the general attitude toward pedestrianization, and season-related mobility behavior which will be covered in the present study.

2.3.3 Pedestrian-cyclist interaction satisfaction indices

As the interaction and mutual behavior of cyclists and pedestrians are the main concern of pedestrianized streets (Mastora et al., 2023), this section, discuss the indices related to the interactions between pedestrians and cyclists. Earlier research categorized the interactions between cyclists and pedestrians in shared spaces according to the angles at which they come into conflict. These interactions are typically grouped into three types (Alsaleh et al., 2020) :

- 1- Same Direction Interaction: This occurs when cyclists and pedestrians are moving in the same direction with a small angle difference, typically around $0^\circ \pm 30^\circ$.
- 2- Opposite Direction Interaction: In this case, cyclists and pedestrians are traveling in opposite directions, with an angle difference of approximately $180^\circ \pm 30^\circ$.
- 3- Crossing Interaction: This category encompasses all other cyclist-pedestrian interactions, where the angle differences do not fit the criteria for same or opposite direction interactions.

Studies related to interaction is an important aspect of pedestrian streets' studies as it helps to understand how well pedestrians and cyclists coexist in shared spaces, and what factors influence

the interaction satisfaction. Despite the increasing number of pedestrian streets worldwide, there have been relatively few studies addressing the assessment of interactions between pedestrians and cyclists and the potential issues that can lead to dissatisfaction with these spaces. The following paragraphs in this section will discuss some objective and subjective indices related to the coexistence of pedestrians and cyclists' interaction satisfaction.

Speed of Cyclists: In the study by Gkekas et al., pedestrians stated that bikers' speed was a factor in contributing to accidents in high-volume non-motorized shared spaces. Both bikers and pedestrians saw a decrease in satisfaction as a result of these collisions (Gkekas et al., 2020). Research of Beitel et al. argued that the speed of cyclists is the key factor in the frequency of accidents with pedestrians (Beitel et al., 2018). This study was conducted on university campuses, where bike and pedestrian traffic is heavy, and speeds are slow.

Intersection Crosswalks: According to a study by Singh et al., pedestrian satisfaction can be affected by intersection crossings, right-turn traffic lights, channelization islands, pedestrian markings (zebra crossings), turning vehicles, motorist behavior, and pedestrian bicycle interactions at crosswalks where there may be a conflict between pedestrians and vehicles traveling faster than pedestrians (Singh Yadav et al., 2015).

Comfort: Based on the study by Lee et al. the three components of pedestrian satisfaction are pedestrian safety, walking convenience, and pedestrian comfort; in a way that indicators of such satisfaction include sidewalks, illegal parking, sufficient walking space, cleanliness, etc. (S. Lee et al., 2021). Study conducted by Calvey et al. have indicated that comfort has a direct impact on overall route satisfaction. This suggests that the comfort experienced by both cyclists and pedestrians during their interactions could potentially influence their overall satisfaction of their interaction (Calvey et al., 2015).

Perceived safety: A study conducted in Dublin City regarding cyclists' safety perceptions revealed that cyclists favor streets perceived as safe (Lawson et al., 2013). Study by Calvey et al. indicated that there is potential link between perceived safety of both pedestrians and cyclists' satisfaction with interaction in shared spaces (Calvey et al., 2015). Perceived safety is how street users personally gauge safety levels. It's based on their own feelings, shaped by experiences and observations on the streets. These experiences and observations come from interacting with street users, their own observations, and information from others (Drs. W.P. Vlakveld, 2008).

2.3.4 Pedestrians' satisfaction studies by Machine Learning

During the last few decades, machine learning and deep learning have added new methodological options to the study of pedestrians and walking satisfaction (Kandt et al., 2021). As an example, Google and other platforms now offer street view services that allow users to navigate vast amounts of street data through online mapping services. Accordingly, computer vision techniques are being used to batch-process visual image data (Gong et al., 2019; Kandt et al., 2021; Meng et al., 2020). Therefore, the use of advanced technologies allows researchers to carry out studies by machine learning and Artificial Intelligence that were previously impossible.

The following paragraphs in this section will introduce some studies that have been conducted using Machine Learning techniques and are related to pedestrian satisfaction in walking environments.

- 1- Lee et al. was developed a Multinomial Logit model to evaluate pedestrian satisfaction based on five themes: personal, walk-facilities, land-use, pedestrian volumes, and weather in Seoul, Korea. Results indicated that personal variability attributes such as spatial cognition, travel purpose, and environmental effects were the most significant features in the model. Additionally, land-use types, crowdedness, and residential information were examined in this study (W. Do Lee et al., 2018).
- 2- Dong et al. incorporated impact-asymmetry analysis and Gradient-Boosting Decision Trees (GBDT) for assessing walkability measures and satisfaction among residents of gated and open communities in Harbin, China. Results showed that in gated communities, neighborhood leisure activities, space, and street furniture play the most significant role in walkability. It also showed that gated communities place the highest priority on social interaction when integrating attribute classification and performance. Moreover, persuading people to walk in open communities is largely dependent on the quality of sidewalks, the safety of the neighborhood, and the availability of leisure activities (W. Dong et al., 2019).
- 3- An analysis by Nag et al. in Kolkata, India which was done by Ordinal Logarithmic regression model measured the satisfaction with the current walking environment of residents on 14 factors such as quality of infrastructure, comfort, safety, design, and others. Analysis results revealed that buffers, ease of walking, zebra crossings, continuity of footpaths, and nighttime safety are the factors most likely to significantly influence overall satisfaction with pedestrian facilities. Following that, stated preferences were collected and the results of a Negative

Binomial regression model revealed that improving footpath continuity had the most significant impact on user behavior change (Nag et al., 2020).

- 4- The study of Lee et al. using Binary and Ordered logit models, examined the effects of land use and street type on pedestrian satisfaction in the city of Changwon in South Korea. The analysis results showed that the less illegal parking, the more pedestrian space, existence of pedestrian guidance facilities, and green space have higher impact on pedestrian satisfaction. Factors positively affecting the satisfaction of pedestrian paths according to land use were physical environmental variables, such as the separated sidewalk variable. The cleanliness of streets and pedestrian guidance facilities are dominant factors influencing commercial areas, suggesting that land use influences these factors differently. In both separated and non-separated sidewalk cases, a common factor affecting satisfaction was also mentioned as the sufficiency of walking space (S. Lee et al., 2021).
- 5- Based on a study conducted in Seoul by Lee et al., Logistic Regression Random Forest, XGBoost (Extreme Gradient Boosting) and SHAP algorithm were utilised. The results showed that pedestrian satisfaction is influenced by the physical and visual characteristics of the street environment. SHAP (Shapley additive explanations) examines how different features will influence the probability of event occurrence. Thus, every visual feature of the streetscape, such as the visible area, urban design quality, and street environmental factors had a greater effect on pedestrian satisfaction than any physical features. Further analysis showed that XGBoost had a higher Accuracy and Precision than Random Forests and Logistic Regressions (J. Lee et al., 2022).

2.4 Importance of the Subject

In recent years, transportation planners and decision-makers realized that implementing exclusive lanes for different categories of road users is extremely difficult due to the limited public space (Nikiforiadis et al., 2021) and streets need to play a more active role in social interactions (Hamilton-Baillie, 2008). Pedestrianization is considered one of the feasible methods to improve mobility and enhance environmental livability in congested areas. In fact, pedestrianization is a very effective, low-cost, and sustainable solution to tackle the lack of space and socioeconomic system (Soni et al., 2016).

A pedestrian-friendly zone known as Strøget, which was converted to a car-free zone in 1962, is a global successful example of pedestrianization designed by Jan Gehl (Dextre et al., 2013) which

demonstrated that a healthier society can be achieved with a prosperous economy. Strøget proved to be a huge success, with businesses realizing that traffic-free environments provide increased financial revenue due to a different range of activities (Global Designing Cities Initiative, 2020). Among Canadian cities Montreal was one of the cities to create temporary or permanent pedestrian streets and proposed pedestrian zones which cyclists were allowed to circulate. Soon after the need for pedestrian zones came into play.

Over the past few years, several countries have begun measuring non-motorized user satisfaction alongside satisfaction with public transport (Rissel et al., 2016). There are many factors that reflect users' experience and satisfaction within pedestrian streets, including the user's socioeconomic characteristics, the purpose of the visit, and overall attitudes towards pedestrian streets and mutual behavior of cyclists and pedestrians. Additionally, it is important to consider other factors such as the weather in some cities like Montreal due to the long winter months which will affect people's mobility. Moreover, given that pedestrian streets function as shared spaces, the possible conflict between pedestrians and cyclists not only impacts their overall satisfaction with the environment but also this interaction poses a potential safety hazard for both groups. Despite the growing number of pedestrian streets around the world, collisions between pedestrians and cyclists have received limited academic attention, even though they can lead to serious injuries and death (O'Hern et al., 2019). However, the investigation of pedestrian-cyclist interaction satisfaction within pedestrian streets remains an unexplored area of study. Consequently, this study aimed to gain insight into examining interaction satisfaction between pedestrians and cyclists on pedestrian streets and assess the potential factors that may affect the overall satisfaction of those places.

Considering the research background, an extensive spectrum of personal and environmental factors that contribute to satisfaction on pedestrianized streets were proposed. Indeed, in the realm of urban transportation management, considering qualitative and quantitative factors simultaneously could provide a more accurate insight for satisfaction assessment. Numerous factors are included in studies on satisfaction, which might make the model redundant. Furthermore, when it comes to issues like determining the variables that influence satisfaction, conventional statistical methods did not provide adequate insights. Machine Learning (ML) approaches, on the other hand, are often better at managing and examining the relationship between multiple factors, particularly in studies where there may be complicated interactions

between factors (Liew et al., 2022) . Furthermore, past research has primarily focused on the high severity of injuries often incurred when motor vehicles collide with pedestrians and cyclists but there is an increasing concern from pedestrians about the threats they perceive from cyclists in spaces where cyclists and pedestrians have interactions (Mueller et al., 2015) . In fact, assessing the interaction satisfaction of pedestrians and cyclists not only brings urban planners the opportunity to modify the current car-free zones but also could give policymakers a new approach to designing safer and more satisfying pedestrian streets.

2.5 Research gaps

Previous studies have predominantly focused on analyzing the impact of a limited number of street design factors on pedestrian satisfaction through mathematical models. However, some of these studies only considered a few subjective and objective factors (Ujjwal et.al, 2021). Additionally, the research targets have often been limited in scope, overlooking the comprehensive spectrum of factors that can affect user satisfaction. In addition, despite the growing number of pedestrian streets worldwide, few academic research focused on the interaction satisfaction of pedestrians and cyclists (Philippe Brodeur-Ouimet, 2023). Therefore, there is a need for a more comprehensive and inclusive approach to assessing pedestrian-cyclist interaction satisfaction that incorporates a broader range of factors.

The summarized research gaps are as follows,

- 1- Previous research has not considered a wide range of relevant factors simultaneously in studying pedestrianized streets, including both objective and subjective variables related to the built environment, individual trip information, season-related impacts, mutual behavior of pedestrians and cyclists, and socioeconomic status, especially in Montreal when some segments of commercial streets are converted to car-free zones seasonally.
- 2- The evaluation of the interaction behavior of pedestrians and cyclists has not been sufficiently investigated before, and the contribution of each influencing factor to the performance of the interaction satisfaction model on pedestrian streets has been less addressed.
- 3- Previous studies have not considered the contribution of each respondent's class (ranging from very dissatisfied to very satisfied) in classification models for interaction satisfaction assessment with pedestrian streets.

By considering both the built environment and individual factors, it can be possible to gain insights into the aspects that significantly impact satisfaction with the existing interaction of pedestrians and cyclists on pedestrianized streets. Understanding these factors is crucial for improving the design and functionality of pedestrianized streets to enhance user experiences and create more satisfying and safe environments for users.

In this study, various factors of both subjective and objective will be simultaneously evaluated to determine the pedestrian-cyclist interaction satisfaction on pedestrian streets. The evaluated factors include socio-economic factors, urban design, aesthetic and infrastructure elements, trip-specific information, perception of mutual behavioral features between cyclists and pedestrians, general attitude toward pedestrianization, and season-related mobility behavior. Moreover, due to the complexity and large number of factors involved in this study, ML methods will employ to effectively uncover intricate patterns and relationships that may not be apparent through traditional statistical methods (Liew et al., 2022).

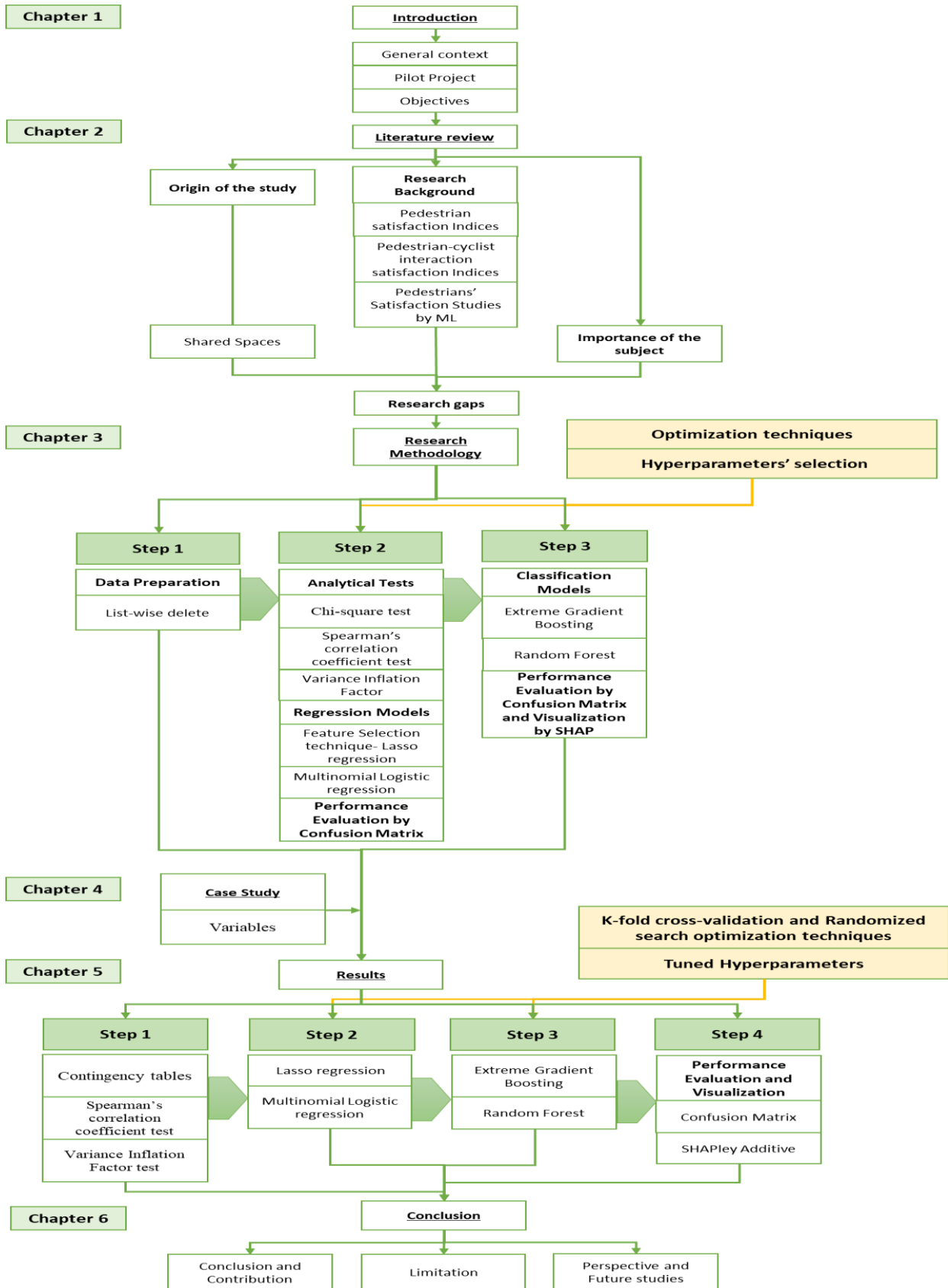


Figure 2.1: Research diagram

CHAPTER 3. METHODOLOGY

This chapter describes the methodological approach undertaken in this Study. In this chapter, a methodology which has three steps will be presented. The first step is centered around data preparation techniques and analytical tests. The second step focuses on the regression and classification methods in Machine Learning (ML). Following the implementation of these methods, an evaluation of accuracy performance is presented.

As a *first* step, the data preparation technique is explained which is used in this study. Through this step, the data is prepared for analysis through a cleaning process based on the data type. The data preparation method employed was list-wise delete, which involved retaining only the complete cases and removing incomplete records of responses.

The *second* step consists of analytical tests and regression models. In this step, before applying regression models, some analytical tests are suggested. Chi-squared and Spearman's rank correlation tests are conducted regarding the relationship among the data. A Chi-squared test is performed in order to assess the association or dependencies between variables of the study. Spearman's correlation coefficient test is performed to determine the strength and direction of relationship between the variables. According to the literature review, as satisfaction assessment studies include applying numerous factors, the possibility of the presence of high correlation among variable may exist. In order to assess multicollinearity which means high correlation between two variables, before conducting the Lasso regression which is the first regression model of the study, it is important to assess multicollinearity and homoscedasticity among the variables of the study. Lasso regression is employed as a feature selection technique in this study to identify influential features in assessing the interaction satisfaction level of users in pedestrian streets. By examining the relationships and dependencies among the variables by Lasso regression the most influential features of the study can be determined. However, it is crucial to test for multicollinearity among the variables prior to conducting the Lasso regression, as the presence of multicollinearity can lead to biased results (Herawati et al., 2018). To assess multicollinearity, this study utilizes the Variance Inflation Factor (VIF) test. The VIF test aids in identifying highly correlated variables, which can affect the reliability of the Lasso regression (Pavlyuk, 2019). As a second regression model, Multinomial Logistic Regression (MNL) is proposed to analyze the most relevant selected features derived from Lasso regression to figure out how these factors'

contribution may affect interaction satisfaction of pedestrians and cyclists with the pedestrian streets.

As the *last* step, two classification models, Extreme Gradient Boosting (XGB) and Random Forest (RF), are presented to assess the importance of all features on the interaction satisfaction assessment of pedestrian streets. By classification models the weight of each feature of the model could be calculated and ranked; thus, the contribution of each feature to assessing the satisfaction level of users in pedestrian streets could be determined. Afterward, some common methods are discussed to evaluate the accuracy performance of both classification and regression techniques used in this study.

3.1 Handling missing values and data preparation technique

Missing values are a common issue encountered in various domains involving data analysis, leading to performance degradation, data analysis challenges, and biased outcomes due to the differences between incomplete and complete values (Ayilara et al., 2019). The severity of missing values depends on various factors such as the extent of missing data, the pattern of missingness, the underlying mechanism causing the missing data, and the important one which is the data type (Kang, 2013). Several techniques can be employed to handle missing values, including the deletion of instances, or replacing them with estimated or potential values (Zhang, 2015). This process, known as imputation, helps to fill in the missing values. Likewise, it is important to exercise caution when utilizing imputation techniques for handling missing values when the data type is categorical, as there is evidence suggesting that this approach can potentially result in biased outcomes (Goeij et al., 2013). Imputing missing values in categorical data introduces the risk of bias and leading to distorted results and potentially misleading conclusions (Nakagawa, 2015). Deletion of missing values can be performed in two ways: pairwise deletion and list-wise deletion (Ferris, 2009). List-wise deletion involves removing any case that has one or more missing values and is commonly used as the default choice for analyzing data in statistical software packages (Soley-Bori, 2013). Analyses without missing values are called complete case analyses (Nakagawa, 2015). This approach was chosen for this study for handling the missing values.

3.1.1 Transforming Categorical Data into Numeric Data

As the dataset is categorical, a technique should be used to transform categorical variables into a numeric format that can be easily processed by machine learning algorithms. Categorical variables must be changed in the data ready step since machine learning models require numeric input variables. There are many techniques to convert categorical variables to numerical ones. One of the approaches is Label Encoding. This technique is appropriate when encoding target variables with no inherent order. This approach ensures that the encoded labels are used as identifiers without inherent order or value. It is useful when the classes have no meaningful numerical relationship or when it is intended to treat them equally without introducing any unintended biases based on their encoded values.

3.2 Analytical tests

In this section some analytical tests considering the type of datasets are introduced in order to figure out the relationship and dependency of the variables of the study.

3.2.1 Chi-squared

It is important to know the relationship among variables while analyzing a dataset, as it provides valuable insight into how factors interact and influence each other. Understanding the relationships between variables helps to uncover patterns, dependencies, and associations in the data, which can lead to meaningful interpretations and predictions. In order to determine if there is a significant association or dependency between the categorical variables, two types of tests are used. The first one is Chi-squared test (χ^2) which is a commonly used technique employed to analyze the relationship among categorical variables (Singhal et al., 2015). By comparing the observed frequencies in a contingency table (crosstab) with the expected frequencies, this test helps identify if there is a meaningful relationship between the variables. In general, the χ^2 test can only be used when the expected values in each cell of the contingency table is equal to or greater than 5. The frequency of more than 5 in each cell can minimize the risk of inaccurate results and statistical errors (Kroonenberg et al., 2018).

If this assumption is violated the Chi-squared test's result is not reliable due to Type I and Type II errors. Type I error involves falsely concluding that there is a significant effect or relationship among variables when there isn't, while Type II error involves failing to detect a true effect or

relationship among variables. Statistical errors, and other statistical tests should be used (Singhal et al., 2015).

The formula of Chi-squared test is as follows.

$$\chi^2 = \sum \left(\frac{O_i - E_i}{E_i} \right)^2 \quad \text{3-1: Equation 1}$$

Where;

- χ^2 is the Chi-squared.
- O_i is the observed frequency of each variable of the study.
- E_i is the expected frequency of each variable of the study.

3.2.2 Spearman's correlation coefficient

The second test for assessing the strength of the correlation among the data is Spearman's correlation coefficient test. Unlike the Chi-squared test, Spearman's rank correlation test measures the strength and direction of the monotonic relationship between variables. This test evaluates how well the ranks of the variables correspond, indicating if there is a significant monotonic relationship between them. It is useful to determine the extent of a monotonic association between the variables of interest. It is suitable where the exact numerical differences between values may not be important, but their relative order is. Spearman's rank correlation coefficient ranges from -1 to 1, with positive values indicating a positive monotonic relationship, negative values indicating a negative monotonic relationship, and values close to zero indicating no monotonic relationship (Bolboaca et.al, 2006). The formula of Spearman's correlation coefficient is as follows.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad \text{3-2: Equation 2}$$

Where;

- ρ is the correlation coefficient
- n is the number of observations in the sample
- d_i is the difference in paired ranks.

3.2.3 Variance Inflation Factor (VIF)

Prior to conducting a regression analysis, it is recommended to assess multicollinearity. Multicollinearity occurs when variables of the study have a high correlation, which can lead to

unreliable coefficient estimates and make interpretation difficult (Pavlyuk, 2019). In a statistical model, multicollinearity occurs when more than one predictor variable is linearly related (Dormann et al., 2013). A high degree of multicollinearity can result in inflated standard errors, which makes it difficult to determine the true relationship between predictors and outcomes. There are some ways to test multicollinearity. In this study multicollinearity is checked by Variance Inflation Factor (VIF). By examining the VIF values of the predictor variables, it can be possible to identify and address multicollinearity issues prior to fitting the regression model. This helps and improves the interpretability of the results. By assessing multicollinearity using VIF beforehand, highly correlated variables can be identified and appropriate technique, such as removing one of the correlated variables could be chosen.

Regarding the accepted value for VIF, there is no correlation between an independent variable and any independent variable in the model when $VIF = 1$. There is a likelihood of multicollinearity with one or more independent variables in the model when the VIF value of an independent variable exceeds 1, causing the variance of the estimated coefficient for that independent variable to be inflated. In general, a VIF value greater than 10 is considered problematic multicollinearity (Singh et al., 2021). A variance inflation factor (VIF) formula is as follows.

$$VIF = \frac{1}{1 - R_i^2} \quad \text{3-3: Equation3}$$

Where;

- R^2 (coefficient of determination) is the proportion of the variation in the dependent variable that is predictable from the independent variable.

One technique that can be used to address multicollinearity is Lasso regression. It can reduce the impact of multicollinearity by shrinking or eliminating coefficients of less important variables and improve the model's stability and interpretability which will be discussed thoroughly in the following sections.

3.3 Supervised Machine Learning methods

Based on the literature review, it is acknowledged that satisfaction assessment is widely affected by extensive range of factors which is addressed in this study. However, incorporating all these factors into a model can make it overly complicated to interpret. To address this issue, researchers

have turned to Machine Learning (ML) methods, which have proven to be effective in simplifying complex models (Sarker, 2021).

When it comes to identifying the important factors that significantly impact user satisfaction with a place, simple analysis methods like bi-variate analysis may fall short of providing comprehensive insights. These methods can only reveal the relationships and dependencies among factors but may not reveal which factors are the most influential (Liew et al., 2022). ML methods, on the other hand, offer a more sophisticated approach. Instead of manually constructing mathematical models, ML algorithms automatically generate models based on available data. This enables ML methods to better handle and explore the relationships between multiple variables, particularly in situations where complex interactions between factors may present (Liew et al., 2022).

Two types of Supervised Machine Learning algorithms: Regression Algorithms and Classification Algorithms are used in this study. In supervised learning algorithms, relationships are modeled between target predictions and input features so that new data can be predicted based on the relationships developed over time (Sarker, 2021).

While the objective of the study is determining the most impactful features of interaction satisfaction of pedestrian streets where cyclists and pedestrians co-exist, regression and classification techniques in machine learning (ML) can be used. On the one hand, regression models can provide insights into how each feature contributes to the prediction of target variables. In a way that by examining the coefficients or feature importance scores by ML, the magnitude, direction, and impact of each feature on the outcome variable can be assessed.

ML techniques can perform feature selection or feature ranking, which helps identify the most relevant features for the prediction or classification of the study. By employing algorithms that assess feature importance, such as Lasso regression, which is performed in this research, it can be possible to extract the most influential features of the research.

On the other hand, classification models such as decision trees, random forests, and gradient boosting can predict which class of feature is the most impactful in categorical datasets. Meaning that by analyzing the feature importance or the weights assigned by the model, which features contribute the most to more accurate predictions can be defined (Feng et al., 2020)

3.3.1 Overfitting and Underfitting

In Machine Learning it is important to avoid two common problems: underfitting and overfitting while finding the right balance between model complexity and data patterns. In order to make accurate predictions on unseen data, it is important to ensure that the model can capture the underlying patterns in the data. Some techniques such as feature selection, ensemble methods, and regularization can be used to prevent overfitting and underfitting in machine learning models (Zhou, 2012).

Overfitting occurs when a model becomes overly complex and fits the training data well but fails to perform well on new and unseen data. Underfitting, on the other hand, happens when a model is overly simple and cannot identify the underlying patterns in the data. Neither overfitting nor underfitting is desirable for machine learning models (Zhou, 2012) .

3.3.2 Optimization techniques and Hyperparameters

A hyperparameter is a configuration setting for a model or algorithm that is not learned from the data and must be set manually or automatically. A model's performance and behavior during training are controlled by these parameters, which can significantly affect the model's overall performance. Finding the optimal values for hyperparameters can be complicated, and time consuming especially when it is not clear which configuration will work best for better performance. Instead of manually trying out different combinations of hyperparameter values, machine learning can automate this process. By utilizing hyperparameter tuning, machine learning algorithms can systematically search and evaluate different combinations of hyperparameter values to automatically discover the optimal model architecture. Two tuning or optimization methods are commonly used in analysis. Randomized search and K-fold cross-validation.

Randomized search is the first tuning method used for choosing the most suitable hyperparameters for machine learning models. In random search, only a randomly selected combination of parameters is tested, meaning that to find the best model, not every combination is evaluated. By using this method, instead of searching exhaustively for hyperparameters, samples are selected randomly from a grid. The range of values and a long list of hyperparameters can therefore be assessed within a shorter period (Gan et al., 2021).

The second method to tune hyperparameters is K-fold cross-validation. In cross-validation, the original sample is divided into a training and a test set to evaluate predictive models. Meaning that, in k-fold cross-validation, a limited dataset is divided into k non-overlapping folds. Among the k folds, each is given the opportunity to be used as a held-back test set, while the other folds are combined to be used as a training dataset. The mean performance is reported for k models that were fitted and evaluated on k holdout test sets. By calculating the mean performance, a single value can be obtained that represents the average performance of the model across multiple test sets. This gives a more reliable estimate of how well the model can generalize and perform on new, unseen data. In many machine learning models, 10-fold cross-validation is a good choice because it balances bias and variance in the model's performance, and it is computationally efficient (Refaeilzadeh et al., 2009) .

3.4 Regression Models

In this section, two regression models, Lasso regression and Multinomial Logistic (MNL) regression, will be introduced. A comprehensive explanation of these models is provided in the subsequent sections of this study.

3.4.1 Feature Selection technique

Processing high-dimensional datasets can be challenging due to complexity, redundant or irrelevant features and overfitting. The Feature Selection (FS) process removes noisy and redundant features from high-dimensional datasets to overcome these problems (Fonti et. al, 2017). This boosts the accuracy and computational speed of a classification algorithm as well. In feature selection, the goal is to identify a subset of relevant model inputs (features) that simplify the model structure and estimation procedure while still producing accurate forecasts (Pavlyuk, 2019)

In order to overcome these problems, several methods have been proposed and established to select only relevant variables from a model. A variety of approaches can be used to select variables, including subset selection, dimension reduction and shrinkage (Muthukrishnan et. al, 2017). Only the shrinkage method of Lasso, which is one of the kinds of regularized linear regression variants, is used in this study.

Regularization is a technique used to prevent overfitting and improve the generalization ability of the model (Patil et. al, 2020). In machine learning, regularization is used to prevent overfitting

and improve the performance of models, specifically in Lasso regression. In the context of Lasso regression, regularization includes adding a L1 penalty term to the loss function when training the model. A loss function is a tool in machine learning to help how good or bad the model's predictions are and compared to the actual values (Yihuan et.al, 2020).

3.4.2 Lasso Regression

The acronym “LASSO” stands for Least Absolute Shrinkage and Selection Operator. In regression analysis, Lasso regression is a popular supervised machine learning technique, particularly when there are a lot of features or variables in a dataset. Lasso has been extensively used for feature selection and regularization. With Lasso, the input size of the model is reduced by minimizing the sum of the coefficients' absolute values (L1-penalty function) (Naseri et al., 2022). In fact, LASSO model minimizes the Sum of Squared Errors (SSE), with an upper bound on the sum of the absolute values of the parameters.

Lasso was proposed first time by Robert Tibshirani (Tibshirani, 2011). It is a type of linear regression that utilizes a shrinkage approach; in fact; a Shrinkage occurs when data values shrink towards a central point like the mean. As a result of the shrinkage (regularization) process, models can be interpreted more easily with less bias. Shrinkage also allows for the identification of variables strongly associated with target variables. Therefore, the coefficient of unimportant features becomes zero automatically meaning that the features with the assigned coefficient of zero are removed from the model. In fact, the features with non-zero coefficients will be selected for interpreting the model (Muthukrishnan et. al, 2017). In other words, in an effective approach, input variables are eliminated while these coefficients are shrunk according to the optimal alpha value provided after tuning, resulting in a form of automatic feature selection.

Alpha, in the context of Lasso regression, controls the level of regularization applied to the model. Regularization helps prevent overfitting by adding a penalty term to the model. A higher value of alpha increases the regularization strength, leading to eliminate more features and potentially simpler models. On the other hand, a lower value of alpha reduces the regularization strength, allowing the model to include more features in the selection process and more complex models. The optimal value of alpha is determined through cross-validation, where different values are tested to find the one that provides the best balance between model complexity and performance on unseen data. In simpler terms, the key hyperparameter in Lasso regression models is λ or α . When α is set to 0, the model behaves like a standard linear regression, considering all variables

in the model. On the other hand, setting α to 1 is often considered the ideal value for Lasso regression. As α increases, the penalization becomes more aggressive, leading to fewer variables being selected by the model.

The L1-penalty function is used to avoid overfitting and detect selected features. In other words, the penalty parameter prevents the model from selecting significant values for coefficients (Muthukrishnan et. al, 2017) . By adding a penalty term, Lasso regression reduces the impact of irrelevant and less important features (variables that do not contribute much to the predictor variable), making the model more parsimonious and interpretable (Muthukrishnan et. al, 2016)

By conducting a linearity check, specifically using the VIF test, before performing Lasso regression, the linearity assumption can be verified, multicollinearity can be detected, and the reliability of the variable selection process can be enhanced. This ensures that Lasso regression is applied to a dataset that meets the underlying assumptions, resulting in more accurate and meaningful results (Altalbany, 2021). Moreover, taking multicollinearity into account before applying Lasso regression allows variable selection to be more accurate and prevents bias. Overall, conducting a linearity check and addressing multicollinearity through the VIF test increases the validity of Lasso regression.

In case of numerous variables in the model, the possibility of the presence of high correlation among some variables may exist (Herawati et al., 2018). Thus, the multicollinearity among variables of the study should be tested. The high correlation between variables is called multicollinearity. Lasso regression is well-suited feature selection technique in case of the presence of high to middle levels of multicollinearity among the variables (Muthukrishnan et. al, 2017). In the case of multicollinearity among two predictor variables, by selecting just one variable from two highly correlated variables, LASSO shrinks the rest to zero. In this way, some of the coefficients are reduced exactly to zero, producing easily interpretable models. Research communities, particularly those involved in machine learning, may find Lasso useful as an alternative to conventional feature selection methods (Muthukrishnan et. al, 2017).

For Lasso Regression the general equation is as follows.

$$Y_i = \sum_{i=1}^n \left(y_i - \sum_j x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad 3-4: Equation3$$

Where;

- n is the number of observations
- y is the vector of dependent variable values
- X is the design matrix of independent variables
- β is the vector of regression coefficients to be estimated.
- λ is the regularization parameter that controls the strength of the penalty term (mostly known as ' α ')

3.4.2.1 Performance evaluation of Lasso regression

In order to evaluate the performance of predication of Lasso regression, some techniques are used in the literature (Jomthanachai et al., 2022). In regression analysis, the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), R-Square or Coefficient of determination, and Mean Absolute Error (MAE) are used to measure the model's performance. Generally, a regression model with a high R-square or adjusted R-square have high predictive power (Peterson K. Ozili, 2023). Meaning that model predicts the outcome variable well or the model fits the data. Here five of them are introduced.

1- R-squared (R^2): The coefficient of determination (R-squared) is a statistical evaluation metrics that indicates how much of the variance of a dependent variable can be explained by the independent variables in a linear regression model. A higher value indicates a better fit of the model, with a range of 0 to 1. R-squared between 0.50 and 0.99 is acceptable in social science research, especially when most of the explanatory variables are statistically significant (Peterson K. Ozili, 2023).

$$R^2 = 1 - \frac{\text{sum squared regression}}{\text{total sum of squares}} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad 3-5: \text{Equation4}$$

Where;

- y_i is the observed value of the dependent variable for the i^{th} observation.
- $\bar{y}$ is the mean value of the dependent variable across all observations.
- $\hat{y}$ is the predicted value of the dependent variable for the i^{th} observation based on the regression model.
- i is a variable that represents each observation.

2- Mean Absolute Percentage Errors (MAPE): MAPE, also known as Mean Absolute Percentage Deviation (MAPD), is another method to measures the accuracy of a forecast system. It can be calculated as the average absolute percent error divided by the actual values for each

predictor. It is most commonly used to forecast error since its units are percentage units, which makes it easier to understand. The range of acceptable MAPE is below 50%, and more than this indicates inaccurate forecasting (Montaño Moreno et al., 2013) .

$$MAPE = \frac{1}{N} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \quad 3-6: Equation 5$$

Where;

- N is the number of times the summation iteration happens
- A_i is the i^{th} actual value.
- F_i is the i^{th} predicated value.
- n is the total number of observations.
- i is a variable that represents each observation.

3- Mean Squared Error (MSE): Mean Squared Error is the average of the squared difference between the actual and predicted values in the dataset. In other words, it measures the residual variance.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad 3-7: Equation6$$

Where;

- y_i is the observed value of the dependent variable for the i^{th} observation.
- $\hat{y}$ is the predicted value of the dependent variable for the i^{th} observation based on the regression model.
- n is the total number of observations.
- i is a variable that represents each observation.

4- Average absolute error (MAE): refers to the average of the absolute difference between the predicted values and the actual values.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad 3-8: Equation7$$

Where;

- y_i is the observed value of the dependent variable for the i^{th} observation.

- $\hat{y}$ is the predicted value of the dependent variable for the i^{th} observation based on the regression model.
- n is the total number of observations.
- i is a variable that represents each observation.

5- Root Mean Squared Error (RMSE): calculated by taking the square root of the average of the squared differences between the predicted values and the actual values. RMSE is an excellent general-purpose error metric for numerical predictions, and it is very commonly used.

$$RMSE = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{n}} \quad \text{3-9: Equation8}$$

Where;

- y_i is the observed value of the dependent variable for the i^{th} observation.
- $\hat{y}$ is the predicted value of the dependent variable for the i^{th} observation based on the regression model.
- n is the total number of observations.
- i is a variable that represents each observation.

3.4.3 Multinomial Logistic (MNL) Regression

When there are more than two values for the response variable, Multinomial Logistic regression can be used to analyze categorical data type. Due to its ability to predict the probability of each class based on the values of its predictors, it can be a suitable choice for multi-class datasets. MNL regression model can estimate with maximum likelihood; meaning that the model can estimate the probability of different opinions occurring for each variable in a multi-class dataset. Multinomial Logistic regression is an extension of multiple regression that uses categorical variables instead of continuous ones (Coughenour et al., 2015). When the data are unordered and nominal the preferable regression model is MNL regression which is a kind of classification methods and belongs to supervised machine learning models. By this regression, it is possible to figure out that by increasing one unit of the predictor how the dependent variable will be changed.

For a Multinomial Logistic Regression, the general equation is as follows:

$$\log(odds) = \text{logit}(P) = \ln\left(\frac{P}{1-P}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + \dots \quad 3-10: \text{Equation9}$$

$$p = \frac{\exp(a+b_1x_1+b_2x_2+b_3x_3+\dots)}{1+\exp(a+b_1x_1+b_2x_2+b_3x_3+\dots)} \quad 3-11: \text{Equation10}$$

Where;

- p is the probability that a case is in a particular category,
- $\exp('B')$ (sometimes called ' β ') is the exponential or the odds ratio,
- a is the constant of the equation and,
- b is the coefficient of the predictor or independent variables.

In the context of Multinomial Logistic Regression, in order to assess the statistical significance of the estimated parameters and the overall fit of the model, two parameters of ' B ' and ' df ' should be considered. ' B ' refers to the regression coefficient of independent variables and shows the amount of change for each dependent variable, and ' df ' (Degree of Freedom) refers to the number of independent variables that can vary in the model (Klein, 2020).

3.4.3.1 McFadden's pseudo-R squared

In logistic regression analysis, McFadden's pseudo-R-squared measures the goodness of fit of statistical models. In the 1970s, Daniel McFadden developed this measure (Allison, 2014). McFadden's pseudo-R-squared compares log-likelihoods of logistic regression models and null models with intercepts only. A logistic regression model is estimated by maximization of the likelihood function. A model with no predictors will have a likelihood function of L_{cNull} , and a model being estimated will have a likelihood function of L_c . McFadden's R^2 is calculated as follows:

$$R^2_{McFadden} = \frac{1 - \log(L_c)}{\log(L_{cNull})} \quad 3-12: \text{Equation11}$$

Pseudo-R-squared values range from 0 to 1, with higher values indicating a better fit. Acceptable values for the McFadden's pseudo-R-squared is between 0.2 to 0.4, meaning that model predicts the outcome well or model fits the data (Allison, 2014).

3.5 Classification Models

In this section, two widely used supervised machine learning classification models are proposed along with their hyperparameters deployed in the tuning procedure: XGBoost (Extreme Gradient Boosting) and Random Forest (RF). These models are employed in the study to perform

classification tasks and make predictions based on their respective algorithms. XGBoost utilizes a sequential boosting process, where decision trees are created iteratively to correct the errors of previous trees. In contrast, Random Forest employs a parallelized bagging approach, where multiple decision trees are trained independently and aggregated to make predictions. Both XGBoost and Random Forest are popular tree-based ensemble learning techniques that enhance predictive accuracy by leveraging the multiple weak learners and commonly employed to classify the feature importance (Feng et al., 2020).

A branch of machine learning called Ensemble Learning was developed on the basis of the following assumption: whereas it can be difficult to build an extremely accurate classifier (strong learner), it is much easier to find many approximate rules (weak learners) that come together to create models with high generalization capacity (Zhou, 2012). 'High generalization capacity' refers to the ability of a machine learning model to perform well on unseen or new data points beyond the training dataset. In other words, a model with high generalization capacity can recognize the underlying relationships and patterns within data, allowing it to predict or classify accurately previously unseen instances (Zhou, 2012). By utilizing it, the performance and reliability of the statistical models are improved in terms of overfitting or underfitting.

In the following sections, XGB and RF are presented along with their hyperparameters.

3.5.1 Extreme Gradient Boosting

The initial classification model employed in this study is XGBoost (Extreme Gradient Boosting), in which decision trees are constructed sequentially. In the first tree, the entire dataset is used and in subsequent trees, the errors made by previous trees are corrected. Each new tree focuses on examples that had higher residuals or were misclassified in the previous iteration. It is this sequential process known as "boosting," where each tree learns from the mistakes of its predecessors to become a strong learner. To make the final prediction, all the trees' predictions are combined.

Extreme Gradient Boosting, or XGBoost, is a scalable (which means it can handle large datasets efficiently) algorithm that combines gradient boosting (a technique that sequentially builds decision trees by learning from the mistakes of previous trees prediction (Asselman et al., 2021) with regularization to handle regression, classification, and ranking problems. This makes XGBoost well-suited for solving a wide range of problems, including regression (predicting

numerical values), classification (predicting categories or classes), and ranking (ordering items based on importance). (Henriques et al., 2020).

In practice, a single tree is usually not strong and accurate enough to find a better prediction for the model. Ensemble models are used to combine multiple trees' predictions and summarizing the predictions of all trees. By building a sequence of weak models into increasingly powerful ones, boosting algorithms differ from other ensemble learning techniques. XGBoost is a highly efficient, flexible, and accurate application of distributed gradient boosting systems (H. Dong et al., 2018) and for larger datasets or faster training, XGBoost is suitable for developing parallel computing solutions.

XGBoost runs much faster and more accurately than many other machine learning algorithms and was developed to improve the model performance (Chen, T., et al., 2016). Additionally, XGBoost is compatible with categorical as well as numerical data, and able to handle nonlinearity and multicollinearity data, implicit variable selection, and ease of interpretation (Guo et al., 2021). Tree-based models such as Random Forest and XGBoost are inherently immune to multicollinearity and robust to correlated features. Unlike linear models, tree-based models do not rely on specific assumptions about the independence of predictors or the relationship between predictors and the response variable (Guo et al., 2021). This characteristic of tree-based models alleviates the problem of multicollinearity and enhances the model's robustness and performance.

3.5.1.1 XGBoost hyperparameters

A crucial step in the construction of machine learning models is hyperparameter tuning. Hyperparameter optimization can tune the model performance. In fact, a machine learning model's behavior can be controlled by hyperparameter tuning. By tuning hyperparameters, the optimal value set of hyperparameters for a learning algorithm that can be applied to any dataset can be found. Combining hyperparameters maximizes performance of the prediction while minimizing loss functions, fewer error and leading to better results (Prashant Banerjee, 2020).

XGBoost models contain many parameters. Here are some of the most commonly used hyperparameters used for tuning in this study. *The names in the parentheses are standard in the XGBoost library.*

- **Maximum Depth (max_depth):** Determines the depth of the decision trees.

- **Learning Rate (learning_rate):** Controls how quickly the model learns from previous errors.
- **Subsample (subsample):** Fraction of observations to be sampled for each tree.
- **Column Sampling by tree (colsample_bytree):** Ratio of columns to be used when constructing each tree.
- **Number of Estimators (n_estimators):** Number of trees to be built in the ensemble.
- **Number of Classes (num_class):** Number of classes in the multiclass classification problem.
- **Number of Leaves (num_leaves):** Maximum number of leaves for each weak learner.
- **Minimum Data in Leaf (min_data_in_leaf):** Minimum number of data points required in a leaf node.
- **Gamma:** Minimum loss reduction required for further partitioning a leaf node.
- **Regularization Alpha (reg_alpha):** Regularization parameter for controlling to prevent overfitting and control the complexity by adding a penalty term to the loss function during training.

3.5.2 Random Forest

The second classification model of the present study is Random Forest (RF). In RF, multiple decision trees are trained simultaneously with random subsets of data and features. Trees make their own predictions independently, and the final prediction is determined by a majority vote among the trees. The process is called "bagging," where each tree acts as a weak learner, and the collective decision-making improves the overall accuracy and reduces overfitting.

As a supervised machine learning algorithm, Random Forests (RF) are widely used for classification and regression problems. Classification algorithms assign test data to specific categories by identifying specific entities within the dataset and attempting to determine how they should be labeled or defined. RF is one of a common ensemble method that uses tree-type classifiers.

According to Breiman's studies, random forests are classification and regression algorithms that construct a number of decision trees from training data, then predict the classes based on each

tree, then present a final prediction class based on each tree's maximum vote (Breiman, 2001). The algorithm builds decision trees on different samples and takes their majority vote for classification and average for regression (Tim Jones, 2018). In a majority vote, each decision tree in the ensemble makes its own prediction for a given input. When it's time to make a final prediction, the majority vote counts how many trees predict each class. The class with the most votes is chosen as the overall prediction.

In general, RF consists of a group of decision trees trained using the "bagging" method, meaning that the bagging method involves combining different learning models to increase the results performance (Niklas Donges, 2022).

RF Algorithm has the advantage of being able to handle data sets containing continuous variables as in regression as well as categorical variables as in classification. It gives better results when dealing with classification problems, and also composed of uncorrelated decision trees that are then merged to reduce variance and improve predictions (Wan et al., 2022). RF algorithm is more robust than a single decision tree and can handle outliers and missing values better (Dogru et al., 2018).

The development of a Random Forest model consists of four steps. The first step is to select random samples from a given training set or data set. In the second step, this algorithm builds a decision tree based on all the training data. In the third step, voting will be conducted by averaging the decision tree. In the final step, choose the prediction result that received the most votes (Niklas Donges, 2022) .

3.5.2.1 Random Forest hyperparameters

Random forest models' performance is heavily influenced by the hyperparameters values and contains many parameters. Here are some of the most commonly used hyperparameters used for tuning. *The names in the parentheses are standard in the Random Forest library.*

- **Number of Estimators** (n_estimators): The number of trees in the forest, which affects accuracy and overfitting.
- **Maximum Features** (max_features): The maximum number of features used per tree, influencing model performance and diversity.
- **Maximum Depth** (max_depth): The maximum depth or length of a decision tree.

- **Minimum Samples Split** (`min_samples_split`): The minimum number of samples required to split an internal node (an internal node is a condition or rule that determines how the data is split into different branches).
- **Minimum Samples Leaf** (`min_samples_leaf`): The minimum number of samples required in a leaf node.
- **Bootstrap**: Determines whether bootstrap samples or the whole dataset are used to build the trees. (Bootstrap is the process of creating new training datasets from the original data by randomly selecting observations) (Nisha Arya, 2022)

3.5.3 Feature Importance computation

In order to understand what factors, contribute most to the response variable, a technique called feature importance can be used. Feature importance helps to identify the specific features or variables that have the greatest impact on the dependent variable of the study.

In tree ensemble models, such as gradient boosting and random forest, different metrics can be used to measure feature importance such as 'Weight', 'Gain', and 'Cover'. In simpler terms, "weight" indicates how frequently a feature is used, "gain" tells how much a feature contributes to improving the model's performance, and "cover" shows the average coverage or presence of a feature in the tree splits. Comparing these three feature importance computation techniques, 'Gain' is more informative than just measuring the frequency of a feature in a model.

While the models are constructed with scikit-learn library in Python, the common feature importance calculation is 'gain' method. Gain shows the average score across all splits in the tree-based model where the feature was used. The 'gain' of each feature in a tree-based models is computed over all the decision trees in the model, and then normalized by the sum of all the gains.

3.5.4 Performance Evaluation and Visualization

This section aims to evaluate the quality of the regression and classification models used in the research. For interpretation and visualization of the results in order to assess the accuracy and interpretability of the models, common evaluation metrics like the Confusion Matrix, as well as more advanced methods like the SHAP method are used.

3.5.4.1 Confusion Matrix

The Confusion Matrix (CM) is a popular measure for solving classification problems. Multiclass classification problems as well as binary classification problems can be analyzed using this method (Kulkarni et al., 2020). CM is commonly used to describe the performance of machine learning classification models on a set of test data for which the true values are known. It provides a summary of whether the model is correctly predicted or incorrectly predicted by the evaluation of the classification model.

The confusion matrix typically contains four values: True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) (Spiegel et al., 2018). The TP, FP, TN, and FN values represent the following: (Table 3.1)

- True Positive (TP): This refers to the number of instances correctly classified as a specific class. In the context of the confusion matrix, it means that the model predicted an instance to be in a particular class, and in reality, it belongs to that class.
- True Negative (TN): This represents the number of instances correctly classified as not belonging to a specific class. In the confusion matrix, it means that the model correctly predicts an instance as not being in a particular class, and it does not belong to that class.
- False Positive (FP): This indicates the number of instances incorrectly classified as belonging to a specific class when they do not. In the confusion matrix, it means that the model predicted an instance to be in a particular class, but it does not belong to that class.
- False Negative (FN): This represents the number of instances incorrectly classified as not belonging to a specific class when they do. In the confusion matrix, it means that the model predicted an instance as not being in a particular class, but it belongs to that class. (Kulkarni et al., 2020)

Table 3.1: Confusion matrix components

		Predicted Failure	
		True	False
Actual failure	True	TP	FN
	False	FP	TN

Various model performance metrics, such as accuracy, precision, recall, and F1, are calculated using these values.

Accuracy: In literature, the result of accuracy is the most common value used to compare models. The accuracy of a prediction is the ratio of the correct predictions to the total predictions. In other words, it measures how well the model is able to predict both positive and negative outcomes. Calculating accuracy is as follows:

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad 3-13: Equation12$$

Accuracy measures the overall performance of the model but does not measure the accuracy of prediction for different classes of the model. If most of the observations choose one class more than the other classes, the overall accuracy of the model can be biased. For the purpose of handling this issue and avoiding misinterpretation of prediction accuracy, the sensitivity, F1, and precision are calculated for the negative and positive classes of the model.

Recall (Sensitivity or TPR (True Positive Rate): In a test data set, recall is the ratio of true positive predictions to the actual number of positive instances. It is a measure of how well the model can identify all positive instances. The recall is calculated as follows:

$$Recall = \frac{TP}{(TP + FN)} \quad 3-14: Equation13$$

F1 score (F score): F1-scores have several advantages over accuracy. F1-scores also consider false positives and false negatives, whereas accuracy indicators focus solely on true positives and true negatives. F1 is the harmonic mean of precision and recall. This metric provides a balance between precision and recall. F1 scores are calculated as follows:

$$F1\ score = \frac{2*(precision * recall)}{(precision + recall)} \quad 3-15: Equation14$$

Precision: Precision is the ratio of true positive predictions to total positive predictions. This parameter measures how accurate the positive predictions of the model are. The precision is calculated as follows:

$$Precision = \frac{TP}{(TP + FP)} \quad 3-16: Equation15$$

3.5.4.2 SHAPley Additive

SHAP or SHAPley Additive exPlanations make machine learning models more understandable by visualizing their output. This method can be used to explain any model's prediction by computing the contribution of each feature.

A Shapley additive explanation (SHAP, Shapley Additive explanations) interprets the predictions of machine learning models. It is based on the concept of the Shapley value, which is a solution concept in cooperative game theory (Shapley, 1953).

A SHAP prediction is based on assigning a "Shapley value" to each feature, which represents the feature's contribution to the final prediction. In order to calculate the Shapley value, all possible subsets of features are compared, and the marginal contribution of each feature is determined (Li, 2022).

CHAPTER 4.

CASE STUDY

This chapter provides an overview of the research context and the survey employed in the research. It consists of a concise description of the geographical location and site selection of the pilot project of CCP (Cohabitation Cyclists et Piétons). This chapter discusses the conducted survey information, and suggested categories of the features taken from the survey by the author as well.

4.1 Survey type, Questionnaire and Categorization

The primary data source for this study is taken from the CCP project (Cohabitation Piétons et Cyclistes). The data obtained from CCP project is utilized to evaluate interaction satisfaction levels regarding the coexistence of pedestrians with "traditional" cyclists and various types of non-electric two-wheeled riders.

In the summer of 2021, the borough of Mont-Royal conducted two types of surveys: an on-street survey and a web-based survey. Both on-street and web-based surveys were conducted during Covid-19 infection issues when the population of Montreal was 4,247,000 (macrotrends, 2022). Data from respondents on the web and on the street were collected from a questionnaire. Pr. Marie-Soleil Cloutier (INRS), Pr. Owen Waygood (Polytechnique Montréal), and Francesco Ciari (Polytechnique Montréal) were primarily responsible for the development of the survey. The selected sites by the City for the pilot project are located in Mont Royal and Wellington. The research team conducted the survey, and the borough of Mont Royal and Wellington promoted the survey. The promotion of the survey was undertaken jointly by the boroughs, indicating a collaborative effort to engage the local community. The survey aimed to inquire about the preferences of the residents and identify the places they frequent the most, thereby contributing valuable data to the pilot project's objectives. PR Nicolas Saunier conducted complementary research using video analysis of the streets (INRS, 2021). The on-street survey schedule was conducted each week depending on the weather, but all sites were visited twice between 12 p.m. and 3 p.m. or 2 p.m. and 5 p.m. (3-hour block each time) between August 6 and August 19 (INRS, 2021)

The web-based questionnaire is made up of 4 sections including, Name of The Frequently Visited Streets (3 questions), Habitual Use of Pedestrianized Streets (41 questions), Perception

of Pedestrianization Projects (5 questions), and Socio-Demographic Information (9 questions) (INRS, 2021). The online survey was conducted between July and August 2021 on the Survey Monkey platform on social media focusing on the coexistence of pedestrians and cyclists in 12 streets of Montreal. The web-based survey received responses from approximately 1,490 participants, and these responses were considered for analysis. The on-street survey was not included in the analysis, as 4 questions were different, and the sample size were very small (INRS, 2021). Out of the 1,490 web-based survey respondents, 1,190 (which is almost 80% of all observations) provided complete responses (answered all the questions from start to finish) and were selected for data analysis. Thus, just the complete cases were selected for further analysis.

In the survey, participants were provided with the option to choose 'I prefer not to answer.' Since the primary objective of the research was to evaluate the satisfaction levels of pedestrians and cyclists, responses marked with this option were excluded from the analysis due to the lack of specific insights. Additionally, respondents who did not complete the survey were also filtered out, given the categorical nature of the dataset. After these exclusions, a total of 1190 respondents out of the initial 1490 were selected for further analysis, forming the basis for a more focused examination of the data and its implications for the objectives of the research. This process of excluding certain responses did not influence the overall sample size and did not potentially introduce bias into the analysis, as the Likert scale included a comprehensive range of options (very satisfied, rather satisfied, neither satisfied nor dissatisfied, dissatisfied, and very dissatisfied). Thus, respondents who chose "I prefer not to answer" were excluded, it seems that the available responses covered a spectrum of satisfaction levels. In this case, the exclusion of those responses could not introduce significant bias, given the richness of the Likert scale.

In the data preparation stage, incomplete cases were removed using the list-wise deletion method. This approach was chosen because the data type in this study is categorical and falls into different options/classes with unequal values and imputing missing data could bias the result and provided integrity.

The survey questionnaire, developed by the CCP research project team during the COVID-19 pandemic, consisted of 32 questions encompassing multiple-choice, yes/no, open-ended, and Likert scale questions.

Among the 32 questions, 16 were rated on a Likert scale ranging from 1 (very unfavorable) to 5 (very favorable). From this questionnaire, 44 variables are selected for further analysis based on the literature review. To improve comprehension of the research question and reduce complexity, these variables were categorized and named into 7 main categories by the author. These seven categories are: ‘*Trip Specific Information*’, ‘*Urban Design and Infrastructure*’, ‘*Socio-economic Characteristics*’, ‘*Season-Related Mode of Transport*’, ‘*Mutual Perception of Cyclists and Pedestrians*’, ‘*General Opinion and Attitude Toward Pedestrianization*’ and ‘*Favorite Pedestrian Street*’. This categorization is done to enhance the understanding and interpretation of the variables, allowing for a more organized and structured analysis. By grouping the variables into specific categories, it becomes easier to identify patterns, relationships, and trends between categories, thereby facilitating a comprehensive analysis of the data. The variables of this study were categorical, except for the household numbers, which were divided into four numerical categories: 1, 2, 3, and 4 or more. The web-based survey information summary is shown in Table 4.1 along with data type and proposed categories of variables.

Table 4.1: Information derived from the survey employed in this research.

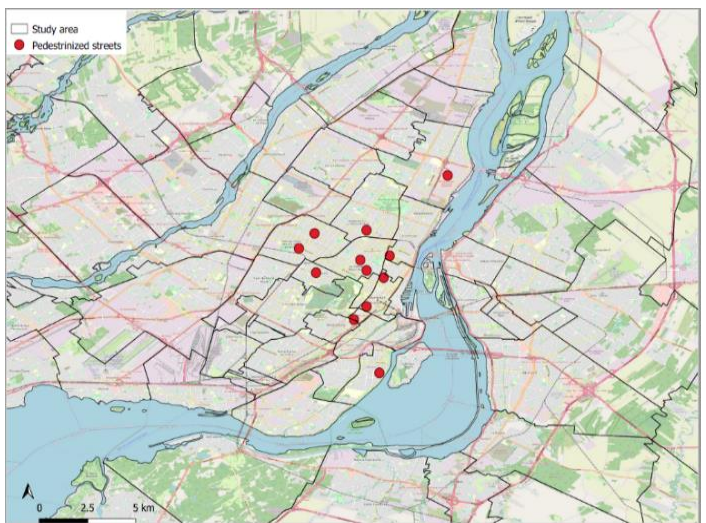
Study area: Great Montreal	
Study target: The co-existence of cyclists and pedestrians' interaction satisfaction assessment on pedestrianized streets	
Study period: Summer 2021	
Type of survey: Web-based	
Number of variables: 44	
Data type: Categorical and Numerical	
Suggested categories of variables: Favorite Pedestrian Street, Trip-specific Information, Mutual perception of cyclists and pedestrians, Season-Related Mode of Transport, Urban Design and Infrastructure, Socioeconomic Characteristics, General Opinion, and attitude toward pedestrianization	

Figure 4.1: Location of pedestrianized streets on Montreal map

4.2 Variables of the study

In this study, the primary focus is on assessing the interaction satisfaction of pedestrians and cyclists with their coexistence on pedestrian streets. Two dependent variables were taken from two questions of the CCP survey. These questions with their possible responses are as follows.

1- On the chosen pedestrian street, how satisfied are you with the coexistence of pedestrians and traditional cyclists? (Very Dissatisfied / Rather Dissatisfied / Rather Satisfied / Very Satisfied)

2- On the chosen pedestrian street, how satisfied are you with the cohabitation of pedestrians and other non-electric two-wheeled riders (skates, scooters, bikes, etc.)? (Very Dissatisfied / Rather Dissatisfied / Rather Satisfied / Very Satisfied)

The satisfaction levels were measured using two dependent or response variables taken from above questions, called Y1 and Y2. Y1 represents the satisfaction level with the coexistence of pedestrians and "traditional" cyclists, while Y2 represents the satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders. In both Y1 and Y2, the satisfaction levels are assessed using a categorical Likert scale consisting of four classes: Very Dissatisfied, Rather Dissatisfied, Rather Satisfied, and Very Satisfied. These variables were used as key indicators for evaluating the interaction satisfaction of pedestrians and cyclists on pedestrian streets.

Table 4.2 depicts the utilized independent variables taken from the web-based questionnaire and related categories for easily comprehension. In addition, the study's dataset underwent a categorical-to-numeric transformation, and the specific code assigned to each variable's potential option can be found beside each one.

To review the survey questions, refer to APPENDIX A.

Table 4.2: Variables taken from the survey, possible answers, and suggested categories (cont'd)

Variables taken from the survey and possible answers with numerical coded assigned in the analysis	Suggested Category
Streets Visiting time Before 9 a.m. (1) / Between 9 a.m. and 11 a.m. (2) / From 11 a.m. to 1 p.m. (3) / From 1 p.m. to 5 p.m. (4) / From 5 p.m. to 7 p.m. (5) / After 7 p.m. (6)	Trip-specific Information
Utilitarian walking During Covid (including Mobility Aid) Every day (5) / Few times in a week (4) / Few times in a month (3) / Few times in a year (2) / Never (1) / I prefer not to answer. (0)	
Cycling (personal, Bixi, electric or non) During Covid-19 Every day (5) / Few times in a week (4) / Few times in a month (3) / Few times in a year (2) / Never (1) / I prefer not to answer. (0)	
Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19 Every day (5) / Few times in a week (4) / Few times in a month (3) / Few times in a year (2) / Never (1) / I prefer not to answer. (0)	
Adaptive transport using during Covid-19 Every day (5) / Few times in a week (4) / Few times in a month (3) / Few times in a year (2) / Never (1) / I prefer not to answer. (0)	
Paratransit using (Uber, taxi, etc.) During Covid-19 Every day (5) / Few times in a week (4) / Few times in a month (3) / Few times in a year (2) / Never (1) / I prefer not to answer. (0)	
Car Sharing (Communauto) using During Covid-19 Every day (5) / Few times in a week (4) / Few times in a month (3) / Few times in a year (2) / Never (1) / I prefer not to answer. (0)	
Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19 Every day (5) / Few times in a week (4) / Few times in a month (3) / Few times in a year (2) / Never (1) / I prefer not to answer. (0)	

Table 4.2: Variables taken from the survey, possible answers, and suggested categories (cont'd)

Variables taken from the survey and possible answers with numerical coded assigned in the analysis	Suggested Category
Preferred mode of transport during summer without pedestrianization Walking (including Mobility Aids) (1) / Cycling (personal, Bixi, electric or non) (2) / Public transport (3) / Personal motor vehicles (automobile, truck, motorcycle, etc.) (4) / Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates) (5) / Adaptive transport (6) / Paratransit (Uber, taxi, etc.) (7) / Car sharing (Communauto) (8) / I do not like to pass pedestrian streets at this time (9)	Mutual behavior of cyclists and pedestrians
Preferred mode of transport during Winter without pedestrianization Walking (including Mobility Aids) (1) / Cycling (personal, Bixi, electric or non) (2) / Public transport (3) / Personal motor vehicles (automobile, truck, motorcycle, etc.) (4) / Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates) (5) / Adaptive transport (6) / Paratransit (Uber, taxi, etc.) (7) / Car sharing (Communauto) (8) / I do not like to pass pedestrian streets at this time (9)	
Behavior of cyclists toward pedestrians regarding speed unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Behavior of cyclists toward pedestrians regarding priority crossing unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Behavior of pedestrians toward cyclists regarding sharing the road unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Behavior of pedestrians toward cyclists regarding priority crossing unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Hangout more often on pedestrianized streets unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	General Opinion and Attitude Toward Pedestrianization
Hangout long time in pedestrianized streets unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Hangout more in shops on pedestrianized streets unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	

Table 4.2: Variables taken from the survey, possible answers, and suggested categories (cont'd)

Variables taken from the survey and possible answers with numerical coded assigned in the analysis	Suggested Category
Changing the route for passing pedestrianized streets unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	General Opinion and Attitude Toward Pedestrianization
Changing the route to avoid pedestrianized streets unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
General satisfaction of pedestrianization unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	General Opinion and Attitude Toward Pedestrianization
Continuation of pedestrianization for the future on commercial streets unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Extend pedestrianization into other streets unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Impact of pedestrianization on mobility unfavorable (1) / Rather unfavorable (2) / Neither favorable nor unfavorable (3) / Rather favorable (4) / Very favorable (5)	
Gender Female (1) / Male (2) / Other (3)	Socioeconomic Characteristics
Number of household members 1-member household (1) / 2-member household (2) / 3-member household (3) / 4-member household and more (4)	
Mobility restriction No (1) / Yes (2)	
Disability – Walking No (1) / Yes (2)	
Disability – Cycling No (1) / Yes (2)	
Bus Riding Disability No (1) / Yes (2)	

Table 4.2: Variables taken from the survey, possible answers, and suggested categories (cont'd)

Variables taken from the survey and possible answers with numerical coded assigned in the analysis	Suggested Category
Being a regular cyclist Yes (1) / No (2)	Socioeconomic Characteristics
Accessibility Transportation Options Driver's license ownership (1) / Car, truck and motorcycle ownership (2) / Transit pass ownership (3) / Bike ownership (4) / Carsharing users (Communauto) (5) / BIXI users (6) / Other mobility options (7)	
Motorized vehicle ownership Automobile (1) / Truck (2) / Motorcycle (3)	

CHAPTER 5.

RESULTS

This chapter illustrates the analytical outputs and discusses the results of the utilized models in this Study. It also provides an overview of the findings obtained from the analytical tests and machine learning (ML) methods employed in the study. The chapter discusses the accuracy of the results obtained through these techniques and presents the evaluation of their performance as well.

The result of the present study is divided into four steps. In the first step, descriptive data analysis by contingency tables (crosstab) for each variable is performed, followed by the results of Spearman's correlation coefficient and Variance Inflation Factor tests. In this step some important features related to *favorite Pedestrian Street*, the *socio-economic status*, and *trip specific information* of respondents are analyzed. The Second step is divided into two sub-sections of Lasso regression and MNL regression. In this step, first, by applying Lasso regression as a Feature Selection (FS) technique, the effective independent variables of the dataset are selected. Next, the important features taken from Lasso regression are used in MNL regression to estimate how dependent variable will be changed by increasing one unit of the independent variable. The Third step is divided into two sub-sections of XGB and RF machine learning methods. By deploying two ML classification methods (XGB, RF) each of independent variables' weight is calculated as a feature importance. Each ML method comes with their hyperparameters tuning results as well. The Fourth step is divided into two subsections of Confusion Matrix and Shapley Additive results for the purpose of performance evaluation. The accuracy of the models' prediction is based on the Confusion Matrix, F1 score, Sensitivity, precision and recall. Finally, for XGB and RF the contribution of important features for each dependent variable will be visualized by SHAPley additive.

In the final section of this chapter, a comparison and discussion of the results from XGBoost (XGB), Random Forest (RF), and Lasso regression models are presented. Specifically, this section analyzes the performance and findings of each model in relation to interaction satisfaction assessment with pedestrian streets.

5.1 Descriptive Data Analysis

As mentioned in Chapter3, complete-case data analysis was used by removing the cases with missing values for two reasons. Firstly, considering the categorical type of data, and secondly, handling missing values of categorical data may bias the results.

To have a comprehensive understanding of the dataset derived from a survey, Cross tabulation as a descriptive data analysis was performed for all independent variables. Cross tabulation, also known as contingency table or crosstab, is used to analyze and present the relationship between the two dependent variables and the independent variables of the study. Each dependent variable's values are shown in columns, and the independent variable's values are shown in rows. For each combination of variables, the table shows how many cases fit into each category.

The frequency tables of some of the variables used in this study are shown in this section. This included the frequencies of respondents' favorite pedestrian streets, their trip companions, gender, preferred visiting time, trip purpose, and preferred mode of transport (cycling or walking). The rest of the variables used in the study can be found in APPENDIX C

The abbreviations used in the graph are as follows: Very Dissatisfied (VD), Rather Dissatisfied (RD), Rather Satisfied (RS) and Very Satisfied (VS).

According to Table 5.1, it was found that among the respondents, 254 people out of 814 (almost 31%) chose Mont Royal as their favorite street, are 'Rather Satisfied' with the co-existence of pedestrians and traditional cyclists (Y1), while 21% of them (172 people) are 'Very Dissatisfied'. Furthermore, 299 people out of 814 respondents (36.7%) who chose Mont Royal as their favorite street, are 'Rather Satisfied' with the co-existence of pedestrians and non-electric 2-wheeled riders, while just 159 of them (19.5%) are 'Very Dissatisfied'.

In the case of Wellington, the second favorite street among respondents, 48 individuals out of 110 (35%) expressed being 'Very Satisfied' with the coexistence of pedestrians and traditional cyclists on this street (Y1). Additionally, 39 respondents (44%) who chose Wellington as their favorite street reported being 'Very Satisfied' with the coexistence of pedestrians and non-electric any type of two-wheeled riders (Y2).

Table 5.1: Frequency of people with different satisfaction levels who have chosen their favorite pedestrian street.

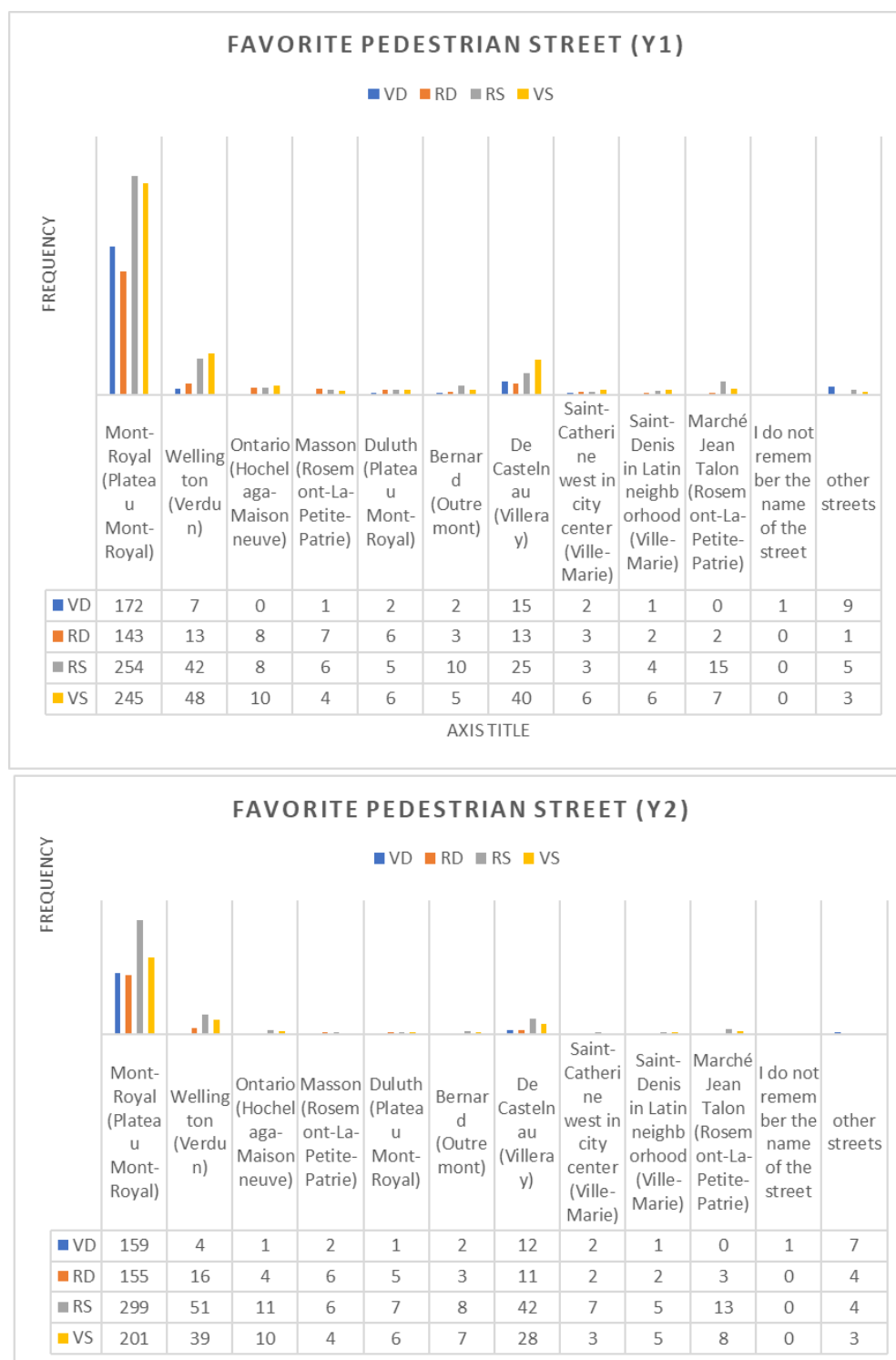
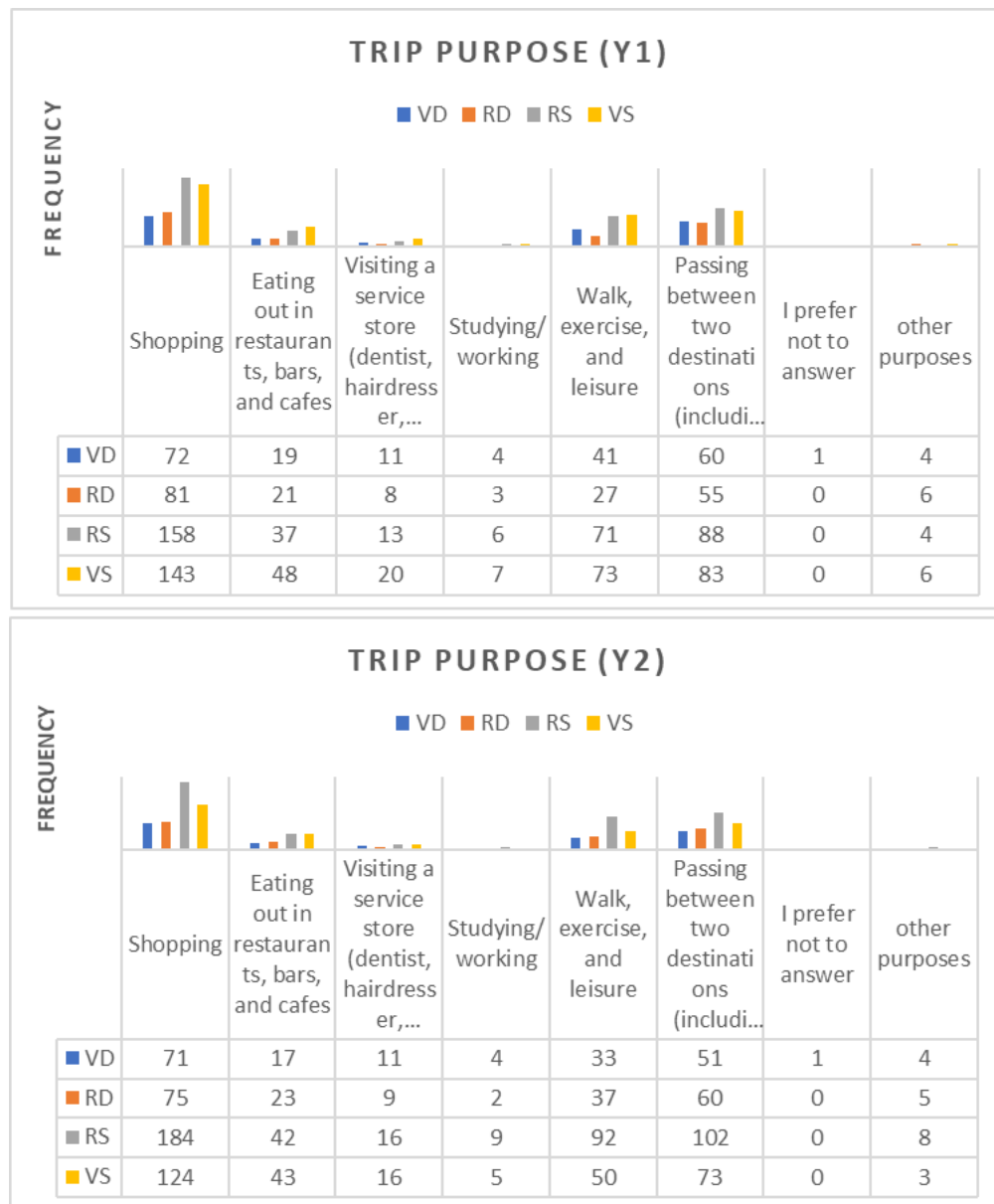


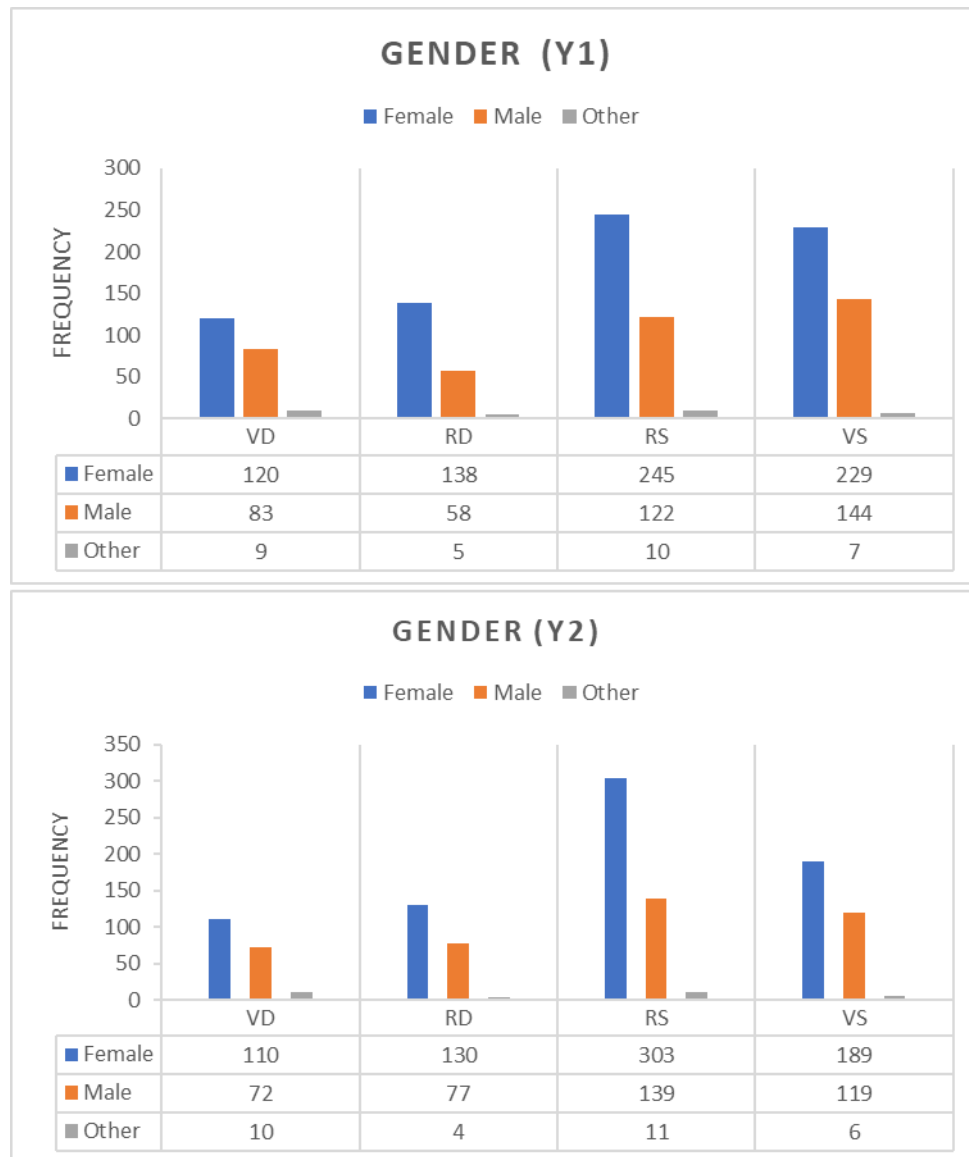
Table 5.2: Frequency of trip purpose of people with different satisfaction levels



As Table 5.2 shows, 158 respondents out of 454 people (34.8%) who go to pedestrianized street for shopping purposes are ‘Rather Satisfied’ with the co-existence of pedestrians and traditional cyclists (Y1), while almost 16% of them (72 people) are ‘Very Dissatisfied’. This table also shows that 48 respondents out of 125 people who go to pedestrianized street (38.4%) wanted to eat out in bar, restaurant and café are ‘Very Satisfied’, while 19 of them are ‘Very Dissatisfied’.

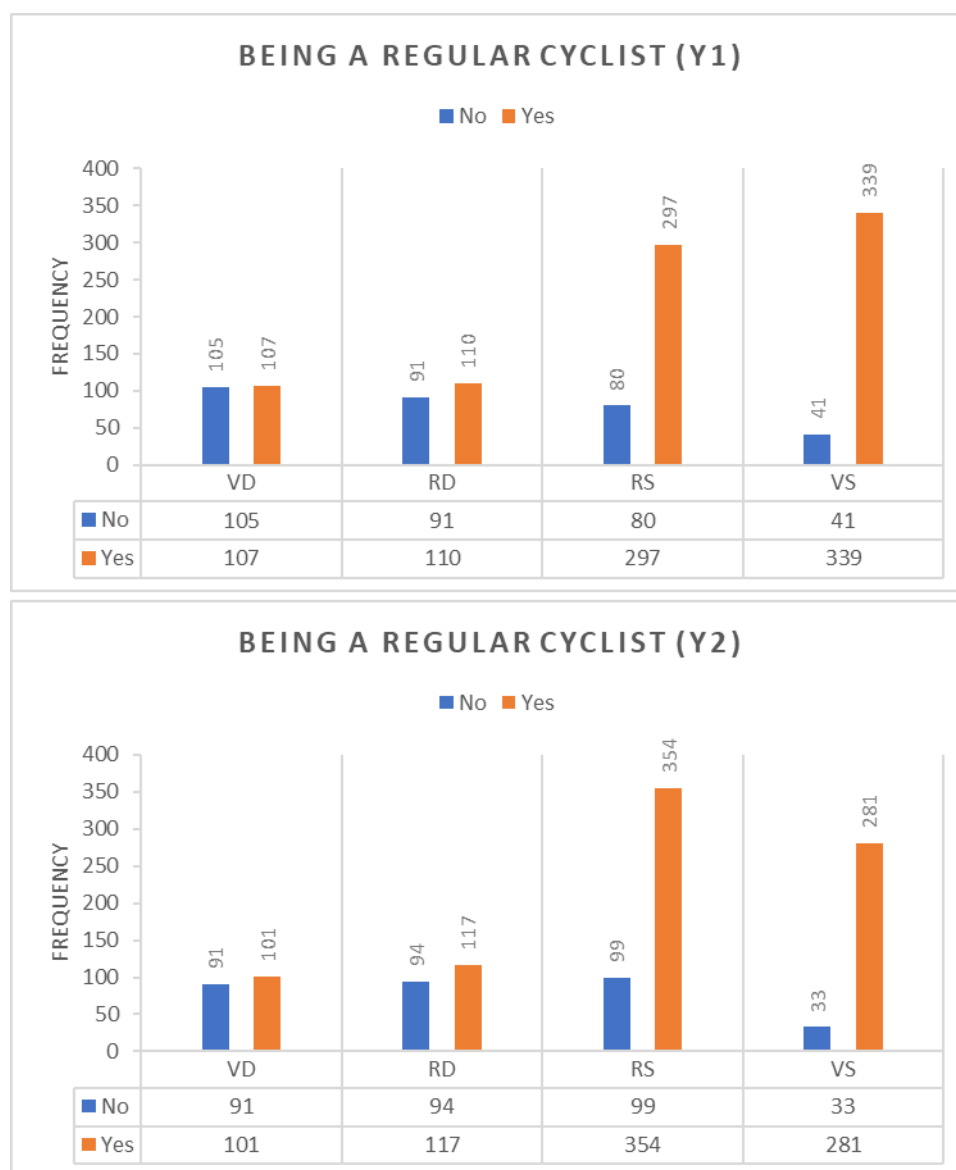
As Table 5.3 shows, 299 out of 733 female (40.8%), and 144 out of 407 male respondents (35.3%) are ‘Very Satisfied’, while 120 of female (16.4%), and 122 of male respondents (30%) are ‘Very Dissatisfied’ with the co-existence of pedestrians and traditional cyclists (Y1).

Table 5.3: Gender frequency of people with different satisfaction levels



As Table 5.4 shows, 339 out of 853 cyclists (almost 40%), are ‘Very Satisfied’ with the co-existence of pedestrians and traditional cyclists (Y1), while 107 of them (12.5%) are ‘Very Dissatisfied’.

Table 5.4: Frequency of Being a Regular cyclist and level of satisfaction.



As Table 5.5 shows, 126 out of 364 respondents (34.6%) visit pedestrian streets with family are ‘Very Satisfied’, while just 63 out of 364 of them (17.3%) are ‘Very Dissatisfied’ with the coexistence of pedestrians and traditional cyclists (Y1). This table also shows that 70 out of 366 respondents (19%) who visit pedestrian streets alone are ‘Very Dissatisfied’ with the coexistence of pedestrians and traditional cyclists (Y1). In the case of the coexistence of pedestrians and any type of non-electric two-wheeled riders (Y2), 143 out 366 respondents who visit those streets alone are ‘Very Dissatisfied’.

Table 5.5: Frequency of Trip Companionship

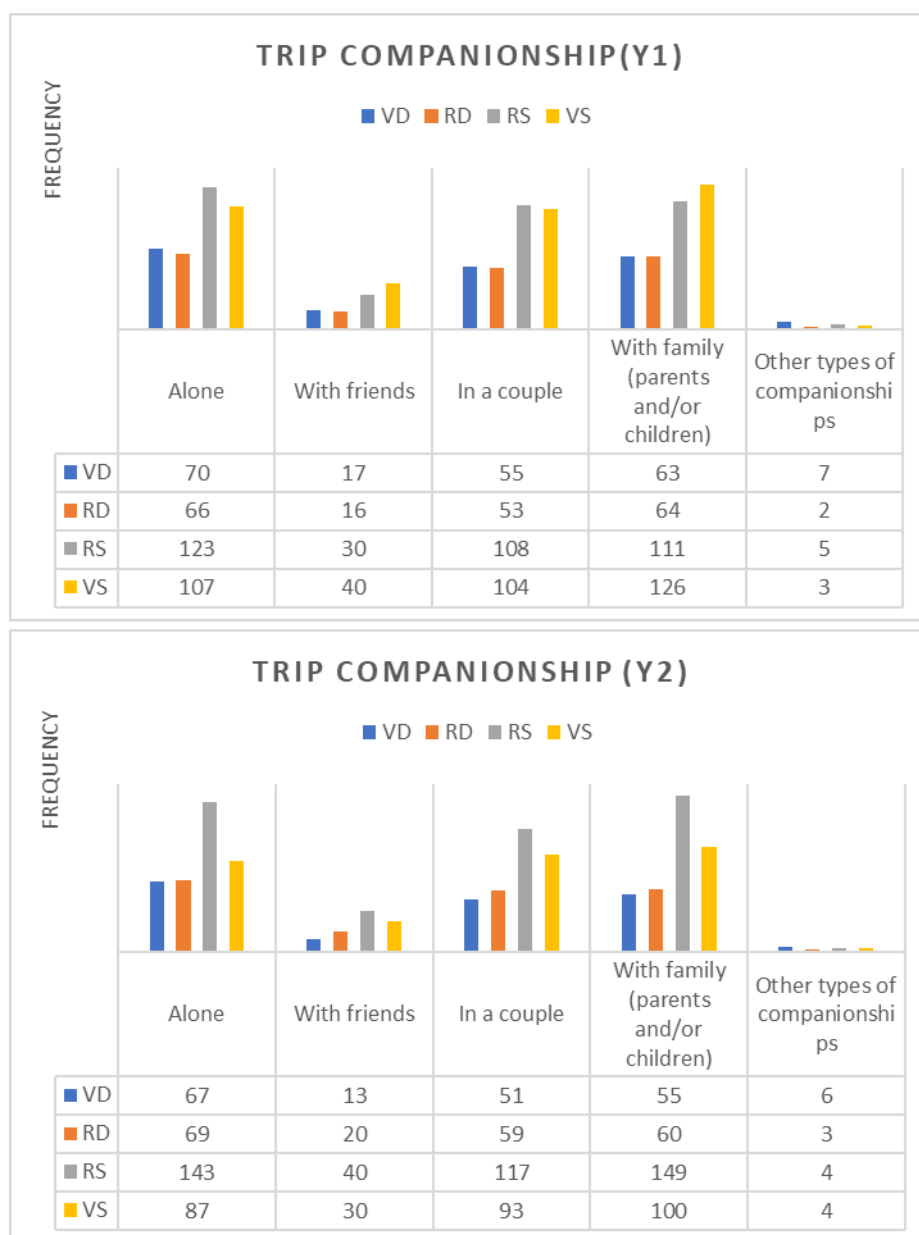
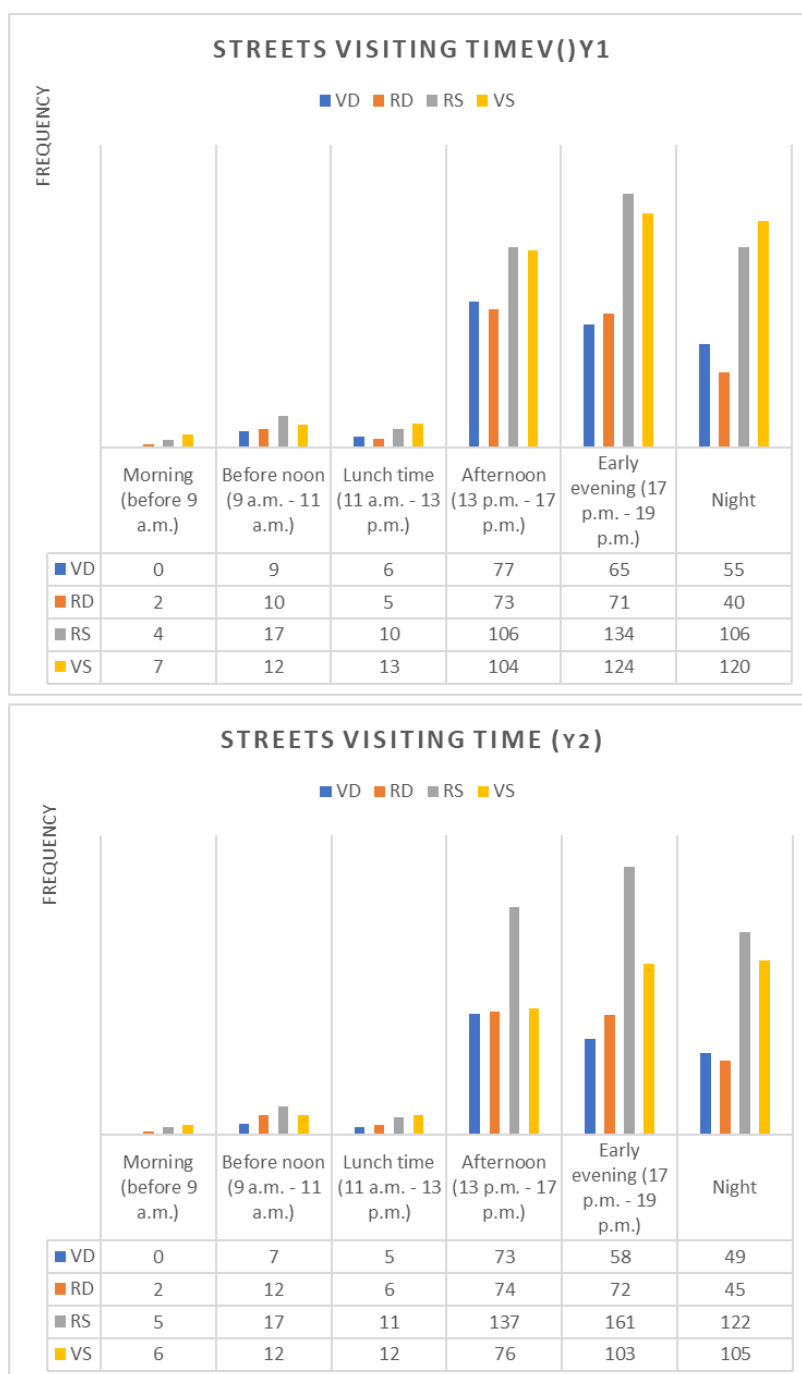


Table 5.6: Frequency of Street Visiting Time

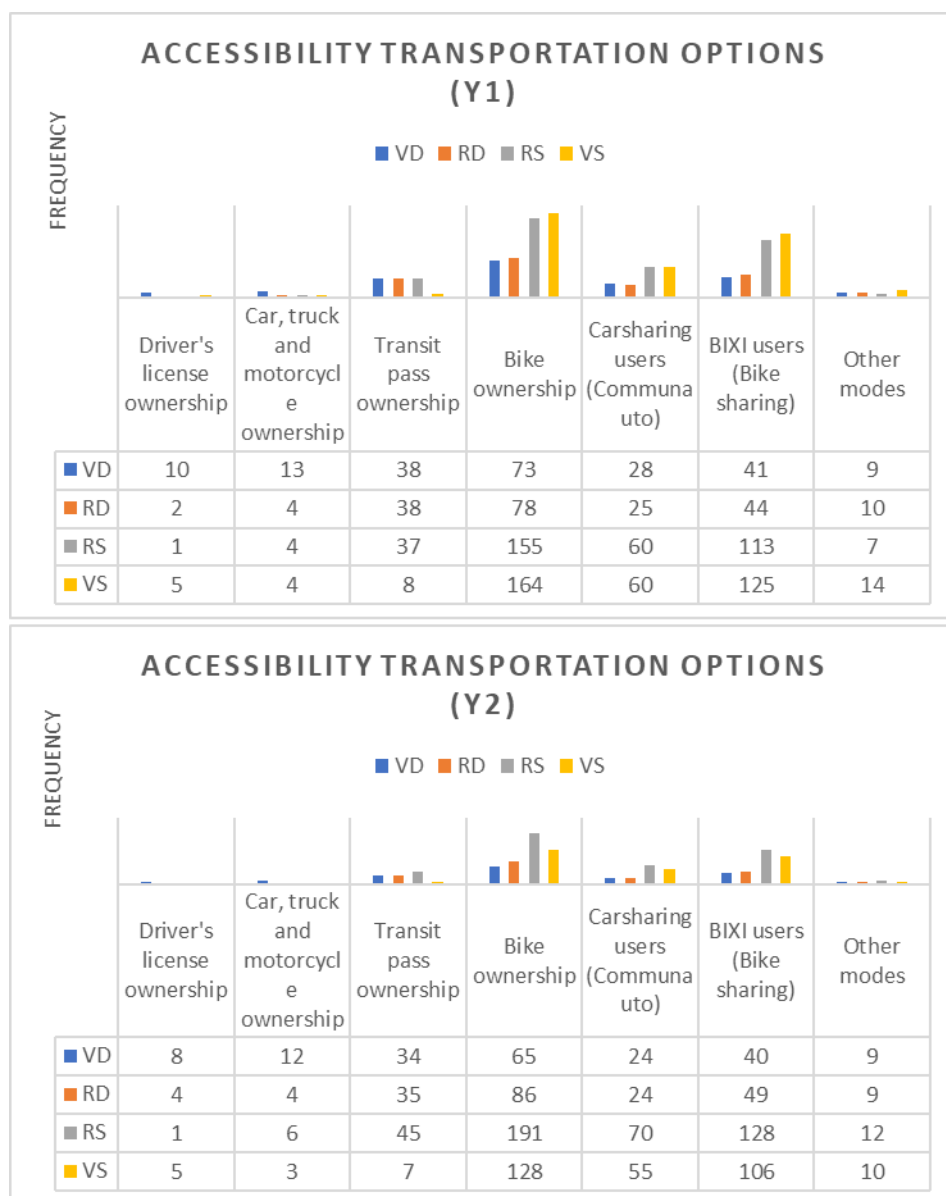


As Table 5.6 indicates, 134 out of 394 respondents (34%) visit the pedestrian street in early evening (17 p.m. - 19 p.m.) are 'Rather Satisfied', while just 65 of them (16.5%) are 'Very Dissatisfied' with the co-existence of pedestrians and traditional cyclists (Y1). In the case of the coexistence of pedestrians and any type of non-electric two-wheeled riders (Y2), 161 out of 394

respondents (almost 41%) are ‘Rather Satisfied’ while visit the pedestrian street in early evening (17 p.m. - 19 p.m.)

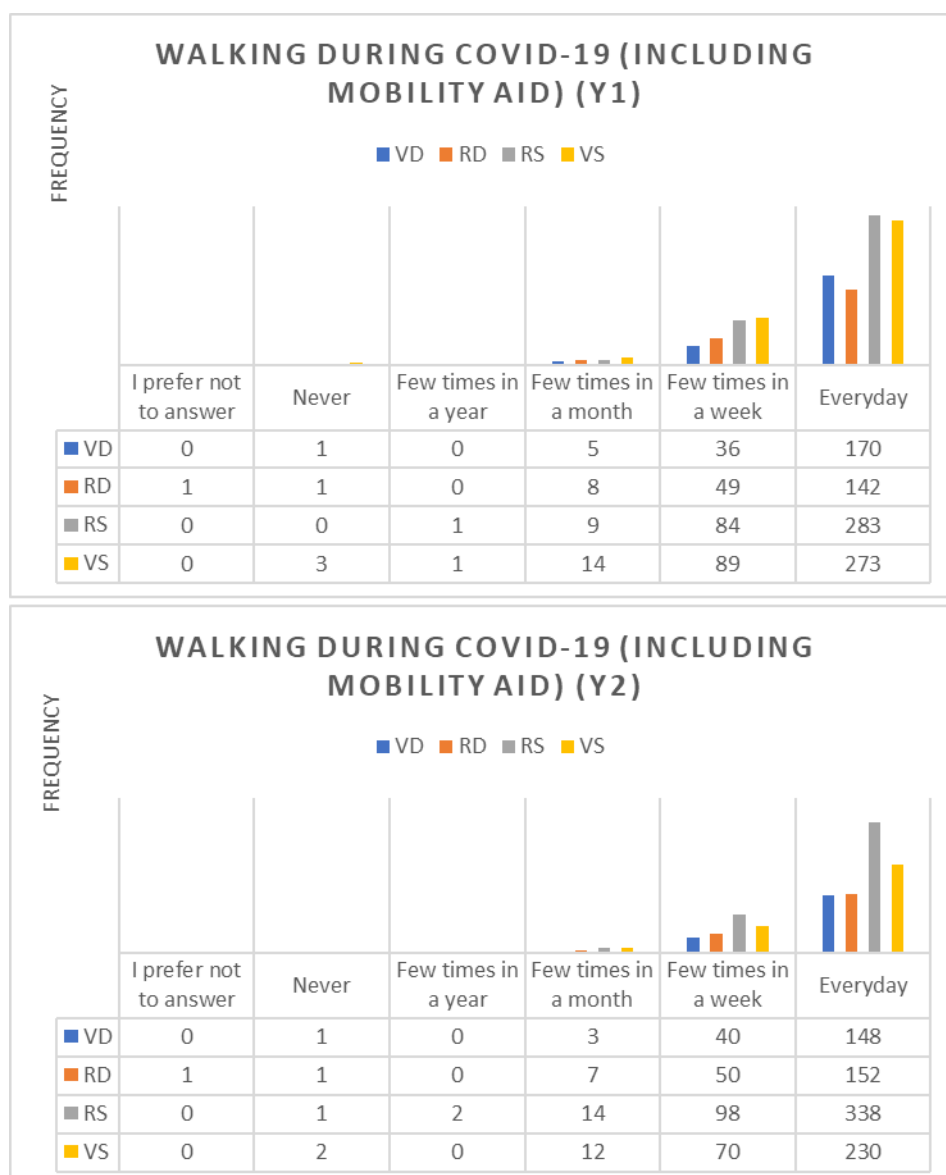
As Table 5.7 shows, 164 out of 470 (almost 35%) bike owners are ‘Very Satisfied’, while 41 of them (8.7%) are ‘Very Dissatisfied’ with the co-existence of pedestrians and traditional cyclists (Y1). In the case of the coexistence of pedestrians and any type of non-electric two-wheeled riders (Y2), 191 out of 470 bike owners (40%) are ‘Rather Satisfied.’

Table 5.7: Frequency of Accessibility Transportation Options



As Table 5.8 indicates, 273 out of 868 respondents (31.5%) who walk everyday during the survey time (Covid-19), are ‘Very Satisfied’ with the co-existence of pedestrians and traditional cyclists (Y1). In the case of the coexistence of pedestrians and any type of two-wheeled riders, 338 out of 868 respondents (almost 40%) who walk every day during the survey time (Covid-19), are ‘Rather Satisfied’.

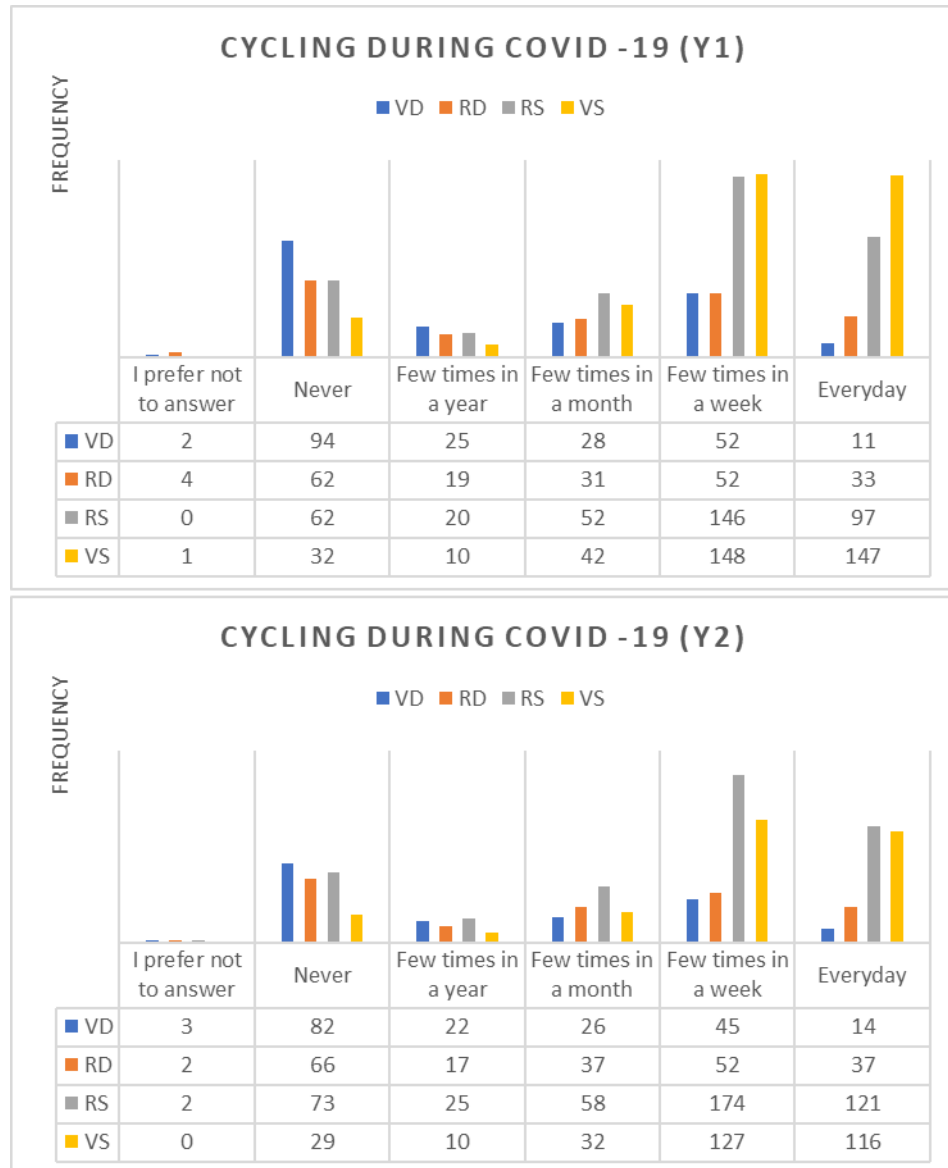
Table 5.8: Frequency of Walking during Covid-19



As Table 5.9 indicates 148 out of 398 (37%) respondents ride a bicycle at least few times in a week during Covid-19 are ‘Very Satisfied’ with the co-existence of pedestrians and traditional cyclists (Y1). In the case of the coexistence of pedestrians and any type of non-electric two-

wheeled riders (Y2), 174 out of 398 (almost 44%) respondents ride a bicycle at least a few times in a week during Covid-19 are ‘Rather Satisfied’.

Table 5.9: Frequency of Cycling during Covid-19



5.2 Statistical tests' result

Due to the categorical nature of the dependent variables in this study, a Chi-squared test was run on the dataset. The results show several cells less than the expected frequency of 5. Considering that the value of each cell in contingency table should be at least 5, the result is not reliable; therefore, the Spearman's correlation test is performed instead. The results of Chi-squared can be found in APPENDIX D.

Spearman's coefficient correlation was run to determine the degree of association with the dependent and independent variables. The variables who are strongly correlated with the satisfaction with the coexistence of pedestrians, cyclists are selected and are shown in Table 5.10. This table shows there is a strong correlation between '*Behavior of cyclists toward pedestrians regarding speed*', '*Behavior of cyclists toward pedestrians regarding priority crossing*', '*Personal Security*' and '*Hangout more often on pedestrianized streets*', and the satisfaction level with the coexistence of pedestrians and "traditional" cyclists (Y1), and the satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders (Y2). Meaning that the mentioned independent variables have a positive impact on the satisfaction of existence of pedestrians and cyclists in pedestrian streets. The result of Spearman's test suggests that factors related to mutual behavior of cyclists and pedestrians and personal security issues have the strongest positive impact on satisfaction levels while studying coexistence of cyclists and pedestrians in pedestrian streets.

The whole table of Spearman's rank coefficient correlation result is shown in APPENDIX B. The result shows that a positive relationship exists between 33 predictors and 11 negative associations with the response variable.

Table 5.10: Spearman correlation result for highly correlated variables

Variables	Satisfaction level with the coexistence of pedestrians and "traditional" cyclists(Y1)	Variables	Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders (Y2)
Behavior of cyclists toward pedestrians regarding speed	0.76	Behavior of cyclists toward pedestrians regarding speed	0.75
Behavior of cyclists toward pedestrians regarding priority crossing	0.74	Behavior of cyclists toward pedestrians regarding priority crossing	0.73
Hangout more often on pedestrianized streets	0.46	Personal Security	0.45
Personal Security	0.45	Hangout more often on pedestrianized streets	0.43

Table 5.10: Spearman correlation result for highly correlated variables (cont'd)

Variables	Satisfaction level with the coexistence of pedestrians and "traditional" cyclists(Y1)	Variables	Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders (Y2)
Impact of pedestrianization on mobility	0.44	The behavior of pedestrians toward cyclists regarding sharing the road	0.42
The behavior of pedestrians toward cyclists regarding sharing the road	0.44	Impact of pedestrianization on mobility	0.42

5.2.1 Variance Inflation Factor

As it was stated in Chapter 3, prior to running regression models, it is recommended to assess multicollinearity. The VIF test was run in Python by using Variance Inflation Factor function from Stats models library to examine the multicollinearity among the variables. According to the VIF test results in Table 5.11, ‘The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets’ and ‘The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets’ exhibit high correlation, as indicated by their identical VIF scores of 14.17, and 13.22 respectively. This suggests a strong multicollinearity between these two variables while assessing the satisfaction with the co-existence of pedestrian and cyclists in pedestrian streets. The result, shown in the mentioned table, also shows 5 variables are moderately correlated which means the presence of multicollinearity among predictor variables (Singh et.al, 2021).

There is no statistical method for "fixing" multicollinearity (Dormann et al., 2013). Nevertheless, some methods have been developed to mitigate the effects of multicollinearity. Perhaps the most effective method for dealing with multicollinearity is to analyze the relationships between predictors beforehand. This is to eliminate or shrink predictors that have known correlations such as Lasso regression which outperforms other methods (Altalbany, 2021) . The result of VIF test for all predictor variables is in APPENDIX E.

Table 5.11: Moderately to highly correlated predictor variables of the model

Selected Independent Variables for Y1 and Y2 (Y1: Satisfaction level with the coexistence of pedestrians and "traditional" cyclists), (Y2: Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	VIF
The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets	14.17
The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets	13.22
Being a regular cyclist	7.25
Hangout more often on pedestrianized streets	7.21
Continuation of pedestrianization for the future on commercial streets	6.72
Hangout long time in pedestrianized streets	6.48
General satisfaction of pedestrianization in pedestrianized streets	5.59

5.3 Lasso regression results

Since the VIF test indicates multicollinearity among variables, it is difficult to decide which features should be selected over the two variables that are highly correlated in order to mitigate the multicollinearity. To simplify the model, Lasso regression is used as the first method of ML with the aim of feature selection of influential variables on the satisfaction level of cyclists and pedestrians in pedestrian streets. Lasso is able to pick one variable over other variables that are highly correlated.

By using Lasso regression, the weight of important variables was determined, and the least important features were removed. The process includes the selection of the covariates that have the most correlation with the response variable.

The best value of alpha after tuning the model for Y1 and Y2 are 0.032, and 0.035 respectively.

In Table 5.12, Lasso selected 17 predictor variables for Y1, and 14 variables for Y2. In APPENDIX F, the coefficient of all selected features by Lasso regression for both dependent variables are shown. The importance of feature for each predictor variable is the absolute value of its coefficient, and -/+ sign indicates negatively/positively related with the dependent variable.

The results for Lasso regression in Table 5.12 indicates that '*Behavior of cyclists toward pedestrians regarding speed*', '*Behavior of cyclists toward pedestrians regarding priority crossing*' and '*Intersection Safety*' have the highest selection probability and are positively related

to the satisfaction of coexistence of pedestrian, cyclists, or any type of two wheelers. It shows that a cyclist's behavior and safe intersections can affect the satisfaction with the coexistence of pedestrians, cyclists, or any type of two-wheelers.

'Trip purpose' and 'Changing the route to avoid pedestrianized streets' have negatively related to the satisfaction of coexistence of pedestrian, cyclists, or any type of two wheelers. This implies that if people are traveling for specific purposes or actively avoid pedestrianized streets, it may negatively affect the satisfaction of coexistence with pedestrians, cyclists, or any type of non-electric two-wheelers. The whole result of this part is in Error! Reference source not found.

Table 5.12: Lasso best coefficient for both dependent variables

Selected variables by Lasso	Lasso best coefficient (Y1: Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)	Lasso best coefficient (Y2: Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)
Behavior of cyclists toward pedestrians regarding speed	0.516	0.852
Behavior of cyclists toward pedestrians regarding priority crossing	0.119	0.177
Intersection Safety	0.099	0.139
Impact of pedestrianization on mobility	0.068	0.088
Attractiveness	0.062	0.087
Changing the route for passing pedestrianized streets	0.059	0.085
Universal Accessibility'	0.052	0.004
Preferred mode of transport during summer with pedestrianization	0.042	0.03
Favorite Pedestrian Streets	0.022	0.002
The behavior of pedestrians toward cyclists regarding sharing the road	0.012	0.005
Trip purpose	-0.016	-0.001

Table 5.12: Lasso best coefficient for both dependent variables (cont'd)

Selected variables by Lasso	Lasso best coefficient (Y1: Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)	Lasso best coefficient (Y2: Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)
Changing the route to avoid pedestrianized streets	-0.026	-0.024

5.3.1 Performance evaluation results for Lasso regression

Table 5.13 is the performance result of Lasso regression. The accuracy for testing and training datasets shows better performance, indicating that the model performs better than before optimization. From the table, it can be seen that the model's accuracy improved slightly on the Y1 dataset after tuning, but its accuracy decreased on the Y2 dataset. However, the model's accuracy on the testing set improved significantly after tuning for both Y1 and Y2 datasets.

A comparison of R-squared before and after tuning shows improved performance, and both values fall acceptable limits for social science research (Peterson K. Ozili, 2023). The R-squared values for both datasets are quite good, indicating that the model can explain a significant portion of the variance in the dependent variable. The R-squared value is higher on the testing set after tuning, indicating that the model generalizes well on unseen data.

The Mean Absolute Error (MAE) is relatively low, indicating that the model's predictions (Peterson K. Ozili, 2023) are close to the actual values on average. MAPE value is quite low, indicating that the model's predictions are accurate for the majority of instances. Based on Lasso's MAPE value, regression prediction is reasonable because below 50% falls within the range of reasonable forecasting (Montaño Moreno et al., 2013). Regarding interpreting the MAPE value, a lower value means a smaller percentage difference between the predicted values and the actual values. In this case, the MAPE value is below 50%, which suggests that the regression model's predictions have a reasonable level of accuracy. This means that, on average, the predicted values are within a 50% margin of error compared to the actual values.

Table 5.13: Lasso regression evaluation metrics for both dependent variables

Evaluation metrics for Lasso	Y1	Y2
Training Accuracy Before tuning	0.680	0.618
Testing Accuracy Before tuning	0.740	0.683
Training Accuracy After tuning	0.672	0.642
Testing Accuracy After tuning	0.746	0.744
R ² without splitting the dataset into test and train	0.695	0.673
R ² training set	0.672	0.641
R ² test set	0.746	0.744
MSE	0.630	0.578
RMSE	0.894	0.760
MAE	0.491	0.509
MAPE value	44.02 %	42.68 %
Overall Accuracy of Lasso Regression	56%	57%

5.4 Multinomial Logistic (MNL) Regression results

Multinomial Logistic regression is the second supervised ML technique deployed in this research. Since there are 44 predictors in the model, only the predictors which were selected by Lasso regression for both Y1 and Y2 were entered into the model. This part determines how selected independent variables taken from Lasso regression predict the satisfaction of the co-existence of pedestrians and cyclists on pedestrianized streets.

In MNL regression of the present study, the reference category is ‘Very Dissatisfied’, and all other classes are compared with the reference category. Selecting this class as a reference category allows for a clearer interpretation of how satisfaction levels change from the lowest to the highest, enabling easier interpretation of the results. Table 5.14, is the parameter estimation of the MNL for top 5 important predictors selected by Lasso regression, and ‘ β ’ is the regression coefficient of independent variables.

Table 5.14: MN Logit Regression Results using Very Dissatisfied as the reference variable.

MN Logistic Regression Results							
Predictors of the model		Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)			Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)		
		β	Std. Error	p-value	β	Std. Error	p-value
Rather Dissatisfied	Intercept	-2.521	0.874	0.004	-1.963	0.810	0.015
	Behavior of cyclists toward pedestrians regarding speed	1.399	0.525	0.008	0.888	0.436	0.042
	Behavior of cyclists toward pedestrians regarding priority crossing	-0.147	0.455	0.747	0.010	0.436	0.982
	Intersection Safety	-0.057	0.187	0.761	0.308	0.186	0.097
	Impact of pedestrianization on mobility	0.149	0.214	0.487	0.131	0.212	0.537
	Attractiveness	0.295	0.246	0.229	-0.012	0.237	0.960
Rather Satisfied	Intercept	-10.627	2.056	0.000	-6.491	1.379	0.000
	Behavior of cyclists toward pedestrians regarding speed	2.649	0.590	0.000	1.497	0.451	0.001
	Behavior of cyclists toward pedestrians regarding priority crossing	0.002	0.504	0.996	0.543	0.447	0.224
	Intersection Safety	-0.115	0.292	0.694	0.265	0.250	0.289
	Impact of pedestrianization on mobility	1.005	0.406	0.013	0.416	0.311	0.180
	Attractiveness	0.982	0.401	0.014	0.698	0.336	0.038

Table 5.14: MN Logit Regression Results using Very Dissatisfied as the reference variable (cont'd)

MN Logistic Regression Results							
Predictors of the model		Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)			Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)		
		β	Std. Error	p-value	β	Std. Error	p-value
Very Satisfied	Intercept	-20.548	3.015	0.000	-18.932	2.891	0.000
	Behavior of cyclists toward pedestrians regarding speed	3.380	0.661	0.000	2.548	0.577	0.000
	Behavior of cyclists toward pedestrians regarding priority crossing	0.512	0.552	0.353	0.936	0.519	0.072
	Intersection Safety	0.800	0.395	0.043	1.196	0.386	0.002
	Impact of pedestrianization on mobility	1.773	0.502	0.000	1.468	0.475	0.002
	Attractiveness	1.332	0.526	0.011	0.863	0.497	0.083
	The reference category is 'Very Dissatisfied'.						

The following paragraphs are going to discuss the five important predictors for two dependent variables of Y1 and Y2.

- ***The "Behavior of cyclists towards pedestrians regarding speed"***

The "Behavior of cyclists towards pedestrians regarding speed" is statistically significant and positively associated with the satisfaction level with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians and any type of two-wheelers (Y2) across all dependent variable classes ('Rather Dissatisfied', 'Rather Satisfied' and 'Very Satisfied'). It means that the probability of being 'Rather Dissatisfied', 'Rather Satisfied' and 'Very Satisfied' with the satisfaction level with the coexistence of pedestrians and traditional cyclists (Y1) and

with the coexistence of pedestrians and any type of two-wheelers (Y2) increase in comparison to people who are "Very Dissatisfied". This implies that improving the behavior of cyclists, especially in terms of speed when interacting with pedestrians, is likely to lead to higher satisfaction levels with the coexistence of pedestrian and cyclists among respondents in this context. Furthermore, this predictor has the strongest effect observed in the 'Very Satisfied' class in Y1 as the ' β ' value is 3.380.

- ***The "Behavior of cyclists toward pedestrians regarding priority crossing"***

This independent variable is not statistically significant for any of the satisfaction level classes ('Rather Dissatisfied,' 'Rather Satisfied,' and 'Very Satisfied') in both Y1 and Y2. The p-values for all classes are greater than 0.05, indicating that this predictor does not have a significant effect on satisfaction levels with the coexistence of pedestrians and cyclists in this context. This suggests that there is insufficient evidence to conclude that this predictor has a significant impact on the satisfaction level with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians and any type of two-wheelers (Y2).

- ***The "Intersection Safety"***

The predictor 'Intersection Safety' is not statistically significant for the classes 'Rather Dissatisfied' and 'Rather Satisfied,' for both Y1 and Y2. Their p-value indicates no significant association between this predictor and the satisfaction level with the coexistence of pedestrians and cyclists among those who are 'Rather Satisfied' and 'Rather Dissatisfied'. In the 'Very Satisfied' class in Y1 and Y2, "Intersection Safety" is statistically significant, and positively associated with the satisfaction level with the coexistence of pedestrians and cyclists (as the ' β ' value is 0.800 and 1.196 for Y1 and Y2 respectively). It means that the probability of being 'Very Satisfied' with the satisfaction level with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians and any type of two-wheelers (Y2) increase compared to those who are "Very Dissatisfied". In other words, individuals who have a more favorable view of intersection safety compared to those who are very dissatisfied are more likely to be highly satisfied with the coexistence of pedestrians and cyclists.

- ***The "Impact of pedestrianization on mobility"***

The predictor 'Impact of pedestrianization on mobility' is not statistically significant for the 'Rather Dissatisfied' class in Y1 and Y2. This suggests that this predictor has no significant impact

on the satisfaction levels with the coexistence of pedestrians and cyclists for individuals who are 'Rather Dissatisfied'. In the 'Rather Satisfied' class, this predictor is statistically significant in Y1, and positively associated with the satisfaction level with the coexistence of pedestrians and cyclists (p -value of 0.013 and a ' β ' value of 1.005) but is not statistically significant in Y2. This indicates no significant association between the "Impact of pedestrianization on mobility" and the satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders compared to those who are 'Very Dissatisfied'. In the 'Very Satisfied' class in Y1 and Y2, this predictor is statistically significant, and positively associated with the satisfaction level with the coexistence of pedestrians and cyclists (' β ' value is 1.773 and 1.468 for Y1 and Y2 respectively). This implies a strong positive association between this predictor and the satisfaction levels with the coexistence of pedestrians and cyclists among those who are 'Very Satisfied' compared those who are 'Very Dissatisfied'. In other words, improving the impact of pedestrianization on mobility of pedestrian streets positively impact the satisfaction of pedestrians and cyclists coexisting in the same environment.

- ***The "Attractiveness"***

The predictor 'Attractiveness' is not statistically significant for those who are 'Rather Dissatisfied' for both Y1 and Y2, and 'Very Satisfied' for Y2. This indicates that this predictor has no significant impact on the satisfaction levels with the coexistence of pedestrians and cyclists for individuals who are 'Rather Dissatisfied', and 'Very Satisfied'. In the 'Rather Satisfied' class, this predictor is statistically significant for both Y1 and Y2, and positively associated with the satisfaction level with the coexistence of pedestrians and cyclists (' β ' value is 0.982 and 0.698 for Y1 and Y2 respectively). This indicates that the possibility of being 'Rather Satisfied' with the attractiveness of the pedestrian streets in comparison to the reference category will increase and have a direct relationship with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians and any type of two-wheelers (Y2). This suggests that when pedestrian streets are perceived as attractive, people are more likely to be rather satisfied with how pedestrians and cyclists interact in those spaces.

5.4.1 McFadden's pseudo-R-squared

The value of McFadden's pseudo-R-squared for Y1 and Y2, shows the model fits the data. For Y1, the McFadden's pseudo-R-squared value of 0.43 suggests that the model explains

approximately 43% of the variation in satisfaction with the co-existence of pedestrians and "traditional" cyclists. This indicates a good fit, meaning that the model captures a significant portion of the factors influencing satisfaction in this context.

Similarly, for Y2, the McFadden's pseudo-R-squared value of 0.39 indicates that the model explains about 39% of the variation in satisfaction with the co-existence of pedestrians and non-electric two-wheeled riders (such as skates, scooters, bikes, etc.). This value also suggests a good fit, indicating that the model captures a substantial portion of the factors affecting satisfaction in this scenario.

5.4.2 Classification evaluation performance for MNL

In this section, the classification performance of the Multinomial Logit (MNL) model for predicting satisfaction levels with the co-existence of pedestrians and different types of cyclists is evaluated. Various metrics such as precision, recall, F1-score, and the confusion matrix are analyzed to assess the model's performance in classifying different satisfaction classes (Very Satisfied, Rather Satisfied, Rather Dissatisfied, Very Dissatisfied).

From Table 5.15, it is observed that the model has the highest precision percentage for the 'Very Dissatisfied' class, where 70% and 77% for Y1 and Y2, respectively. Meaning that these percentages were correctly predicted out of the selected samples of the test dataset. Based on Table 5.15, the precision percentages for 'Rather Dissatisfied' for Y1 is relatively low (53%) indicating that the model has more difficulty in accurately predicting this class.

The model has higher recall scores for 'Very Satisfied' and 'Very Dissatisfied' classes in Y1, 92% and 76% respectively. The recall scores for the 'Rather Satisfied' and 'Rather Dissatisfied' classes are relatively low in Y1, indicating that the model has more difficulty for identifying these classes correctly. But the model correctly identifies 92% of the actual samples in 'Very Satisfied' class. The lowest recall value is for the 'Rather Satisfied' class at 31% and 29% for Y1 and Y2 respectively, indicating that the model is struggling to identify these classes.

F1-score of Y1 and Y2 in 'Very Satisfied' class is 76% and 71% respectively, meaning that the model made more than 70% of all predictions of this class correctly for the entire test dataset. Furthermore, the 'Rather Satisfied' class has the lowest F1-score of just 42% and 39% for Y1 and Y2 respectively, which indicates that the model did not perform well in predicting these classes.

Based on the diagonal elements of the Confusion Matrix (CM) (Figure 5.1, Figure 5.2), it can be concluded that out of the 234 test samples, 148 samples for Y1 and 150 samples for Y2 (sum of diagonal elements) were correctly predicted, resulting in an accuracy of 63% and 64%, respectively. In confusion CM maps, the diagonal elements show the well predicted instances for each class. The values in each row except the diagonal values, show the level of misclassification for each class prediction. Consequently, the matrix also indicates that the 'Rather Satisfied' class has the highest misclassification among the other classes; this means that the model struggles to predict people who were 'Rather Satisfied' with the coexistence of pedestrians and cyclists.

Table 5.15: Classification evaluation performance for MNL

Model performance score for MNL regression	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)				Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)			
	precision	recall	f1-score	support	precision	recall	f1-score	support
Very Satisfied	0.65	0.92	0.76	48	0.6	0.87	0.71	47
Rather Satisfied	0.64	0.31	0.42	51	0.61	0.29	0.39	48
Rather Dissatisfied	0.53	0.52	0.53	63	0.61	0.72	0.66	82
Very Dissatisfied	0.70	0.76	0.73	72	0.77	0.63	0.69	57
Accuracy	0.63				0.64			

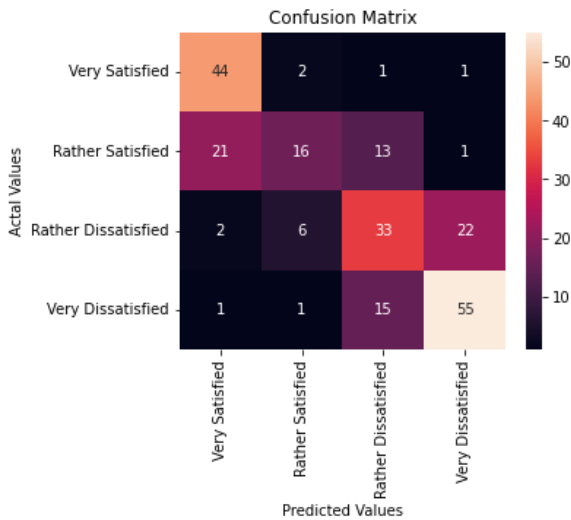


Figure 5.1: Confusion matrix for MNL_Y1

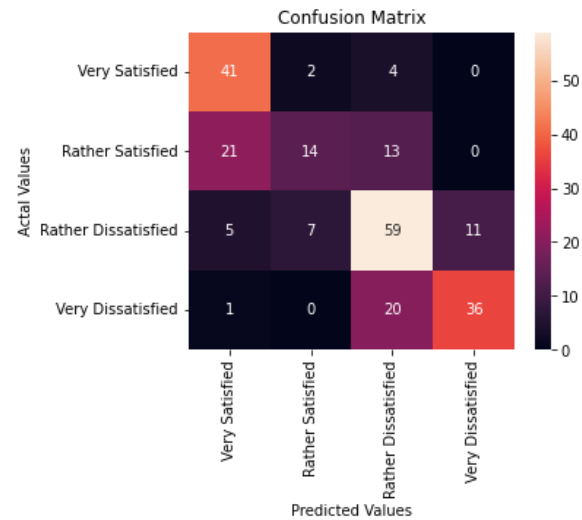


Figure 5.2: Confusion matrix for MNL_Y2

5.5 XGBoost results

XGBoost was introduced in chapter 3 of this thesis. XGBoost, as a first classification technique, was performed on the data in order to determine the weight of each feature and rank the importance of the features. Through XGBoost it is possible to determine how each independent variable contributes to predication of dependent variables.

The best hyperparameters of the model after tuning are shown in APPENDIX G. To evaluate the predictions made by the model, the confusion matrix function from scikit-learn library in Python were used, which creates a table that shows the number of true positives, true negatives, false positives, and false negatives. Finally, the output of the model is visualized by creating a bar plot for each value of the response variable. This plot can help to understand how well the model is able to predict each class of dependent variable in the dataset. The Table 5.16 is the output from XGBoost feature importance.

The feature importance plots derived from XGBoost are in APPENDIX H.

Table 5.16: Feature Importance of each independent variable in XGB

Feature Description	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)		Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	
	Weight (Gain)	Percentage	Weight (Gain)	Percentage
Behavior of cyclists toward pedestrians regarding speed	42.75	10.51%	39.88	10.20%
Behavior of cyclists toward pedestrians regarding priority crossing	33.17	8.16%	30.32	7.75%
The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets	14.78	3.64%	21.2	5.42%
Personal Security (fear of strangers, lack of lighting, etc.)	14.33	3.52%	11.85	3.03%
The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets	14.21	3.49%	14.84	3.80%
Favorite Pedestrian Streets	13.35	3.28%	10.5	2.69%
Preferred mode of transport during summer with pedestrianization	13.09	3.22%	9.8	2.51%
Extend pedestrianization into other streets	11.84	2.91%	8.29	2.12%
Hangout long time in pedestrianized streets	11.83	2.91%	8.5	2.17%
Intersection Safety	11.38	2.80%	11.68	2.99%
Impact of pedestrianization on mobility	11.27	2.77%	11.57	2.96%
Hangout more often on pedestrianized streets	10.03	2.47%	11.33	2.90%
Do Cycling (personal, Bixi, electric or non) During Covid-19	10.03	2.47%	6.05	1.55%
Hangout more in shops on pedestrianized streets	9.95	2.45%	8.28	2.12%
Cleanliness	9.89	2.43%	8.01	2.05%
General satisfaction of pedestrianization	9.25	2.27%	10.8	2.76%

Table 5.16: Feature Importance of each independent variable in XGB (cont'd)

Feature Description	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)		Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	
	Weight (Gain)	Percentage	Weight (Gain)	Percentage
Continuation of pedestrianization for the future on commercial streets	8.85	2.18%	8.19	2.09%
Changing the route for passing pedestrianized streets	8.82	2.17%	7.8	1.99%
Trip purpose	8.79	2.16%	9.61	2.46%
Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19	8.26	2.03%	8.48	2.17%
Preferred mode of transport during summer without pedestrianization	7.9	1.94%	6.16	1.58%
Universal Accessibility	7.77	1.91%	10.24	2.62%
Preferred mode of transport during Winter without pedestrianization	7.76	1.91%	6.33	1.62%
Accessibility Transportation Options	7.74	1.90%	9.35	2.39%
Number of Household members	7.57	1.86%	4.27	1.09%
Attractiveness	7.22	1.77%	7.26	1.86%
Changing the route to avoid pedestrianized streets	7.04	1.73%	7.86	2.01%
Street Furniture	6.89	1.69%	7.54	1.93%
Public Transport using during covid-19	6.54	1.61%	8.41	2.15%
Car Sharing using During Covid-19	6.54	1.61%	6.71	1.72%
Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19	5.99	1.47%	4.31	1.10%
Utilitarian walking During Covid-19 (including Mobility Aid)	5.87	1.44%	3.86	0.99%
Being a regular cyclist	5.65	1.39%	9.12	2.33%
Trip Companion	5.41	1.33%	6.94	1.77%

Table 5.16: Feature Importance of each independent variable in XGB (cont'd)

Feature Description	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)		Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	
	Weight (Gain)	Percentage	Weight (Gain)	Percentage
Streets Visiting Time	4.76	1.17%	5.41	1.38%
Motorized vehicle ownership	4.6	1.13%	4.57	1.17%
Adaptive transport using during Covid-19	4.52	1.11%	4.12	1.05%
Trip frequency	4.41	1.08%	4.42	1.13%
Paratransit using (Uber, taxi, etc.) During Covid-19	4.27	1.05%	4.39	1.12%
Disability - Walking	3.35	0.82%	3.39	0.87%
Gender	3.13	0.77%	2.63	0.67%
Bus Riding Disability	2.22	0.55%	3.01	0.77%
Disability - Cycling	1.87	0.46%	0.88	0.22%
Mobility Restriction	1.82	0.45%	2.93	0.75%

The output from XGBoost feature importance (Table 5.16) shows that ‘*The behavior of cyclists toward pedestrians regarding speed*’, ‘*The behavior of cyclists toward pedestrians regarding priority crossing*’, ‘*The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets*’, ‘*Personal Security*’, ‘*The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets*’, and ‘*Intersection Safety*’ are the most prominent features in satisfaction assessment of coexistent environment of pedestrians, cyclists and any type of two-wheelers. Meaning that mutual behavior of cyclists and pedestrians, personal and intersection safety are the most impactful factors for predicting Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists) and Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders).

5.5.1 Classification evaluation performance by Confusion Matrix for XGB

In this section, the classification performance of the XGB (Extreme Gradient Boosting) model for predicting satisfaction levels with the co-existence of pedestrians and different types of cyclists is evaluated. The performance is assessed using metrics such as precision, recall, F1-

score, and the confusion matrix. These metrics provide insights into how well the model performs in classifying different satisfaction classes (*Very Satisfied*, *Rather Satisfied*, *Rather Dissatisfied*, *Very Dissatisfied*).

From Table 5.17, it is observed that the model has the highest precision percentage for the 'Very Satisfied' and 'Very Dissatisfied' classes, where 76% and 72% for Y1 and 70% and 65% for Y2, respectively; meaning that these percentages were truly predicted out of the selected samples of the test dataset.

Table 5.17: Classification performance for XGB

Model performance score for XGBoost	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)				Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)			
	precision	recall	f1-score	support	precision	recall	f1-score	support
Very Satisfied	0.76	0.84	0.80	44	0.70	0.92	0.80	36
Rather Satisfied	0.35	0.28	0.31	40	0.43	0.29	0.35	41
Rather Dissatisfied	0.59	0.61	0.60	74	0.62	0.66	0.64	98
Very Dissatisfied	0.72	0.74	0.73	76	0.65	0.61	0.63	59
Accuracy	0.64				0.62			

F1-score of both Y1 and Y2 in class 'Very Satisfied' is 80% meaning that the model made 80% of all predictions right for the entire test dataset. Based on this table, the precision percentages for 'Rather Dissatisfied' and 'Rather Satisfied' classes are relatively low, with 59% and 35% for Y1 and 62% and 43% for Y2, respectively. This suggests that the model struggled to correctly identify and predict these classes. Furthermore, the 'Rather Satisfied' class has the lowest F1-score of just 31% and 35% for Y1 and Y2, respectively, which indicates that the model did not perform well in predicting these classes overall.

In Y1, the recall for class "Very Satisfied" is 0.84, and for Y2 is 0.92 which means that 84% and 92% of the actual instances of "Very Satisfied" were correctly identified for Y1 and Y2 respectively by the model. Similarly, the recall for class of "Rather Satisfied" is 0.28 in Y1 and 0.29 for Y2, which means that only 28% of the actual instances of "Rather Satisfied" for Y1 and 29% for Y2 were correctly identified by the model.

5.5.2 Visualization of XGBoost results by SHAP

In this section, the results of the XGBoost model are visualized using SHAP (Shapley Additive Explanations) plots. These plots provide insights into the important features that contribute to the model's predictions for satisfaction levels with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians, cyclists, and any types non-electric two-wheeled riders (Y2). As mentioned in chapter 4, the dataset underwent categorical to numeric transformation, and each variable has a specific code associated with it. To review the code created for each variable to understand, refer to Table 4.2.

Figure 5.3 and Figure 5.4 show the SHAP additive explainer visualization result for Y1 and Y2 respectively. As can be seen 'The behavior of cyclists toward pedestrians regarding speed' and 'The behavior of cyclists toward pedestrians regarding priority crossing' were the most impactful features for the prediction of interaction satisfaction assessment of pedestrians and cyclists in pedestrian streets. The respondents who chose the option 'Very Unfavorable' (1) for these features had the first and second greatest contribution to the final prediction for the class of 'Very Dissatisfied' with the coexistence of pedestrians, and traditional cyclists (Y1), and with the coexistence of pedestrians, cyclists, and any types of two wheelers (Y2). This means that respondents who are very dissatisfied with the mutual behavior of cyclists towards pedestrians regarding 'speed' and 'priority crossing' have the greatest contribution to the prediction of being very unsatisfied with the coexistence of pedestrians and cyclists in pedestrian streets.

Surprisingly the respondents who were 'Rather Favorable' (4) and 'Very Favorable' (5) with the behavior of cyclists toward pedestrians regarding 'speed' and 'priority crossing' also can be seen in the first ten impactful features in both Y1 and Y2 SHAP summary plot. Furthermore, respondents who are 'Rather Unfavorable' (2) with the behavior of cyclists toward pedestrians regarding 'speed' also can be seen in the first six impactful features in both Y1 and Y2 SHAP summary plot as well. This means that these respondents had a significant contribution to the prediction of satisfaction assessment with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians, cyclists, and any types of two wheelers (Y2) in pedestrian streets. In other words, even respondents with favorable opinions about pedestrians' and cyclists' mutual behavior in terms of speed and priority crossings can significantly influence the prediction of interaction satisfaction on pedestrian streets, suggesting that these features and satisfaction are closely related.

Following those respondents who were 'Rather Unfavorable' (2) with the 'Intersection safety' and 'Personal security' also have contribution to the prediction for the class of 'Very Satisfied' with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians, cyclists, and any types of two wheelers (Y2) in pedestrian streets. Furthermore, those individuals who were 'Rather Unfavorable' (2) with the behavior of cyclists toward pedestrians regarding 'speed' also have contribution to the prediction for the class of 'Rather Dissatisfied' with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians, cyclists, and any types of two wheelers (Y2) in pedestrian streets. In other words, individuals who expressed dissatisfaction with the intersection safety and personal security aspects played a significant role in predicting 'Very Satisfied' class for the interaction satisfaction assessment. It highlights that addressing concerns related to intersection safety and personal security can potentially contribute to higher levels of satisfaction among pedestrians and cyclists.

In summary, this suggests that behavioral aspects and safety issues significantly impact the interaction satisfaction of pedestrians and various types of two-wheelers on pedestrian streets. Factors such as cyclists' behavior towards pedestrians regarding speed and priority crossings, as well as intersection safety and personal security concerns like fear of the unknown and inadequate lighting, play key roles in shaping individuals' perception and experience in these environments. Thus, addressing these factors are essential for creating safer and more satisfying pedestrian street.

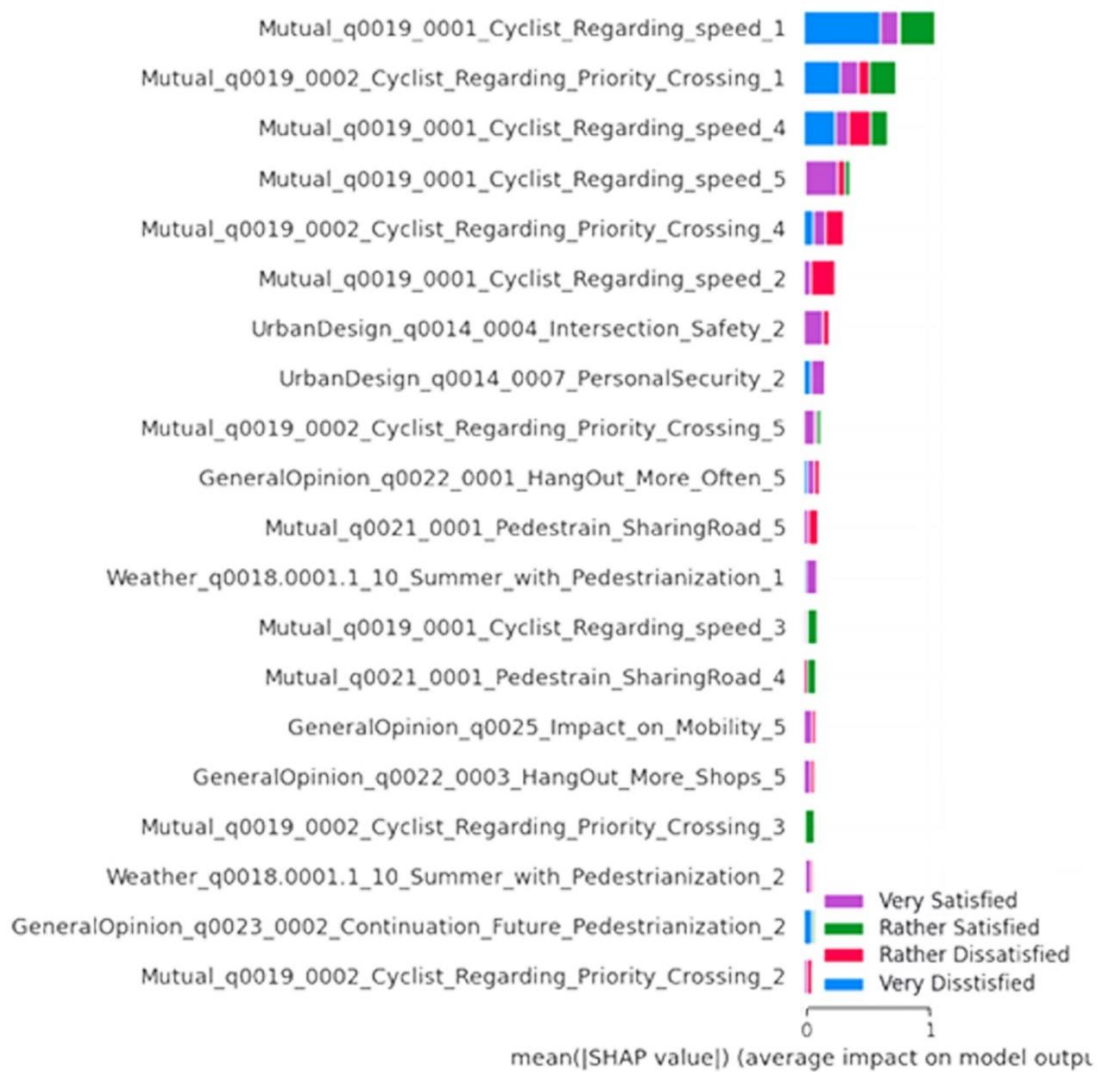


Figure 5.3: SHAP value for Y1_XGB

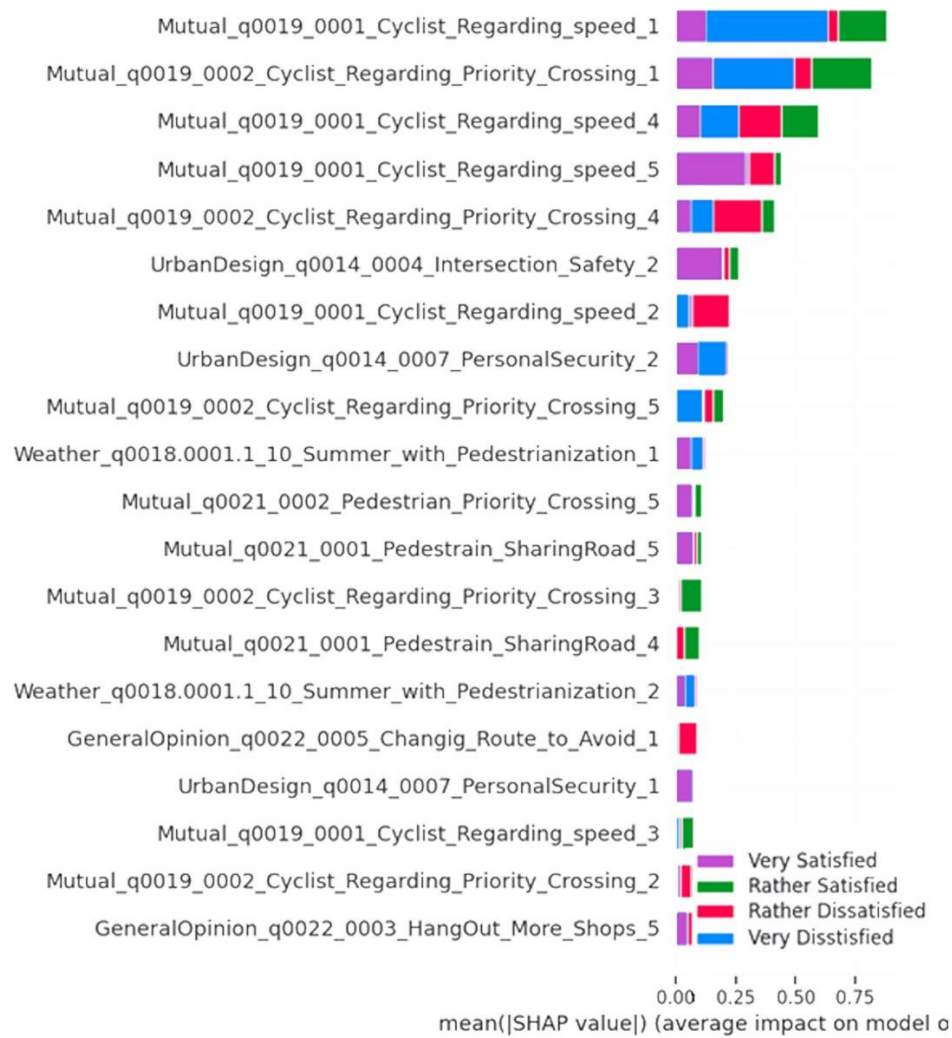


Figure 5.4: SHAP value for Y2_XGB

5.6 Random Forest results

Random Forest (RF) was used as the second ML classification algorithm model, introduced in chapter 3 of this thesis. RF as a second classification technique, was performed on the data in order to determine the weight of each feature and rank the importance of the features. By RF same as XGBoost, it can be possible to determine how each independent variable predict the satisfaction of the coexistence of pedestrians and cyclists in pedestrian streets.

The best hyperparameters of the models after tuning are shown Table 5.18. To evaluate the predictions made by the model, the confusion matrix function from scikit-learn is used, which creates a table that shows the number of true positives, true negatives, false positives, and false negatives. Finally, the output of the model is visualized by creating a bar plot for each value of the response variable. This plot can help to understand how well the model is able to predict each class of dependent variable in the dataset and can be found in APPENDIX I.

Table 5.18: Feature Importance of each independent variable in RF

Feature Description	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)		Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	
	Gain (weight)	percentage	Gain (weight)	percentage
Behavior of cyclists toward pedestrians regarding speed	0.3165	48.00%	0.3525	35.25%
Behavior of cyclists toward pedestrians regarding priority crossing	0.1177	24.18%	0.2097	20.97%
Personal Security (fear of strangers, lack of lighting, etc.)	0.0422	4.23%	0.0293	2.93%
Intersection Safety	0.0266	2.66%	0.0233	2.33%
Hangout more often on pedestrianized streets	0.0062	2.29%	0.0209	2.09%
The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets	0.0035	1.88%	0.0462	4.62%

Table 5.18: Feature Importance of each independent variable in RF (cont'd)

Feature Description	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)		Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	
	Gain (weight)	percentage	Gain (weight)	percentage
Preferred mode of transport during summer with pedestrianization	0.0068	1.78%	0.017	1.70%
The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets	0.0051	1.66%	0.0213	2.13%
Changing the route for passing pedestrianized streets	0.0011	1.20%	0.0097	0.97%
Cleanliness	0.0002	1.11%	0.0083	0.83%
Impact of pedestrianization on mobility	0.0063	1.07%	0.0181	1.81%
Being a regular cyclist	0.0033	1.02%	0.0026	0.26%
Hangout more in shops on pedestrianized streets	0.0022	0.88%	0.0169	1.69%
Hangout long time in pedestrianized streets	0.0005	0.87%	0.0097	0.97%
Extend pedestrianization into other streets	0.0032	0.85%	0.0046	0.46%
Street Furniture	0.0067	0.67%	0.0116	1.16%
Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19	0.0066	0.66%	0.0082	0.82%
Continuation of pedestrianization for the future on commercial streets	0.0055	0.55%	0.0066	0.66%
Do Cycling (personal, Bixi, electric or non) During Covid-19	0.005	0.50%	0.0137	1.37%
Universal Accessibility	0.0043	0.43%	0.0143	1.43%

Table 5.18: Feature Importance of each independent variable in RF (cont'd)

Feature Description	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)		Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	
	Gain (weight)	percentage	Gain (weight)	percentage
Preferred mode of transport during summer without pedestrianization	0.0034	0.34%	0.006	0.60%
Paratransit using (Uber, taxi, etc.) During Covid-19	0.0028	0.28%	0.0065	0.65%
Preferred mode of transport during Winter without pedestrianization	0.0013	0.28%	0.0088	0.88%
Accessibility Transportation Options	0.0028	0.28%	0.0079	0.79%
Trip frequency	0.0028	0.28%	0.01	1.00%
Changing the route to avoid pedestrianized streets	0.0002	0.27%	0.0073	0.73%
Gender	0.0009	0.22%	0.0045	0.45%
Motorized vehicle ownership	0.0016	0.20%	0.0083	0.83%
Trip purpose	0.0005	0.20%	0.0088	0.88%
Car Sharing (Communauto) using During Covid-19	0.002	0.20%	0.006	0.60%
Utilitarian walking During Covid-19 (including Mobility Aid)	0.0018	0.18%	0.0034	0.34%
Trip Companion	0.0008	0.17%	0.0111	1.11%
Public Transport using during covid-19	0.0016	0.16%	0.0157	1.57%
General satisfaction of pedestrianization	0.0003	0.13%	0.0146	1.46%
Attractiveness	0.0006	0.11%	0.0106	1.06%
Favorite Pedestrian Streets	0.0001	0.06%	0.0038	0.38%

Table 5.18: Feature Importance of each independent variable in RF (cont'd)

Feature Description	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)		Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)	
	Gain (weight)	percentage	Gain (weight)	percentage
Disability - Walking	0.0005	0.05%	0.0012	0.12%
Adaptive transport using during Covid-19	0.0004	0.04%	0.0011	0.11%
Disability - Cycling	0.0002	0.02%	0.0001	0.01%
Number of Household members	0.0002	0.02%	0.0066	0.66%
Streets Visiting Time	0.0001	0.01%	0.0121	1.21%
Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19	0	0.00%	0.0009	0.09%
Mobility Restriction	0	0.00%	0	0.00%
Bus Riding Disability	0	0.00%	0	0.00%

The output from RF feature importance (Table 5.18) shows that ‘*The behavior of cyclists toward pedestrians regarding speed*’, ‘*The behavior of cyclists toward pedestrians regarding priority crossing*’, ‘*Personal Security*’, ‘*Intersection Safety*’, ‘*The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets*’, and ‘*The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets*’ are the most prominent features in both dependent variables across all independent variables of the study. This suggests that the overall satisfaction of people using pedestrianized streets is significantly influenced by the mutual behavior of cyclists and pedestrians, and personal and intersection safety.

5.6.1 Classification evaluation performance by confusion matrix for RF

In this section, the classification performance of the Random Forest (RF) model for predicting satisfaction levels with the coexistence of pedestrians and different types of cyclists or non-electric two-wheeled riders is examined. Various metrics such as precision, recall, and F1-score

to assess how well the model performs in classifying different satisfaction classes are evaluated (Very Satisfied, Rather Satisfied, Rather Dissatisfied, Very Dissatisfied).

From Table 5.19, it is observed that the model has the highest precision percentage for the 'Very Satisfied' class, where 76% for Y1 and %70 for Y2; meaning that these percentages were truly predicted out of the selected samples of the test dataset.

Table 5.19: Classification performance for RF

Model performance score for Random Forest	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)				Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)			
	precision	recall	f1-score	support	precision	recall	f1-score	support
Very Satisfied	0.76	0.84	0.80	44	0.70	0.92	0.80	36
Rather Satisfied	0.36	0.12	0.19	40	0.48	0.29	0.36	41
Rather Dissatisfied	0.51	0.68	0.58	74	0.61	0.65	0.63	98
Very Dissatisfied	0.70	0.67	0.68	76	0.61	0.59	0.60	59
Accuracy	0.61				0.62			

F1-score of both Y1 and Y2 in class 'Very Satisfied' is 80% meaning that the model made 80% of all predictions right for the entire test dataset. Based on Table 5.19, the precision percentages for 'Rather Satisfied' class are relatively low, with 36% and 48% for Y1 and Y2, respectively. This suggests that the model struggled to correctly identify and predict this class comparing to other classes. Furthermore, the 'Rather Satisfied' class has the lowest F1-score of just 19% and 36% for Y1 and Y2, respectively, which indicates that the model did not perform well in predicting these classes. The recall for "Very Satisfied" class is 0.84 for Y1 and 0.92 for Y2, which means that 84% and 92% of the actual instances of "Very Satisfied" class were correctly identified by the model. Similarly, the recall for "Rather Satisfied" is 0.12 and 0.29 for Y1 and Y2 respectively, which means that only 12% and 29% of the actual instances of "Rather Satisfied" were correctly identified by the model.

5.6.2 Visualization of Random Forest results by SHAP

In this section, the results of the Random Forest model are visualized using SHAP (Shapley Additive Explanations) plots. These plots provide insights into the important features that contribute to the model's predictions for satisfaction levels with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians, cyclists, and any types non-electric two-wheeled riders (Y2). *To review the code created for each independent variable which explained in chapter 4, refer to Table 4.2.*

Figure 5.5, and Figure 5.6 show the SHAP additive explainer visualization result for Y1 and Y2 respectively. As can be seen 'The behavior of cyclists toward pedestrians regarding speed' and 'The behavior of cyclists toward pedestrians regarding priority crossing' were the most impactful features for the prediction of interaction satisfaction assessment of pedestrians and cyclists in pedestrian streets. The respondents who chose the option 'Very Unfavorable' (1) for these features had the first and second greatest contribution to the prediction for the class of 'Very Dissatisfied' with the coexistence of pedestrians, and traditional cyclists (Y1), and with the coexistence of pedestrians, cyclists, and any types of two wheelers (Y2). Meaning that respondents who express strong dissatisfaction with the behavior of cyclists towards pedestrians in terms of speed and priority crossing have the most significant impact on predicting the 'Very Dissatisfied' class. In other words, respondents who are very dissatisfied with the mutual behavior of cyclists towards pedestrians regarding 'speed' and 'priority crossing' have the greatest contribution to the prediction of being very unsatisfied with the coexistence of pedestrians and cyclists in pedestrian streets.

Surprisingly the respondents who were 'Rather Favorable' (4) and 'Very Favorable' (5) with the behavior of cyclists toward pedestrians regarding 'speed' and 'priority crossing' also can be seen in the first ten impactful features in both Y1 and Y2 SHAP summary plot but with different order. Furthermore, respondents who are 'Rather Unfavorable' (2) with the behavior of cyclists toward pedestrians regarding 'speed' also can be seen in the first ten impactful features in the SHAP summary plot in both Y1 and Y2 as well, but with different order. This means that these respondents had a significant contribution to the prediction of satisfaction assessment with the coexistence of pedestrians and traditional cyclists (Y1) and with the coexistence of pedestrians, cyclists, and any types of two wheelers (Y2) in pedestrian streets. In other words, even respondents with favorable opinions about pedestrians' and cyclists' mutual behavior in terms of

speed and priority crossings can significantly influence the prediction of interaction satisfaction on pedestrian streets, suggesting that these features and satisfaction are closely related.

Following those respondents who were 'Rather Unfavorable' (2) with the 'Intersection safety' and 'Personal security' also have contribution to the prediction of the class 'Very Satisfied' with the coexistence of pedestrians and traditional cyclists (Y1). Furthermore, the respondents who were 'Very Favorable' (5) with 'The behavior of pedestrians toward cyclists regarding sharing the road' also have contribution to the prediction of the class 'Rather Dissatisfied' with the coexistence of pedestrians, cyclists, and any types of two wheelers (Y2) in pedestrian streets.

In conclusion, the mutual behavior of cyclists towards pedestrians regarding 'speed' and 'priority crossing' emerges as influential factors. This implies that the way cyclists interact with pedestrians in terms of their speed and giving priority at crossings greatly influences overall satisfaction with pedestrian streets. Furthermore, factors such as intersection safety and personal security, which encompass aspects like fear of the unknown and inadequate lighting, also play a significant role in this context. These findings suggest that ensuring safe and secure conditions, such as well-lit areas and addressing concerns related to personal security, can contribute to higher satisfaction levels among pedestrians, cyclists, and other two-wheeler users on pedestrian streets.

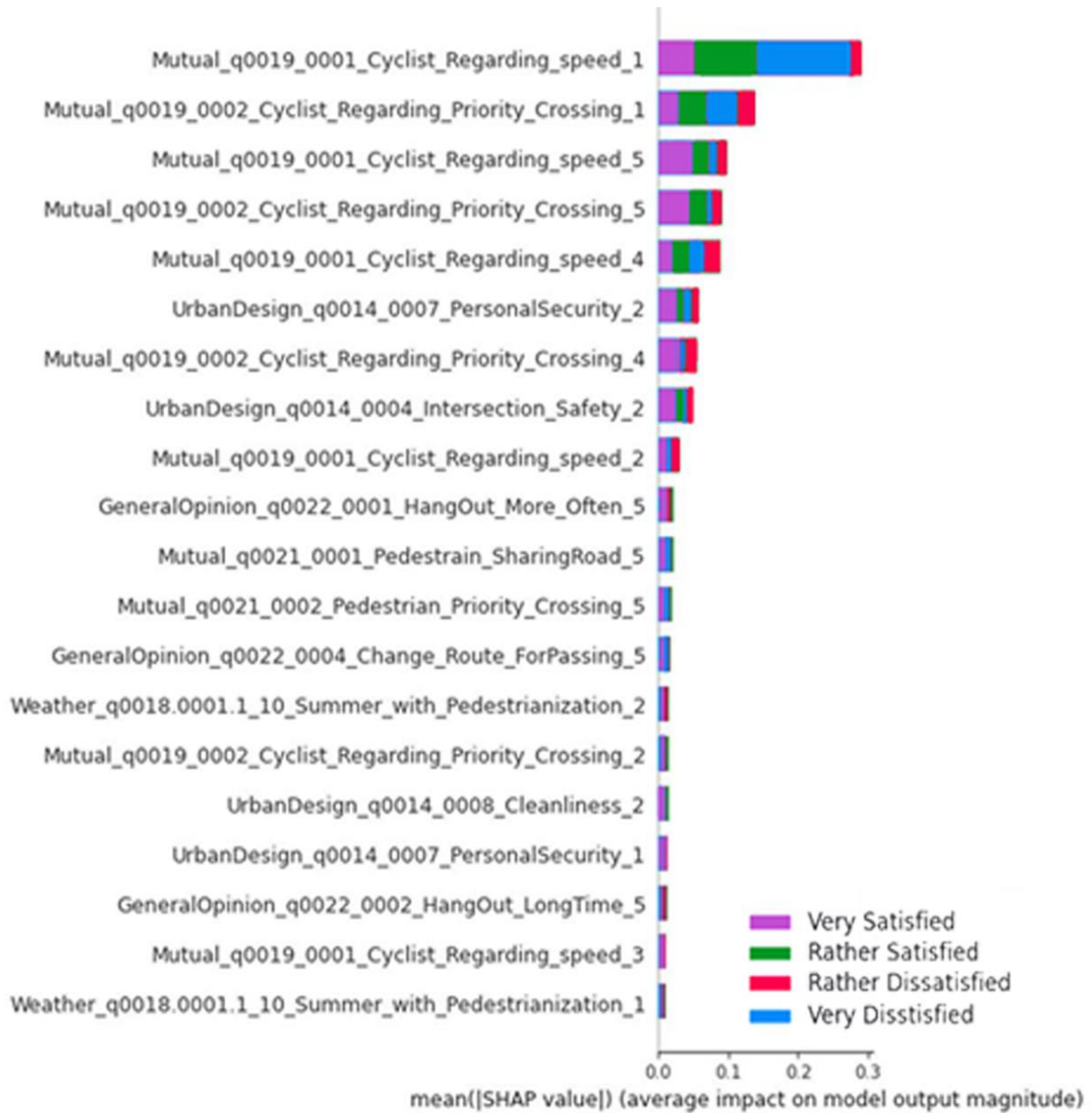


Figure 5.5: SHAP value for Y1_RF

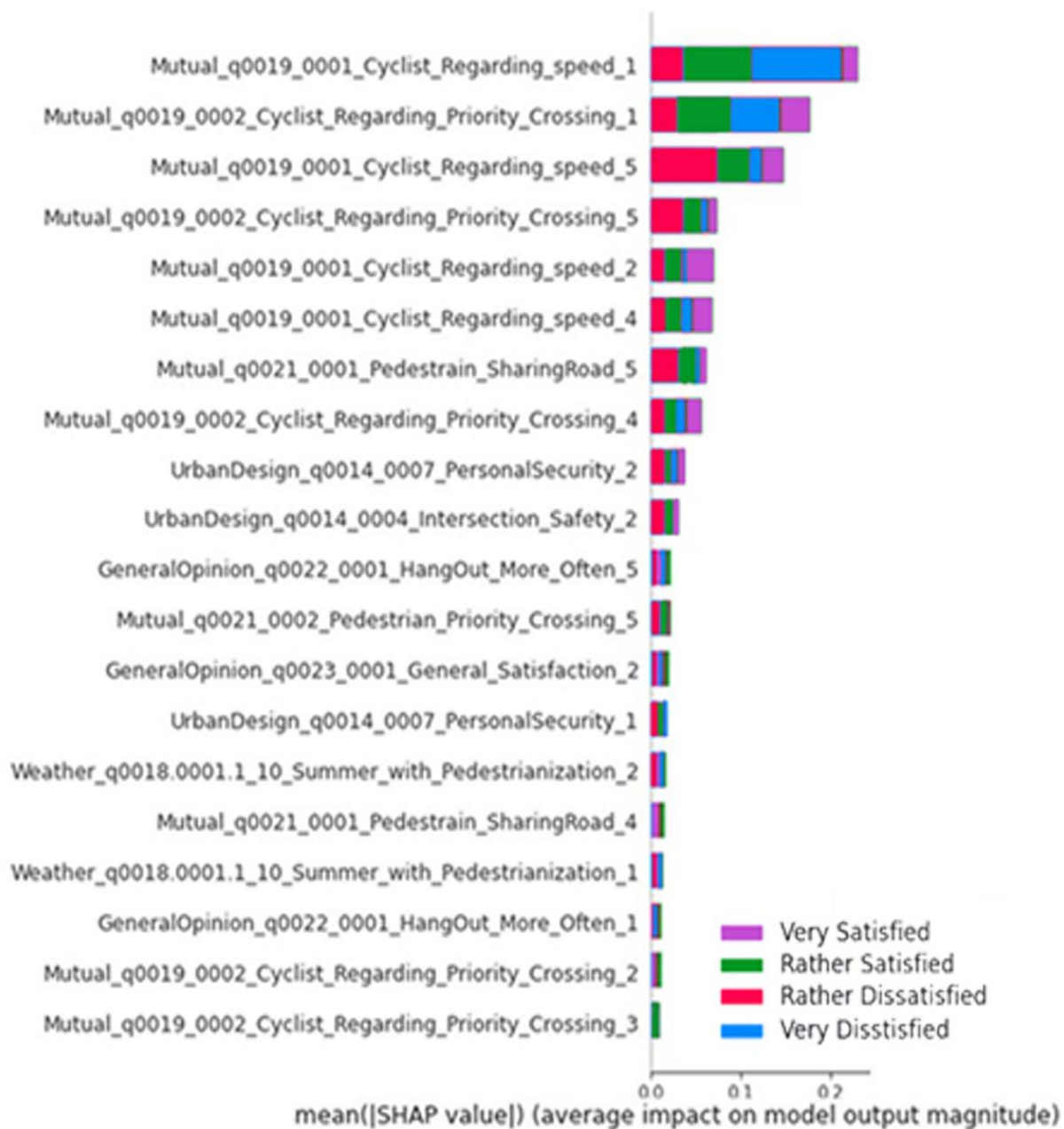


Figure 5.6: SHAP value for Y2_RF

5.7 Comparison of top five features in XGBoost, Random Forest and Lasso regression

The following tables compare the results of the top five important features as determined by XGBoost (XGB), Random Forest (RF), and Lasso regression.

Given that the MNL model only incorporates predictors selected by Lasso regression for both Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists) and Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders), other features were excluded from its result, the comparison is limited to Lasso, XGB, and RF, as these models utilize all available features.

Table 5.20: Top-five important features of XGB

Top five important features (XGB)	Percentage (Y1)	Rank (Y1)	Percentage (Y2)	Rank (Y2)
Behavior of cyclists toward pedestrians regarding speed	10.51%	1	10.20%	1
Behavior of cyclists toward pedestrians regarding priority crossing	8.16%	2	7.75%	2
The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets	3.64%	3	5.42%	3
Personal Security (fear of strangers, lack of lighting, etc.)	3.52%	4	3.03%	5
The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets	3.49%	5	3.80%	4

Table 5.21: Top-five important features of RF

Top five important features (RF)	Percentage (Y1)	Rank (Y1)	Percentage (Y2)	Rank (Y2)
Behavior of cyclists toward pedestrians regarding speed	48.00%	1	35.25%	1
Behavior of cyclists toward pedestrians regarding priority crossing	24.18%	2	20.97%	2
Personal Security (fear of strangers, lack of lighting, etc.)	4.23%	3	2.93%	4

Table 5.21: Top-five important features of RF (cont'd)

Top five important features (RF)	Percentage (Y1)	Rank (Y1)	Percentage (Y2)	Rank (Y2)
Intersection Safety	2.66%	4	2.33%	5
Preferred mode of transport during summer with pedestrianization	1.78%	5	1.70%	9
The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets	1.88%	14	4.62%	3

Table 5.22: Top-five selected features by Lasso

Top five selected features (Lasso)	Best coefficient (Y1)	Rank (Y1)	Best coefficient (Y2)	Rank (Y2)
Behavior of cyclists toward pedestrians regarding speed	0.516	1	0.852	1
Behavior of cyclists toward pedestrians regarding priority crossing	0.119	2	0.177	2
Intersection Safety	0.099	3	0.139	3
Impact of pedestrianization on mobility	0.068	4	0.088	4
Attractiveness	0.062	5	0.087	5

According to Table 5.20, the top five features in XGB for both Y1 and Y2 are the same, with only a difference in the order of features 4 and 5. Furthermore, neither Y1 nor Y2 shows a significant difference between the first two features and the subsequent features.

The common top five features for both Y1 and Y2 in XGB include 'Behavior of cyclists toward pedestrians regarding speed,' 'Behavior of cyclists toward pedestrians regarding priority crossing,' 'The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets,' 'Personal Security (fear of strangers, lack of lighting, etc.),' and 'The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets'.

In contrast, Table 5.21 shows that RF yields different top five features for Y1 and Y2, except for the first two. For Y1, the top five features in RF are 'Behavior of cyclists toward pedestrians regarding speed,' 'Behavior of cyclists toward pedestrians regarding priority crossing,' 'Personal

Security', 'Intersection Safety,' and 'Preferred mode of transport during summer with pedestrianization'. On the other hand, for Y2, the top five features in RF are 'Behavior of cyclists toward pedestrians regarding speed,' 'Behavior of cyclists toward pedestrians regarding priority crossing,' 'The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets,' 'Personal Security' and 'Intersection Safety'. Notably, the top two features in the RF table are significantly more important than the subsequent ones.

Based on Table 5.22, Lasso regression reveals consistent top five features for both Y1 and Y2. These include 'Behavior of cyclists toward pedestrians regarding speed,' 'Behavior of cyclists toward pedestrians regarding priority crossing,' 'Intersection Safety,' 'Impact of pedestrianization on mobility,' and 'Attractiveness'. Particularly, 'Behavior of cyclists toward pedestrians regarding speed' is deemed the most influential feature in the Lasso regression model.

Overall, the three tables demonstrate that 'Behavior of cyclists toward pedestrians regarding speed' and 'Behavior of cyclists toward pedestrians regarding priority crossing' consistently appear among the top features in RF, XGB, and Lasso regression. Further, 'Personal Security' appears in both RF and XGB as a top five feature, while 'Intersection Safety' only appears in both RF and Lasso regressions.

5.8 Category importance comparison for XGB and RF

This research also aims to determine pedestrianized street categories' importance (which is named and categorized by the author for easy comprehension and presented in chapter 4). For two dependent variables of the study (Y1, Y2).

According to the two classification techniques used in this study, the percentage weights of each category are shown in the Table 5.23. In the following paragraphs, these categories are going to be discussed.

Table 5.23: Comparison of categories importance in XGB and RF

Category Description	XGB		RF	
	Y1	Y2	Y1	Y2
Mutual Perception of Cyclists and Pedestrians	26.33%	27.17%	75.71%	62.98%
General Opinion and Attitude Toward Pedestrianization	22.31%	21.12%	9.21%	9.74%
Trip Specific Information	16.85%	18.59%	8.11%	10.85%

Table 5.23: Comparison of categories importance in XGB and RF (cont'd)

Category Description	XGB		RF	
	Y1	Y2	Y1	Y2
Urban Design and Infrastructure	14.42%	14.46%	2.69%	9.74%
Socio-economic Characteristics	9.53%	10.26%	2.41%	3.18%
Season-Related Mode of Transport	7.21%	5.70%	1.81%	3.12%
Favorite Pedestrian Streets	3.35%	2.69%	0.06%	0.38%

When comparing the importance of categories in assessing satisfaction levels, it is evident that both XGB and RF models provide different weightings.

Firstly, the percentage change in the weight of categories in both RF and XGB models for Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists) and Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders) is relatively similar. However, there are some noticeable differences between the two models.

In XGB, there is not a significant difference between the first important category and the subsequent categories. While in RF there is a significant percentage drop from the top category, 'Mutual Perception of Cyclists and Pedestrians,' to the subsequent categories. This suggests that XGB does not prioritize one category significantly over others. For the category of 'Mutual Perception of Cyclists and Pedestrians,' RF assigns it a weight almost three times greater than the weight assigned by XGB. This indicates that RF considers this category to be significantly more important in the assessment of satisfaction compared to XGB.

Secondly, the table highlights the category 'Favorite Pedestrian Streets' has relatively low importance when assessing satisfaction with pedestrian streets. This category encompasses all 12 pedestrian streets and was established to investigate whether the selection of a particular pedestrian street has any influence on the interaction satisfaction of pedestrians and cyclists in pedestrian streets. In XGB, this category receives a weight of at least 3%, indicating some level of importance. However, in Random Forest (RF), the importance of this category accounts for less than 1% of the overall weight. The low weight assigned to this category suggests that RF does not consider it to be a significant factor in determining overall satisfaction with pedestrian streets.

Furthermore, it is notable that the category 'Socio-economic Characteristics' exhibits varying levels of importance in the XGB and RF models. In XGB, this category gains approximately 10% weight for both Y1 and Y2. On the other hand, in RF, the weight assigned to the 'Socio-economic Characteristics' category is around 3% for both Y1 and Y2, indicating a relatively lower influence.

In conclusion, improving the mutual perception and respect between pedestrians and cyclists, as well as considering general opinions and attitudes toward pedestrianization, are key factors for enhancing interaction satisfaction of cyclists and pedestrians in pedestrian streets. These findings emphasize the importance of fostering positive interactions and addressing the needs and preferences of pedestrians and cyclists to create a harmonious and satisfactory environment for all users.

5.9 Comparison of XGBoost and Random Forest by Confusion Matrix

In this section, the performance of two popular machine learning algorithms, XGBoost, and Random Forest, is compared for the classification of satisfaction levels in pedestrian street studies.

As Confusion Matrix (CM) for XGBoost shows (Figure 5.7, and Figure 5.8), on the diagonal elements, out of the 234 test samples, 149 samples for Y1 and 146 samples for Y2 were correctly predicted. Figure 5.9, and Figure 5.10 show the Confusion Matrix (CM) for Random Forest and out of the 234 test samples, 143 samples for Y1 and 144 samples for Y2 were correctly predicted. As the CM results show, the lowest prediction is for class 'Rather Satisfied' for both dependent variables. This means that for both RF and XGBoost, models have difficulty in predicting this class, and the amount of misclassification is the highest compared to other classes.



Figure 5.7: Confusion matrix for XGB_Y1

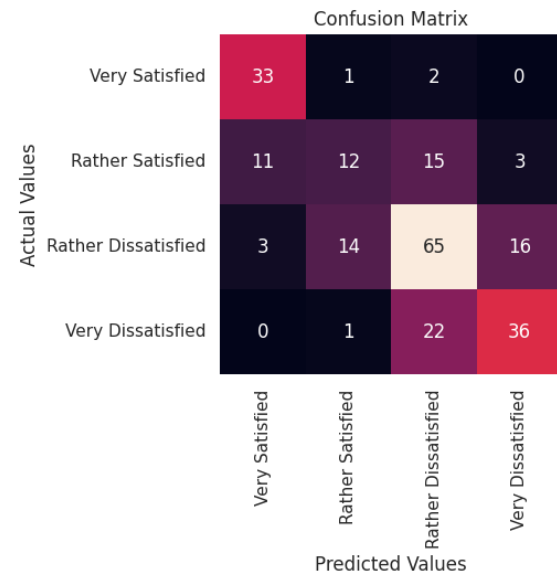


Figure 5.8: Confusion matrix for XGB_Y2

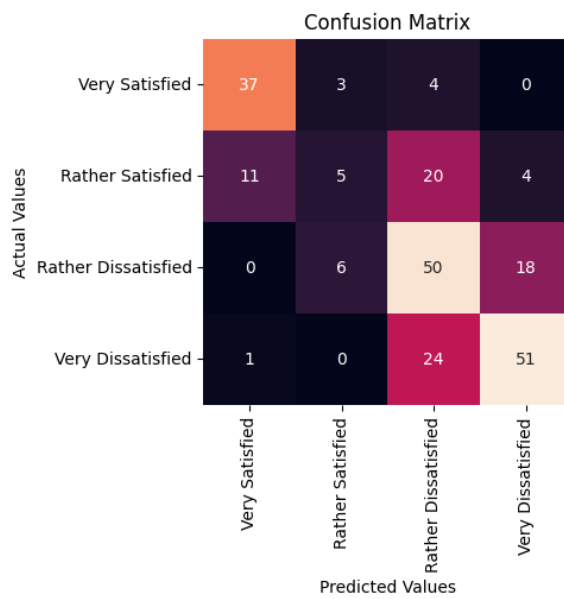


Figure 5.9: Confusion matrix for RF_Y1

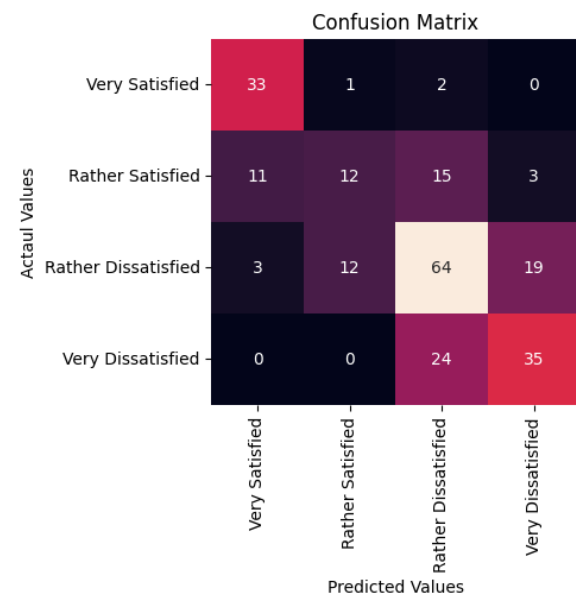


Figure 5.10: Confusion matrix for RF_Y2

Table 5.24: F1-score and Accuracy comparison for XGB and RF

	Y1 (Satisfaction level with the coexistence of pedestrians and "traditional" cyclists)				Y2 (Satisfaction level with the coexistence of pedestrians and non-electric two-wheeled riders)			
	Very Satisfied	Rather Satisfied	Rather Dissatisfied	Very Dissatisfied	Very Satisfied	Rather Satisfied	Rather Dissatisfied	Very Dissatisfied
XGB								
F1-score	0.80	0.31	0.60	0.73	0.8	0.35	0.64	0.63
Accuracy	64%				62%			
RF								
F1-score	0.80	0.19	0.58	0.68	0.80	0.36	0.63	0.60
Accuracy	61%				62%			

Regarding the Table 5.24, the higher F1-score indicates better performance of the model. Thus, the class 'Very Satisfied' has the highest well prediction of the classifier among all the classes. Furthermore, the value of F1-score in 'Rather Satisfied' class has the lowest among all classifiers; meaning that the class 'Rather Satisfied' has the highest misclassification rate among all the classes in two ML methods.

As for accuracy in this table, the XGBoost model performed slightly better than the Random Forest model in terms of classification performance of satisfaction assessment in pedestrian streets' studies. The accuracy of the XGBoost model is 64% in Y1 and 62% in Y2, while the accuracy of the Random Forest model is 61% in Y1 and 62% in Y2. In conclusion, the comparison of XGBoost and Random Forest models for the classification of satisfaction levels in pedestrian street studies reveals some notable findings. Both models struggle to accurately predict the 'Rather Satisfied' class, indicating a higher misclassification rate for this category compared to others. However, the models perform relatively well in predicting the 'Very Satisfied' class, as reflected by their higher F1-scores.

CHAPTER 6. CONCLUSIONS, CONTRIBUTIONS and PERSPECTIVES

This chapter summarizes the findings, addresses limitations, and provides recommendations for future research. This research presents a three-step methodology for estimating interaction satisfaction of pedestrians and cyclists on pedestrianized streets consisting of: 1) Data Preparation, 2) Analytical Test, running Regression models, and their Performance Evaluation and 3) running Classification models along with their Performance Evaluation.

The primary source of data for this study is the Cohabitation Piétons et Cyclistes (CCP) survey, which was conducted during the summer of 2021. A description of the research process is presented below.

In the *first* step, since the data consisted of categorical variables, in the data ready process, the list-wise deletion method was employed to mitigate potential bias. This involved conducting a complete case analysis, resulting in a selection of 1170 respondents out of the 1491 participants in the web based CCP survey. Furthermore, a subset of 44 features pertaining to interaction satisfaction of pedestrians and cyclists was extracted from the survey data for subsequent analysis. Based on insights from existing literature, the extracted variables were classified into seven distinct categories. Considering the scattered nature of the factors applied in this study, categorization provided a clearer understanding, as well as facilitated identifying patterns, relationships, and trends. The categories devised include *'Trip Specific Information'*, *'Urban Design and Infrastructure'*, *'Socio-economic Characteristics'*, *'Season-Related Mode of Transport'*, *'Mutual Perception of Cyclists and Pedestrians'*, *'Favorite Pedestrian Street'* and *'General Opinion and Attitude Toward Pedestrianization'*. To determine the relative significance of each category in the context of pedestrian streets, the weight of each feature along with the suggested categories were estimated with respect to feature importance computation based on 'Gain' metric.

In the *second* step, the correlation between the predictor and response variables was assessed using Spearman's correlation coefficient and the VIF (Variance Inflation Factor) test. The VIF test revealed that certain predictors exhibited high correlation with each other, indicating the presence of multicollinearity. To address this issue, one approach was to utilize Lasso regression, which helps in identifying the most significant features while disregarding those with lesser importance. This technique allows for the selection of key features while mitigating the effects of multicollinearity. In this step, the Lasso regression technique was employed to identify the most

significant features for the research. The results revealed that factors such as 'Behavior of cyclists toward pedestrians regarding speed', 'Behavior of cyclists toward pedestrians regarding priority crossing', 'Intersection Safety', 'Impact of pedestrianization on mobility', and 'Attractiveness' had a strong positive relationship with the satisfaction of coexistence between pedestrians, cyclists, and any type of two-wheelers. These influential features, selected through Lasso regression, were then utilized in a Multinomial Logistic (MNL) regression model to predict satisfaction levels regarding the coexistence of pedestrians and cyclists on pedestrian streets. The findings indicated that instances of negligent cyclist behavior in terms of speed and priority crossings, unsafe intersections, a lack of personal safety, and less attractive pedestrian streets were associated with lower satisfaction levels concerning the coexistence of pedestrians, cyclists, and two-wheelers.

In the *last* step, two classification methods, Extreme Gradient Boosting (XGB) and Random Forest (RF), were utilized. These methods were chosen to predict the weight of each feature and determine which class of respondents had the greatest impact on the satisfaction assessment of the coexistence of pedestrians and cyclists.

The results revealed that respondents who expressed strong dissatisfaction with the behavior of cyclists towards pedestrians in terms of 'speed' and 'priority crossing' made the most significant contribution to the final prediction of satisfaction assessment regarding pedestrianized streets.

Furthermore, there were subtle differences between XGB and RF in terms of estimating personal and intersection safety. In the XGB results, it was found that respondents who were highly satisfied with intersection safety while assessing their satisfaction were more likely to be 'Rather Satisfied' with the coexistence of pedestrians and two-wheelers on pedestrian streets, but 'Very Dissatisfied' with personal security matters. On the other hand, in the RF results, respondents who were highly satisfied with both personal and intersection safety were more inclined to rate the coexistence of pedestrians, cyclists, and two-wheelers on pedestrian streets as 'Very Satisfied'.

Finally, this study also compared the results and accuracy of XGB, RF, and Lasso regression. Since in the MNL, all the features were not considered (just that important ones taken from Lasso regression were analyzed), only the results of XGB, RF and Lasso regression were compared. Based on the findings from XGB, there were five important features selected by XGB that were the same when assessing satisfaction levels between pedestrians, traditional cyclists (Y1), and pedestrians and any other non-electric two-wheeled riders (Y2) on pedestrian streets. These include: 'Behavior of cyclists toward pedestrians regarding speed,' 'Behavior of cyclists toward

pedestrians regarding priority crossing,' 'The behavior of pedestrians toward cyclists regarding priority crossing in pedestrianized streets,' 'Personal Security (fear of strangers, lack of lighting, etc.),' and 'The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets'. However, the results for RF showed different top five features. In RF, the top five features differed for Y1 and Y2, except for the first two: "Behavior of cyclists towards pedestrians regarding speed," and "Behavior of cyclists towards pedestrians regarding priority crossing." The third, fourth, and fifth top five features in RF for Y1 were 'Personal Security', 'Intersection Safety,' and 'Preferred mode of transportation during summer with pedestrianization'. Furthermore, the third, fourth, and fifth top-five features in RF for Y2 were 'The behavior of pedestrians toward cyclists regarding sharing the road in pedestrianized streets,' 'Personal Security', and 'Intersection Safety'.

Regarding the overall accuracy, precision, recall, and F1 score were considered and measured for the classification and regression models as evaluation performance. Findings demonstrated that the overall accuracy of XGB was slightly better than RF, and in both models, the class 'Rather Satisfied' had the highest misclassification rate among all the classes in the two Machine Learning (ML) methods. In other words, ML models had difficulty for correctly predicting this type of class. Among the analyzed models, Lasso regression exhibited the lowest accuracy.

Some findings have been drawn from this study, as follows,

One of the key findings of the present research indicated that mutual perception of pedestrians and cyclists regarding speed, priority crossings, and sharing the road are the most important factors that affect respondents' satisfaction of interaction between pedestrians and cyclists while assessing the pedestrian streets. This implies that mutual understanding and respect between cyclists and pedestrians is one of the key determinants of satisfaction assessment.

Another key finding of the present study indicated that intersection safety and personal security played significant roles in the satisfaction level of pedestrians coexisting with two-wheelers on pedestrian streets. The results of this research indicated that cyclists who behaved respectfully and safely towards pedestrians could improve the satisfaction of pedestrians and other riders in pedestrian streets.

Here is what we have learned from this study that we could not learn as compared to simple analysis,

In this research, as a preliminary step, two examples of bivariate analysis were performed to identify the relationship between dependent and independent variables: the Chi-squared test and the Spearman correlation coefficient. These analyses aimed to determine the dependency between the study variables. The Chi-squared test was employed to analyze relationships among categorical variables. However, as the results were not reliable, a second bivariate analysis was conducted. The Spearman test was employed to determine the strength and direction of the relationship between the study variables; however, the primary objective, which was the assessment of pedestrian-cyclist interaction satisfaction, was not achieved. As discussed before, a methodology involving two regression models (Lasso regression and Multinomial Logistic Regression) and two classification models (Extreme Gradient Boosting and Random Forest) was suggested. Here are some key aspects learned from this research aligned with the objectives compared to simple bivariate analysis:

- Identification of influential factors that goes beyond simple statistical analyses (such as bivariate analysis) by considering the joint impact of multiple variables on interaction satisfaction.
- Understanding the relative importance of influential factors on the interaction satisfaction of pedestrians and cyclists in pedestrian streets provides a nuanced view that simple bivariate analysis might not capture.
- As the dataset was non-linear, the use of XGB and RF could capture nonlinear relationships and identify the most influential factors along with their feature importance, while bivariate analysis typically assumes linear relationships.
- Bivariate analyses lack the ability to accurately assess the performance of models in predicting satisfaction levels.
- Simple statistical analyses are unable to clarify the contribution of each class of the dependent variable, making it difficult to understand how different classes of respondents influence the overall findings.

6.1 Practical implications

Based on the findings, behavior of cyclists toward pedestrians regarding speed and priority crossings, intersection safety and personal security (fear of unknowns, lack of lighting, etc.) are the influential factors affecting interaction satisfaction of pedestrians and cyclists in pedestrian streets. Drawing from the insights gathered in the study, the City, in its efforts to redesign existing

pedestrian streets or plan new ones, could consider these various interventions. These interventions have the potential to enhance the overall satisfaction of pedestrians and cyclists' coexistence within these spaces. Examples of such interventions include:

Improving Intersection Safety: Enhance safety features at intersections through measures such as clear signage, warning signs, the implementation of traffic calming techniques, and physical modifications to correct street layouts. These improvements can increase visibility, eliminate blind spots, and facilitate better communication between pedestrians and cyclists. By promoting compliance with right-of-way rules, these measures can contribute to safer interactions, allowing sufficient reaction time for both pedestrians and cyclists and regulating cycling speeds.

Regulating Cyclist Behavior: Addressing specific behaviors of cyclists, particularly in terms of speed and priority crossings, is crucial. Measures to encourage cyclists to adhere to speed limits and prioritize pedestrian crossings can positively influence interaction satisfaction.

Enhancing Lighting Conditions: Considering personal security concerns identified in the study, improving lighting conditions emerges as a vital intervention. Installing adequate lighting infrastructure can address fears related to the unknown, lack of visibility, and overall enhance the sense of personal security.

6.2 Linking Research Gaps with Findings

Moving through the different parts of our study, a strong connection between the aspects intended to be investigated (research gaps) and the findings made is becoming evident. These connections will be discussed in the following paragraphs.

1- Comprehensive Evaluation of Factors

Research Gap: Previous studies lacked a simultaneous evaluation of a wide range of subjective and objective variables related to pedestrianized streets.

Findings: The study filled this gap by assessing factors such as socio-economic characteristics, urban design, trip-specific information, and the impact of season-related behavior. The findings highlighted the importance of mutual perception, safety, and attitudes toward pedestrianization.

2- Understanding Interaction Behavior

Research Gap: Limited investigation into the interaction behavior of pedestrians and cyclists, and the contribution of each influencing factor.

Findings: The study confirmed the critical role of mutual perception, emphasizing the influence of factors like speed, priority crossings, and road sharing. Understanding these interactions emerged as a key determinant of satisfaction between pedestrians and cyclists.

3- Inclusion of Respondent Class in Models

Research Gap: Previous studies did not consider the contribution of respondent class in classification models for interaction satisfaction assessment.

Findings: The study incorporated respondent satisfaction classes into models, revealing that respondents expressing strong dissatisfaction significantly contributed to the prediction of interaction satisfaction assessment regarding pedestrianized streets.

6.3 Contribution

This research study makes significant contributions to understanding the factors that influence user interaction satisfaction in pedestrianized spaces. The findings emphasize the importance of safety considerations in evaluating the interaction satisfaction of pedestrians and cyclists in pedestrian streets. Therefore, it is crucial for any pedestrianization project to prioritize and assess the safety aspects of the pedestrian environment.

The outcomes of this study can provide practical insights for both public and private stakeholders involved in pedestrianization projects. The results can serve as a valuable resource for urban planners and policymakers when designing and implementing pedestrian-friendly initiatives at the neighborhood, district, or city scales. The findings can be particularly useful in the development of active mode projects, including full-time pedestrian streets, part-time closures, and traffic calming measures.

The research underscores the significance of incorporating safety measurements as a fundamental guideline in the assessment of pedestrian streets and car-free zones. These measurements should be implemented prior to any intervention, ensuring that safety concerns are addressed and managed effectively. The results also highlight the need for robust management strategies and enforcement mechanisms to foster a safe and harmonious coexistence between pedestrians and

cyclists in larger urban transportation interventions. *A segment of this research, entitled 'Using Machine Learning XGBoost to Evaluate Levels of User Satisfaction on Pedestrianized Streets,' has been accepted at the Transportation Research Board 2023 conference in Washington, USA, on January 8, 2023.*

6.4 Limitations

The present study has certain limitations that should be acknowledged.

Firstly, it did not take into account two significant socio-economic factors of the respondents: income level and precise age. These variables could potentially shed light on the relationship between socioeconomic characteristics and interaction satisfaction of pedestrians and cyclists in pedestrian streets.

Secondly, the study did not incorporate spatial analysis, including the examination of land use patterns and physical features of the built environment. Factors such as universal access characteristics, footpath quality, and amenities like canopies and benches could have an impact on the overall interaction satisfaction of pedestrians and cyclists.

Thirdly, it is important to acknowledge that the present research focused on surveys conducted during a specific period of time, namely the summer season. However, weather conditions can significantly impact active transport trips, and this factor may vary across different cities. Therefore, to obtain a more comprehensive understanding of interaction satisfaction of pedestrians and cyclists in pedestrianized streets, it is recommended to conduct seasonal surveys that account for the weather variations throughout the year. This would provide complementary results and offer insights into how weather influences user satisfaction with the coexistence of pedestrians and cyclists in different seasons.

Lastly, despite having an adequate overall sample size, the scarcity of data for specific variable's classes can adversely affect the accuracy of predictions obtained through the deployed regression and classification machine learning of the study. Thus, increasing the sample size could potentially enhance the models' accuracy by providing more data for training. This could enable the algorithms to better capture the underlying patterns and improve predictions, particularly in classes with low frequency.

Furthermore, it would be beneficial to thoroughly examine interactions between pedestrians and cyclists through video analysis methods. With the help of the video cameras that have already

been installed, various safety indicators could have been extracted and the results could be integrated with the present study's results. Integrating these video-based findings into the present study would provide valuable insights into the dynamics and potential conflicts between pedestrians and cyclists in pedestrian zones.

It is worth noting that the specific context of the survey location was Montreal, Canada. Therefore, the recommendations mentioned above can be applied in the context of Montreal or other cities with similar characteristics, taking into account their unique weather patterns and local considerations.

6.5 Perspectives

Since Montreal is one of the cities in Canada that has long and harsh winters, considering interaction satisfaction assessment of pedestrians and cyclists in pedestrian streets in cold weather is one of the important parts of future studies. One crucial aspect that future studies on pedestrianized streets should address is the inclusion of interaction satisfaction assessment in cold weather conditions, considering that Montreal experiences long cold months. Understanding how user satisfaction with the coexistence of pedestrians and cyclists is affected by weather variations, including extremely cold temperatures, would provide valuable insights for designing and managing pedestrian streets to ensure year-round high satisfaction.

In this study, a wide range of influential features were investigated, and the results indicated that respect for the right of way and speed between cyclists and pedestrians are the most important factors. Intersection security and personal security are two other effective factors that should be considered in future studies of pedestrian streets. Specifically, the impact of concerns such as fear of strangers and inadequate lighting on the overall perception of safety should be thoroughly investigated. By understanding and addressing these aspects, the design and management of pedestrian streets can be improved to enhance the safety and comfort of users.

Future research can build upon these findings and focus on the mutual behavior of pedestrians and cyclists as well as safety issues. Conducting in-depth studies on the right of way and speed limitations would contribute to improving safety measures and ensuring a harmonious coexistence between pedestrians and cyclists.

Future research should focus more on pedestrian safety by deploying safety-related designs. In fact, personal safety and intersection safety considerations need to be intervened in the long and

short term in pedestrian areas. These considerations include future extensive studies on street design characteristics, like lighting, transparency, boosting intersections' safety by implementing protective equipment, dedicating routes for bicycles or clear signage to minimize potential conflicts between pedestrians and cyclists in pedestrian streets.

Hence, one of the most important aspects in future studies of pedestrian streets will be establishing secure areas before making any alterations and combining this study's result with video analysis results. This would increase the interaction satisfaction of pedestrians and cyclists in terms of safety. In other words, safety studies could improve the satisfaction level with the coexistence of pedestrians and cyclists and establish policies on securing pedestrian streets.

APPENDIX A SURVEY

Survey Questions no.	Suggested Main category	Variables taken from the survey, including options and numerical coded in the analysis
q0004- What street do you go to most often?	Favorite Pedestrian Streets	Favorite Street : Avenue du Mont-Royal (Plateau Mont-Royal) (1)/ Wellington (Verdun) (2) / Rue Ontario (Hochelaga-Maisonneuve) (3)/ Rue Masson (Rosemont-Petite-Patrie) (4)/ Rue Duluth (Plateau Mont-Royal) (5)/ Rue Bernard (Outremont) (6)/ Rue Castelnau (Villeray) (7)/ Rue Saint-Catherine west in downtown (Ville-Marie) (8)/ Rue Crescent in downtown (Ville-Marie) (9)/ Rue Sainte-Catherine East in Marie Village (Ville-Marie) (10)/ Rue Saint-Denis in Quartier Latin (Ville-Marie) (11)/ Rues du Marché Jean-Talon (Rosemont Petite-Patrie) (12)
q0014_0004- How satisfied are you with the safety at crossings with streets where vehicles are allowed?	Urban Design and Infrastructure	Intersection safety: Very unfavorable (-2)/ Unfavorable (-1)/ I prefer not to answer (0)/ Rather favorable (1)/ Very favorable (2)
q0014_0005- How satisfied are you with the universal accessibility to facilities and shops?		Universal accessibility: Very unfavorable (-2)/ Unfavorable (-1)/ I prefer not to answer (0)/ Rather favorable (1)/ Very favorable (2)
q0014_0006- How satisfied are you with the attractiveness (shops, atmosphere, layout)?		Attractiveness: Very unfavorable (-2)/ Unfavorable (-1)/ I prefer not to answer (0)/ Rather favorable (1)/ Very favorable (2)
q0014_0007- How satisfied are you with the personal security (fear of unknowns, lack of lighting, etc.)?		Personal security: Very unfavorable (-2)/ Unfavorable (-1)/ I prefer not to answer (0)/ Rather favorable (1)/ Very favorable (2)
q0014_0008- How satisfied are you with the Cleanliness		Cleanliness: Very unfavorable (-2)/ Unfavorable (-1)/ I prefer not to answer (0)/ Rather favorable (1)/ Very favorable (2)
q0014_0009- How satisfied are you with the Street furniture (aesthetic, comfort, positioning, number)?		Street furniture: Very unfavorable (-2)/ Unfavorable (-1)/ I prefer not to answer (0)/ Rather favorable (1)/ Very favorable (2)
q0005- What is the main reason for going to the pedestrian streets?	Trip-specific Information	Trip purpose: Shopping (1)/ Eating out in restaurants, bars, and cafes (2)/ Visiting a service store (dentist, hairdresser, daycare, etc.) (3)/ Studying/ working (4)/ Walk, exercise, and leisure (5)/ Passing between two destinations (including access to the residence) (6)/ I prefer not to answer (7)

APPENDIX A.1 SURVEY (cont'd)

Survey Questions no.	Suggested Main category	Variables taken from the survey, including options and numerical coded in the analysis
q0006- How many times do you go to the pedestrian street?		Trip frequency: Everyday (4)/ Few times a week (3)/ Few times a month (2)/ Only once (1) / I prefer not to answer (0)
q0007.0001-6- Who do you hang out with in the pedestrian street?		Trip companion while visiting the street: Alone (1)/ With friends (2)/ In a couple (3)/ With family (parents and / or children) (4)/ Other types of companions (5)
q0008.0001-7- At what time of day do you go to the pedestrian street?		Streets Visiting time: Before 9 a.m. (1) / Between 9 a.m. and 11 a.m. (2)/ From 11 a.m. to 1 p.m. (3)/ From 1 p.m. to 5 p.m. (4)/ From 5 p.m. to 7 p.m. (5) / After 7 p.m. (6)
q0017-0001- How many times do you use walking during Covid-19 (including mobility aid)?		Utilitarian walking During Covid (including Mobility Aid): Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)
q0017-0002- How many times do you use bike (personal, Bixi, electric or non) during Covid-19?		Cycling (personal, Bixi, electric or non) During Covid-19: Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)
q0017-0003- How many times do you use two-wheelers (scooter, skateboard electric or non) using during Covid-19?		Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19: Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)
q0017-0004- How many times do you use public transport (bus, metro) during Covid-19?		Public Transport using During Covid-19: Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)
q0017-0005- How many times do you use adaptive transport (bus, metro) during Covid-19?		Adaptive transport using during Covid-19: Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)
q0017-0006- How many times do you use Taxi, including Uber and other services during Covid-19?		Paratransit using (Uber, taxi, etc.) During Covid-19: Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)

APPENDIX A.2 SURVEY (cont'd)

Survey Questions no.	Suggested Main category	Variables taken from the survey, including options and numerical coded in the analysis
q0017-0007- How many times do you use Car Sharing (Communauto) during Covid-19?		Car Sharing (Communauto) using During Covid-19: Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)
q0017-0008- How many times do you drive personal motor vehicle (automobile, truck, motorcycle, etc.) during Covid-19?		Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19: Every day (5)/ Few times in a week (4)/ Few times in a month (3)/ Few times in a year (2)/ Never (1)/ I prefer not to answer. (0)
q0018.1.1-10- What mode of transport (out of these options) do you prefer during summer with pedestrianization?	Season-Related Mode of Transport	Preferred mode of transport during summer with pedestrianization: Walking (including Mobility Aids) (1)/ Cycling (personal, Bixi, electric or non) (2)/ Public transport (3)/ Personal motor vehicles (automobile, truck, motorcycle, etc.) (4)/ Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates) (5)/ Adaptive transport (6) Paratransit (Uber, taxi, etc.) (7)/ Car sharing (Communauto) (8) / I do not like to pass pedestrian streets at this time (9)
q0018.2.1-10- What mode of transport (out of these options) do you prefer during summer without pedestrianization?		Preferred mode of transport during summer without pedestrianization: Walking (including Mobility Aids) (1)/ Cycling (personal, Bixi, electric or non) (2)/ Public transport (3)/ Personal motor vehicles (automobile, truck, motorcycle, etc.) (4)/ Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates) (5)/ Adaptive transport (6) / Paratransit (Uber, taxi, etc.) (7)/ Car sharing (Communauto) (8) / I do not like to pass pedestrian streets at this time (9)
q0018.3.1-10- What mode of transport (out of these options) do you prefer during winter without pedestrianization?		Preferred mode of transport during Winter without pedestrianization: Walking (including Mobility Aids) (1)/ Cycling (personal, Bixi, electric or non) (2)/ Public transport (3)/ Personal motor vehicles (automobile, truck, motorcycle, etc.) (4)/ Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates) (5)/ Adaptive transport (6) / Paratransit (Uber, taxi, etc.) (7)/ Car sharing (Communauto) (8) / I do not like to pass pedestrian streets at this time (9)

APPENDIX A.3 SURVEY (cont'd)

Survey Questions no.	Suggested Main category	Variables taken from the survey, including options and numerical coded in the analysis
q0019-0001- As a pedestrian, what do you think of cyclists' behavior regarding riding to an acceptable speed?	Mutual behavior of cyclists and pedestrians	Behavior of cyclists toward pedestrians regarding speed: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0019-0002- As a pedestrian, what do you think of cyclists' behavior regarding priority during their passage?		Behavior of cyclists toward pedestrians regarding priority crossing: Very unfavorable (1)/Rather unfavorable (2)/ Neither favorable nor unfavorable (3) / Rather favorable (4)/ Very favorable (5)
q0021-0001- As a cyclist, what do you think of pedestrians' behavior regarding sharing the road?		Behavior of pedestrians toward cyclists regarding sharing the road: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0021-0002- As a cyclist, what do you think of pedestrians' behavior regarding being respectful during their passage?		Behavior of pedestrians toward cyclists regarding priority crossing: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0022-0001-Do you hang out more often due to pedestrianization and installation?	General Opinion and attitude toward pedestrianization	Hangout more often on pedestrianized streets: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0022-0002- Do you hang out long time due to pedestrianization and installation?		Hangout long time in pedestrianized streets: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0022-0003-Do you hang out more in shops due to pedestrianization and installation?		Hangout more in shops on pedestrianized streets:/ Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0022-0004- When you pass the pedestrian street, do you change your route to undertake activities on the street, or because it's nicer there to pass?		Changing the route for passing pedestrianized streets: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0022-0005- When you pass the pedestrian street, do you change your route to avoid the street (you would pass normally, but the pedestrianization makes other roads better)?		Changing the route to avoid pedestrianized streets:/ Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)

APPENDIX A.4 SURVEY (cont'd)

Survey Questions no.	Suggested Main category	Variables taken from the survey, including options and numerical coded in the analysis
q0023_0001- Are you more favorable or unfavorable to pedestrianization projects in general?		General satisfaction of pedestrianization: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0023-0002- Are you more favorable or unfavorable to redo the pedestrianization of the commercial artery every year?		Continuation of pedestrianization for the future on commercial streets: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0023-0003- Are you more favorable or unfavorable to the extension of pedestrianization to other arteries or neighborhood?		Extend pedestrianization into other streets: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0025- What is the impact of the pedestrianization on your mobility?		Impact of pedestrianization on mobility: Very unfavorable (1)/ Rather unfavorable (2)/ Neither favorable nor unfavorable (3)/ Rather favorable (4)/ Very favorable (5)
q0026- What is your gender?	Socioeconomic Characteristics	Gender: Female (1)/ Male (2)/ Other (3)
q0027.0001_3- Please indicate the number of your household		Number of household members: 1-member household (1)/ 2-member household (2)/ 3-member household (3)/ 4-member household and more (4)
q0030- Do you use a mobility aid?		Mobility restriction: No (1)/ Yes (2)
q0031-0001- Do you have mobility restriction for walking?		Disability - Walking: No (1)/ Yes (2)
q0031-0002- Do you have mobility restriction for riding a bike?		Disability - Cycling: No (1)/ Yes (2)
q0031-0003- Do you have mobility restriction for bus riding?		Bus Riding Disability: No (1)/ Yes (2)
q0020- Are you a regular cyclist?		Being a regular cyclist: Yes (1)/ No (2)
q0015.0001- 8- Which of these options do you own or have access to?		Accessibility Transportation Options: Driver's license ownership (1)/ Car, truck and motorcycle ownership (2) /Transit pass ownership (3)/ Bike ownership (4)/ Carsharing users (Communauto) (5)/ BIXI users (6)/ Other mobility options (7)
q0016.0001-3- Please indicate the motor vehicle you have		Motorized vehicle ownership: Automobile (1)/ Truck (2)/ Motorcycle (3)

APPENDIX B SPEARMAN'S RANK COEFFICIENT CORRELATION TEST

Variables	Y1	Variables	Y2
q0014_0001_ Satisfaction with the coexistence (pedestrians with "traditional" cyclists)	1	q0014_0002_ Satisfaction with the coexistence of pedestrians and non-electric 2-wheeled riders (skates, scooters, bikes, etc.)	1
q0019_0001_Behavior of cyclists toward pedestrians regarding speed	0.76	q0019_0001_Behavior of cyclists toward pedestrians regarding speed	0.75
q0019_0002_Behavior of cyclists toward pedestrians regarding priority crossing	0.74	q0019_0002_Behavior of cyclists toward pedestrians regarding priority crossing	0.73
q0022_0001_Hangout more often on pedestrianized streets	0.46	q0014_0007_Personal Security	0.45
q0014_0007_Personal Security	0.45	q0022_0001_Hangout more often on pedestrianized streets	0.43
q0025_Impact of pedestrianization on mobility	0.44	q0021_0001_The behavior of pedestrians toward cyclists regarding sharing the road	0.42
q0021_0001_The behavior of pedestrians toward cyclists regarding sharing the road	0.44	q0025_Impact of pedestrianization on mobility	0.42
q0022_0002_Hangout long time in pedestrianized streets	0.43	q0022_0002_Hangout long time in pedestrianized streets	0.41
q0023_0003_Extend pedestrianization into other streets	0.42	q0023_0003_Extend pedestrianization into other streets	0.4
q0022_0003_Hangout more in shops on pedestrianized streets	0.42	q0023_0002_Continuation of pedestrianization for the future on commercial streets	0.39
q0022_0004_Changing the route for passing pedestrianized streets	0.41	q0021_0002_The behavior of pedestrians toward cyclists regarding priority crossing	0.39
q0021_0002_The behavior of pedestrians toward cyclists regarding priority crossing	0.41	q0014_0004_Intersection Safety	0.39
q0023_0002_Continuation of pedestrianization for the future on commercial streets	0.41	q0022_0004_Changing the route for passing pedestrianized streets	0.39
q0017_0002_ Do Cycling (personal, Bixi, electric or non) During Covid-19	0.4	q0022_0003_Hangout more in shops on pedestrianized streets	0.39
q0023_0001_General satisfaction of pedestrianization	0.4	q0023_0001_General satisfaction of pedestrianization	0.39
q0014_0006_Attractiveness	0.4	q0014_0006_Attractiveness	0.38
q0014_0004_Intersection Safety	0.38	q0017_0002_ Do Cycling (personal, Bixi, electric or non) During Covid-19	0.36
q0014_0008_Cleanliness	0.36	q0014_0009_Street Furniture	0.35
q0014_0009_Street Furniture	0.35	q0014_0008_Cleanliness	0.34
q0020_Being a regular cyclist	0.34	q0020_Being a regular cyclist	0.32
q0018.0002.1_10_Preferred mode of transport during summer without pedestrianization	0.25	q0014_0005_Universal Accessibility	0.25

APPENDIX B.1 SPEARMAN'S RANK COEFFICIENT CORRELATION TEST (cont'd)

Variables	Y1	Variables	Y2
q0018.0001.1_10_Preferred mode of transport during summer with pedestrianization	0.25	q0018.0001.1_10_Preferred mode of transport during summer with pedestrianization	0.24
q0014_0005_Universal Accessibility	0.25	q0018.0002.1_10_Preferred mode of transport during summer without pedestrianization	0.24
q0015.0001_8_Accessibility_Transportation_Options	0.18	q0015.0001_8_Accessibility_Transportation_Options	0.17
q0017_0007_ Car Sharing (Communauto) using During Covid-19	0.17	q0017_0007_ Car Sharing (Communauto) using During Covid-19	0.16
q0018.0003.1_10_Preferred mode of transport during Winter without pedestrianization	0.13	2.2 Favorite Pedestrian Streets	0.12
q0031_0002_Disability - Cycling	0.12	q0031_0002_Disability - Cycling	0.12
2.2 Favorite Pedestrian Streets	0.1	q0018.0003.1_10_Preferred mode of transport during Winter without pedestrianization	0.11
q0027_Number of Household members	0.1	q0027_Number of Household members	0.11
q0008.0001_7_Streets Visiting time	0.07	q0031_0001_Disability -Walking	0.09
q0031_0001_Disability - Walking	0.07	q0008.0001_7_Streets Visiting time	0.08
q0017_0003_ Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19	0.06	q0017_0003_ Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19	0.06
q0031_0003_Bus Riding Disability	0.03	q0007.0001_6_Trip_Companionship	0.03
q0007.0001_6_Trip_Companionship	0.02	q0031_0003_Bus Riding Disability	0.02
q0026_Gender	-0.02	q0017_0006_ Paratransit using (Uber, taxi, etc.) During Covid-19	-0.01
q0017_0006_ Paratransit using (Uber, taxi, etc.) During Covid-19	-0.01	q0017_0001_ Utilitarian walking During Covid (including Mobility Aid)	-0.02
q0017_0004_Public Transport using during covid	-0.02	q0017_0004_Public Transport using during covid	-0.02
q0017_0005_Adaptive transport using during Covid-19	-0.04	q0026_Gender	-0.02
3.1.1 Trip purpose	-0.05	q0017_0005_Adaptive transport using during Covid-19	-0.03
q0017_0001_ Utilitarian walking During Covid-19(including Mobility Aid)	-0.05	3.1.1 Trip purpose	-0.05
q0006_Trip frequency	-0.07	q0016_0001_0003_Motorized_vehicle_ownership	-0.07
q0016_0001_0003_Motorized_vehicle_ownership	-0.08	q0006_Trip frequency	-0.07
q0017_0008_Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19	-0.09	q0017_0008_Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19	-0.08

APPENDIX B.2 SPEARMAN'S RANK COEFFICIENT CORRELATION TEST (cont'd)

Variables	Y1	Variables	Y2
q0030_Mobility Restriction	-0.11	q0030_Mobility Restriction	-0.12
q0022_0005_Changing the route to avoid pedestrianized streets	-0.29	q0022_0005_Changing the route to avoid pedestrianized streets	-0.29

APPENDIX C CROSS-TABULATION (CONTINGENCY TABLE)

Variable Category	Variables	Count	Y1				Y2			
Category: Favorite Pedestrian Street	Streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Mont-Royal (Plateau Mont-Royal)	Number	172	143	254	245	159	155	299	201
	Wellington (Verdun)	Number	7	13	42	48	4	16	51	39
	Ontario (Hochelaga-Maisonneuve)	Number	0	8	8	10	1	4	11	10
	Masson (Rosemont-La-Petite-Patrie)	Number	1	7	6	4	2	6	6	4
	Duluth (Plateau Mont-Royal)	Number	2	6	5	6	1	5	7	6
	Bernard (Outremont)	Number	2	3	10	5	2	3	8	7
	De Castelnau (Villeray)	Number	15	13	25	40	12	11	42	28
	Saint-Catherine west in city center (Ville-Marie)	Number	2	3	3	6	2	2	7	3
	Saint-Denis in Latin neighborhood (Ville-Marie)	Number	1	2	4	6	1	2	5	5
	Marché Jean Talon (Rosemont-La-Petite-Patrie)	Number	0	2	15	7	0	3	13	8
	I do not remember the name of the street	Number	1	0	0	0	1	0	0	0
	other streets	Number	9	1	5	3	7	4	4	3
Category : Urban Design And Infrastructure	Intersection Safety	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	48	19	9	3	47	20	9	3

APPENDIX C.1 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
			VD	RD	RS	VS	VD	RD	RS	VS
Category: Urban Design And Infrastructure	Rather dissatisfied	Number	34	40	45	22	35	33	58	15
	I prefer not to answer	Number	2	0	0	0	2	0	0	0
	Rather satisfied	Number	100	107	225	167	83	119	271	126
	Very satisfied	Number	28	35	98	188	25	39	115	170
	Universal Accessibility	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	37	17	9	8	37	16	13	5
	Rather dissatisfied	Number	26	25	28	20	26	24	32	17
	I prefer not to answer	Number	17	23	72	56	13	24	77	54
	Rather satisfied	Number	95	74	142	100	83	87	175	66
	Very satisfied	Number	37	62	126	196	33	60	156	172
	Attractiveness	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	24	4	6	1	23	5	6	1
	Rather dissatisfied	Number	38	18	14	5	33	20	17	5
	I prefer not to answer	Number	2	1	1	0	2	1	1	0
	Rather satisfied	Number	104	103	155	107	94	109	179	87
	Very satisfied	Number	44	75	201	267	40	76	250	221
Category: Urban Design And Infrastructure	Personal Security	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	26	4	5	2	25	5	5	2
	Rather dissatisfied	Number	30	12	11	4	27	15	11	4
	I prefer not to answer	Number	8	3	2	4	5	6	3	3
	Rather satisfied	Number	103	93	159	60	99	99	173	44
	Very satisfied	Number	45	89	200	310	36	86	261	261

APPENDIX C.2 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Urban Design And Infrastructure	Cleanliness	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	32	15	8	4	32	15	7	5
	Rather dissatisfied	Number	37	27	31	22	37	30	31	19
	I prefer not to answer	Number	1	1	1	2	1	1	1	2
	Rather satisfied	Number	117	106	220	145	102	114	247	125
	Very satisfied	Number	25	52	117	207	20	51	167	163
	Street Furniture	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	36	8	9	3	36	12	7	1
	Rather dissatisfied	Number	40	35	38	26	38	32	43	26
	I prefer not to answer	Number	5	5	0	3	3	6	2	2
	Rather satisfied	Number	102	102	190	147	93	108	222	118
	Very satisfied	Number	29	51	140	201	22	53	179	167
Category: Trip-specific Information	Trip purpose	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Shopping	Number	72	81	158	143	71	75	184	124
	Eating out in restaurants, bars, and cafes	Number	19	21	37	48	17	23	42	43
	Visiting a service store (dentist, hairdresser, daycare, etc.)	Number	11	8	13	20	11	9	16	16
	Studying/ working	Number	4	3	6	7	4	2	9	5
	Walk, exercise, and leisure	Number	41	27	71	73	33	37	92	50

APPENDIX C.3 CROSS-TABULATION (CONTINGENCY TABLE) (cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Trip-specific Information	Passing between two destinations (including access to the residence)	Number	60	55	88	83	51	60	102	73
	I prefer not to answer	Number	1	0	0	0	1	0	0	0
	other purposes	Number	4	6	4	6	4	5	8	3
	Trip frequency	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	3	0	0	0	3	0	0	0
	Only once	Number	5	3	4	4	3	4	5	4
	Few times a month	Number	25	48	90	99	23	51	109	79
	Few times a week	Number	103	101	187	175	92	96	233	145
	Everyday	Number	76	49	96	102	71	60	106	86
	Trip Companion	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Alone	Number	70	66	123	107	67	69	143	87
	With friends	Number	17	16	30	40	13	20	40	30
	In a couple	Number	55	53	108	104	51	59	117	93
	With family (parents and/or children)	Number	63	64	111	126	55	60	149	100
	Other types of companionships	Number	7	2	5	3	6	3	4	4
	Streets Visiting time	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Morning (before 9 a.m.)	Number	0	2	4	7	0	2	5	6
	Before noon (9 a.m. - 11 a.m.)	Number	9	10	17	12	7	12	17	12
	Lunch time (11 a.m. - 13 p.m.)	Number	6	5	10	13	5	6	11	12
	Afternoon (13 p.m. - 17 p.m.)	Number	77	73	106	104	73	74	137	76

APPENDIX C.4 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Trip-specific Information	Early evening (17 p.m. - 19 p.m.)	Number	65	71	134	124	58	72	161	103
	Night	Number	55	40	106	120	49	45	122	105
	Walking during Covid-19 (including Mobility Aid)	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	0	1	0	0	0	1	0	0
	Never	Number	1	1	0	3	1	1	1	2
	Few times in a year	Number	0	0	1	1	0	0	2	0
	Few times in a month	Number	5	8	9	14	3	7	14	12
	Few times in a week	Number	36	49	84	89	40	50	98	70
	Everyday	Number	170	142	283	273	148	152	338	230
	Cycling During Covid -19	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	2	4	0	1	3	2	2	0
	Never	Number	94	62	62	32	82	66	73	29
	Few times in a year	Number	25	19	20	10	22	17	25	10
	Few times in a month	Number	28	31	52	42	26	37	58	32
	Few times in a week	Number	52	52	146	148	45	52	174	127
	Everyday	Number	11	33	97	147	14	37	121	116

APPENDIX C.5 CROSS-TABULATION (CONTINGENCY TABLE) (cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Trip-specific Information	Two-wheelers during Covid-19	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	1	2	5	3	1	3	5	2
	Never	Number	196	183	342	330	179	192	403	277
	Few times in a year	Number	1	3	10	18	1	3	15	13
	Few times in a month	Number	8	5	9	12	7	7	10	10
	Few times in a week	Number	4	7	11	14	4	3	19	10
	Everyday	Number	2	1	0	3	0	3	1	2
	Public Transport using During Covid-19	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	1	1	1	1	1	1	1	1
	Never	Number	52	53	85	93	49	49	113	72
	Few times in a year	Number	54	42	103	97	49	52	109	86
	Few times in a month	Number	51	63	124	117	45	58	154	98
	Few times in a week	Number	40	36	53	62	37	42	65	47
	Everyday	Number	14	6	11	10	11	9	11	10
	Adaptive transport using during Covid-19	No.	VD	RD	RS	VS	VD	RD	RS	VS

APPENDIX C.6 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Trip-specific Information	I prefer not to answer	Number	4	5	6	14	4	5	9	11
	Never	Number	199	190	366	355	182	197	438	293
	Few times in a year	Number	7	3	2	3	5	5	2	3
	Few times in a month	Number	1	2	1	5	0	3	2	4
	Few times in a week	Number	0	0	2	3	0	1	1	3
	Everyday	Number	1	1	0	0	1	0	1	0
	Paratransit using during Covid-19	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	1	4	1	4	2	2	3	3
	Never	Number	118	115	210	212	104	118	263	170
	Few times in a year	Number	58	60	140	124	57	68	149	108
	Few times in a month	Number	33	21	24	39	27	22	36	32
	Few times in a week	Number	2	1	2	1	2	1	2	1
	Car Sharing (Communauto) using During Covid-19	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	2	5	2	8	3	2	7	5
	Never	Number	175	147	255	223	154	162	298	186

APPENDIX C.7 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Trip-specific Information	Few times in a year	Number	11	27	60	59	12	23	71	51
	Few times in a month	Number	19	17	50	69	19	17	62	57
	Few times in a week	Number	4	5	8	21	3	7	13	15
	Everyday	Number	1	0	2	0	1	0	2	0
	Motor vehicle driving during Covid-19	No.	VD	RD	RS	VS	VD	RD	RS	VS
	I prefer not to answer	Number	3	5	6	7	4	4	7	6
	Never	Number	77	84	181	170	68	96	205	143
	Few times in a year	Number	3	13	22	22	1	13	30	16
	Few times in a month	Number	44	33	80	88	44	33	97	71
	Few times in a week	Number	71	55	77	72	62	54	98	61
	Everyday	Number	14	11	11	21	13	11	16	17
	Mode of transport during summer with pedestrianization	No.	VD	RD	RS	VS	VD	RD	RS	VS
Category: Season-Related Mode of Transport	Walking (including Mobility Aids)	Number	153	120	159	100	137	126	187	82

APPENDIX C.8 CROSS-TABULATION (CONTINGENCY TABLE)

(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Season-Related Mode of Transport	Cycling (personal, Bixi, electric or non)	Number	18	45	161	217	18	51	194	178
	Public transport (bus, metro)	Number	13	16	26	22	11	16	31	19
	Personal motor vehicle (automobile, truck, motorcycle, etc.)	Number	13	16	18	22	14	11	24	20
	Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates)	Number	3	0	5	10	2	1	7	8
	Adaptive transport	Number	2	1	0	0	2	1	0	0
	Paratransit (Uber, taxi, etc.)	Number	2	1	2	2	1	2	3	1
	Car sharing (Communauto)	Number	0	2	6	5	0	2	6	5
	I do not like to pass pedestrian streets at this time	Number	8	0	0	2	7	1	1	1
	Mode of transport during summer without pedestrianization	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Walking (including Mobility Aids)	Number	168	134	191	147	151	139	227	123
	Cycling (personal, Bixi, electric or non)	Number	16	41	129	173	18	44	155	142

APPENDIX C.9 CROSS-TABULATION (CONTINGENCY TABLE) (cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Season-Related Mode of Transport	Public transport (bus, metro)	Number	19	20	33	34	16	21	40	29
	Personal motor vehicle (automobile, truck, motorcycle, etc.)	Number	1	0	5	5	0	1	7	3
	Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates)	Number	2	1	0	0	2	1	0	0
	Adaptive transport	Number	3	1	4	2	2	2	4	2
	Paratransit (Uber, taxi, etc.)	Number	0	3	9	14	1	2	11	12
	Car sharing (Communauto)	Number	3	1	6	5	2	1	9	3
	I do not like to pass pedestrian streets at this time	Number	8	0	0	2	7	1	1	1
	Mode of transport during Winter without pedestrianization	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Walking (including Mobility Aids)	Number	139	118	197	154	125	122	233	128
	Cycling (personal, Bixi, electric or non)	Number	6	7	46	74	6	8	54	65
	Public transport (bus, metro)	Number	20	31	49	60	18	34	59	49

APPENDIX C.10 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Season-Related Mode of Transport	Personal motor vehicle (automobile, truck, motorcycle, etc.)	Number	37	36	61	63	33	39	75	50
	Riding Electric or non-electric 2-wheelers (scooter, skateboard, inline skates)	Number	0	0	1	0	0	0	1	0
	Adaptive transport	Number	2	1	1	1	2	1	1	1
	Paratransit (Uber, taxi, etc.)	Number	5	1	5	4	4	2	6	3
	Car sharing (Communauto)	Number	0	3	11	18	1	2	14	15
	I do not like to pass pedestrian streets at this time	Number	3	4	6	6	3	3	10	3
Category: Mutual behavior of cyclists and pedestrians	Behavior of cyclists toward pedestrians regarding priority crossing	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	187	63	4	6	165	76	15	4
	Rather dissatisfied	Number	18	85	64	15	18	85	67	12
	Neither satisfied nor dissatisfied	Number	3	21	57	7	3	22	57	6
	Rather satisfied	Number	1	21	210	177	4	20	253	132
	Very satisfied	Number	3	11	42	175	2	8	61	160

APPENDIX C.11 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Mutual behavior of cyclists and pedestrians	Behavior of cyclists toward pedestrians regarding speed	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	183	92	20	7	167	102	28	5
	Rather dissatisfied	Number	23	67	93	35	20	69	100	29
	Neither satisfied nor dissatisfied	Number	2	20	84	23	1	20	93	15
	Rather satisfied	Number	2	17	155	184	2	14	200	142
	Very satisfied	Number	2	5	25	131	2	6	32	123
	Behavior of pedestrians toward cyclists regarding sharing the road	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Not cyclist	Number	107	93	81	41	93	96	100	33
	Very dissatisfied	Number	5	5	1	3	5	5	2	2
	Rather dissatisfied	Number	9	16	17	8	9	16	19	6
	Neither satisfied nor dissatisfied	Number	24	30	31	20	24	27	41	13
	Rather satisfied	Number	37	40	173	129	35	42	203	99
	Very satisfied	Number	30	17	74	179	26	25	88	161
	Behavior of pedestrians toward cyclists regarding priority crossing	No.	VD	RD	RS	VS	VD	RD	RS	VS

APPENDIX C.12 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Mutual behavior of cyclists and pedestrians	Not cyclist	Number	107	93	83	93	96	102	33	33
	Very dissatisfied	Number	4	3	1	4	3	2	1	2
	Rather dissatisfied	Number	11	13	21	9	13	22	10	6
	Neither satisfied nor dissatisfied	Number	19	31	50	19	29	62	15	13
	Rather satisfied	Number	40	42	160	39	45	190	109	99
	Very satisfied	Number	31	19	62	28	25	75	146	161
Category: General Opinion and attitude toward pedestrianization	Hangout more often on pedestrianized streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	52	19	9	5	54	18	8	5
	Rather dissatisfied	Number	29	8	3	3	22	12	7	2
	Neither satisfied nor dissatisfied	Number	39	37	39	21	32	42	43	19
	Rather satisfied	Number	43	59	81	50	38	55	98	42
	Very satisfied	Number	49	78	245	301	46	84	297	246
	Hangout long time in pedestrianized streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	52	17	11	6	52	17	10	7
	Rather dissatisfied	Number	28	8	9	2	23	11	12	1

APPENDIX C.13 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: General Opinion and attitude toward pedestrianization	Neither satisfied nor dissatisfied	Number	40	40	47	23	34	46	51	19
	Rather satisfied	Number	40	63	74	67	36	57	96	55
	Very satisfied	Number	52	73	236	282	47	80	284	232
	Hangout more often on pedestrianized streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	52	19	9	5	54	18	8	5
	Rather dissatisfied	Number	29	8	3	3	22	12	7	2
	Neither satisfied nor dissatisfied	Number	39	37	39	21	32	42	43	19
	Rather satisfied	Number	43	59	81	50	38	55	98	42
	Very satisfied	Number	49	78	245	301	46	84	297	246
	Hangout long time in pedestrianized streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	52	17	11	6	52	17	10	7
	Rather dissatisfied	Number	28	8	9	2	23	11	12	1
	Neither satisfied nor dissatisfied	Number	40	40	47	23	34	46	51	19
	Rather satisfied	Number	40	63	74	67	36	57	96	55
	Very satisfied	Number	52	73	236	282	47	80	284	232

APPENDIX C.14 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: General Opinion and attitude toward pedestrianization	Hangout more in shops on pedestrianized streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	70	29	21	10	68	31	20	11
	Rather dissatisfied	Number	25	17	24	9	20	19	29	7
	Neither satisfied nor dissatisfied	Number	51	65	101	62	45	67	114	53
	Rather satisfied	Number	32	54	91	84	28	51	113	69
	Very satisfied	Number	34	36	140	215	31	43	177	174
	Changing the route for passing pedestrianized streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	64	29	20	6	60	29	22	8
	Rather dissatisfied	Number	23	19	18	7	20	18	24	5
	Neither satisfied nor dissatisfied	Number	35	42	48	30	33	43	51	28
	Rather satisfied	Number	46	51	119	88	40	61	132	71
	Very satisfied	Number	44	60	172	249	39	60	224	202
	Changing the route to avoid pedestrianized streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	68	73	211	246	62	73	252	211

APPENDIX C.15 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: General Opinion and attitude toward pedestrianization	Rather dissatisfied	Number	25	35	62	47	21	42	72	34
	Neither satisfied nor dissatisfied	Number	31	38	43	39	28	41	54	28
	Rather satisfied	Number	41	37	42	32	34	40	50	28
	Very satisfied	Number	47	18	19	16	47	15	25	13
	General satisfaction of pedestrianization	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	30	6	2	4	30	6	2	4
	Rather dissatisfied	Number	24	15	5	2	20	17	8	1
	Neither satisfied nor dissatisfied	Number	3	3	1	1	3	3	1	1
	Rather satisfied	Number	66	52	62	30	59	56	71	24
	Very satisfied	Number	89	125	307	343	80	129	371	284
	Continuation of pedestrianization for the future on commercial streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	45	13	6	4	42	16	6	4
	Rather dissatisfied	Number	19	13	8	2	16	14	10	2
	Neither satisfied nor dissatisfied	Number	3	5	1	2	3	5	1	2
	Rather satisfied	Number	57	41	44	25	50	42	57	18

APPENDIX C.16 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: General Opinion and attitude toward pedestrianization	Very satisfied	Number	88	129	318	347	81	134	379	288
	Extend pedestrianization into other streets	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	56	16	7	5	53	17	10	4
	Rather dissatisfied	Number	23	25	14	10	22	23	18	9
	Neither satisfied nor dissatisfied	Number	13	16	13	6	11	15	18	4
	Rather satisfied	Number	51	47	72	39	42	54	83	30
	Very satisfied	Number	69	97	271	320	64	102	324	267
	Impact of pedestrianization on mobility	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Very dissatisfied	Number	38	7	3	0	38	7	2	1
	Rather dissatisfied	Number	29	14	11	4	24	20	12	2
	Neither satisfied nor dissatisfied	Number	45	38	31	20	41	38	38	17
	Rather satisfied	Number	45	64	102	59	38	59	125	48
	Very satisfied	Number	55	78	230	297	51	87	276	246
Category: Socioeconomic Character	Gender	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Female	Number	120	138	245	229	110	130	303	189

APPENDIX C.17 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Socioeconomic Characteristics	Male	Number	83	58	122	144	72	77	139	119
	Other	Number	9	5	10	7	10	4	11	6
	Number of household members	No.	VD	RD	RS	VS	VD	RD	RS	VS
	1-member household	Number	67	63	113	87	59	75	127	69
	2-member household	Number	78	74	130	133	78	69	153	115
	3-member household	Number	30	32	66	67	26	34	78	57
	4-member household and more	Number	37	32	68	93	29	33	95	73
	Mobility Restriction	No.	VD	RD	RS	VS	VD	RD	RS	VS
	No	Number	193	193	369	371	173	205	439	309
	Yes	Number	19	8	8	9	19	6	14	5
	Disability - Walking	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Yes	Number	19	4	9	13	19	5	12	9
	No	Number	193	197	368	367	173	206	441	305
	Disability - Cycling	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Yes	Number	25	16	17	14	23	18	19	12
	No	Number	187	185	360	366	169	193	434	302

APPENDIX C.18 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Socioeconomic Characteristics	Bus Riding Disability	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Yes	Numb er	8	6	5	9	8	5	6	9
	No	Numb er	204	195	372	371	184	206	447	305
	Being a regular cyclist	No.	VD	RD	RS	VS	VD	RD	RS	VS
	No	Numb er	105	91	80	41	91	94	99	33
	Yes	Numb er	107	110	297	339	101	117	354	281
	Accessibility Transportation Options	No.	VD	RD	RS	VS	VD	RD	RS	VS
	Driver's license ownership	Numb er	10	2	1	5	8	4	1	5
	Car, truck and motorcycle ownership	Numb er	13	4	4	4	12	4	6	3
	Transit pass ownership	Numb er	38	38	37	8	34	35	45	7
	Bike ownership	Numb er	73	78	155	164	65	86	191	128
	Carsharing users (Communauto)	Numb er	28	25	60	60	24	24	70	55
	BIXI users (Bike sharing)	Numb er	41	44	113	125	40	49	128	106
	Other modes	Numb er	9	10	7	14	9	9	12	10
	Motorized vehicle ownership	No.	VD	RD	RS	VS	VD	RD	RS	VS

APPENDIX C.19 CROSS-TABULATION (CONTINGENCY TABLE)
(cont'd)

Variable Category	Variables	Count	Y1				Y2			
Category: Socioeconomic Characteristics	I do not own a motorized vehicle	Number	83	101	202	202	75	111	233	169
	Automobile	Number	120	94	164	161	110	92	203	134
	Truck	Number	9	6	11	15	7	8	16	10
	Motorcycle	Number	0	0	0	2	0	0	1	1

APPENDIX D CHI-SQUARED TEST

Y1: Satisfaction with the coexistence (pedestrians with "traditional" cyclists),				
Question No.	Description	Chi-square statistic	P-value	Df
q0004	Favorite Pedestrian Street	80.016	0.000	33
q0005	Trip purpose	20.258	0.505	21
q0006	Trip frequency	36.487	0.000	12
q0007.0001_6	Trip companion while visiting the street	11.604	0.478	12
q0008.0001_7	Streets Visiting time	20.580	0.151	15
q0017_0001	Utilitarian walking During Covid (including Mobility Aid)	15.636	0.407	15
q0017_0002	Cycling (personal, Bixi, electric or non) During Covid-19	212.364	0.000	15
q0017_0003	Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19	18.419	0.241	15
q0017_0004	Public Transport using During Covid-19	16.536	0.347	15
q0017_0005	Adaptive transport using during Covid-19	23.017	0.084	15
q0017_0006	Paratransit using (Uber, taxi, etc.) During Covid-19	22.636	0.031	12
q0017_0007	Car Sharing (Communauto) using During Covid-19	56.458	0.000	15
q0017_0008	Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19	35.156	0.002	15
q0018.1.1_10	Preferred mode of transport during summer with pedestrianization	220.910	0.000	24
q0018.2.1_10	Preferred mode of transport during summer without pedestrianization	142.280	0.000	21
q0018.3.1_10	Preferred mode of transport during Winter without pedestrianization	90.303	0.000	24
q0014_0004	Intersection safety	237.645	0.000	12
q0014_0005	Universal accessibility	153.924	0.000	12
q0014_0006	Attractiveness	229.648	0.000	12
q0014_0007	Personal security	291.885	0.000	12

APPENDIX D.1 CHI-SQUARE TEST (cont'd)

Y1: Satisfaction with the coexistence (pedestrians with "traditional" cyclists),				
Question No.	Description	Chi-square statistic	P-value	Df
q0014_0008	Cleanliness	184.667	0.000	12
q0014_0009	Street furniture	190.422	0.000	12
q0019_0001	Behavior of cyclists toward pedestrians regarding speed	1134.242	0.000	12
q0019_0002	Behavior of cyclists toward pedestrians regarding priority crossing	925.313	0.000	12
q0021_0001	The behavior of pedestrians toward cyclists regarding sharing the road	297.689	0.000	15
q0021_0002	The behavior of pedestrians toward cyclists regarding priority crossing	259.269	0.000	15
q0022_0001	Hangout more often on pedestrianized streets	331.843	0.000	12
q0022_0002	Hangout long time in pedestrianized streets	290.086	0.000	12
q0022_0003	Hangout more in shops on pedestrianized streets	258.798	0.000	12
q0022_0004	Changing the route for passing pedestrianized streets	248.507	0.000	12
q0022_0005	Changing the route to avoid pedestrianized streets	130.497	0.000	12
q0023_0001	General satisfaction of pedestrianization	240.868	0.000	12
q0023_0002	Continuation of pedestrianization for the future on commercial	254.106	0.000	12
q0023_0003	Extend pedestrianization into other streets	278.242	0.000	12
q0025	Impact of pedestrianization on mobility	315.322	0.000	12
q0015.0001_8	Accessibility Transportation Options	110.767	0.000	18
q0016.0001_3	Motorized vehicle ownership	18.693	0.028	9
q0020	Being a regular cyclist	145.372	0.000	3
q0026	Gender	10.950	0.090	6
q0027.0001_.3	Number of household members	14.305	0.112	9
q0030	Mobility restriction	20.707	0.000	3
q0031_0001	Disability - Walking	19.232	0.000	3
q0031_0002	Disability - Cycling	18.586	0.000	3
q0031_0003	Bus Riding Disability	3.869	0.276	3

APPENDIX D.2 CHI-SQUARE TEST (cont'd)

Y2: Satisfaction with the coexistence of pedestrians and non-electric any type of 2-wheeled riders (skates, scooters, bikes, etc.)				
Question No.	Description	Chi-square statistic	P-value	Df
q0004	Favorite Pedestrian Street	59.361	0.003	33
q0005	Trip purpose	19.871	0.529	21
q0006	Trip frequency	37.879	0.000	12
q0007.0001_6	Trip companion while visiting the street	10.645	0.560	12
q0008.0001_7	Streets Visiting time	24.422	0.058	15
q0017_0001	Utilitarian walking During Covid (including Mobility Aid)	11.402	0.724	15
q0017_0002	Cycling (personal, Bixi, electric or non) During Covid-19	169.258	0.000	15
q0017_0003	Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19	20.218	0.164	15
q0017_0004	Public Transport using During Covid-19	16.296	0.363	15
q0017_0005	Adaptive transport using during Covid-19	19.502	0.192	15
q0017_0006	Paratransit using (Uber, taxi, etc.) During Covid-19	8.271	0.764	12
q0017_0007	Car Sharing (Communauto) using During Covid-19	41.084	0.000	15
q0017_0008	Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19	30.881	0.009	15
q0018.1.1_10	Preferred mode of transport during summer with pedestrianization	187.421	0.000	24
q0018.2.1_10	Preferred mode of transport during summer without pedestrianization	121.629	0.000	21
q0018.3.1_10	Preferred mode of transport during Winter without pedestrianization	84.743	0.000	24
q0014_0004	Intersection safety	265.163	0.000	12
q0014_0005	Universal accessibility	177.897	0.000	12
q0014_0006	Attractiveness	214.261	0.000	12
q0014_0007	Personal security	291.032	0.000	12

APPENDIX D.3 CHI-SQUARE TEST (cont'd)

Y2: Satisfaction with the coexistence of pedestrians and non-electric any type of 2-wheeled riders (skates, scooters, bikes, etc.)				
Question No.	Description	Chi-square statistic	P-value	Df
q0014_0008	Cleanliness	179.054	0.000	12
q0014_0009	Street furniture	201.325	0.000	12
q0019_0001	Behavior of cyclists toward pedestrians regarding speed	1044.154	0.000	12
q0019_0002	Behavior of cyclists toward pedestrians regarding priority crossing	917.819	0.000	12
q0021_0001	The behavior of pedestrians toward cyclists regarding sharing the road	283.242	0.000	15
q0021_0002	The behavior of pedestrians toward cyclists regarding priority crossing	245.113	0.000	15
q0022_0001	Hangout more often on pedestrianized streets	311.780	0.000	12
q0022_0002	Hangout long time in pedestrianized streets	277.108	0.000	12
q0022_0003	Hangout more in shops on pedestrianized streets	242.635	0.000	12
q0022_0004	Changing the route for passing pedestrianized streets	225.872	0.000	12
q0022_0005	Changing the route to avoid pedestrianized streets	147.293	0.000	12
q0023_0001	General satisfaction of pedestrianization	239.650	0.000	12
q0023_0002	Continuation of pedestrianization for the future on commercial	236.424	0.000	12
q0023_0003	Extend pedestrianization into other streets	262.409	0.000	12
q0025	Impact of pedestrianization on mobility	309.469	0.000	12
q0015.0001_8	Accessibility Transportation Options	88.385	0.000	18
q0016.0001_3	Motorized vehicle ownership	13.714	0.133	9
q0020	Being a regular cyclist	122.628	0.000	3
q0026	Gender	12.096	0.060	6

APPENDIX D.4 CHI-SQUARE TEST (cont'd)

Y2: Satisfaction with the coexistence of pedestrians and non-electric any type of 2-wheeled riders (skates, scooters, bikes, etc.)				
Question No.	Description	Chi-square statistic	P-value	Degrees of freedom
q0027.0001_.3	Number of households	19.047	0.025	9
q0030	Mobility restriction	25.099	0.000	3
q0031_0001	Disability - Walking	22.815	0.000	3
q0031_0002	Disability - Cycling	19.315	0.000	3
q0031_0003	Bus Riding Disability	5.101	0.165	3

APPENDIX E VIF (VARIANCE INFLATION FACTOR)

	Y1	Y2
Variables	VIF	
q0021_0001_The behavior of pedestrians toward cyclists regarding sharing the road	14.17	14.17
q0021_0002_The behavior of pedestrians toward cyclists regarding priority crossing	13.22	13.22
q020_Being a regular cyclist	7.25	7.25
q0022_0001_Hangout more often on pedestrianized streets	7.21	7.21
q0023_0002_Continuation of pedestrianization for the future on commercial streets	6.72	6.72
q0022_0002_Hangout long time in pedestrianized streets	6.48	6.48
q0023_0001_General satisfaction of pedestrianization	5.59	5.59
q0019_0001_Behavior of cyclists toward pedestrians regarding speed	5.43	5.43
q0019_0002_Behavior of cyclists toward pedestrians regarding priority crossing	5.13	5.13
q0025_Impact of pedestrianization on mobility	3.78	3.78
q0023_0003_Extend pedestrianization into other streets	3.74	3.74
q0022_0004_Changing the route for passing pedestrianized streets	3.37	3.37
q0022_0003_Hangout more in shops on pedestrianized streets	3.2	3.2
q0016_0001_0003_Motorized_vehicle_ownership	2.48	2.48
q0017_0008_Motor vehicle driving (automobile, truck, motorcycle, etc.) During Covid-19	2.47	2.47
q0014_0006_Attractiveness	2.46	2.46
q0017_0002_Do cycling (personal, BIXI, electric or non) During Covid-19	2.23	2.23
q0018.0002.1_10_Preferred mode of transport during summer without pedestrianization	2.08	2.08
q0018.0003.1_10_Preferred mode of transport during Winter without pedestrianization	1.97	1.97
q0014_0009_Street_Furniture	1.94	1.94

APPENDIX E.1 VIF (cont'd)

	Y1	Y2
Variables	VIF	
q0014_0007_Personal Security	1.92	1.92
q0031_0001_Disability - Walking	1.88	1.88
q0030_Mobility Restriction	1.71	1.71
q0014_0008_Cleanliness	1.7	1.7
q0022_0005_Changing the route to avoid pedestrianized streets	1.64	1.64
q0031_0002_Disability - Cycling	1.62	1.62
q0018.0001.1_10_Preferred mode of transport during summer with pedestrianization	1.62	1.62
q0031_0003_Bus Riding Disability	1.58	1.58
q0027_Number of Household members	1.46	1.46
q0014_0005_Universal Accessibility	1.44	1.44
q0007.0001_6_Trip_Companionship	1.43	1.43
q0014_0004_Intersection Safety	1.43	1.43
q0017_0007_Car Sharing (Communauto) using During Covid-19	1.38	1.38
q0015.0001_8_Accessibility_Transportation_Options	1.29	1.29
q0006_Trip frequency	1.28	1.28
q0017_0005_Adaptive transport using during Covid-19	1.21	1.21
q0017_0004_Public Transport using during Covid-19	1.19	1.19
q0017_0001_Utilitarian walking During Covid-19 (including Mobility Aid)	1.18	1.18
q0008.0001_7_Streets Visiting time	1.17	1.17
2.2 Favorite Pedestrian Street	1.14	1.14

APPENDIX E.2 VIF (cont'd)

	Y1	Y2
Variables	VIF	
q0026_Gender	1.11	1.11
q0017_0006_ Paratransit using (Uber, taxi, etc.) During Covid-19	1.11	1.11
3.1.1 Trip purpose	1.1	1.1
q0017_0003_ Two-wheelers (scooter, board casters, skates in-line wheels, etc. electric or non) using during Covid-19	1.1	1.1

APPENDIX F LASSO REGRESSION COEFFICIENT

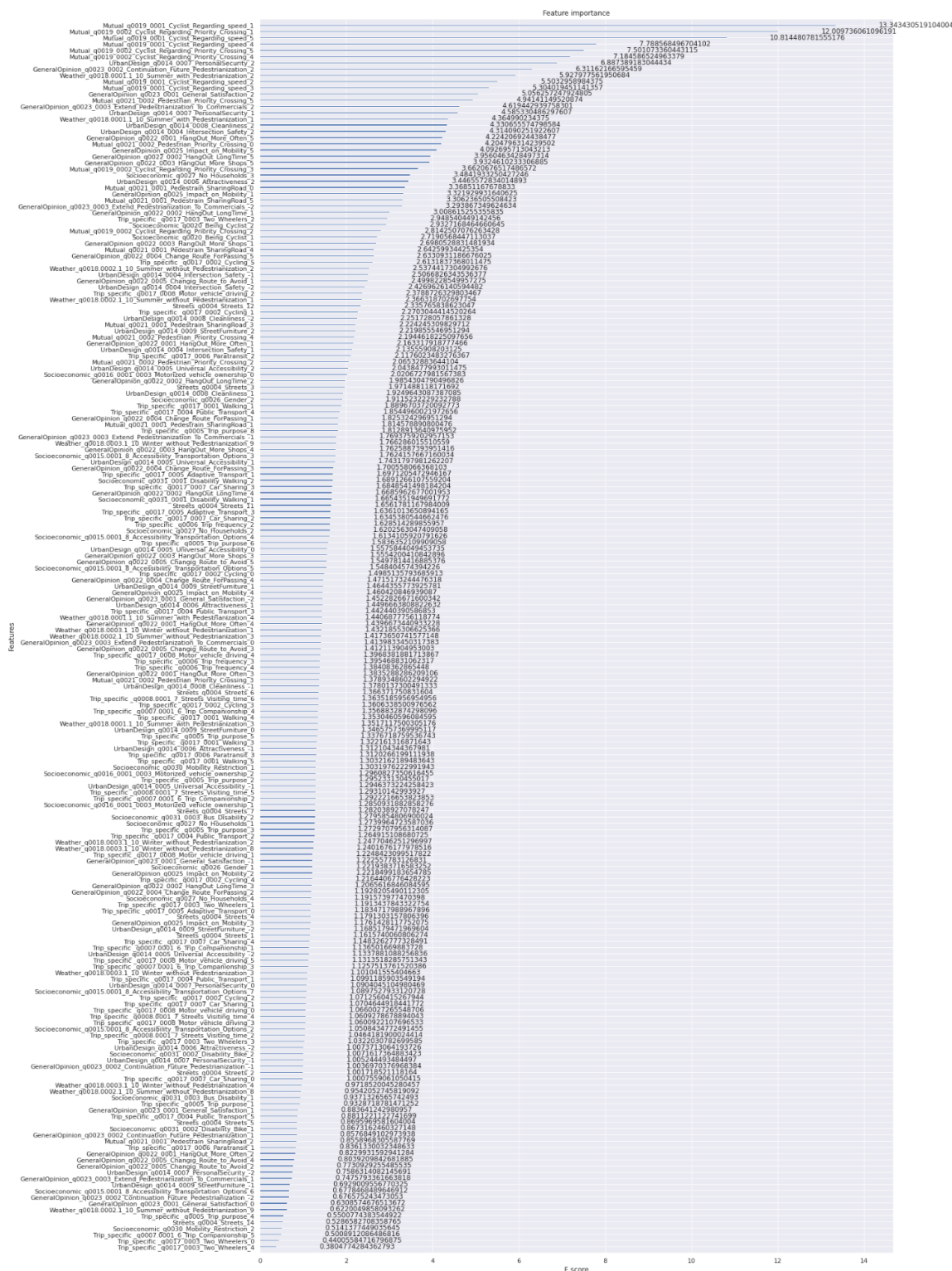
Lasso best coefficient_Y1	Variables
0.516	'q0019_0001_Behavior of cyclists toward pedestrians regarding speed'
0.119	'q0019_0002_Behavior of cyclists toward pedestrians regarding priority crossing'
0.099	'q0014_0004_Intersection Safety'
0.068	'q0025_Impact of pedestrianization on mobility'
0.062	'q0014_0006_Attractiveness'
0.059	'q0022_0004_Changing the route for passing pedestrianized streets'
0.052	'q0014_0005_Universal Accessibility'
0.042	'q0018.0001.1_10_Preferred mode of transport during summer with pedestrianization'
0.031	'q0014_0007_Personal Security'
0.027	'q0014_0008_Cleanliness'
0.022	'q0004_Favorite Pedestrian Streets'
0.012	'q0027_Number of household members '
0.012	'q0021_0001_The behavior of pedestrians toward cyclists regarding sharing the road'
0.006	'q0022_0001_Hangout more often on pedestrianized streets'
-0.015	'q0026_Gender'
-0.016	'q0005_Trip purpose'
-0.026	'q0022_0005_Changing the route to avoid pedestrianized streets'
Lasso best coefficient_Y2	Variables
0.852	'q0019_0001_Behavior of cyclists toward pedestrians regarding speed '
0.177	'q0019_0002_Behavior of cyclists toward pedestrians regarding priority crossing'
0.139	'q0014_0004_Intersection Safety'
0.088	'q0014_0006_Attractiveness'
0.087	'q0025_Impact of pedestrianization on mobility'
0.085	'q0022_0004_Changing the route for passing pedestrianized streets'
0.033	'q0017_0002_Do cycling (personal, BIXI, electric or non) During Covid-19'
0.03	'q0018.0001.1_10_Preferred mode of transport during summer with pedestrianization'
0.005	'q0021_0001_The behavior of pedestrians toward cyclists regarding sharing the road'

APPENDIX F.1 LASSO REGRESSION COEFFICIENT (cont ' d)

Lasso best coefficient_Y2	APPENDIX F LASSO REGRESSION COEFFICIENT
	Variables
0.004	'q0014_0005_Universal Accessibility'
0.002	'q0004_Favorite Pedestrian Streets'
0.001	'q0020_Being a regular cyclist'
-0.001	'q0005_Trip purpose'
-0.024	'q0022_0005_Changing the route to avoid pedestrianized streets'

APPENDIX G BEST HYPERPARAMETERS FOR XGB AND RF

XGBoost Hyperparameters Optimal value after tuning for Y1	"subsample": 0.5, "reg_alpha": 0.4, "num_leaves": 120, "num_class": 6, "n_estimators": 300, "min_data_in_leaf": 30, "max_depth": 3, "learning_rate": 0.01, "gamma": 0.1, "colsample_bytree": 0.6, "colsample_bynode": 0.8999999999999999, "colsample_bylevel": 0.5
XGBoost Hyperparameters Optimal value after tuning for Y2	"subsample": 0.7999999999999999, "reg_alpha": 0.3, "num_leaves": 100, "num_class": 10, "n_estimators": 100, "min_data_in_leaf": 40, "max_depth": 3, "learning_rate": 0.05, "gamma": 0.4, "colsample_bytree": 0.7, "colsample_bynode": 0.5, "colsample_bylevel": 0.5
Random Forest Hyperparameters Optimal value after tuning for Y1	bootstrap': False, 'max_depth': 5, 'max_features': 40, 'min_samples_leaf': 8, 'min_samples_split': 2, 'n_estimators': 20
Random Forest Hyperparameters Optimal value after tuning for Y2	bootstrap': True, 'max_depth': 7, 'max_features': 46, 'min_samples_leaf': 8, 'min_samples_split': 2, 'n_estimators': 40



APPENDIX H.1 XGBOOST FEATURES IMPORTANCE PLOT (cont'd)

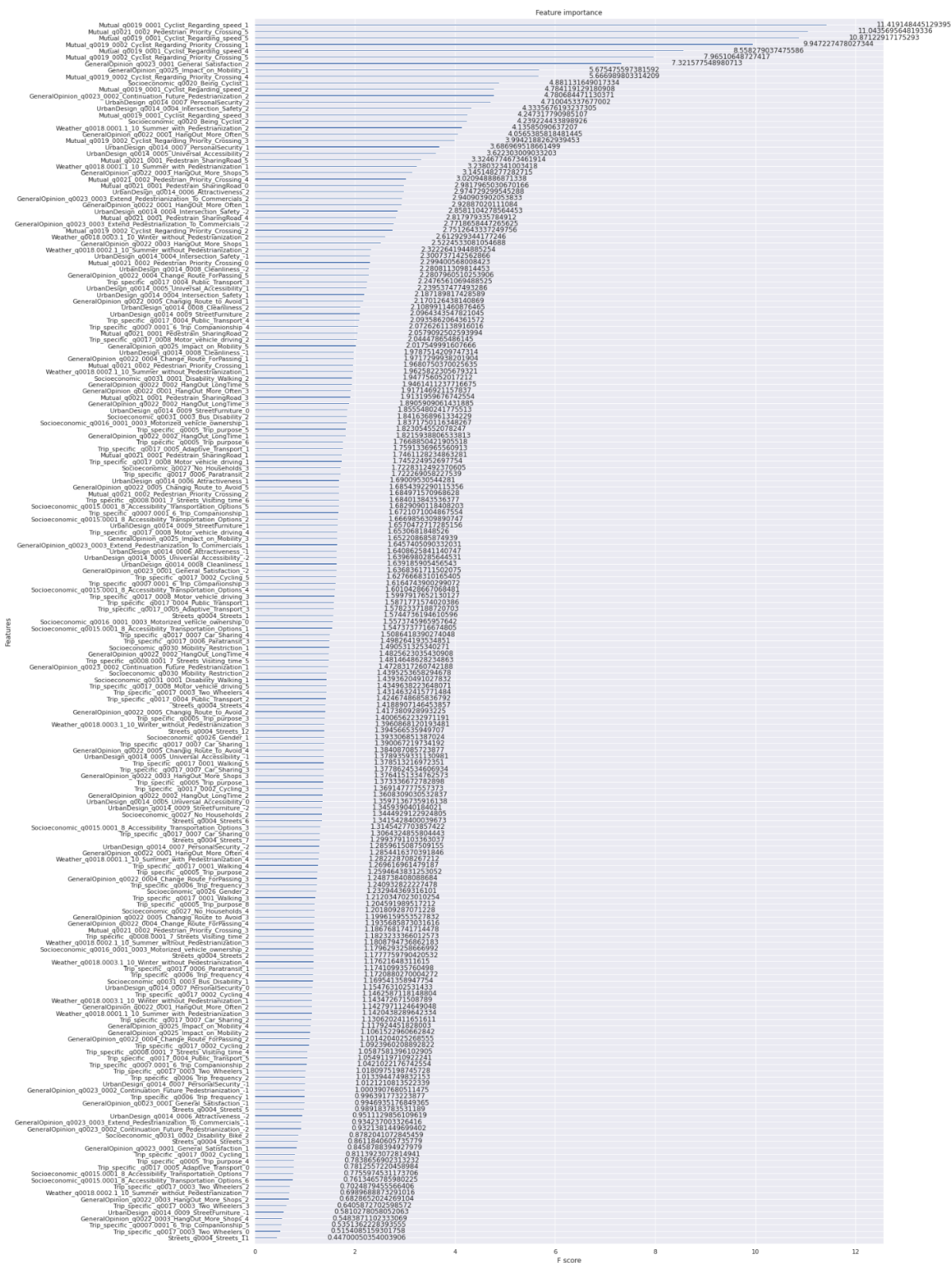


Figure 6.2: XGB Feature importance Y2



APPENDIX I.1 RANDOM FOREST FEATURES IMPORTANCE PLOT

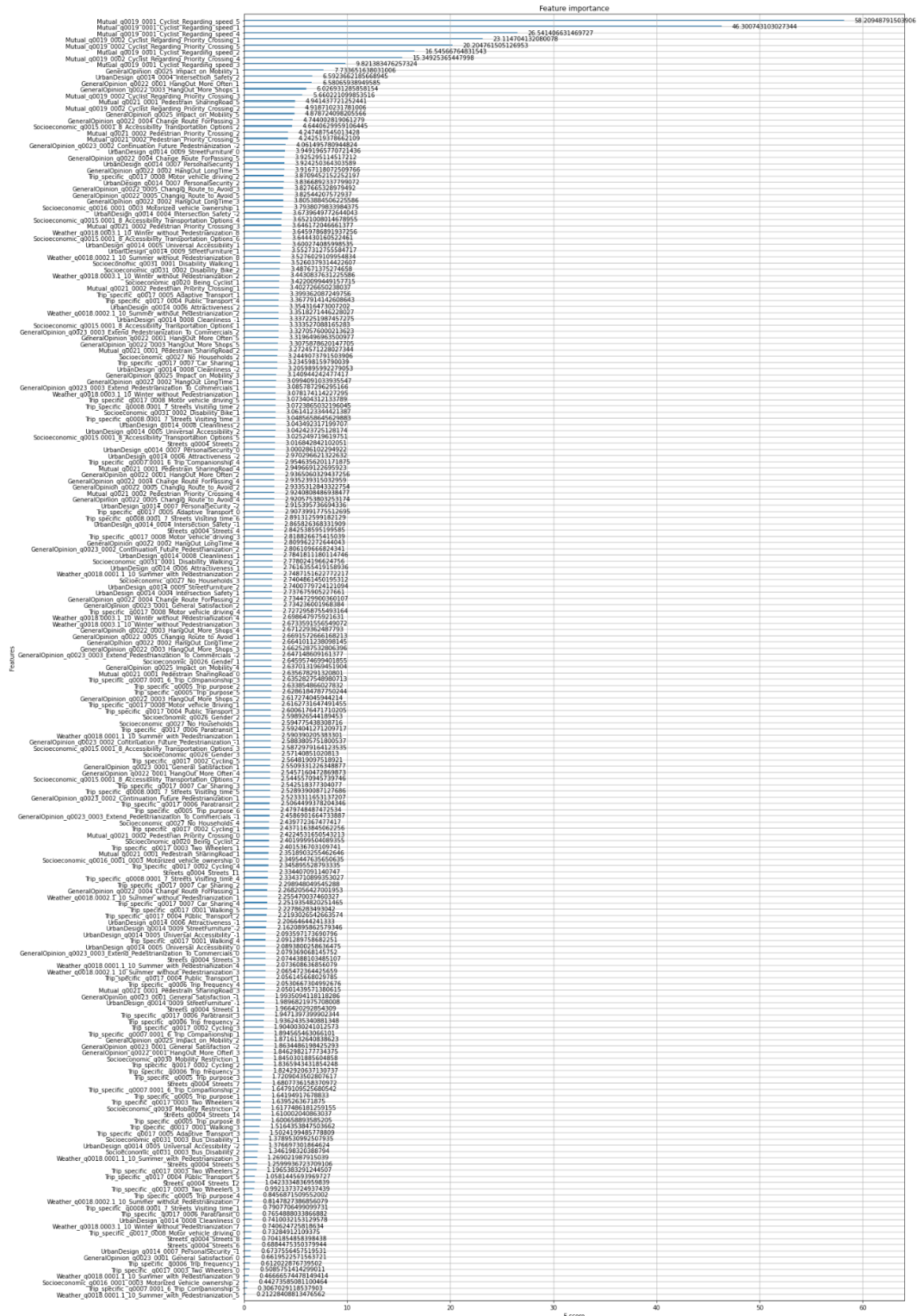


Figure 6.4: RF Feature importance Y2

REFERENCES

- Alasdair Sandford. (2020). *Coronavirus: Half of humanity now on lockdown as 90 countries call for confinement*. Euronews. <https://www.euronews.com/2020/04/02/coronavirus-in-europe-spain-s-death-toll-hits-10-000-after-record-950-new-deaths-in-24-hou>
- Allison, P. D. (2014). Measures of Fit for Logistic Regression. *SAS Global Forum 2014*, 2(1970).
- Alsaleh, R., Hussein, M., & Sayed, T. (2020). Microscopic behavioural analysis of cyclist and pedestrian interactions in shared spaces. *Canadian Journal of Civil Engineering*, 47(1). <https://doi.org/10.1139/cjce-2018-0777>
- Altelbany, S. (2021). Evaluation of Ridge, Elastic Net and Lasso Regression Methods in Precedence of Multicollinearity Problem: A Simulation Study. *Journal of Applied Economics and Business Studies*, 5(1). <https://doi.org/10.34260/jaeb.517>
- Ariffin, R. N. R., & Zahari, R. K. (2013). Perceptions of the Urban Walking Environments. *Procedia - Social and Behavioral Sciences*, 105. <https://doi.org/10.1016/j.sbspro.2013.11.062>
- Asselman, A., Khaldi, M., & Aammou, S. (2021). Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. *Interactive Learning Environments*, 0(0), 1–20. <https://doi.org/10.1080/10494820.2021.1928235>
- Ayilara, O. F., Zhang, L., Sajobi, T. T., Sawatzky, R., Bohm, E., & Lix, L. M. (2019). Impact of missing data on bias and precision when estimating change in patient-reported outcomes from a clinical registry. *Health and Quality of Life Outcomes*, 17(1), 106. <https://doi.org/10.1186/s12955-019-1181-2>
- Beirão, G., & Sarsfield Cabral, J. A. (2007). Understanding attitudes towards public transport and private car: A qualitative study. *Transport Policy*, 14(6). <https://doi.org/10.1016/j.tranpol.2007.04.009>
- Beitel, D., Stipancic, J., Manaugh, K., & Miranda-Moreno, L. (2018). Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis. *Transportation Research Part D: Transport and Environment*, 65. <https://doi.org/10.1016/j.trd.2018.10.001>
- Bergström, A., & Magnusson, R. (2003). Potential of transferring car trips to bicycle during winter. *Transportation Research Part A: Policy and Practice*, 37(8), 649–666. [https://doi.org/https://doi.org/10.1016/S0965-8564\(03\)00012-0](https://doi.org/https://doi.org/10.1016/S0965-8564(03)00012-0)
- Biddulph, M. (2010). Evaluating the English home zone initiatives. *Journal of the American Planning Association*, 76(2). <https://doi.org/10.1080/01944361003622688>
- Biddulph, M. (2012a). Radical streets? the impact of innovative street designs on liveability and activity in residential areas. *Urban Design International*, 17(3). <https://doi.org/10.1057/udi.2012.13>
- Biddulph, M. (2012b). Street Design and Street Use: Comparing Traffic Calmed and Home Zone Streets. *Journal of Urban Design*, 17(2). <https://doi.org/10.1080/13574809.2012.666206>
- Blečić, I., Canu, D., Cecchini, A., Congiu, T., & Fancello, G. (2016). Factors of perceived walkability: A pilot empirical study. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9789. https://doi.org/10.1007/978-3-319-42089-9_9
- Bolboaca, S.-D., & Jäntschi, L. (2006). Pearson versus Spearman, Kendall's tau correlation analysis on structure-activity relationships of biologic active compounds. *Leonardo Journal of Sciences*, 5(9).

- Borst, H. C., de Vries, S. I., Graham, J. M. A., van Dongen, J. E. F., Bakker, I., & Miedema, H. M. E. (2009). Influence of environmental street characteristics on walking route choice of elderly people. *Journal of Environmental Psychology*, 29(4). <https://doi.org/10.1016/j.jenvp.2009.08.002>
- Breiman, L. (2001). Machine Learning, Volume 45, Number 1 - SpringerLink. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Calvey, J. C., Shackleton, J. P., Taylor, M. D., & Llewellyn, R. (2015). Engineering condition assessment of cycling infrastructure: Cyclists' perceptions of satisfaction and comfort. *Transportation Research Part A: Policy and Practice*, 78. <https://doi.org/10.1016/j.tra.2015.04.031>
- Campos, M. B. M. D. A., Chiaraida, A., Smith, A., Stonor, T., & Takamatsu, S. (2003). *Towards a "walkability index."*
- Cervero, R. (1996). Mixed land-uses and commuting: Evidence from the American housing survey. *Transportation Research Part A: Policy and Practice*, 30(5 PART A). [https://doi.org/10.1016/0965-8564\(95\)00033-X](https://doi.org/10.1016/0965-8564(95)00033-X)
- Chapman, D., Nilsson, K., Larsson, A., & Rizzo, A. (2017). Climatic barriers to soft-mobility in winter: Luleå, Sweden as case study. *Sustainable Cities and Society*, 35, 574–580. <https://doi.org/10.1016/j.scs.2017.09.003>
- Chen, T., & Guestrin, C. (2016). A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). *The Journal of the Association of Physicians of India*.
- Chen, Y., Wang, Q. Y., Xi, W. Q., & Mao, J. (2017). Influence of spatial form on pedestrians. *Planners*, 33, 74–80.
- Chong, S., Poulos, R., Olivier, J., Watson, W. L., & Grzebieta, R. (2010). Relative injury severity among vulnerable non-motorised road users: Comparative analysis of injury arising from bicycle-motor vehicle and bicycle-pedestrian collisions. *Accident Analysis and Prevention*, 42(1). <https://doi.org/10.1016/j.aap.2009.08.006>
- Clark, A. F., Scott, D. M., & Yiannakoulis, N. (2014). Examining the relationship between active travel, weather, and the built environment: a multilevel approach using a GPS-enhanced dataset. *Transportation*, 41(2). <https://doi.org/10.1007/s11116-013-9476-3>
- Clarke, P., Hirsch, J. A., Melendez, R., Winters, M., Gould, J. S., Ashe, M., Furst, S., & McKay, H. (2017). Snow and Rain Modify Neighbourhood Walkability for Older Adults. *Canadian Journal on Aging*, 36(2), 159–169. <https://doi.org/10.1017/S071498081700006X>
- Coughenour, C., Paz, A., de la Fuente-Mella, H., & Singh, A. (2015). Multinomial logistic regression to estimate and predict perceptions of bicycle and transportation infrastructure in a sprawling metropolitan area. *Journal of Public Health (Germany)*, 38(4). <https://doi.org/10.1093/pubmed/fdv179>
- de Goeij, M. C. M., van Diepen, M., Jager, K. J., Tripepi, G., Zoccali, C., & Dekker, F. W. (2013). Multiple imputation: dealing with missing data. *Nephrology Dialysis Transplantation*, 28(10), 2415–2420. <https://doi.org/10.1093/ndt/gft221>
- Dextre, J. C., Hughes, M. (Consultant), & Bech, L. (2013). *Cyclists & cycling around the world : creating liveable & bikeable cities*. Fondo Editorial, Pontificia Universidad Católica del Perú.
- Distefano, N., Pulvirenti, G., & Leonardi, S. (2021). Neighbourhood walkability: Elderly's priorities. *Research in Transportation Business and Management*, 40. <https://doi.org/10.1016/j.rtbm.2020.100547>

- Dogru, N., & Subasi, A. (2018). Traffic accident detection using random forest classifier. *2018 15th Learning and Technology Conference, L and T 2018*. <https://doi.org/10.1109/LT.2018.8368509>
- Dong, H., Xu, X., Wang, L., & Pu, F. (2018). Gaofen-3 PolSAR Image Classification via XGBoost and Polarimetric Spatial Information. *Sensors*, 18(2). <https://doi.org/10.3390/s18020611>
- Dong, W., Cao, X., Wu, X., & Dong, Y. (2019). Examining pedestrian satisfaction in gated and open communities: An integration of gradient boosting decision trees and impact-asymmetry analysis. *Landscape and Urban Planning*, 185. <https://doi.org/10.1016/j.landurbplan.2019.02.012>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1). <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Drs. W.P. Vlakveld, dr. Ch. G. & drs. D. A. M. T. (2008). Beleving van verkeersonveiligheid. SWOV.
- DVRPC. (2018). *Curbless Streets - Evaluating Curbless and Shared Space Concepts for use on City of Philadelphia Streets*.
- ECMT. (1996a). *SUSTAINABLE TRANSPORT IN CENTRAL AND EASTERN EUROPEAN CITIES*.
- ECMT. (1996b). *SUSTAINABLE TRANSPORT IN CENTRAL AND EASTERN EUROPEAN CITIES*.
- Feng, W., Ma, C., Zhao, G., & Zhang, R. (2020). FSRF: An Improved Random Forest for Classification. *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 173–178. <https://doi.org/10.1109/AEECA49918.2020.9213456>
- Ferris, J. A. (2009). Missing Data: A Gentle Introduction. *Drug and Alcohol Review*, 28(1). https://doi.org/10.1111/j.1465-3362.2008.00013_4.x
- Flynn, B. S., Dana, G. S., Sears, J., & Aultman-Hall, L. (2012). Weather factor impacts on commuting to work by bicycle. *Preventive Medicine*, 54(2), 122–124. <https://doi.org/10.1016/j.ypmed.2011.11.002>
- Fonseca, F., Ribeiro, P. J. G., Conticelli, E., Jabbari, M., Papageorgiou, G., Tondelli, S., & Ramos, R. A. R. (2022). Built environment attributes and their influence on walkability. *International Journal of Sustainable Transportation*, 16(7), 660–679. <https://doi.org/10.1080/15568318.2021.1914793>
- Fonti, V., & Belitser, E. (2017). Feature selection using lasso. *VU Amsterdam Research Paper in Business Analytics*, 30.
- Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American Customer Satisfaction Index: Nature, purpose, and findings. *Journal of Marketing*, 60(4). <https://doi.org/10.2307/1251898>
- Forsyth, A. (2015). What is a walkable place? The walkability debate in urban design. *Urban Design International*, 20(4). <https://doi.org/10.1057/udi.2015.22>
- Forsyth, A., & Southworth, M. (2008). Guest editorial: Cities afoot - Pedestrians, walkability and urban design. In *Journal of Urban Design* (Vol. 13, Issue 1). <https://doi.org/10.1080/13574800701816896>
- Friman, M., & Fellesson, M. (2009). Service Supply and Customer Satisfaction in Public Transportation: The Quality Paradox. *Journal of Public Transportation*, 12(4). <https://doi.org/10.5038/2375-0901.12.4.4>

- Gan, M., Pan, S., Chen, Y., Cheng, C., Pan, H., & Zhu, X. (2021). Application of the machine learning lightgbm model to the prediction of the water levels of the lower columbia river. *Journal of Marine Science and Engineering*, 9(5). <https://doi.org/10.3390/jmse9050496>
- Gkekas, F., Bigazzi, A., & Gill, G. (2020). Perceived safety and experienced incidents between pedestrians and cyclists in a high-volume non-motorized shared space. *Transportation Research Interdisciplinary Perspectives*, 4. <https://doi.org/10.1016/j.trip.2020.100094>
- Global Designing Cities Initiative. (2020). *Pedestrian Only Streets: Case Study Stroget, Copenhagen*. GDCl. <https://globaldesigningcities.org/publication/global-street-design-guide/streets/pedestrian-priority-spaces/pedestrian-only-streets/pedestrian-streets-case-study-stroget-copenhagen/>
- Gong, Z., Ma, Q., Kan, C., & Qi, Q. (2019). Classifying Street Spaces with Street View Images for a Spatial Indicator of Urban Functions. *Sustainability*, 11(22). <https://doi.org/10.3390/su11226424>
- Grigoroudis, E., & Siskos, Y. (2010). Customer satisfaction evaluation: methods for measuring and implementing service quality. In *International series in operations research management science* (Issue 139).
- Guo, M., Yuan, Z., Janson, B., Peng, Y., Yang, Y., & Wang, W. (2021). Older Pedestrian Traffic Crashes Severity Analysis Based on an Emerging Machine Learning XGBoost. *Sustainability*, 13(2). <https://doi.org/10.3390/su13020926>
- Hamilton-Baillie, B. (2008a). Built Environment Vol 34 (2) Shared Space: Reconciling People, Places and Traffic. *Built Environment*, 34(2).
- Hamilton-Baillie, B. (2008b). *SHARED SPACE: RECONCILING PEOPLE, PLACES AND TRAFFIC 161 BUILT ENVIRONMENT VOL 34 NO 2 Shared Space: Reconciling People, Places and Traffic*.
- Hatfield, J., & Prabhakaran, P. (2016). An investigation of behaviour and attitudes relevant to the user safety of pedestrian/cyclist shared paths. *Transportation Research Part F: Traffic Psychology and Behaviour*, 40. <https://doi.org/10.1016/j.trf.2016.04.005>
- Henriques, J., Caldeira, F., Cruz, T., & Simões, P. (2020). Combining K-Means and XGBoost Models for Anomaly Detection Using Log Datasets. *Electronics*, 9(7). <https://doi.org/10.3390/electronics9071164>
- Herawati, N., Nisa, K., & Setiawan, E. (2018). Regularized Multiple Regression Methods to Deal with Severe Multicollinearity. *International Journal of Statistics and Applications*, 8(4).
- Hussein, N. (2018). The Pedestrianisation and Its Relation with Enhancing Walkability in Urban Spaces. *Journal of Contemporary Urban Affairs*, 2(1). <https://doi.org/10.25034/ijcua.2018.3666>
- INRS. (2021). *Cohabitation cyclistes-piétons sur les rues piétonnes Projet-pilote sur l'avenue du Mont-Royal et la rue Wellington*. Institute National de la Recherche Scientifique. www.ucs.inrs.ca
- Johnes, J., Gatrell, T., Jefferson, C. W., Haining, R. P., Hepworth, M., Pratt, A. C., Yannopoulos, G. N., Hitchens, D. M. W. N., Read, R., Alden, J. D., Crampton, G., & Judge, E. (1991). Book Reviews. *Regional Studies*, 25(4). <https://doi.org/10.1080/00343409112331346567>
- Jomthanachai, S., Wong, W. P., & Khaw, K. W. (2022). An application of machine learning regression to feature selection: a study of logistics performance and economic attribute. *Neural Computing and Applications*, 34(18), 15781–15805. <https://doi.org/10.1007/s00521-022-07266-6>
- Judge, C. (2013). The experiment of American pedestrian malls: Trends, analysis, necessary indicators for success and recommendations for Fresno's Fulton Mall. Fresno, CA In *City of Fresno, California*.

- Kandt, J., & Batty, M. (2021). Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities*, 109, 102992. <https://doi.org/https://doi.org/10.1016/j.cities.2020.102992>
- Kang, H. (2013). The prevention and handling of the missing data. *Korean J Anesthesiol*, 64(5), 402–406. <https://doi.org/10.4097/kjae.2013.64.5.402>
- Karndacharuk, A., Wilson, D., & Dunn, R. (2013). Analysis of pedestrian performance in shared-space environments. *Transportation Research Record*, 2393. <https://doi.org/10.3141/2393-01>
- Kashani Jou, K. (2011). Pedestrian areas and sustainable development. *World Academy of Science, Engineering and Technology*, 77.
- Kim, S., Park, S., & Lee, J. S. (2014). Meso- or micro-scale? Environmental factors influencing pedestrian satisfaction. *Transportation Research Part D: Transport and Environment*, 30, 10–20. <https://doi.org/https://doi.org/10.1016/j.trd.2014.05.005>
- Kim, U., Lee, J., & He, S. Y. (2021). Pedestrianization impacts on air quality perceptions and environment satisfaction: The case of regenerated streets in downtown Seoul. *International Journal of Environmental Research and Public Health*, 18(19). <https://doi.org/10.3390/ijerph181910225>
- Klein, D. (2020). MIMRGNS: Stata module to run margins after mi estimate. *Statistical Software Components*.
- Kroonenberg, P. M., & Verbeek, A. (2018). The Tale of Cochran's Rule: My Contingency Table has so Many Expected Values Smaller than 5, What Am I to Do? *American Statistician*, 72(2). <https://doi.org/10.1080/00031305.2017.1286260>
- Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). 5 - Foundations of data imbalance and solutions for a data democracy. In F. A. Batarseh & R. Yang (Eds.), *Data Democracy* (pp. 83–106). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-818366-3.00005-8>
- Lawson, A. R., Pakrashi, V., Ghosh, B., & Szeto, W. Y. (2013). Perception of safety of cyclists in Dublin City. *Accident Analysis and Prevention*, 50. <https://doi.org/10.1016/j.aap.2012.05.029>
- Lee, W. Do, Ectors, W., Bellemans, T., Kochan, B., Janssens, D., Wets, G., Choi, K., & Joh, C. H. (2018). Investigating pedestrian walkability using a multitude of Seoul data sources. *Transportmetrica B*, 6(1). <https://doi.org/10.1080/21680566.2017.1325783>
- Lee, H., & Kim, S. N. (2021). Perceived safety and pedestrian performance in pedestrian priority streets (Ppss) in Seoul, Korea: A virtual reality experiment and trace mapping. *International Journal of Environmental Research and Public Health*, 18(5). <https://doi.org/10.3390/ijerph18052501>
- Lee, J., Kim, D., & Park, J. (2022). A Machine Learning and Computer Vision Study of the Environmental Characteristics of Streetscapes That Affect Pedestrian Satisfaction. *Sustainability*, 14(9). <https://doi.org/10.3390/su14095730>
- Lee, S., Han, M., Rhee, K., & Bae, B. (2021). Identification of factors affecting pedestrian satisfaction toward land use and street type. *Sustainability (Switzerland)*, 13(19). <https://doi.org/10.3390/su131910725>
- Li, Z. (2022). Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. *Computers, Environment and Urban Systems*, 96, 101845. <https://doi.org/https://doi.org/10.1016/j.compenvurbsys.2022.101845>
- Liew, B. X. W., Kovacs, F. M., Rügamer, D., & Royuela, A. (2022). Machine learning versus logistic regression for prognostic modelling in individuals with non-specific neck pain. *European Spine Journal*, 31(8). <https://doi.org/10.1007/s00586-022-07188-w>

- Litman, T. A. (2017). Economic Value of Walkability. In *of World Transport Policy & Practice* (Vol. 10, Issue 1). Transportation Research Board. www.vtpi.org/Info@vtpi.org/www.ecologica.co.uk/WTPPhome.html
- Litman, T. A. (2021). Developing Indicators for Sustainable and Livable Transport Planning. *Transportation Research Record*, 116. www.vtpi.org
- Lorraine Carpenter. (2022, April). 10 Montreal streets are going pedestrian-only this summer. *Cult Mtl*. <https://cultmtl.com/2022/06/10-montreal-streets-are-going-pedestrian-only-this-summer-pedestrianization-projects/#:~:text=The%20city%20is%20investing%20%2412,over%20a%20three%2Dyear%20period.>
- macrotrends. (2022). *Montreal, Canada Metro Area Population 1950-2022*. Macrotrends. <https://www.macrotrends.net/cities/20384/montreal/population#:~:text=The%20current%20metro%20area%20population,a%200.62%25%20increase%20from%202020.>
- Majumdar, B. B., Sahu, P. K., Patil, M., & Vendotti, N. (2021). Pedestrian Satisfaction-Based Methodology for Prioritization of Critical Sidewalk and Crosswalk Attributes Influencing Walkability. *Journal of Urban Planning and Development*, 147(3). [https://doi.org/10.1061/\(asce\)up.1943-5444.0000718](https://doi.org/10.1061/(asce)up.1943-5444.0000718)
- Mäkinen, T. M., Raatikka, V. P., Rytönen, M., Jokelainen, J., Rintamäki, H., Ruuhela, R., Näyhä, S., & Hassi, J. (2006). Factors affecting outdoor exposure in winter: Population-based study. *International Journal of Biometeorology*, 51(1). <https://doi.org/10.1007/s00484-006-0040-0>
- Mastora, C., Paschalidis, E., Nikiforiadis, A., & Basbas, S. (2023). Pedestrian Crossings as a Means of Reducing Conflicts between Cyclists and Pedestrians in Shared Spaces. *Sustainability*, 15(12). <https://doi.org/10.3390/su15129377>
- Matt Worona. (2020, August). *Pedestrianization – a data-driven approach in Kelowna, Canada*. Eco-Counter. <https://www.eco-counter.com/blog/the-summer-of-giving-space-to-pedestrians-a-data-driven-approach-in-kelowna-canada/>
- Medical Officer of Health, the. (2014). *Healthy Streets Evidence Review Copies*. <http://www.toronto.ca/health>
- Meng, L., Wen, K.-H., Zeng, Z., Brewin, R., Fan, X., & Wu, Q. (2020). The Impact of Street Space Perception Factors on Elderly Health in High-Density Cities in Macau—Analysis Based on Street View Images and Deep Learning Technology. *Sustainability*, 12(5). <https://doi.org/10.3390/su12051799>
- Mesimäki, J., & Luoma, J. (2021). Near accidents and collisions between pedestrians and cyclists. *European Transport Research Review*, 13(1). <https://doi.org/10.1186/s12544-021-00497-z>
- Methorst, R., Ing Jürgen Gerlach, U., & Jens Leven, cand-I. (2007). *Shared Space: Safe or Dangerous? A contribution to objectification of a popular design philosophy*.
- Methorst, R., & Van Der Horst, R. (2010). *Pedestrians' performance and satisfaction. Pedestrians' performance and satisfaction*.
- Mike Biddulph. (2001). *Home zones A planning and design handbook*. Policy Press.
- Montaño Moreno, J. J., Palmer Pol, A., Sesé Abad, A., & Cajal Blasco, B. (2013). Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema*, 25(4). <https://doi.org/10.7334/psicothema2013.23>
- Mouratidis, K., Ettema, D., & Næss, P. (2019). Urban form, travel behavior, and travel satisfaction. *Transportation Research Part A: Policy and Practice*, 129. <https://doi.org/10.1016/j.tra.2019.09.002>
- Mueller, N., Rojas-Rueda, D., Cole-Hunter, T., de Nazelle, A., Dons, E., Gerike, R., Götschi, T., Int Panis, L., Kahlmeier, S., & Nieuwenhuijsen, M. (2015). Health impact assessment of active transportation: A systematic review. In *Preventive Medicine* (Vol. 76). <https://doi.org/10.1016/j.ypmed.2015.04.010>

- Muthukrishnan, R., & Rohini, R. (2016). LASSO: A feature selection technique in predictive modeling for machine learning. *2016 IEEE International Conference on Advances in Computer Applications (ICACA)*, 18–20. <https://doi.org/10.1109/ICACA.2016.7887916>
- Muthukrishnan, R., & Rohini, R. (2017). LASSO: A feature selection technique in predictive modeling for machine learning. *2016 IEEE International Conference on Advances in Computer Applications, ICACA 2016*. <https://doi.org/10.1109/ICACA.2016.7887916>
- Nag, D., Bhaduri, E., Kumar, G. P., & Goswami, A. K. (2020). Assessment of relationships between user satisfaction, physical environment, and user behaviour in pedestrian infrastructure. *Transportation Research Procedia*, 48. <https://doi.org/10.1016/j.trpro.2020.08.284>
- Nakagawa, S. (2015). Missing data: mechanisms, methods, and messages. *Ecological Statistics: Contemporary Theory and Application*.
- Narelle Haworth, A. S. A. K. D. (2014). An observational study of conflicts between cyclists and pedestrians in the city centre. *Journal of the Australasian College of Road Safety*, 25(4).
- Naseri, H., Waygood, E. O. D., Wang, B., Patterson, Z., & Daziano, R. A. (2022). A Novel Feature Selection Technique to Better Predict Climate Change Stage of Change. *Sustainability*, 14(1). <https://doi.org/10.3390/su14010040>
- Nguyen, L. M., & Mertens, L. (2021). Psychosocial and social environmental factors as moderators in the relation between the objective environment and older adults' active transport. *International Journal of Environmental Research and Public Health*, 18(5). <https://doi.org/10.3390/ijerph18052647>
- Nikiforiadis, A., Basbas, S., Mikiki, F., Oikonomou, A., & Polymeroudi, E. (2021). Pedestrians-Cyclists Shared Spaces Level of Service: Comparison of Methodologies and Critical Discussion. *Sustainability*, 13(1). <https://doi.org/10.3390/su13010361>
- Nikiforiadis, A., Chatzali, E., Ioannidis, V., Kalogiros, K., Paipai, M., & Basbas, S. (2023). Investigating factors that affect perceived quality of service on pedestrians-cyclists shared infrastructure. *Travel Behaviour and Society*, 31, 323–332. <https://doi.org/10.1016/j.tbs.2023.01.006>
- Nikitas, A., Tsigdinos, S., Karolemeas, C., Kourmpa, E., & Bakogiannis, E. (2021). Cycling in the Era of COVID-19: Lessons Learnt and Best Practice Policy Recommendations for a More Bike-Centric Future. *Sustainability*, 13(9). <https://doi.org/10.3390/su13094620>
- Niklas Donges, M. U. (2022). Random Forest Classifier: A Complete Guide to How It Works in Machine Learning. *Built In*.
- Nisha Arya. (2022, September). *Tuning Random Forest Hyperparameters*. KDnuggets. https://www.kdnuggets.com/2022/08/tuning-random-forest-hyperparameters.html#:~:text=Hyperparameters%20of%20a%20Random%20Forest&text=n_estimators%3A%20This%20is%20the%20number,looking%20for%20the%20best%20split
- Oakes, J. M., Forsyth, A., & Schmitz, K. H. (2007). The effects of neighborhood density and street connectivity on walking behavior: The Twin Cities walking study. *Epidemiologic Perspectives and Innovations*, 4. <https://doi.org/10.1186/1742-5573-4-16>
- O'Hern, S., & Oxley, J. (2019). Pedestrian injuries due to collisions with cyclists Melbourne, Australia. *Accident Analysis and Prevention*, 122. <https://doi.org/10.1016/j.aap.2018.10.018>
- Olgun, R. (2019). *Evaluation of Pedestrianized Streets in Terms of Landscape Design and User Satisfaction: The Case of Antalya-Şarampol Street*.

- Parkhurst, G. (2003). Regulating cars and buses in cities: The case of pedestrianisation in Oxford. *Economic Affairs*, 23(2). <https://doi.org/10.1111/1468-0270.00410>
- Patil, A. R., & Kim, S. (2020). Combination of ensembles of regularized regression models with resampling-based lasso feature selection in high dimensional data. *Mathematics*, 8(1). <https://doi.org/10.3390/math8010110>
- Pavlyuk, D. (2019). Feature selection and extraction in spatiotemporal traffic forecasting: a systematic literature review. *European Transport Research Review*, 11(1). <https://doi.org/10.1186/s12544-019-0345-9>
- Payne, R. (2021). Will the COVID-19 outbreak propel the demand for active spaces or scare the public away? *Cities & Health*, 5(sup1), S200–S203.
- Peterson K. Ozili. (2023). The Acceptable R-Square in Empirical Modelling for Social Science Research. *Electronic Journal*.
- Philippe Brodeur-Ouimet, A.-A. L. P.-M. L. and M.-S. C. (2023). Cyclist-Pedestrian Cohabitation: Lessons to Learn from a Pilot Project on Pedestrian Streets in Montréal (Canada). *EasyChair Preprint No. 10408*.
- Poulos, R. G., Hatfield, J., Rissel, C., Flack, L. K., Shaw, L., Grzebieta, R., & McIntosh, A. S. (2017). Near miss experiences of transport and recreational cyclists in New South Wales, Australia. Findings from a prospective cohort study. *Accident Analysis and Prevention*, 101. <https://doi.org/10.1016/j.aap.2017.01.020>
- Prashant Banerjee. (2020). *A Guide on XGBoost hyperparameters tuning*. Wwww.Kaggle.Com. <https://www.kaggle.com/code/prashant111/a-guide-on-xgboost-hyperparameters-tuning/notebook>
- Qiu, L.-Y., & He, L.-Y. (2018). Bike Sharing and the Economy, the Environment, and Health-Related Externalities. *Sustainability*, 10(4). <https://doi.org/10.3390/su10041145>
- Rahman, M. T., & Nahiduzzaman, K. M. (2019). Examining the walking accessibility, willingness, and travel conditions of residents in Saudi cities. *International Journal of Environmental Research and Public Health*, 16(4). <https://doi.org/10.3390/ijerph16040545>
- Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-Validation. In L. LIU & M. T. ÖZSU (Eds.), *Encyclopedia of Database Systems* (pp. 532–538). Springer US. https://doi.org/10.1007/978-0-387-39940-9_565
- Rissel, C., Crane, M., Wen, L. M., Greaves, S., & Standen, C. (2016). Satisfaction with transport and enjoyment of the commute by commuting mode in inner Sydney. *Health Promotion Journal of Australia*, 27(1). <https://doi.org/10.1071/HE15044>
- Sahani, R., & Bhuyan, P. K. (2020). Modelling Pedestrian Perspectives in Evaluating Satisfaction Levels of Urban Roadway Walking Facilities. *Transportation Research Procedia*, 48. <https://doi.org/10.1016/j.trpro.2020.08.289>
- Sahani, R., Praveena, D., & Bhuyan, P. K. (2016). Use of Multinomial Logit Model in Evaluation of Service Levels of Pedestrian Facilities. *Tpmdc 2016*, 1965.
- Sandford, A. (2020). Coronavirus: Half of humanity now on lockdown as 90 countries call for confinement. *Euronews*, 3(04), 2020.
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. In *SN Computer Science* (Vol. 2, Issue 3). <https://doi.org/10.1007/s42979-021-00592-x>

- Shapley, L. S. (1953). 17. A Value for n-Person Games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the Theory of Games (AM-28), Volume II* (pp. 307–318). Princeton University Press.
<https://doi.org/doi:10.1515/9781400881970-018>
- Sheng, Q., Jiao, J., & Pang, T. (2021). Understanding the Impact of Street Patterns on Pedestrian Distribution: A Case Study in Tianjin, China. *Urban Rail Transit*, 7(3). <https://doi.org/10.1007/s40864-021-00152-9>
- Simone Caron. (2023). *Avenue du Mont-Royal, a lively and unifying space*. La Presse.
<https://www.lapresse.ca/actualites/grand-montreal/2023-07-07/les-rues-pietonnes-de-montreal/l-avenue-du-mont-royal-un-espace-vivant-et-rassembleur.php>
- Singh, S. G., & Kumar, S. V. (2021). Dealing with Multicollinearity Problem in Analysis of Side Friction Characteristics Under Urban Heterogeneous Traffic Conditions. *Arabian Journal for Science and Engineering*, 46(11). <https://doi.org/10.1007/s13369-020-05213-y>
- Singh Yadav, J., Jaiswal, A., & Nateriya, R. (2015). Modelling Pedestrian Overall Satisfaction Level at Signalised Intersection Crosswalks. *International Research Journal of Engineering and Technology*.
- Singhal, R., & Rana, R. (2015). Chi-square test and its application in hypothesis testing. *Journal of the Practice of Cardiovascular Sciences*, 1(1). <https://doi.org/10.4103/2395-5414.157577>
- Soley-Bori, M. (2013). *Dealing with missing data: key assumptions and methods for applied analysis*.
- Soni, N., & Soni, N. (2016). Benefits of pedestrianization and warrants to pedestrianize an area. *Land Use Policy*, 57. <https://doi.org/10.1016/j.landusepol.2016.05.009>
- Spiegel, S., Müller, F., Wiesmann, D., & Bird, J. (2018). Cost-Sensitive Learning for Predictive Maintenance. *ArXiv*, *abs/1809.10979*.
- Teddy Elliot. (2020, March 15). *Montreal Is Creating 327 km Of Bike & Pedestrian Paths So You Can Still Enjoy The Summer*. MTL Blog. <https://www.mtlblog.com/montreal/montreal-is-creating-327-km-of-bike-and-pedestrian-paths-so-you-can-still-enjoy-the-summer>
- Thomas, C. (2008). Discussion: Shared space—safe space? *Proceedings of the Institution of Civil Engineers - Municipal Engineer*, 161(1), 59–60. <https://doi.org/10.1680/muen.2008.161.1.59>
- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 73(3). <https://doi.org/10.1111/j.1467-9868.2011.00771.x>
- Tim Jones. (2018). Supervised learning models. *IBM Developer*.
- Ujjwal, J., & Bandyopadhyaya, R. (2021). Development of Pedestrian Level of Service (PLOS) model and satisfaction perception rating models for pedestrian infrastructure for mixed land-use urban areas. *Transportation*. <https://doi.org/10.1007/s11116-021-10247-8>
- Ville de Montreal. (2021). *Montreal plan*. Agir maintenant pour préparer la relance. Ville de Montréal.
https://portailm4s.s3.montreal.ca/pdf/ville_montreal_plan_relance_phase_2_2020_reduit.pdf
- Voss, C. (2018). 1 - Public Health Benefits of Active Transportation. In R. Larouche (Ed.), *Children's Active Transportation* (pp. 1–20). Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-12-811931-0.00001-6>
- Wan, M., Wu, Q., Yan, L., Guo, J., Li, W., Lin, W., & Lu, S. (2022). Application of Random Forest and Shap for Real-Time Taxi Drivers' Traffic Violations Detection and Feature Analysis. *SSRN Electronic Journal*.
- Watson AY, B. R. K. D. editors. (1988). *Air Pollution, the Automobile, and Public Health*. Washington (DC): National Academies Press (US). <https://www.ncbi.nlm.nih.gov/books/NBK218144/>

- Westphal, J. M., Schweitzer, J., & Mullins, L. (2012). *Special Issue*.
- Winters, M., Davidson, G., Kao, D., & Teschke, K. (2011). Motivators and deterrents of bicycling: comparing influences on decisions to ride. *Transportation*, 38(1), 153–168. <https://doi.org/10.1007/s11116-010-9284-y>
- Yassin, H. H. (2019). Livable city: An approach to pedestrianization through tactical urbanism. *Alexandria Engineering Journal*, 58(1). <https://doi.org/10.1016/j.aej.2019.02.005>
- Yihuan Huang and Amanda Montoya. (2020). *Lasso and Group Lasso with Categorical Predictors: Impact of Coding Strategy on Variable Selection and Prediction*.
- Zainol, R., Ahmad, F., Nordin, N. A., & Aripin, A. W. M. (2014). Evaluation of users' satisfaction on pedestrian facilities using pair-wise comparison approach. *IOP Conference Series: Earth and Environmental Science*, 18(1). <https://doi.org/10.1088/1755-1315/18/1/012175>
- Zhang, Z. (2015). Missing values in big data research: Some basic skills. *Annals of Translational Medicine*, 3(21). <https://doi.org/10.3978/j.issn.2305-5839.2015.12.11>
- Zhou, Z.-H. (2012). *Ensemble methods: foundations and algorithms*. CRC press.