UNDERSTANDING PEDESTRIAN ROAD CROSSING BEHAVIOR

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
OZYEGIN UNIVERSITY

İrem KÜÇÜKALİ, B.Sc.

S016206

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF DESIGN, TECHNOLOGY AND SOCIETY

JULY 2020

ABSTRACT

While cities tend to grow in terms of both land and population, millions of people from different backgrounds, social classes, and demographics living in a complex system like a city, requires some regularization. Car-centered planning approach is not sustainable anymore considering ever-growing city populations, motor-vehicle-use related environmental, and health issues such as air pollution, climate change, urban traffic congestions, and the obesity epidemic.

The primary aim of this research is to understand the underlying mechanism for the jaywalking behavior and the shortcomings of the current pedestrian signaling in terms of safety, efficiency, and practicality; and to suggest a new way of organizing pedestrian traffic more efficiently and safely.

To answer my research questions, observational data on pedestrian behavior at signalized intersections with pedestrian signal phases were required along with descriptive characteristics of pedestrian, intersection, signalization, and other environmental elements. Primary and secondary data is collected for the research. The thesis research is conducted at two signalized intersections, Beşiktaş Square and Fatih Malta, in Istanbul.

The pedestrian violation rate is found 32.9%. Resulting statistical models suggest the importance of the number of approaching vehicles, the number of people crossing, time of the day that crossing has happened, and signalization cycle length on the jaywalking behavior.

ÖZET

Şehirler hem arazi hem de nüfus açısından büyüme eğilimindeyken, farklı geçmişlere, sosyal sınıflara ve demografik özelliklere sahip milyonlarca insanın bir arada yaşabilmesi için bazı düzenlemeler gerekir. Araba merkezli planlama yaklaşımı, sürekli artan şehir nüfusu, motorlu taşıt kullanımının sebep olduğu çevre ve hava kirliliği, iklim değişikliği, kentsel trafik sıkışıklıkları ve obezite salgını gibi sağlık konuları göz önüne alındığında artık sürdürülebilir değildir.

Bu çalışmanın temel amacı, yayaların karşıdan karşıya geçiş davranışının altında yatan mekanizmayı ve mevcut yaya sinyalizasyonunun güvenlik, verimlilik ve pratiklik açısından eksikliklerini anlamak ve yaya trafiğini daha verimli ve güvenli bir şekilde organize etmenin yeni bir yolunu önermektir.

Araştırma sorularını cevaplamak için, yaya sinyal fazları ile sinyalize kavşaklarda yaya davranışı ile ilgili gözlemsel veriler ile; yaya, kavşak, sinyalizasyon ve diğer çevresel unsurların tanımlayıcı özellikleri kullanılmıştır. Araştırma için birincil ve ikincil veriler toplanmıştır. Tez araştırması, İstanbul'da Beşiktaş Meydanı ve Fatih Malta olmak üzere iki sinyalize kavşakta yürütülmüştür.

Yaya ihlal oranı %32.9 olarak bulunmuştur. Sonuçta ortaya çıkan istatistiksel modeller, yaklaşan araç sayısının, geçiş yapan insan sayısının, geçişin gerçekleştiği günün saatinin ve sinyalizasyon döngüsü uzunluğunun; karşıdan karşıya kural ihlali yaparak geçme davranışındaki önemini ortaya koymaktadır.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisors Dr.Ali O. İlhan and Dr.Fatih Gündoğan for their comments, remarks, and guidance throughout this study. I am also grateful to Dr.Ayşe Hazar Köksal for academic support. Furthermore, I would like to thank my husband Hüseyin, for being always there for me and encouraging me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

City is where more than half of the world's population live, work, and participate in society (Angel, et al., 2011, p. 78). Most of the cities tend to grow in terms of both land and population (Angel, et al., 2011, p. 91). By 2030, global urban land is expected to triple as it was in 2000 (Seto et al., 2012, p. 16083). Millions of people from different backgrounds, social classes, and demographics living in a complex system like a city, requires some regularization.

City and traffic planning is done with a car/driver focus for about a century. Car-centered planning dominated government policies (Sheller and Urry, 2000, p. 746). However, this approach is not sustainable anymore considering ever-growing city populations, motor-vehicle-use related environmental, and health issues such as air pollution, climate change, urban traffic congestions, and the obesity epidemic. Cities need to provide and prioritize pedestrian mobility more than ever. Another important reason for improving pedestrian mobility is that all people whether they drive cars or ride bikes are pedestrians in some parts of their urban travel. Also, some people, because of their socioeconomic conditions, do not have access to private cars. This makes improving pedestrian mobility is a way to create urban equity (Mavoa et al., 2012, p. 15).

Mobility advantage of cars over pedestrians becomes apparent when citizens of both modes need to use the city space at the same time: in road-crossings. Crossings are regulated for both car and pedestrian use with the help of pedestrian signaling. Yet jaywalking, illegal road crossing of pedestrians is a global phenomenon, showing that pedestrian signalization falls short for regulating pedestrian traffic.

Pedestrian violations in signalized intersections are frequently encountered phenomena in Istanbul, as in many other cities. However, in article 68 of the Highway Traffic Law, it is clearly stated that pedestrians are obliged to comply with traffic signals while crossing at signalized intersections.

> *Pedestrians who want to cross the road to the opposite side of the road are forbidden to cross the road from any place other than the pedestrian and school crossing and intersection entrances and exits. When pedestrians pass through these places, where pedestrian crossings or school crossings are located, if there are traffic light signals for pedestrians in the crossing, they must comply with these signs… However, in places where there are no pedestrian crossings or crossroads at a distance of about 100 meters, pedestrians can cross the vehicle road in the shortest direction and as soon as possible, provided that they do not constitute an obstacle for vehicle traffic and by controlling the road* (Karayolları Trafik Kanunu, 1983).

Having a broader understanding of the pedestrians' frequent rule-breaking behavior at signalized intersections might help to understand the assumptions made about and shortcomings of the way of regulating pedestrian traffic.

1.2. Research Objective

The primary aim of this research is to understand the underlying mechanism for the jaywalking behavior and the shortcomings of the current pedestrian signaling in terms of safety, efficiency, and practicality; and to suggest a new way of organizing pedestrian traffic more efficiently and safely.

1.3. Provisional Research Questions

1.3.1. Overall Question

How does pedestrian behavior at signalized crossings change vary by individual conditions of pedestrian and surrounding conditions of the environment, intersection, and place in general?

1.3.2. Specific Questions

1. What are the factors that affect pedestrian rule-complying and/or rule-breaking behavior while crossing at signalized intersections?
2. What are the factors that affect crossing duration and speed at signalized intersections?
3. Which conditions lead to varying crossing patterns of pedestrians and how?

1.4. Significance of the Study

Previous studies on pedestrian behavior examined the jaywalking phenomenon by looking to factors of age, gender, luggage, operating conditions of signal timing, pedestrian group, upcoming vehicles, oncoming pedestrians, density, space, pedestrian flow directions, and special days. Rule-breaking traffic behavior was associated with risk-taking, impatience, anger, hostility characteristics; cognitive skills at detecting hazardous situations; and time-saving attempts (Dallmeyer et al., 2016, p. 83; Diaz, 2002, p. 174). The literature on pedestrian behavior has some explanations for the conditions of jaywalking; however, they lack a wider explanation for causes and applicable solutions to current inefficient pedestrian signaling.

CHAPTER 2

LITERATURE REVIEW

2.1. Land-use Policies

People's transportation behavior is directly affected by neighborhood design and land-use (Frank, 2000, p. 13). Outcomes of the relationship between land-use and travel behavior affect various global issues such as physical activity, air quality, traffic congestion, and traffic safety.

Land-use is found to be correlated with air quality through vehicle emissions. Oxides of nitrogen, volatile organic compounds, and carbon monoxide are generated less in the neighborhoods with more transit-supportive development (Frank, 2000, p. 8). Physical activity and pedestrian-oriented environments are correlated. Health policy as well as land-use policy should include regulations to reduce the need and convenience of car use and promote pedestrian activity (King et al., 2002, p. 23). Low-density development patterns result in traffic congestion (King et al., 2002, p. 22).

The main factors of land-use policy that affect travel behavior are proximity and connectivity (Frank, 2000, p. 16).

Proximity determines travel behavior in terms of **density** and **land-use mix**. In-neighborhood travels, such as grocery shopping or visiting a neighbor or friend results in different transportation choices in **high-density** and **low-density** areas. Also being able to access basic needs in the neighborhood is determined by **land-use mix** which means the level of combination of residential and business uses of land together in a neighborhood. Research shows that dense urban areas, and areas with land use mix have so much more walking and cycling rates for travel than low density and single-land-use areas (Saelens et al., 2003, p. 84).

Connectivity states the convenience of movement for transportation (Saelens et al., 2003, p. 81). Grid urban layout, more blocks and more intersections, and fewer barriers create a variety of routes between transportation origins and destinations with shorter distances, therefore increases connectivity as opposed to long blocks or cul de sacs (Frank, 2000, p. 15; Saelens et al., 2003, p. 82). A variety of routes means an opportunity for pedestrians to choose for safety or convenience.

Mixed-use areas, narrow streets, short blocks, buildings fronting streets, and fewer parking spaces off would create a pedestrian-oriented environment and encourage pedestrian activity (King et al., 2002, p. 23).

Whether for traffic safety, preventing congestion, reducing emissions, or increasing physical activity, changes in the built environment are more effective to create more permanent and important travel behavior change than other programs based on the environmental circumstances and targeting an entire population in a neighborhood (Saelens et al., 2003, p. 86). Also, embedding technological strategies to land-use and other behavior modification attempts would create promising work in the field (Frank, 2000, p. 18).

2.1.1. Cars over Pedestrians: Right to the City

Land-use approaches in the US distinguished as before and after World War II. After WW2, urban development policy in the US became auto-dependent because of the increased income that led to more private vehicle ownership, investment into highway infrastructure, and the land use policy that favors vehicles over walkers and bike users. Curvilinear street patterns, cul-de-sacs, lack of sidewalks are signs of auto-dependent urban design (Frank, 2000, p. 6).

Prioritizing cars over pedestrians is a problematic policy for economic and social reasons. An explanation of the economic implications of prioritizing cars lays behind traffic signalization.

Traffic signalization systems favor cars over pedestrians no matter what (Noland, 1996, p. 28) as an extension of land-use policies. What determines the travel times of motor vehicles is the cycle length of the signal and the time amount of the green phase. Any favor for cars in signalization costs to pedestrians. However, for measuring transportation performance, the key, and sometimes the only variable accepted is vehicle speeds. Especially in large street crossings, pedestrians need to wait a long time for their safety. This may often mean more cost to society in general (Noland, 1996, p. 29). Noland examined the relative travel time costs of both motor vehicle drivers and pedestrians to infer an optimal signalization timing that minimizes the total cost of waiting for society. As a result, he suggests to prioritize pedestrians in high-density areas, and even eliminate vehicle traffic in some (Noland, 1996, p. 28).

The social issue that is raised by prioritizing cars is equity. Accessibility (Mavoa et al., 202, p. 17) and the value of one’s time (Noland, 1996, p. 30) for users of other transport modes (public transit, walking, cycling) is vital because of the equity reasons. There are segments of society that do not have access to private cars and reliant on public transit (Mavoa et al., 2012, p. 17).

2.2. Travel Mode Choice

Street planning, signalization planning (Noland, 1996, p. 28), demographics and socioeconomic status of travelers (Kadali and Vedagiri, 2013, p. 679), availability of other modes (Antic et al., 2016, p. 8), land-use characteristics in terms of proximity (land-use mix and residential/employment densification) and connectivity (Frank, 2000, p. 10), travel cost, environmental quality, convenience, presence and quality of transportation infrastructure (for example sidewalks, pedestrian signals, mid-street islands), crime, weather (temperature, wind, rainfall) (Saelens et al., 2003, p. 81) affect behaviors of people towards using automobiles instead of walking.

2.3. Traffic Signalization

Despite the first signal system was installed to protect pedestrians at intersections, today signalization is mostly used to reduce the travel times of motor vehicles. In a signalized intersection with 60-second cycle time and 10-second pedestrian green phase, for a pedestrian coming across green light for them to cross happens in only one-sixth of the time. In all other times, a pedestrian is required to wait, unless they sacrifice their preferred walking speed or safety by running to catch the green light or jaywalking (Noland, 1996, p. 29).

A simple calculation of the delay that a signalized intersection costs to a pedestrian would be dependent total cycle length, pedestrian green phase, and the time interval that pedestrian arrives at a crossing. In a more realistic calculation, queuing delay resulting from waiting for other pedestrians to enter the crossing can be included for more congested situations (Noland, 1996, p. 29).

Average pedestrian delays for some cycle lengths are given in Noland's study as in Table 1.

Traffic signalization operates in a way that aims to maximize vehicle traffic flow while mostly neglecting pedestrian traffic. However, a successful system should take all travelers into account and minimize the cost of delay for all. Long cycles with short pedestrian green durations are accepted only where there is no pedestrian traffic. Still, in many cases, short pedestrian green times may be the reason behind no pedestrian traffic. Moreover, in urban centers with pedestrian density, it is recommended to set green durations for pedestrians much higher at the cost of delays for vehicles, or even bans to motor vehicle traffic (Noland, 1996, p. 28).

Table 1: Average pedestrian delays for various cycle lengths (Noland, 1996, p. 29)

Cycle Lengths (seconds)	Green Phase for Pedestrians (seconds)	Average Delay (seconds)
60	15	17.25
60	10	21.25
90	15	31.67
90	10	36.00
120	15	46.38
120	10	50.88

2.4. Importance of Pedestrian Activity

Pedestrian activity is required for every citizen while they are traveling in an urban setting. The importance of public transit and walking, in times when environment and city life suffers from car use because of emissions, congestion, traffic accidents, and lack of physical activity, has become prominent (Mavoa et al., 2012, p. 15).

Certain health benefits of walking for transportation purposes is also a subject of scientific study. Walking or cycling to work is associated with lower weight and less weight gain over time. What makes walking and cycling for transportation more important than health behavior programs is; in an urban form, these types of physical activity target an entire population in a neighborhood so the potential effects are much higher (Saelens et al., 2003, p. 86).

2.5. Pedestrian Behavior

Pedestrian activity is required for every citizen while they are traveling in an urban setting. There are different reasons for people to travel in the city. Therefore, there may be different behavioral patterns of pedestrians while they travel for different reasons (Tahmasbi and Haghshenas, 2019, p. 165).

Pedestrian activity on event-days is found 20% slower than non-event days in one study (Sukhadia, et al., 2016, p. 653). Walking because of necessity and/or no possibility to use other transportation causes pedestrians to more likely to make errors than walking for pleasure (Antic et al., 2016, p. 1). Finer grids with more blocks and intersections lead safer and more convenient pedestrian travel (Frank, 2000, p. 7). These examples show that environmental variables, pedestrian characteristics, land-use, and many other factors can change pedestrian behavior in certain ways.

2.5.1. Pedestrian Crossing Behavior and Jaywalking

Jaywalking is an example of pedestrian aggressiveness as a non-compliance or rule-breaking traffic behavior (Dallmayer et al., 2012, p. 83).

Pedestrian crossing behavior is driven by two main factors: time and distance (Malenje et al., 2018, p. 650). Longer waiting and crossing times and detours are what pedestrians try to avoid in an attempt to save time and energy. To save time and energy, pedestrians are required to make some decisions about crossing the road. Three prominent decisions about crossing are location, rule-complying, and pattern of the crossing. In some cases, the first two are intertwined as some locations are always prohibited or always rightful to cross. Examples of outcomes of the crossing decisions are commonly seen mid-block crossings and jaywalking. Crossing patterns are chosen by pedestrians according to the gap, the number of lanes, speed of the vehicle, and stream directions of traffic (Kadali and Vedagiri, 2013, p. 678; Malenje et al., 2018, p. 651).

2.5.1.1. Perception of Risk

In Beliefs and the perception of risks and accidents of Kouabenan (1998, p. 243), it is discussed how perceived risks and causation for accidents are formed based on one’s beliefs and social practices. His study shows that fatalistic subjects are more likely to take bigger risks because of their belief on either that inevitable cannot be prevented or that practicing certain rites protects from danger. Fatalism is revealed as admitting a lack of power and control of things but also as a denial of responsibility on one’s own actions. It may result in neglecting precautions and prevention.

2.5.1.2. Determinants of Pedestrian Crossing Behavior

Mid-block crossings and intersection crossings are two locations of pedestrian crossings. Midblock crossings are much more convenient than intersection crossings since the number of directions and the number of lanes is fewer and simpler in mid-blocks. Pedestrians are hardly crossing at intersections unless their destination is very close to the intersection (Kadali and Vedagiri, 2013, p. 679), or traffic density and signalization allow no other way.

Location of bus stops, condition of daily traffic, pedestrian volume, vehicular volume, the distance between crosswalks, crossing distance (Zheng et al., 2015, p. 144), pedestrian's decision for safe crossing and level of driver aggression, the number of lanes, traffic speed, oncoming vehicle speed and its distance, and the gap between cars in time and/or distance, level of risk acceptance, trust to the driver to slow down or stop, and perception of gap, and demographic characteristics of pedestrian (Malenje et al., 2018, p. 650) determine pedestrian crossing decision.

Errors in traffic behavior are related to walking motivation and daily walking distances. If people walk because of necessity and no possibility to use other transportation, they are more likely to make errors than people who walk for pleasure. Since they are probably walking because of their obligations, they hurry more and choose the shortest path. Besides, people who walk more distances have a decline in concentration (Antic et al., 2016, p. 1).

Pedestrian density also changes crossing speed and crossing duration. Single-staged, two-staged, and rolling patterns are chosen according to the gap, the number of lanes, speed of the vehicle, and stream directions of traffic (Malenje et al., 2018, p. 650).

Some research used gap selection among other variables to understand crossing behavior and some research focused specifically on understanding factors determining gap selection itself.

Pedestrians' demographics, behavioral characteristics, vehicle behavioral characteristics (speed and distance) (Petzoldt, 2014, p. 128), road geometry, the number of lanes, and platooning with other pedestrians are found to change pedestrian gap selection behavior (Kadali and Vedagiri, 2013, p. 679).

Mobile phone usage while crossing the street is proposed to be included in further research (Antic et al., 2016, p. 1).

2.5.1.3. Consequences of Jaywalking

Dallmayer et al. (2012) analyzed the time cost of rule-breaking pedestrian agents at the crossroads to other road users. They used simulation software to calculate time costs. Their results show that aggressive behavior does not always cause disadvantages to remaining road users (Dallmayer et al., 2012, p. 83). Pedestrians save up to 22% of their time when they ignore red lights. Time gap estimation of pedestrians when the crossing is less than the actual time gap by 0.58s on average (Dallmayer et al., 2012, p. 84). Simulation results show that when crossing, aggressive behavior does not create considerable disadvantages for other road users (Dallmayer et al., 2012, p. 92). Simulation software used realistic road maps but did not use realistic traffic composition and origin-destination information. The efficiency of the aggressive behavior for the rule-breaking pedestrian and its effects on the traffic accidents are still in question.

2.5.1.4. Recommendations to Regulate Pedestrian Mobility

Malenje et al. recommended to put an additional provision for pedestrian crossings or prevent it (2018, p. 658). One of the most important conclusions of Frank (2000, p. 18) is highlighting the potential great gain of applying technological solutions for travel behavior modification.

2.6. Overview of the Studies on Pedestrian Crossing Behavior at Signalized Intersections

Studies published in Web of Science and Scopus indexes in the last five years are included in the overview.

Table 2: Overview of the Studies on Pedestrian Crossing Behavior at Signalized Intersections

Author's Name	**Focus of Investigation**	**Published**	**Period and Sites**	**Data Collection Methods**	**Conclusions**
Ma, Lu, Zhang	"analyze the characteristics of illegal crossings and their impact on behavior choice"	2020	1 hour per 3 intersections	Primary	"waiting time has the strongest influence on the behavior of crossing at a red light"
Pelé, Deneubourg, Sueur	"highlight the decision-making process of pedestrians crossing the road at signalized crossings in France and in Japan"	2019	6 hours per 7 sites	Primary	"pedestrians do not like to wait for too long at the red signal, and this increases their probability of crossing illegally", "crossing behavior is also influenced by the number of pedestrians waiting to cross"
Xiong, Ma, Ning, Zhao, Weng	"predict pedestrian waiting time"	2019	6 hours total for 2 intersections	Primary	"waiting times vary between peak and off-peak hours", "pedestrians were hesitating between waiting at the curbside and violating the signal, and some of them were readily influenced by surrounding pedestrians"
Mukherjee, Mitra	"investigate if the pedestrian signal violation is a good surrogate for fatal pedestrian crashes at	2019	6 hours per 24 intersections	Primary and Secondary	"signalized intersections with marked crosswalk facilities are strongly linked to pedestrians' violation tendency, which in turn is highly associated with planning and design issues as well

	signalized intersections, and if so, what behavioral, spatial, and built environment-related factors influence such violations”				as risky human behaviors”
Wang, Yuan, Liu, Yang, Yang	“analyze and summarize the red-light running behavior of pedestrians and cyclists”	2019	4 hours per 3 intersections	Primary	“no. of people waiting upon arrival, no. of people crossing upon arrival, pedestrians vs. electric bicycle riders, arrival phase, and speed change upon arrival have significant effects on the red-light running behavior”
Shaaban	“investigate the pedestrian crossing speeds at signalized intersections in Qatar”	2018	12 hours per 2 intersections	Primary	“a high percentage of pedestrians (31.9%) crossing during the flash sign and stop sign intervals; the distributions of pedestrian characteristics and trends observed in the speed analysis during the Walk, Flash, and Stop intervals were similar; the speeds of crossing during the Flash and Stop intervals were higher”
Zhou, Liu, Xu, Pu, Zhang, Zhou	“examine the influences of phone use on pedestrian crossing behavior at signalized intersections in China”	2019	4 hours per 4 sites	Primary	“pedestrian behaviors tend to differ by gender and age; mobile phone using during crossing implies more signal violation, starting late, crossing slowly”
Bansal, Goyal, Sharma	“explore the variation, examine the influencing factors and formulate a	2019	1 hour total for 16 sites	Primary	“pedestrian characteristics, traffic factors, geometric conditions, and operability characteristics exhibit

	model for the pedestrian crossing speed at signalized intersections”				significant effect on the crossing speed”
Russo, James, Aguilar, Smaglik	“gain new insights into factors associated with distracted walking, pedestrian violations, and walking speed”	2018	12 hours total for 4 sites	Primary	“talking and texting while walking was not statistically significantly associated with walking speed, indicating that pedestrians may be accustomed to walking while talking or texting and do not significantly slow (or increase) their walking speeds while engaging in these activities”
Marisamynathan, Vedagiri	“examine how various factors affect pedestrian crossing behavior”	2018	2.5 hours per 6 intersections	Primary	“higher levels of education were found to be key determinants of pedestrian compliance”; “pedestrians whose purpose was working were found to have more noncompliant behavior than those with other trip purposes”; “noncompliance rate increases as the pedestrian red-light phase time increases”
Xie, Wong, Ng, Lam	“pedestrian jaywalking at signalized crosswalks”	2017	1.5 hours per 7 intersections	Primary	“pedestrian characteristics (gender, walking speed), the behavior of surrounding pedestrians (total number of pedestrians and the proportion of jaywalkers), the vehicle traffic, the timing of arrival and the length of signal, and the location of the crosswalk were found to significantly determine the probability of pedestrian jaywalking”

Wu, Lu, Zhu, Chen	"the pedestrians crossing behavior at signalized intersections"	2017	2 hours per 2 intersections	Primary	"factors of pedestrian gender, age, group size, cell phone use, pedestrian signal type, green ratio, and pedestrian volume were significantly affected the likelihood of a pedestrian crossing against a red light"
Onelcin, Alver	"investigate the pedestrians' crossing speeds, delays and gap perceptions at six signalized intersections in Izmir, Turkey"	2017	2 hours per 6 intersections	Primary	"Individuals walked faster than pedestrians who crossed in groups. Males walked faster than females. Pedestrians who did not carry any items walked faster than pedestrians who carried items. Young pedestrians walked faster than adults and adults walked faster than senior pedestrians"

CHAPTER 3

DATA AND METHODS

3.1. Data Collection Methods

To answer this research questions, observational data on pedestrian behavior at signalized intersections with pedestrian signal phases were required along with descriptive characteristics of pedestrian, intersection, signalization system, and other environmental elements.

Primary and secondary data is collected for the research.

Primary data collection included getting permissions and obtaining intersection camera footage from Istanbul Metropolitan Municipality Department of Transportation Directorate of Traffic.

The first criterion used to select the traffic cameras for getting related footages is the existence of a nearby intersection, which has a signalized pedestrian crossing. Clear sight for the crossed road and color of traffic signals during daylight were demanded. This way timing and rule-complying of crossings could be observed. These requirements limited the number of available data sources significantly. Especially colors of the traffic lights from video footages are challenging to detect with human eyes or a computerized system during midday hours. To be able to compensate for nonvisible traffic light colors, signal timing information is used. This is because intersections with adaptive signal timing were not chosen as data sources to be able to calculate traffic light color at a specific time. Other criteria are to create a variety of road characteristics such as the number of lanes, crossing distance in meters, quality of public transportation.

Data collected for the research contains video footage from two signalized intersections across Istanbul. Video footages include one or more pedestrian crosswalks for each signalized intersection. To observe pedestrian behavior in alternating conditions such as daylight, vehicular density, pedestrian density, signal timings, etc. 24-hour-long video recordings of each intersection are collected and observed.

Secondary data collection included environmental variables like air pollution levels, hourly temperature and wind levels at intersections sites; and population densities of districts that are collected from online service providers and the Turkish Statistical Institute.

To be able to analyze the data, video footages were transformed into a tabular data form where each row represents a pedestrian crossing.

Relevant permissions were granted by Istanbul Metropolitan Municipality.

3.2. Study Areas

3.2.1. Beşiktaş Square – Beşiktaş

3.2.1.1. Description of Place

Beşiktaş Square is the center of the Beşiktaş district of İstanbul. Beşiktaş has a population of 182649 inhabitants. Beşiktaş Square is a highly busy and dense transit area since it is next to Beşiktaş pier where ships and boats to Kadıköy, Üsküdar, İstinye, Küçüksu, and Bosphorus rings use. It is also at the intersection of three arterial roads, Barbaros Boulevard, Beşiktaş Street and Çırağan Street. It is directly connected to Zincirlikuyu and Büyükdere Street which is the main economic and business center of contemporary İstanbul. Bosphorus Bridge is also connected to the place via Barbaros Boulevard. It is near to numerous tourist destinations such as Dolmabahçe Palace and Ortaköy. It is a historical and lively commercial place. In terms of resident population and area, Beşiktaş is a small district of İstanbul; however, it is a transit hub and tourist destination so it is very busy and lively at any hour. It is also a cultural and social center with its theaters, cafès, bars, and scenic view. Although it is a small part of the users of the Square, residents of Beşiktaş are a mostly educated and high-income population.

Figure 1: Beşiktaş Square Satellite View

3.2.1.2. Intersection Geometry

Research data is gathered from video footage of pedestrian crossings in Beşiktaş Square that are marked with the color blue in the below plan. Crossing 1 is formed of a two-directional segmented road which both includes three lanes. Crossing 2 is formed of a two-directional segmented road where one direction has two and the other one has three lanes.

Figure 2: Beşiktaş Square Traffic Camera Caption

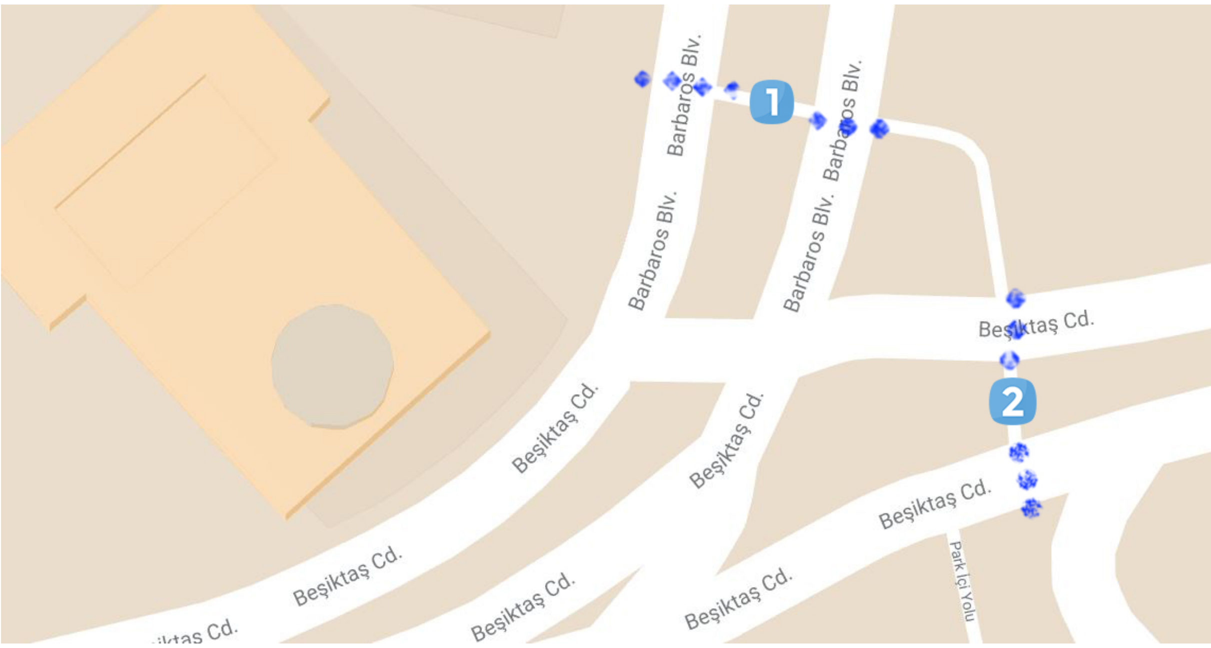

Figure 3: Beşiktaş Square Marked Crossings

3.2.1.3. Signalization

Beşiktaş Square is a busy intersection. Crossings 1 and 2 are a combination of three pedestrian phases of a signalization group of phases that contains five motor vehicle phases, 1 bicycle phase, and 4 pedestrian phases. Crossing 1 is a pedestrian signalization phase that is affected by the movements of two vehicle phases in two directions. Crossing 2 is a combination of two signalization pedestrian phases that is affected by the movements of two vehicle phases in two directions. Crossings are marked based on possible pedestrian movements at the intersection. This is because two signalization phases are considered as one combined pedestrian phase. For Crossing 2, although the movement split into two pedestrian signalization phases for auto-oriented signalization purposes, actually only one pedestrian route exists. In terms of pedestrian behavior, two signalization phases combined actually have one origin and one destination. The reason that this movement is programmed as two pedestrian signalization phases since signalization phases are created vehicle movements in mind. Since the objective of this study is to understand pedestrian movement and behavior, these phases are considered as one pedestrian crossing.

3.2.2. Malta – Fatih

3.2.2.1. Description of Place

Malta intersection is on the Fevzipaşa street where draws the line between Akşemseddin and Ali Kuşcu neighborhoods of Fatih District. It is right next to the Fatih Mosque and Complex that is built by Sultan Mehmed II. It is an important historical and religious center of the city. It is also a cultural place near to Atpazarı that hosts many cafés. Fatih is a district that attracts migrant populations. Since the last migrant wave happened in İstanbul was of Syrian migrants running away from the Syrian war, this neighborhood had its share from the Syrian migrant population. Malta intersection is also a place where adjacent neighborhoods are formed as Syrian ghettos. Fatih is a population-dense district of Istanbul with 443.090 inhabitants and an ongoing rise in the last three years.

Figure 4: Malta Satellite View

3.2.2.2. Intersection Geometry

Research data is gathered from video footage of pedestrian crossings in Malta intersection that are marked with the color blue in the below plan.

Crossing 3 includes one direction of a segmented road. It is formed of two lanes.

3.2.2.3. Signalization

Crossing 3 is a part of a signalization group of phases with 4 vehicle phases and 2 pedestrian phases. Part of one pedestrian phase that is affected by one direction of vehicle traffic is named as Crossing 3 since the angle of traffic camera only allows to record pedestrian movements for limited space in the intersection.

Figure 5: Malta Traffic Camera Caption

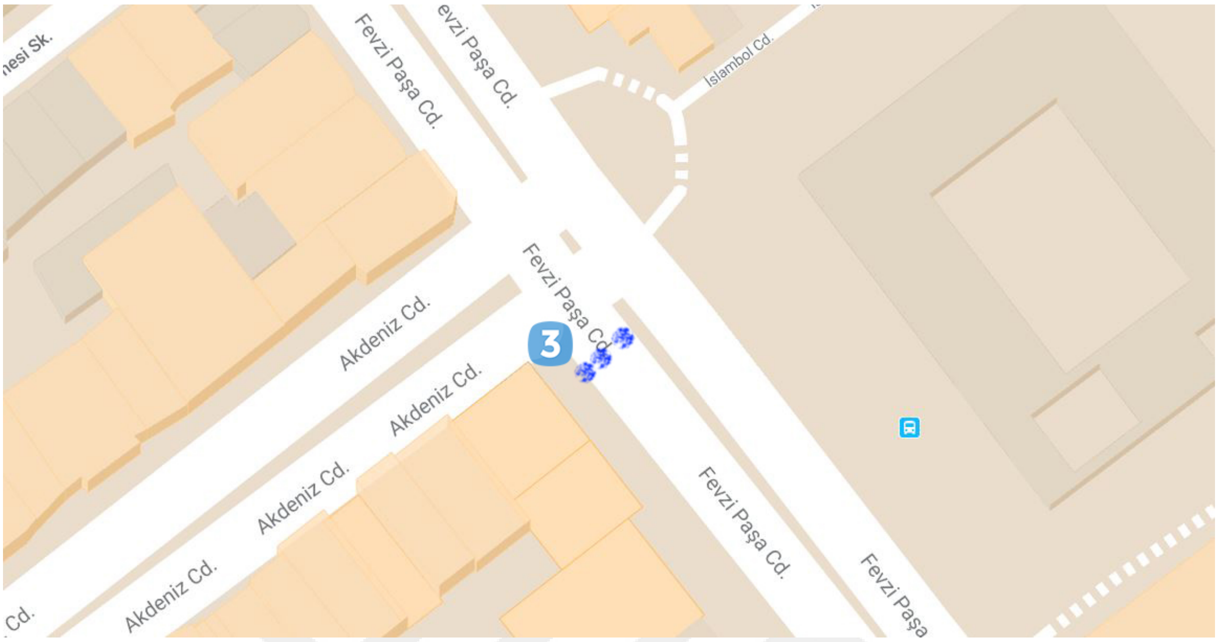

Figure 6: Malta Marked Crossing

3.3. Data Analysis Methods

Correlation is used to understand the association between two continuous variables. Pearson test was used for variables that conform to the normal distribution and the Spearman test was applied for those which do not.

The differences between the groups were evaluated with the Chi-Square test for categorical variables. For continuous variables, in cases where there are two groups, the T-Test was used. In cases where there were more than two groups the differences between the groups were evaluated with the analysis of variance (ANOVA) test. In cases where there is a significant difference between the groups, the Tukey test was applied as a Post Hoc test to determine which groups the differences existed.

Signal color as the dependent variable, binary logistic regression was performed to evaluate the effect of independent variables on jaywalking behavior. Also, multinomial logistic regression was performed to evaluate the effects of independent variables on the crossing pattern. In addition to binary logistic regression, supervised machine learning algorithms were applied to the dataset for the same purpose. Because of the model explainability, decision trees and ensemble tree models were preferred.

3.4. Operationalization: Variables and Indicators

To understand pedestrian behavior, specifically pedestrian jaywalking behavior in signalized crossings, the dataset presented in this study is collected from the universe in Figure 7.

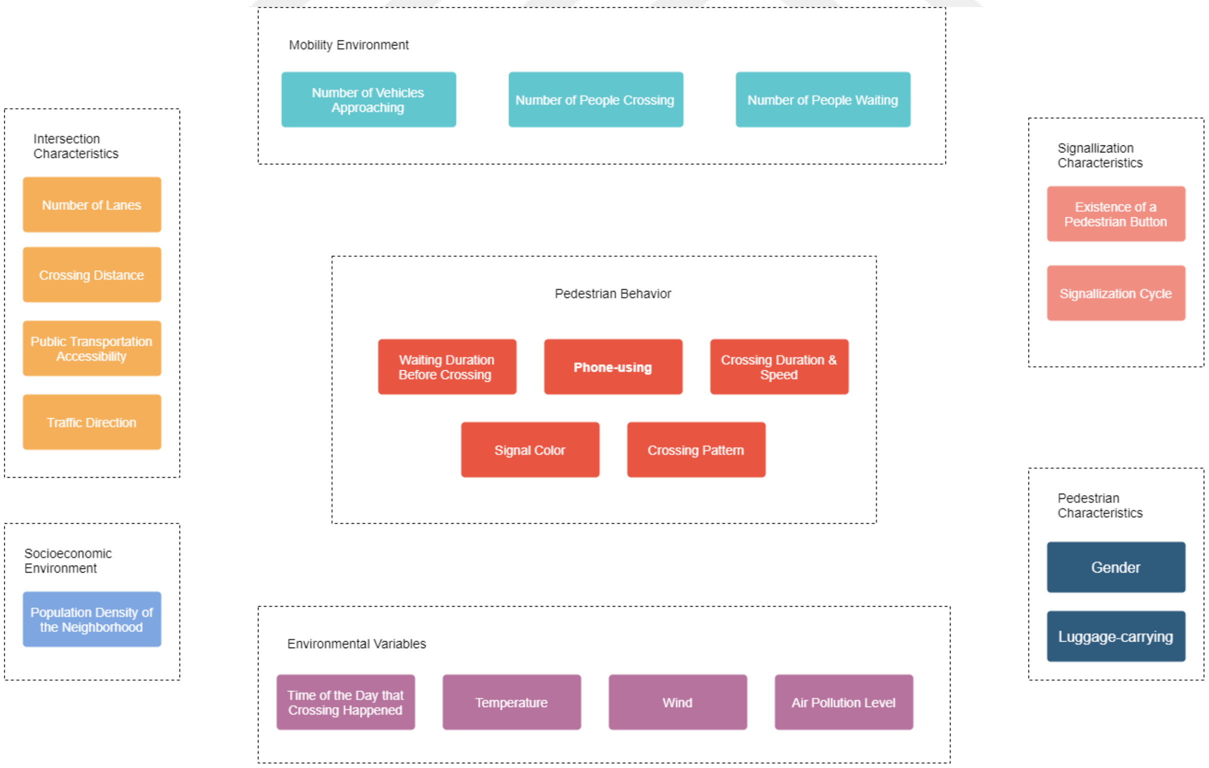

Figure 7: Variables

Each pedestrian crossing has a start and end time which respectively represents the time of pedestrian to step on the carriageway and step off to the sidewalk. In addition to the start and end times of the crossing, the duration of the crossing also used in the dataset as a variable. As marked in below maps, each intersection and each road crossing numbered separately from 1 to 2 and from 1 to 3 respectively. The direction of pedestrian movement for each crossing is also gathered from the video footage and included in the analysis. Because of the assumption that pedestrian behavior might change according to the minimum required carriageways that a pedestrian should cross to get their destination in an intersection, the total time and distance of a pedestrian in the intersection are also observed as far as the camera angles allowed. Since it is not possible to calculate the specific distance covered for each pedestrian, linear distances of crossings are used. Crossing speed for the specific crossing and also for the whole movement in the intersection is calculated and added into the dataset. The total number of lanes for each crossing and whether the crossing is unidirectional or bidirectional are included in the data.

Another assumption made is that waiting time before stepping to the carriageway might be an important signal about pedestrian behavior. Based on the assumption this timing also included in the dataset.

The existence of a pedestrian button in the crossings and (when existent) the model information was complementary variables in the dataset. Although it was not possible to observe all pushes to the pedestrian buttons because of the distance between cameras and the buttons, the quality of the video footages, and the crowd near the polls that buttons are placed; it is noted when a push to pedestrian button clearly noticed in the footages.

For each observation, the reference time is determined to be the starting moment of the crossing movement. The number of people waiting in the sidewalk to cross and the number of people that are already in the carriageway crossing are counted for every crossing according to the mentioned reference time.

For every crossing, the pattern of crossing movement is also included in the dataset as one of the single-staged, two-staged, or rolling. The gender of the pedestrian, whether or not they carry a piece of luggage, and use their phones during crossing are also included.

To be able to decide if the crossing is a legit one or jaywalking, pedestrian signal color for the reference time. In Istanbul, pedestrian signal behavior has three stages. One is green, as an indicator of the “WALK” sign. Other is red, as an indicator of the “DON’T WALK” sign. And the last one is red flash which is the transition phase from green to red and often used to tell pedestrians to speed up if they are already in the crossing, or wait if they did not start crossing yet. In this study red flash and green phases are joined into one as green. One of the reasons is detecting the red flash is not possible using video footages. Also, it is observed that red flash is ineffective for stopping pedestrians on the sidewalk but only makes them speed up; and it is seen as an extension of green light by pedestrians. Another important reason is that crossing during a red light is considered as jaywalking, but crossing during green or red flash is not. Pedestrian green time for the respective signal cycle for each crossing is also added to the dataset as a complementary variable.

The number of approaching vehicles to the crossing for the reference time of a pedestrian cross is counted in a 20-meter distance as the video footages allowed.

Basic weather data including temperature and wind is also used for analysis. Since in the day of video recording (December 17, 2019), any rainfall is not observed in the intersection locations, rainfall data is not used.

Public transportation access to the crossings is also considered for data preparation. Distance to nearest public transport stop, frequency of public transport for the nearest stop in hourly intervals, and variety of routes accessible by the nearest stop in hourly intervals are calculated and included in the dataset.

As a complementary demographic data of the intersection neighborhoods, population intensity is included in the dataset.

Distance to the nearest point of interest is also calculated and included.

Recorded air pollution for the nearest air pollution station is also included for each intersection.

When recording crossings using video footage, the initial purpose was to record all pedestrian crossings happening in the intersections using an automated image recognition algorithm. However the data quality did not allow such approach since the position of the cameras are far from the crossings, camera angles do not cover pedestrian movement continuously in a crossing, quality of the video is not high enough, and lighting at the crossings are inadequate for an existing pedestrian tracking model to gather data. Therefore it is chosen to collect the data manually.

Only at Beşiktaş Square, there are more than a hundred thousand crossings happen in twenty-four hours. An excessive number of crossings made recording them manually with a relatively extensive variable set exhausting. Thus crossings are sampled based on the period they happen. To gain a perception of pedestrian behavior for the whole day long; for each hourly segment of the day, one minute-period is selected randomly and crossings happen in that selected minutes are recorded into the dataset. Randomly selected minute-long periods of the day are presented in Table 3.

Because of the aforementioned limitations for tracking pedestrian data; gender of the pedestrian, pushing to the pedestrian button, waiting duration in sidewalk before crossing, and the number of vehicles approaching crossing could not be gathered for some of the pedestrian crossings from video footages, so are missing in the dataset.

Below in Figure 8, Figure 9, and Figure 10, an example of pedestrian crossing journey and how it is coded is presented.

Figure 8: A male pedestrian waiting on the sidewalk to cross

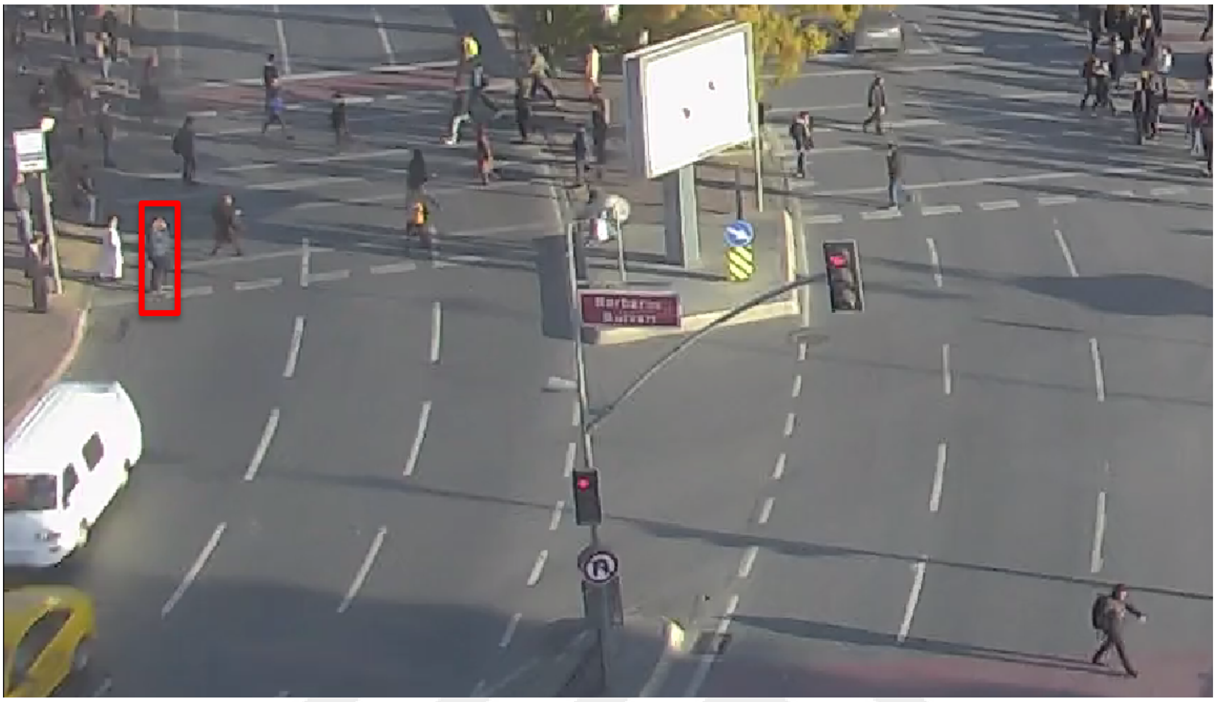

Figure 9: Pedestrian startied to cross the road at 9:53:17 while using his phone

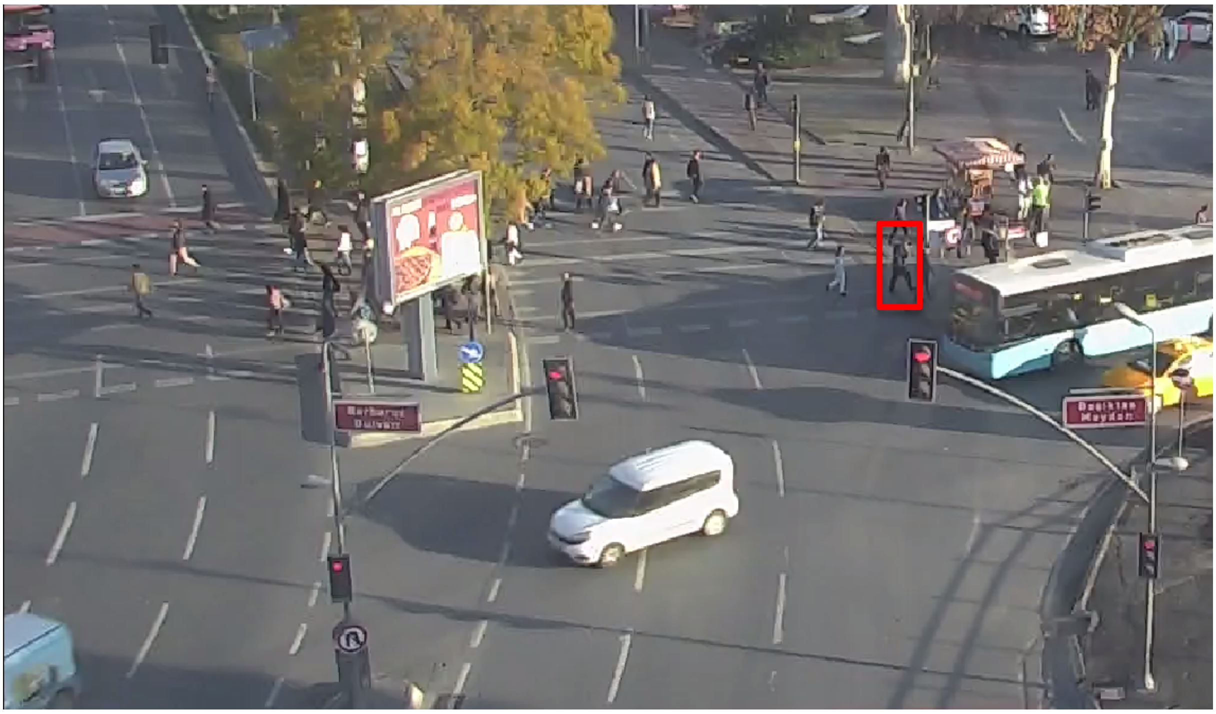

Figure 10: Pedestrian crossing ended at 9:53:37

Table 3: Sampled time intervals

Hour	Minute (Randomly Selected)	Sampled Time Intervals
0	36	00:36:00-00:36:59
1	18	01:18:00-01:18:59
2	29	02:29:00-02:29:59
3	45	03:45:00-03:45:59
4	14	04:14:00-04:14:59
5	17	05:17:00-05:17:59
6	53	06:53:00-06:53:59
7	49	07:49:00-07:49:59
8	59	08:59:00-08:59:59
9	53	09:53:00-09:53:59
10	46	10:46:00-10:46:59
11	14	11:14:00-11:14:59
12	02	12:02:00-12:02:59
13	36	13:36:00-13:36:59
14	19	14:19:00-14:19:59
15	47	15:47:00-15:47:59
16	02	16:02:00-16:02:59
17	07	17:07:00-17:07:59
18	59	18:59:00-18:59:59
19	45	19:45:00-19:45:59
20	36	20:36:00-20:36:59
21	44	21:44:00-21:44:59
22	36	22:36:00-22:36:59
23	41	23:41:00-23:41:59

CHAPTER 4

RESEARCH RESULTS AND ANALYSIS

4.1. Data Analysis

4.1.1. Understanding data: Descriptive Statistics

Dataset consisted of 2376 observations each representing a pedestrian crossing. In this study, the unit of analysis is crossing. In 1859 crossings, pedestrian gender; in 534 crossings, pedestrian button activity; in 1444 crossings waiting time on the sidewalk before crossing; in 92 crossings, the number of vehicles approaching the crossing could not be identified, so are missing in the dataset. It should be noted that missing information about waiting time before crossing tends to be bigger numbers then identified values since the longer pedestrians wait in the sidewalk place get crowded and the identification process gets difficult.

In the sampled recordings throughout the day, 2246 of crossings happened in Beşiktaş Square intersection, and 130 of them happened in Fatih Malta. 1310 crossings happened in Crossing-1, 936 crossings happened in Crossing-2, 130 crossings happened in Crossing-3.

Table 4: Crossing Frequencies

		Frequency	Percent
Beşiktaş Square	Crossing 1	1310	55.1
Beşiktaş Square	Crossing 2	936	39.4
Fatih Malta	Crossing 3	130	5.5
Total		2376	100.0

Frequency of crossings during day changes as presented in Figure 11. The number of crossings is at a peak around 17:00. Most of the crossings happened in the afternoon hours. Pedestrian activity reduced after midnight as expected.

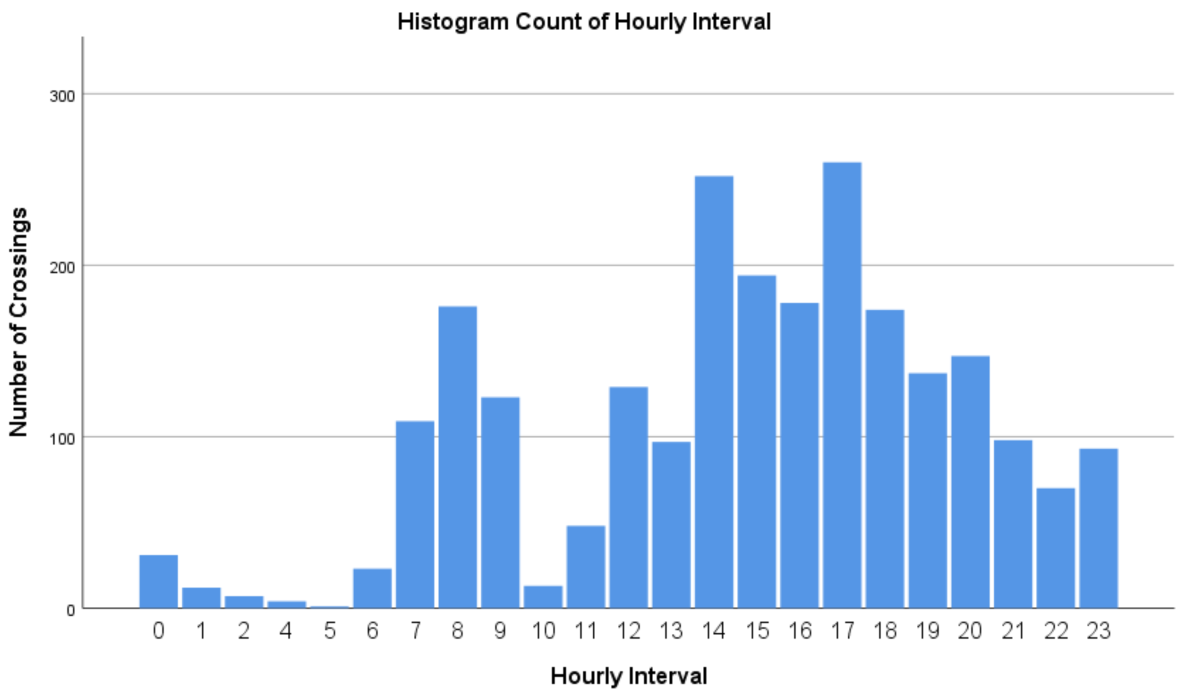

Figure 11 Crossing Distribution During Hours of Day

30.6% of the pedestrians whose gender is identified are female and 69.4% of them are male. Only 0,5% of the pedestrians are carrying luggage and 1.6% are using their phones while crossing.

782 crossings of 2376 recorded happened during the red (DON'T WALK) signal is active for pedestrians. That means 32.9% of crossings recorded are jaywalking and 67.1% of them are rule-abiding crossings.

As shown in Table 5, most of the crossings occurred in a single staged manner.

Table 5: Crossing Pattern Frequencies

Crossing Pattern	**Frequency**	**Percent**
Single Staged	1787	75.1
Two Staged	418	17.6
Rolling	176	7.4
Total	2381	100.0

Continuous variables are described by their means as follows and in Table 6. Crossing speed varies between 0.22-4.29 meters per second with a mean of 1.36 m/s. The number of people waiting on the sidewalk when a crossing happens varies between 0-73 with a mean of 8.53. For each pedestrian crossing, total number of people concurrently crossing the carriageway varies between 1-91 with a mean of 26.33. The number of vehicles approaching the crossing varies between 0-14 with a mean of 3.85.

Table 6: Descriptive analysis of continuous variables

		Waiting Duration Before Crossing	Crossing Speed	Number of People Waiting	Number of People Crossing	Number of Approaching Vehicles
N	Valid	932	2376	2376	2376	2284
	Missing	1444	0	0	0	92
Mean		5.48	1.36	8.53	26.33	3.85
Std. Deviation		14.22	0.47	13.15	19.85	3.41
Minimum		0	0.22	0	1	0
Maximum		108	4.29	73	91	14

4.1.2. Bivariate Analysis

4.1.2.1. Signal Color - Jaywalking

In crosstab in Table 7, the distribution of jaywalking behavior for different intersections are presented.

There is no statistical significance of jaywalking for different intersections. However, when looked closely, it is seen that for different crossings of Beşiktaş Square intersection; jaywalking behavior changes apparently. This can be explained by the pedestrian and vehicle traffic density difference on two roads. The number of vehicles approaching during a pedestrian cross might be used to indicate vehicle traffic. Similarly, the total number of crossing people might be used to indicate pedestrian traffic.

Table 7: Crossings at intersections by rule-complying

Intersection		Signal Color		Total
		Red	Green	
Beşiktaş Square	Count	747	1499	2246
	% within Intersection	33.3%	66.7%	100.0%
Fatih Malta	Count	35	95	130
	% within Intersection	26.9%	73.1%	100.0%
Total	Count	781	1594	2376
	% within Intersection	32.9%	67.1%	100.0%

Table 8: Rule complying comparison for Crossings 1 and 2

Crossing		Signal Color		Total
		Red	Green	
Crossing 1	Count	213	1097	1310
	% within Crossing	16.3%	83.7%	100.0%
Crossing 2	Count	534	402	936
	% within Crossing	57.1%	42.9%	100.0%

For Crossing 1 and Crossing 2, the distribution of the number of approaching vehicles and the number of people crossing behaves differently. Crossing 1 apparently more crowded both in terms of pedestrians and vehicles than Crossing 2 as seen in Table 9.

Table 9: Pedestrian and vehicle densities during crossings for Crossings 1 and 2

		Crossing 1	**Crossing 2**
Number of People Crossing	> Median(24)	785 (59.9%)	332 (35.5%)
	<= Median(24)	525 (40.1%)	604 (64.5%)
Number of Approaching Vehicles	> Median(3)	805 (61.5%)	293 (31.3%)
	<= Median(3)	505 (38.5%)	643 (68.7%)

Another contributing factor to the difference may be the existence of a pedestrian button since Crossing 1 has a pedestrian button, but Crossing 2 does not have one.

Characteristics of the location that intersection exists showed associations with jaywalking behavior. Distance to the nearest point of interest (with a correlation coefficient of 0.39), the variety of public transportation routes (with a correlation coefficient of 0.2) and frequency of public transport (with a correlation coefficient of 0.11) vehicles of the nearest transport stop are found to be correlated with jaywalking behavior. This can be explained by the pedestrian attempts to catch the transport or an event.

The number of people waiting to cross is found to be inversely correlated with jaywalking with a correlation coefficient of -0.18.

Another significant finding from the analysis is that the existence of a pedestrian button is strongly correlated with the number of jaywalkings as in Table 10 ($p<0.001$).

Table 10: Pedestrian Button and Rule-complying crosstab

Pedestrian Button		**Signal Color**		**Total**
		Red	**Green**	
Doesn't Exist	Count	569	497	1066
	Percent	53.4%	46.6%	100.0%
Exists	Count	213	1097	1310
	Percent	16.3%	83.7%	100.0%

As expected and already mentioned in the literature, there is a difference of jaywalking behavior found between genders. Table 11 shows how male pedestrians are more likely to jaywalk (p=0.004).

Luggage carrying did not change pedestrian jaywalking behavior in a significant way. However, the pedestrians who jaywalk are more likely to use their phones while crossing than pedestrians who cross during the green signal. This may be explained with the tendency to rule-breaking habit or lack of attention to traffic. Strikingly, 69.2% of people who used their phones during crossing also happened to be jaywalkers.

Table 11: Gender and Rule-complying Crosstab

Gender		Signal Color		Total
		Red	Green	
Female	Count	56	102	158
	Percent	35.4%	64.6%	100.0%
Male	Count	176	183	359
	Percent	44.9%	55.1%	100.0%

Table 12: Phone-using Rule-complying Crosstab

Phone Using		Signal Color		Total
		Red	Green	
Not using	Count	755	1582	2337
	Percent	32.3%	67.7%	100.0%
Using	Count	27	12	39
	Percent	69.2%	30.8%	100.0%

For pedestrians who cross during WALK sign and DON'T WALK sign, crossing duration, total time spent in crossing (waiting time + crossing time), and also crossing speed show a statistically significant difference ($p<0.001$). Interestingly jaywalkers are seen to cross in longer crossing duration so have lower crossing speeds. Although it is mostly a time-saving attempt, jaywalking does not seem to save any time for who practices it since the vehicle traffic is still proceeding in the carriageway that they are trying to cross.

Wait time before crossing and the signal color is correlated with a significance of $p<0.001$ and a correlation coefficient of -0.12.

Signalization does not save time for pedestrians whether they obey the rule or not. It does not stop many people from rule-breaking. What it does is creating a pseudo-safety for drivers only. By doing this, it creates more danger for pedestrians.

There is a visible effect of the number of vehicles approaching on jaywalking behavior. With the number of approaching vehicles increases, the number of jaywalking cases drops as in Figure 12 ($p<0.001$). This is an indicator of that jaywalking is not blind risk-taking behavior but mostly a conscious decision (for saving one's time or effort).

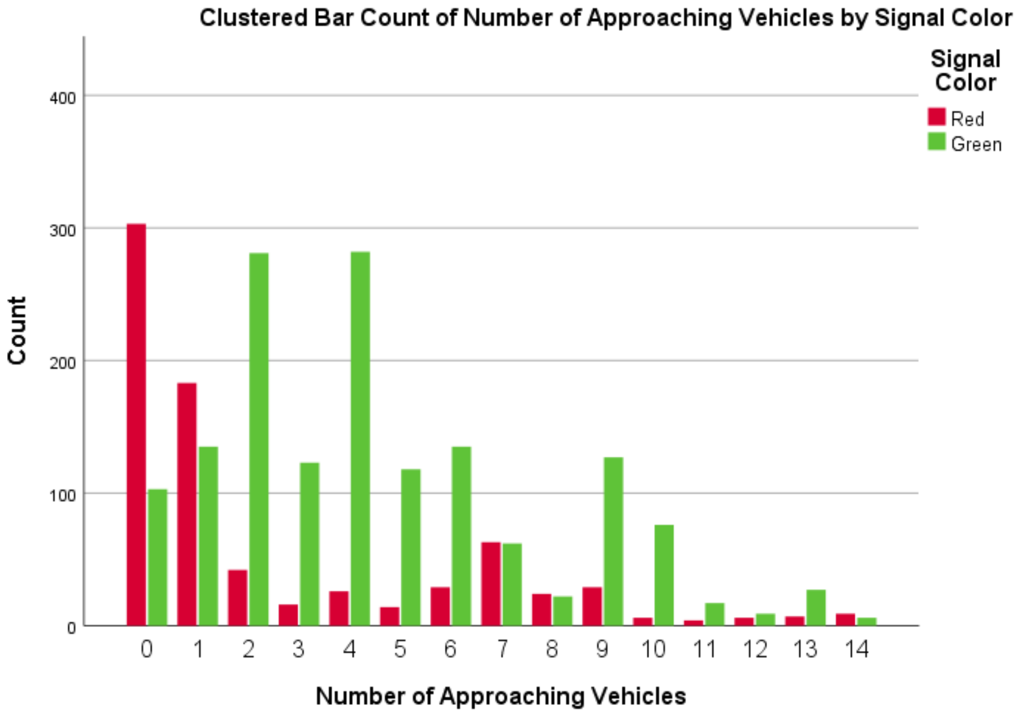

Figure 12 Crossing Counts by Number of Approaching Vehicles

Another variable that affects jaywalking behavior is the duration of the WALK signal for the corresponding signalization cycle to the pedestrian crossing. There is a pairwise correlation with a correlation coefficient of -0.31 between those variables. One possible explanation for the relationship requires a better understanding of a signalization cycle.

In a regular signalized intersection, green time for the pedestrian green phase is a small fraction of the total cycle length. And with pedestrian green time increasing, it is expected for green times for motor vehicle green phase durations to increase too. This is also the case for Beşiktaş Square and Fatih Malta intersections too. Therefore the increase in the duration of WALK sign indicates a longer duration for DON'T WALK sign and cycle length. Cycle programs of intersections vary based on physical geometry (number of lanes, number of directions, number of phases/movements etc.), traffic density, and the time of the day or week.

With this perception, it can be interpreted that when green time is lower than 20 seconds, it is more likely for pedestrians to wait for WALK sign to cross since the total cycle length is also fairly short. In addition, when the pedestrian green signal duration is more than 50 seconds, it is more likely for pedestrians to jaywalk; since in a multi-way intersection, the pedestrian red duration is incrementally longer in practice.

4.1.2.1.1. Jaywalking as a group

The total number of people crossing is found to be correlated with the jaywalking decision with a correlation coefficient of 0.35. People are less likely to jaywalk in overcrowded groups and as a trend number of jaywalking cases decreases when more people are crossing. However, there appears a good size of a group that maximizes jaywalking cases, possibly encouraging people to jaywalk and making them feel safe enough, which is around five or six people.

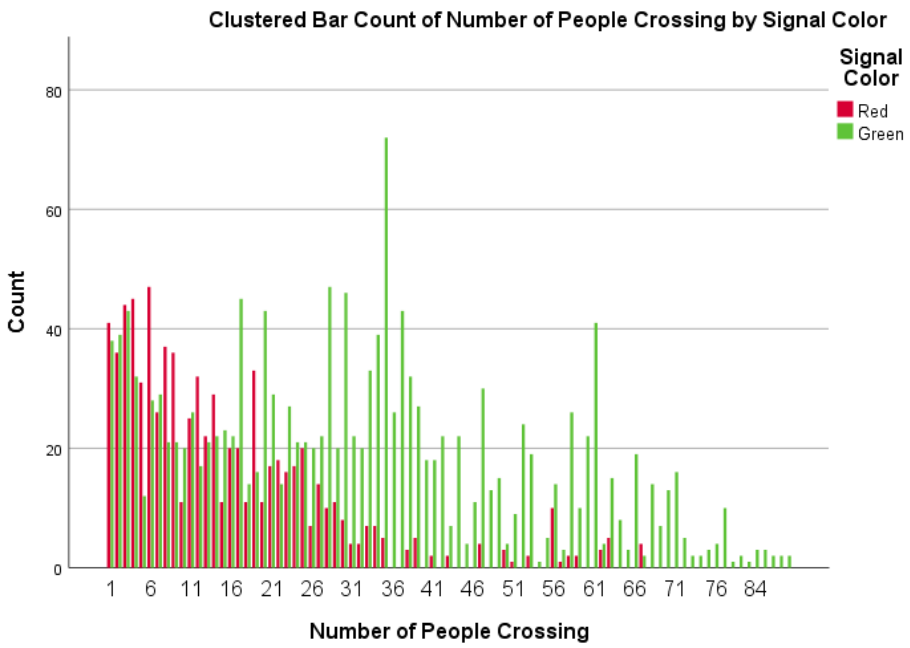

Figure 13 Jaywalking in groups

4.1.2.2. Crossing Pattern

Crossing pattern changes with the number of approaching vehicles with a correlation coefficient of -0.11. With the increasing number of vehicles coming towards crossway, pedestrians are less likely to cross in a single stage.

The relationship between intersections that crossing occurred and crossing patterns are striking as in Table 13. In Beşiktaş people tend to cross in a two-staged or rolling manner much more than they do in Fatih Malta intersection. The reason behind this difference made itself apparent in the crosstab and correlation between the lane number of the road and crossing pattern. When the number of lanes increases, it gets harder for a pedestrian to cross a carriageway in a single staged manner, so they tend to cross in a two-staged or rolling manner.

Table 13: Crossings at Intersections by Crossing Pattern

Intersection		Crossing Pattern			Total
		Single Staged	Two Staged	Rolling	
Beşiktaş Square	Count	1659	418	169	2246
	Percent	73.9%	18.6%	7.5%	100.0%
Fatih Malta	Count	123	0	7	130
	Percent	94.6%	0.0%	5.4%	100.0%

Between the number of lanes and crossing patterns, a correlation found with a statistical significance of $p=0.016$ and a correlation coefficient of 0.05.

Table 14: Crossing Pattern by Number of Lanes

Number of Lanes		Crossing Pattern			Total
		Single Staged	Two Staged	Rolling	
2	Count	123	0	7	130
	Percent	94.6%	0.0%	5.4%	100.0%
5	Count	681	230	25	936
	Percent	72.8%	24.6%	2.7%	100.0%
6	Count	978	188	144	1310
	Percent	74.7%	14.4%	11.0%	100.0%

One of the most important findings of the study is the relationship between jaywalking behavior and crossing pattern. This relationship helps to understand the underlying phenomena in the dataset. Pedestrians' decision to jaywalk or not also affects their crossing pattern. When people cross in WALK sign (green signal for pedestrians) they are more likely to cross the carriageway in a single stage. However, when they are jaywalking, they need to stop in the pedestrian refuge island.

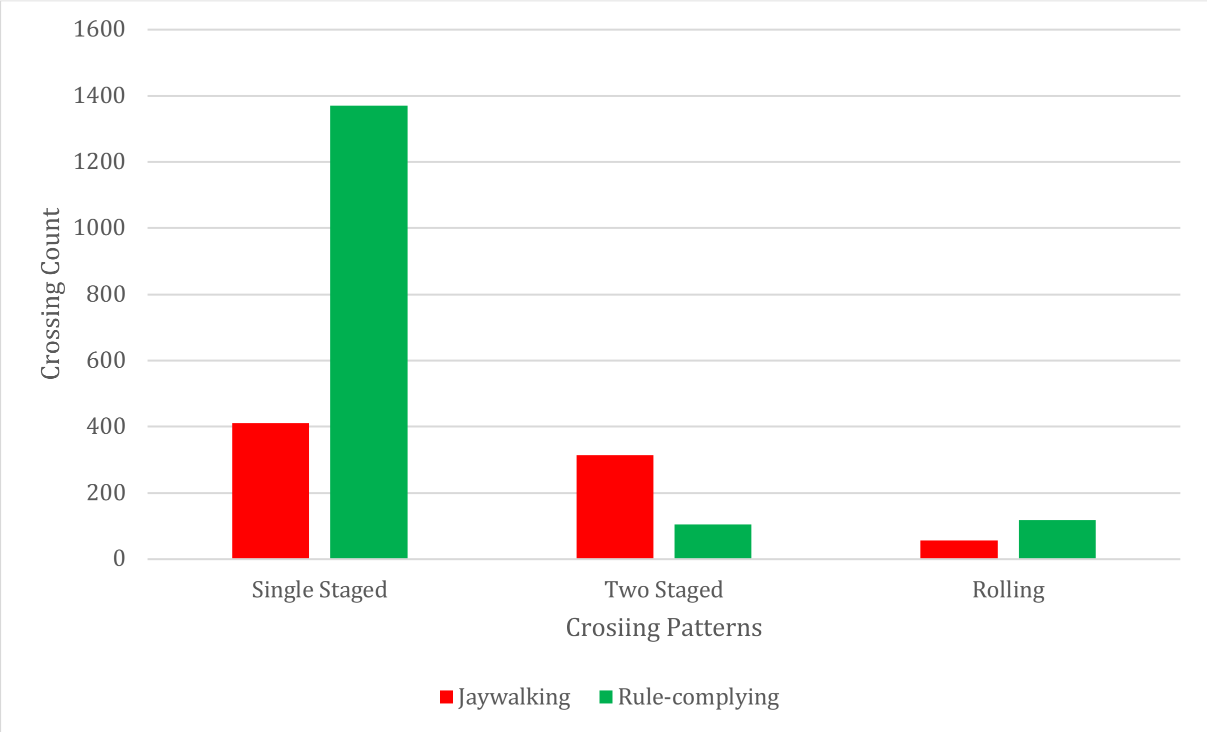

Figure 14: Jaywalking behavior by crossing pattern

Table 15: Crossing Pattern by Rule-complying

Crossing Pattern		Signal Color		Total
		Red	Green	
Single Staged	Count	411	1371	1782
	Percent	23.1%	76.9%	100.0%
Two Staged	Count	314	104	418
	Percent	75.1%	24.9%	100.0%
Rolling	Count	57	119	176
	Percent	32.4%	67.6%	100.0%

The relationship between gender and jaywalking behavior was already mentioned; crossing patterns also show a significant variation between the two genders.

Table 16: Crossing Pattern by Gender

Gender		Crossing Pattern			Total
		Single Staged	Two Staged	Rolling	
Female	Count	138	17	6	161
	Percent	85.7%	10.6%	3.7%	100.0%
Male	Count	261	65	35	361
	Percent	72.3%	18.0%	9.7%	100.0%

However, this difference should not be read as causation. The reason behind this relationship is the correlation between jaywalking behavior and crossing pattern itself. When people jaywalk they are more likely to cross in two stages or as rolling.

The total time that is spent in a crossing is found to be correlated with the crossing pattern with a statistical significance $p<0.001$. Pedestrians who cross in two stages spend more time than pedestrians who cross rolling. And pedestrians who cross rolling spend more time than pedestrians who cross in a single stage. Obviously, it is not possible for the crossing pattern to directly affect the duration of crossing because the duration is set after the crossing is over. This statistical difference shows how a pedestrian comes to a crossing they expect to wait for a period and when their expectation is exceeded they tend to change their behavior (thus their crossing pattern). For most of the pedestrians, they expect to wait for a certain time, they appreciate it when it is as less as it can be, and then cross the carriageway in single-stage and continue their urban journey. When the signalization does not allow this, they tend to look for gaps in the vehicle traffic and cross the carriageway in a rolling pattern. The ones with less risk-taking behavior use islands and turn their crossing pattern into two-staged crossing. This explanation also brings the idea that it is not the demographics or other characteristics of the pedestrians that determines their crossing behavior but the amount of hostility of the urban and traffic planning for pedestrians for a crossing.

The time that a pedestrian waits before crossing shows a statistically significant difference for people who cross in different patterns (p=0.026). When looked at the pedestrians who cross as rolling, it is seen that they tend to wait before crossing less than people who cross in single-stage patterns. Waiting less before crossing and crossing in the rolling pattern are both indications of impatience and time-saving attempts. Homogenous subsets of waiting times before crossing of pedestrians with a certain crossing pattern based on the harmonic mean is below.

Table 17: Homogenous subsets of waiting times by Crossing Patterns

Crossing Pattern	N	Subset for alpha = 0.05	
		1	2
Rolling	81	1.65	
Two Staged	139	4.80	4.80
Single Staged	712		6.05
Sig.		0.15	0.74

4.1.2.3. Timing of the Crossing

The number of lanes and wait time before crossing is correlated with a correlation coefficient of -0.07. When the number of lanes in a carriageway is increased, time that a pedestrian waits before crossing that way is increasing as well. This is because the number of lanes is also related to the traffic volume of the carriageway and signalization cycle length of the intersection. Wider carriageway indicates a longer green time for both vehicles and pedestrians. It also indicates longer crossing time for a pedestrian. Therefore a pedestrian arrives at a crossing on a wider carriageway, they usually need to wait longer for WALK sign or having consecutive gaps in vehicle traffic to jaywalk. Similarly when the number of lanes increases, crossing speed of pedestrians decreases with a correlation coefficient of -0.17.

4.1.3. Statistical Models

Lastly to go beyond bivariate relationships and understand the dataset more holistically, regression analysis, and supervised machine learning algorithms are used. By labeling signal color attribute for pedestrian crossing data as the dependent variable, models are created to understand jaywalking. To understand jaywalking behavior, the signal color variable is set as the dependent variable (label in machine learning algorithms). Since the aim of the study is not to create a prediction model for jaywalking, but understanding the jaywalking phenomena; algorithms that create interpretable models are preferred.

Before creating statistical models, variables that are used to derive new variables were removed from the dataset such as crossing distance.

k-NN, Naive Bayes, Decision Tree, Random Forest, and Gradient Boosted Trees algorithms are iteratively run and their performances were compared. By using 10-fold cross-validation, the dataset split into 10 subsets, and iteratively 9 folds were used as the training set and remaining one used as the test set. This process was repeated for 10 times for different 10 folds. Tuning models this way, overfitting was prevented. Because of the very limited space for interpretation of the k-NN model and lack of accuracies of Naive Bayes and Random Forest models were not investigated any further.

4.1.3.1. Logistic Regression

To evaluate the effects of independent variables on jaywalking behavior at observed signalized intersections, signal color variable is set as the dependent variable. Since signal color is a binary variable with possible values of red and green, logistic regression is applied to the dataset. The performance table of the regression model is in Table 18.

Table 18: Logistic regression performance

Observed		Predicted		
		Signal Color		Percentage Correct
		Green	Red	
Signal Color	Green	1332	191	87.5
	Red	249	512	67.3
Overall Percentage				80.7

This model explains 48% (Nagelkerke R^2) of the variance in jaywalking and correctly classified 80.7% of the cases. As seen in Table 19; hourly interval, phone-using, the number of people crossing, the number of approaching vehicles, green duration in signalization cycle, distance to the nearest public transportation stop, route variety for the nearest public transport stop, temperature and wind contribute to the regression equation.

According to the model, phone-using pedestrians are 3.64 times more likely to jaywalk than pedestrians who do not use their phones while crossing. Increasing wind and the decreasing temperature is associated with an increased likelihood of jaywalking. Since the data is collected in December, this makes sense as pedestrians' patience would reduce in cold and windy weather. It should be noted, wind and temperature influence on jaywalking behavior would be different in summer. Crossings that happened later hours of the day are more likely to be jaywalking than the ones in early hours. Decreasing distance to the nearest public transport stop and increasing route variety in the nearest public transport stop is associated with an increased likelihood of jaywalking. Attempts of pedestrians to catch an event or a public transport vehicle can be the main motivation behind this association. Increased green duration in parallel with increased cycle length is associated with an increased likelihood of jaywalking. Increased vehicular and pedestrian volume is also associated with an increased likelihood of jaywalking.

Table 19: Regression equation variables

							95% C.I.	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Hourly Interval	0.21	0.03	50.18	1	**<0.001**	1.24	1.17	1.31
Luggage Carrying	-1.18	0.75	2.48	1	0.115	0.31	0.07	1.34
Phone Using	1.29	0.43	8.84	1	**0.003**	3.64	1.55	8.52
Number of People Crossing	-0.08	0.01	254.68	1	**<0.001**	0.93	0.92	0.94
Number of Approaching Vehicles	-0.16	0.02	65.84	1	**<0.001**	0.85	0.82	0.88
Green Time	0.05	0.01	12.99	1	**<0.001**	1.05	1.02	1.08
Distance to Nearest Stop	-0.03	0.01	22.01	1	**<0.001**	0.97	0.96	0.98
Bus Frequency for Nearest Stop	-0.04	0.03	1.50	1	0.221	0.97	0.91	1.02
Route Variety for Nearest Stop	0.16	0.07	5.11	1	**0.024**	1.17	1.02	1.34
Distance to POI	0.02	0.01	1.78	1	0.182	1.02	0.99	1.05
Temperature	-0.53	0.08	48.86	1	**<0.001**	0.59	0.51	0.68
Wind	0.55	0.06	72.41	1	**<0.001**	1.73	1.52	1.96
Constant	0.12	0.89	0.02	1	0.891	1.13		

4.1.3.2. Decision Tree

The decision tree algorithm generates a tree-like model with nodes including a root node, interior nodes each corresponding to one attribute, and leaf nodes with class attributes. Nodes are connected with edges that state possible value intervals for their origin nodes one directionally.

To build a decision tree model, the information gain criterion is preferred to find splits and maximal depth is selected as 12. Precision, recall, and F-measure values for the model are in Table 20. The decision tree algorithm obtained a useful model, with an accuracy of 96.43%.

Table 20: Decision Tree Performance

	Precision	Recall	F-Measure
Green	95.05%	99.68%	97.31%
Red	99.35%	90.48%	94.71%

Weights of the most important variables of the decision tree model with corresponding weight percentages are presented in Table 21.

Luggage carrying, distance, speed, phone-using, wind, number of lanes, the existence of a pedestrian button, air pollution, and traffic density attributes were found impractical for the model.

The number of people crossing repeated to appear in the tree model with different intervals at different depth levels. It is highly effective for most of the paths of the tree.

At the root level, the split is dependent on the number of approaching vehicles variable with values higher than one or else. At deeper levels of the tree model; a variety of values for the number of approaching vehicles variable created more splits for internal nodes.

The number of people waiting variable is also repeated for many paths of the tree at relatively deep levels. It is usually positioned at a leaf node or close depth to a leaf node. For most of the splits, it having a value more than splitting value leads to red class labels which indicate crossings happen while more people are waiting to cross is more likely to be jaywalking.

Table 21: Decision Tree Weights of Variables

Variable	Importance Percentage
Number of people crossing	19.03%
Number of approaching vehicles	13.50%
Number of people waiting	11.88%
Total speed (for multiple crossings)	8.89%
Temperature	8.85%
Crossing duration	7.40%
WALK sign duration	7.35%
Crossing id	4.87%
Variety of public transport routes for the nearest stop	4.53%
Gender	3.85%

Crossing id variable is used mostly in leaf nodes of the tree model. Although its weight is important for the model because of the lesser depth of corresponding paths, it is only relevant for a very small number of cases of pedestrian crossings. The model tends to label crossings at Crossing 2 as red, and at Crossing 3 as green if it comes to those splits.

Other variables tend to appear in the model as frequent and in complex ways as their weight values suggest.

4.1.3.3. Gradient Boosting

Gradient Boosting algorithm created the best accuracy among other algorithms with 97.27%.

It is simply an algorithm that forms an ensemble of weak learner models to obtain improved estimations.

Precision, recall, and F-measure values for Green and Red classes are presented in Table 22.

Table 22: Gradient Boosting Performance

	Precision	Recall	F-Measure
Green	96.83%	99.03%	97.92%
Red	98.14%	94.05%	96.05%

The gradient boosted tree algorithm is iteratively tuned to reach the highest possible accuracy. Highest accuracy achieved with the number of trees between 140-150, maximal depth of 5, and learning rate of 0.1. AUC(Area under curve) value for the gradient boosted tree model is 0.99%. The algorithm calculates the importance/weights of variables for the model. The most important variables and their importance percentages are in Table 23.

It is also seen that traffic direction, air pollution, distance to the nearest point of interest, the existence of a pedestrian button, number of lanes, and distance attributes did not contribute to the model.

Table 23: Gradient Boosting Weights of Variables

Variable	Importance Percentage
Number of approaching vehicles	24.26%
Number of people crossing	14.31%
Number of people waiting	11.13%
Crossing pattern	8.21%
Public transport frequency of the nearest stop	6.99%
Crossing speed	6.54%
Crossing duration	4.90%
Total speed (for multiple crossings)	4.65%
WALK sign duration	3.52%
Hour of the day	3.32%

Despite some differences, models suggest the importance of the number of approaching vehicles, number of people crossing, number of people waiting on the sidewalk, crossing duration, crossing speed, and WALK sign duration on the jaywalking behavior. Some of these variables indicate causes of jaywalking, and some of them point out its outcomes or reflections on the pedestrian behavior in general as explained by bivariate relationships. All of them together provide a broad framework to understand pedestrian jaywalking phenomena. Although bivariate relationships of variables showed some other correlations with a statistical significance which are mentioned earlier, they lack meaningful contributions to a holistic view so they are left out.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

This thesis examines the factors that affect pedestrian crossing behavior at signalized intersections. The thesis research is conducted at two signalized intersections, Beşiktaş Square and Fatih Malta, in Istanbul. The study researches the mobility environment, intersection characteristics, socioeconomic environment, pedestrian characteristics, environmental variables, and signalization characteristics.

Road crossings are shared spaces for all road-users regardless of their chosen transportation mode. Yet, the advantage and priority are usually granted to vehicles de facto in signalized urban road-crossings. This results in attempts of pedestrians to save time and effort in these shared spaces. The findings of the study are subject to cultural effects as the other studies in the field. Despite varying percentages or numbers, the pedestrian behavioral phenomenon is similar across cultures.

5.1. Links between Research and Literature

The pedestrian violation rate (32.9%) found in Istanbul is higher than 21.6% found in Lanzhou (China) by Gong et al. (2019, p. 4) and 11.5% found in Guangzhou (China) by Ma et al. (2020, p. 4).

Gender differences in illegal crossing behavior are observed mostly as in the literature. Male pedestrians are more likely to jaywalk than female pedestrians. The difference in violation rate of two genders is observed as 9.5%, less than 24.6% found by Mo and Mo (2017, p. 48) and more than 3% found by Gong et al (2019, p. 4). The findings on the number of people waiting to cross are also corroborative with Mo and Mo's study. When the number of people waiting on the sidewalk to cross increases the rates of jaywalking decreases. People acknowledge the perception of risk of the group.

75.1% of people crossings are observed to be in a single-stage manner, while 17.6% of them are in two-stage and 7.3% are in the rolling pattern. Similarly in Marisamynathan and Vadagiri's (2018, p. 4) study, 69% of crossings happened single-stage and others happened in two-stage.

The average speed of pedestrians is observed as 1.36 m/s. Pedestrian speed when crossing found higher in Istanbul than Izmir where it is observed as 1.31 (Onelcin and Alver, 2017, p. 6) and Mumbai where it is 1.21 m/s (Marisamynathan and Vedagiri, 2018, p. 4).

Cycle length of the signalized intersection is found to be correlated with jaywalking behavior as in Marisamynathan and Vedagiri's (2018, p. 6) research.

5.2. General Conclusions

Returning to the research question and the variables that are used to describe a signalized intersection, it is found that the most effective factors for pedestrian behavior are vehicular and pedestrian density, time of the day that crossing has happened, signalization cycle length, and public transportation accessibility of the place. Figure 15 is a conceptual map of elements that affect pedestrian behavior in signalized crossings.

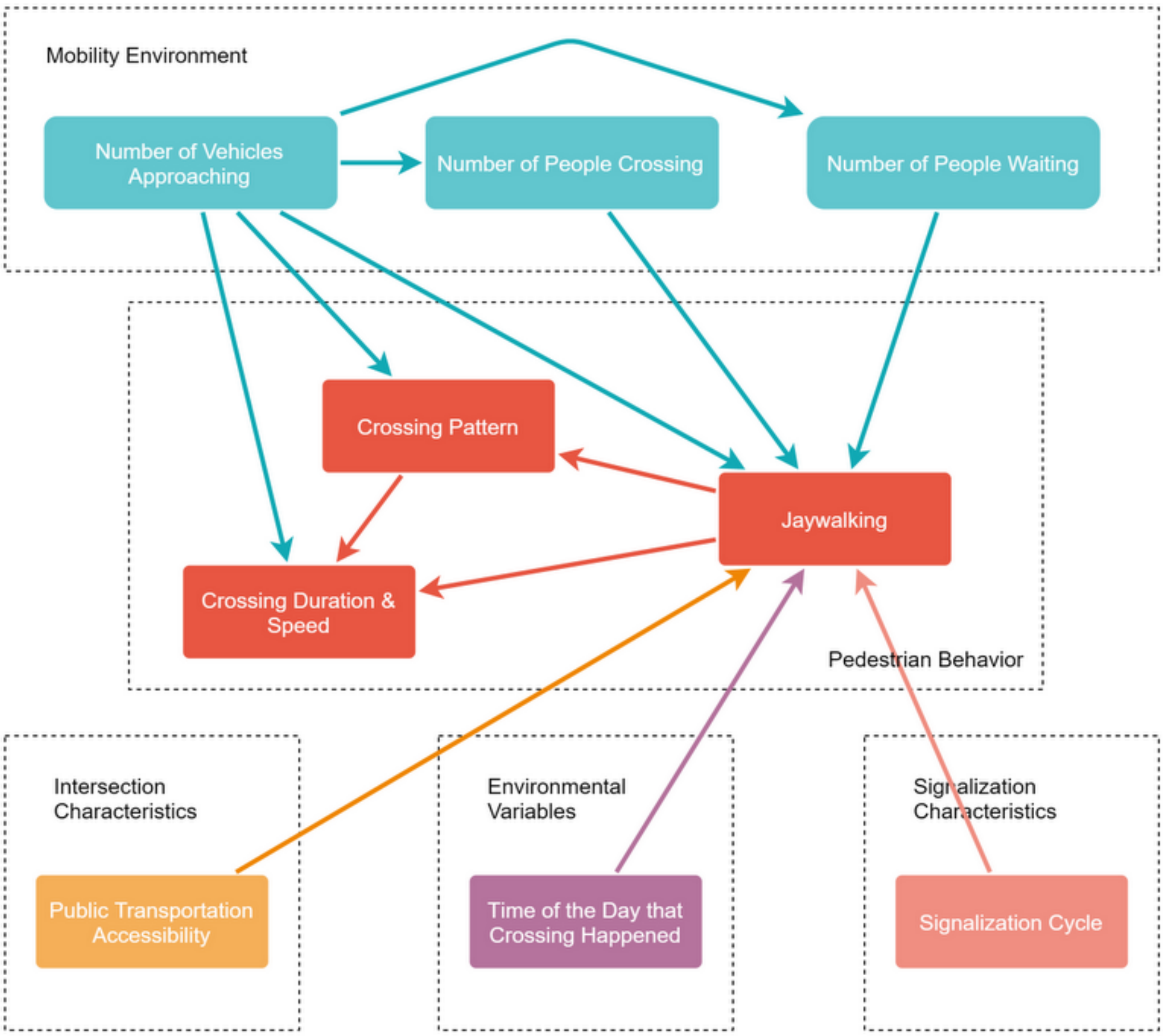

Figure 15 Conceptual Map of Pedestrian Behavior

5.3. Recommendations for Policy and Decision Makers

To reduce illegal behavior and causalities at signalized intersections, prioritizing pedestrians' time and effort at least as it is done for drivers' is a must. Signalization can be more responsive in many ways. Image recognition for traffic management by counting vehicles in a queue at a signalized intersection is an already in-use technology, it is also possible to count the pedestrians and use the information to manage signalization. This way, prioritizing disadvantaged road –users such as elderly, people with young children, or pregnant women is also possible and would be a good step to create equity in urban mobility. Also making sidewalks pleasant places to be, with enough space and thermal comfort; and enhancing public transportation network and accessibility would be simple steps with a quick and high return on investment.

5.3. Recommendations for Future of Mobility Researchers

Not just pedestrian mobility, but the future of mobility in general is currently being shaped by information and communication technologies. Research and studies in the field should consider all transportation modes to create a practical and equal future of mobility. Connected vehicles are already seen as a way to replace signalization technologies. However, there is no adequate study and possible solution for pedestrian mobility issues. Connected vehicles can have the capability of informing each other about road-use and managing traffic by managing car speed based on shared information but digitalizing and connecting pedestrian information is a missing piece of the puzzle yet. Future studies on the issue can be very useful for a healthy future mobility environment.

References

Angel, S., Parent, J., Civco, D. L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Progress in Planning*, *75*(2), 53-107.

Antić, B., Pešić, D., Milutinović, N., & Maslać, M. (2016). Pedestrian behaviours: Validation of the Serbian version of the pedestrian behaviour scale. *Transportation research part F: traffic psychology and behaviour*, *41*, 170-178.

Bansal, A., Goyal, T., & Sharma, U. (2019). Modelling the Pedestrian Speed at Signalised Intersection Crosswalks for Heterogeneous Traffic Conditions. *Promet-Traffic&Transportation*, *31*(6), 681-692.

Dallmeyer, J., Lattner, A. D., & Timm, I. J. (2012, October). Selfish road users–Case studies on rule breaking agents for traffic simulation. In *German Conference on Multiagent System Technologies* (pp. 83-95). Springer, Berlin, Heidelberg.

Dıaz, E. M. (2002). Theory of planned behavior and pedestrians' intentions to violate traffic regulations. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(3), 169-175.

Frank, L. D. (2000). Land use and transportation interaction: implications on public health and quality of life. *Journal of Planning Education and Research*, *20*(1), 6-22.

Kadali, B. R., & Vedagiri, P. (2013). Effect of vehicular lanes on pedestrian gap acceptance behaviour. *Procedia-Social and Behavioral Sciences*, *104*, 678-687.

Karayolları Trafik Kanunu. (1983). *Resmi Gazete* (Sayı: 18195). Retrieved from: https://www.mevzuat.gov.tr/MevzuatMetin/1.5.2918.pdf

King, A. C., Stokols, D., Talen, E., Brassington, G. S., & Killingsworth, R. (2002). Theoretical approaches to the promotion of physical activity: forging a transdisciplinary paradigm. *American journal of preventive medicine*, *23*(2), 15-25.

Kouabenan, D. R. (1998). Beliefs and the perception of risks and accidents. *Risk Analysis*, *18*(3), 243-252.

Ma, Y., Lu, S., & Zhang, Y. (2020). Analysis on illegal crossing behavior of pedestrians at signalized intersections based on Bayesian network. *Journal of advanced transportation*, *2020*.

Malenje, J. O., Zhao, J., Li, P., & Han, Y. (2018). An extended car-following model with the consideration of the illegal pedestrian crossing. *Physica A: Statistical Mechanics and its Applications*, *508*, 650-661.

Marisamynathan, S., & Vedagiri, P. (2018). Modeling pedestrian crossing behavior and safety at signalized intersections. *Transportation research record*, *2672*(31), 76-86.

Mavoa, S., Witten, K., McCreanor, T., & O'sullivan, D. (2012). GIS based destination accessibility via public transit and walking in Auckland, New Zealand. *Journal of Transport Geography*, *20*(1), 15-22.

Mukherjee, D., & Mitra, S. (2019). Impact of road infrastructure land use and traffic operational characteristics on pedestrian fatality risk: a case study of Kolkata, India. *Transportation in developing economies*, *5*(2), 6.

Noland, R. B. (1996). Pedestrian travel times and motor vehicle traffic signals. *Transportation Research Record*, *1553*(1), 28-33.

Onelcin, P., & Alver, Y. (2017). The crossing speed and safety margin of pedestrians at signalized intersections. *Transportation research procedia*, *22*, 3-12.

Pele, M., Deneubourg, J. L., & Sueur, C. (2019). Decision-Making Processes Underlying Pedestrian Behaviors at Signalized Crossings: Part 2. Do Pedestrians Show Cultural Herding Behavior?. *Safety*, *5*(4), 82.

Petzoldt, T. (2014). On the relationship between pedestrian gap acceptance and time to arrival estimates. *Accident Analysis & Prevention*, *72*, 127-133.

Russo, B. J., James, E., Aguilar, C. Y., & Smaglik, E. J. (2018). Pedestrian behavior at signalized intersection crosswalks: observational study of factors associated with distracted walking, pedestrian violations, and walking speed. *Transportation research record*, *2672*(35), 1-12.

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proceedings of the National Academy of Sciences, 109(40), 16083-16088.

Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Annals of behavioral medicine*, *25*(2), 80-91.

Shaaban, K. (2019). Analysis of pedestrian crossing speeds at signalized intersections in Qatar. *Arabian Journal for Science and Engineering*, *44*(5), 4467-4476.

Sheller, M., & Urry, J. (2000). The car and the city. *International Journal of Urban and Regional Research*, *24*(4), 737-757.

Sukhadia, H., Dave, S. M., Shah, J., & Rathva, D. (2016). The effect of events on pedestrian behavior and its comparison with normal walking behavior in CBD area in Indian context. *Transportation Research Procedia*, *17*, 653-663.

Tahmasbi, B., & Haghshenas, H. (2019). Public transport accessibility measure based on weighted door to door travel time. *Computers, Environment and Urban Systems*, *76*, 163-177.

Wang, W., Yuan, Z., Liu, Y., Yang, X., & Yang, Y. (2019). A Random Parameter Logit Model of Immediate Red-Light Running Behavior of Pedestrians and Cyclists at Major-Major Intersections. *Journal of Advanced Transportation*, *2019*.

Wu, Y., Lu, J., Zhu, W., & Chen, Y. (2017, August). Investigation of pedestrians crossing behavior at signalized intersections in China. In *2017 4th International Conference on Transportation Information and Safety (ICTIS)* (pp. 653-657). IEEE.

Xie, S. Q., Wong, S. C., Ng, T. M., & Lam, W. H. (2017). Pedestrian crossing behavior at signalized crosswalks. *Journal of transportation engineering, Part A: Systems*, *143*(8), 04017036.

Xiong, H., Ma, L., Ning, M., Zhao, X., & Weng, J. (2019). The tolerable waiting time: A generalized Pareto distribution model with empirical investigation. *Computers & Industrial Engineering*, *137*, 106019.

Zheng, Y., Chase, T., Elefteriadou, L., Schroeder, B., & Sisiopiku, V. P. (2015). Modeling vehicle–pedestrian interactions outside of crosswalks. *Simulation Modelling Practice and Theory*, *59*, 89-101.

Zhou, Z., Liu, S., Xu, W., Pu, Z., Zhang, S., & Zhou, Y. (2019). Impacts of mobile phone distractions on pedestrian crossing behavior at signalized intersections: An observational study in China. *Advances in Mechanical Engineering*, *11*(4), 1687814019841838.

ANNEXES

ANNEX 1: GLOSSARY

Term	Definition
Adaptive signal timing	A signalization strategy in which durations of signals changes according to actual traffic demand.
Block/Urban block	Urban area that is surrounded by streets of grid urban layout
Crossing pattern	The combination of walking and stopping movements of pedestrian while crossing in motorway
Cul de sac	One-ended short road
Cycle	Total time of signalization for all movements in an intersection
Flash/Pedestrian flash sign	Pedestrian traffic sign that implies "do not begin to cross" for pedestrians on the sidewalk and "complete crossing" for pedestrians who are already in the motorway
Gap	Time taken by the pedestrian to cross the road without conflicting the approaching vehicle
Geometry/Geometric design	Layout of the road including plans for routes, curves, lines, positions of lanes, slopes, etc.
Grid urban layout	A plan that city streets run each other at specific angles and form a grid
Intersection	Place where two or more motorways join or cross
Jaywalking	Illegal street crossing by a pedestrian
Land-use	Use of natural environment to transform into built environment
Mode/Travel Mode	Means of transportation, such as driving or walking
Peak hour/Rush hour	Time of the day that traffic congestion is at its highest
Pedestrian button	Buttons on signalized intersections that functions to create pedestrian demand and bring pedestrian green or walk signal up
Phase/Signalization phase	Intervals for each color sign to be active in a signalization cycle
Platooning	A group of road users to move together
Signalization	System that allows time allocation for conflicting movements at motorways
Stop	Pedestrian traffic sign that implies that pedestrians do not have the right to use the motorway for the moment
Walk	Pedestrian traffic sign that implies that pedestrians have the right to use the motorway for the moment