University of Zurich[UZH]

Comparison of Bike Sharing Systems (BSS) in Different Cities and the Impact of Different Influencing Factors

GEO 511 Master's Thesis

Author
Dario Hug
14-707-798

Supervised by
Prof. Dr. Ross Purves
Ralph Straumann (Ralph.Straumann@ebp.ch)

Faculty representative
Prof. Dr. Ross Purves

10.09.2020
Department of Geography, University of Zurich

Comparison of Bike Sharing Systems (BSS) in Different Cities and the Impact of Different Influencing Factors

MASTER'S THESIS

AUTHOR
DARIO HUG
14-707-798
dario.hug@uzh.ch

SUPERVISOR
PROF. DR. ROSS STUART PURVES
DR. RALPH STRAUMANN
GEOCOMPUTATION
UNIVERSITY OF ZURICH

DATE OF SUBMISSION: 10.09.2020

Abstract

Bike sharing is becoming increasingly more popular all over the world. This thesis investigates which factors influence the usage of bike sharing. The data of nine different bike sharing systems in the USA (New York, Boston, Chicago, and Washington, DC) and in Europe (London, Edinburgh, Oslo, Bergen, and Trondheim) are analysed. The analysis is conducted by using graphs as well as statistical models through which a negative binomial regression was performed. The discussed influencing variables are weather, infrastructure (road types and public transport), and land use. A further aim of this thesis is to determine if all the systems are impacted by the same variables or if there are any differences. Lastly, the thesis discusses if the bike sharing usage is more affected by global or local variables. The results indicate that bike sharing is primarily used for commuting purposes during weekdays and for leisure activities on the weekend. The ridership in the winter is smaller than in the summer. Likewise, the average trip duration is shorter in the winter than in the summer. Subscribers with an annual or monthly membership are a major part of the users. They tend to use the bikes for commuting whereas casual users utilize bike sharing for leisure activities. The average trip duration of subscribers is shorter than that for casual users. The weather variables with the greatest impact on the usage of bike sharing are precipitation and temperature. Precipitation leads to significantly less rides whereas the ridership increases with higher temperatures. The infrastructure analysis delivers insufficient results to make a general statement about the impact for all cities. The variable with the greatest impact in terms of infrastructure is altitude which had a negative impact on bike station usage. The results of the land use analysis suggest that different land use areas have an impact on the usage of bike stations. An example is that many rides start in residential areas in the morning and end in commercial areas. Another example is that casual users tend to use bike sharing more in recreational areas whereas subscribers use it in residential areas.

In summary, the results suggest that all cities are similarly affected by the same global and uncontrollable factors, such as the season, weekends, precipitation, temperature, and altitude. Alternatively, the impact of the local and controllable factors, such as infrastructure and land use, differ from city to city as each has its own infrastructure and land use features.

Acknowledgements

I would like to thank all the people who helped and supported me during the progress of this masters thesis. Especially I want to thank:

- *Prof. Dr. Ross Purves* for his excellent supervision during the whole process of this thesis. The various meetings were always very helpful and inspiring. I would like to thank him especially for his support during the Corona-related lockdown, where his kind and supportive words relieved pressure in a difficult time.
- *Dr. Ralph Straumann* for his precious advice especially in the beginning and at the end of the progress where his inputs were a valuable use.
- *All my friends and family*, especially my comrades from the ≪Masterrümli≫ which were helpfull in any cases, being it programming, writing or just for fun and entertaining breaks and lunches.

Contents

List of Figures

List of Tables

Glossary

BSS	Bike Sharing System
OSM	Open Street Map
GPS	Global Positioning System
UTC	Universal Time Coordinated
CSV	Comma-separated values
DEM	Digital Elevation Model
NED	National Elevation Dataset
ODP	Open Data Portal
ToD	Time of the Day
NOAA	National Oceanic and Atmospheric Administration
WMO	World Meteorological Organization
DWD	Deutscher Wetterdienst (National Weather Service of Germany)
US	United States
USA	United States of America
DC	District of Columbia
NYC	New York City

Chapter 1

Introduction

1.1 Motivation

Cycling is growing in significance for everyday mobility all over the world as it becomes increasingly more popular as an alternative to cars and public transport. In Switzerland, the annual distance covered by bicycle has increased by 13% since 1995 (BFS, 2018), and in the United States, the number of trips through bike sharing has risen from about 320,000 a year in 2010 to approximately 35 million a year in 2017 (National Association of City Transportation Officials, 2017). With the increased number of cyclists, the need for proper cycling infrastructure such as bike paths or cycling streets has also increased. Bike share systems (BSS) are a relatively new feature of the cycling infrastructure in many cities. Parkes et al. define BSS as "... the shared use of a bicycle fleet, which is accessible to the public and serves as a form of public transportation" (Parkes et al., 2013, p. 94). Although many systems have been developed over the last 10–20 years, the first approach for a bike sharing system occurred in 1965 in Amsterdam. Providers offered their bikes for free hoping that they would find a solution for traffic problems. However, the system did not survive long because of thefts and damages to the bikes (Frade and Ribeiro, 2014). The second generation started in 1995 in Copenhagen, Denmark. While the first-generation of shared bikes were distributed randomly in the city, the second generation of bike sharing was station based. To prevent the bikes from being stolen or damaged, each bike had to be unlocked with a coin deposit, which was refunded after returning the bike to a docking station (Shaheen et al., 2010). The third generation of bike sharing started in 1996 at the Portsmouth University in England, and some citie s still use this approach today. The system behind it was that students could rent a bike by using a magnetic stripe card to pay and unlock the bike. This method and other technological improvements like electronic locks, mobile phone access, board computers, and GPS are the characteristics of the smarter third generation (DeMaio, 2009). The fourth generation of bike sharing is more advanced, and it offers e-bikes, solar-powered docking stations, or the possibility to link a bike sharing membership to a public transport card. Moreover, the redistribution of shared bikes is increasingly more important, as many systems have become considerably large (Shaheen et al., 2010). Some systems today belong to either the third or the fourth generation, and others are in a transition phase. An example of a fourth-generation system is PubliBike in Zurich, Switzerland, which offers e-bikes and may be connected with the public transport membership (PubliBike, 2020). The concept behind the third and fourth generation of bike sharing is simple. Bikes are provided either by a city or a private organization, and then they are used by the public. These bikes provide a new way of mobility as they can be used to overcome public transport gaps, which makes them attractive for commuters. Recreational riders and tourists are frequent users of the systems as well (Frade and Ribeiro, 2014). Many contemporary BSS work with a smartphone app, which, for example, may indicate the nearest station or the bike availability of a certain station. The payment system is similar for most of the systems. Users are first required to pay a starting fee before having a free period at their disposal, which is usually between 30–60 minutes. After the free period, the users have to pay for each additional minute or kilometre that they use the bike. Purchasing a membership is also possible. Subscribers do not have to pay the starting fee, and they may have more free minutes at their disposal, depending on the BSS. These are the basic principles that most bike sharing systems have, but many different systems exist which differ from each other. An important distinction is made between "free-floating" (i.e. dockless) bikes and docked bikes. Dockless bikes can be parked anywhere within a certain area, such as a city, whereas docked bikes must be returned to official docking stations. The docked bikes are divided further into those that are either physically docked or virtually docked. On the one hand, physically docked bikes must be returned to actual docking stations where they are docked in and locked. On the other hand, virtually docked bikes are returned to docking areas, which is a place that is marked as a docking station but no physical dock exists; rather, the user parks the bike in the given spot. The difference between these two types is that the physical docking station has a limited number of spaces compared to the virtual docking station. However, the most common type are physically docked bikes. For example, in the US 96% of all shared bike trips were station-based, and only 4% were dockless (National Association of City Transportation Officials, 2017). The reason for this discrepancy is that docked bikes have many benefits. First, controlling where the bikes are is easier for the bike provider, and they do not have to recollect them from time to time. Furthermore, the bikes do not need a GPS tracker as data is recorded through the docking stations and not through GPS. For the users, docked bikes are easier to plan. With dockless bikes, a bike may not be available in the near surroundings, which makes the access more difficult. Alternatively, dockless bikes are convenient, and users can leave the bike in front of their house, for example. In addition, no full-dock problem exists for dockless bikes. This scenario can occur with the docked bikes if a user wants to leave the bike at a certain docking station, but no free dock is available. This thesis focuses on physically docked bikes as they are currently the most common system.

1.2 Objectives and Research Question

The goal of this thesis is to determine what factors influence the usage of bike sharing worldwide. A reproducible approach was used, meaning that only freely available data were used. The thesis focuses on four parts. The first part analyses how bikes are used in different cities. The second part covers the impact of weather events on the usage of shared bikes. In the third part, the influence of infrastructure (roads and public transport) is analysed. In the final part, the impact of different land use types is examined. A further aim of the work is to make a statement about the various data sources and to assess the extent to which they are suitable for an analysis of this kind. The data are from various sources, some are official and governmental data, whereas others are user-generated such as the data from OpenStreetMap (OSM). Comparing different sources is vital as one cannot expect all data to be consistent in terms of data quality and completeness. In this context, it is discussed how the different bike sharing systems and their data are suitable for such an analysis.
However, the main part of this work is the analysis of the bike data as well as the impact of different

influencing factors. Therefore, the questions to be answered with this thesis are the following: Which weather factors influence the usage of shared bikes? Are shared bikes used more often in areas with better cycling infrastructure? Are shared bikes used for commuting or for recreational purposes? Are shared bikes used more in particular land use categories? Stemming from these research questions, the results of different cities are compared to determine if the influencing factors are global or local. One assumption is that the negative impact of rain is greater in a city like New York compared to a city like London, because Londoners are more accustomed to rain than New Yorkers, and thus the former are better equipped. Although all these factors might have an influence, they differ from each other at a certain point. For instance, weather is not controllable, and it is hard to predict over long periods. Additionally, it moves quickly, and the conditions change rapidly. In comparison to the weather, infrastructure is controllable and changes slowly. These differences result in two different types of factors, which might influence the usage of shared bikes in a different way. Following these factors, the research questions are the following:

- Which factors influence the usage of shared bikes?
  - Which weather variables influence the usage of shared bikes?
  - Which infrastructural factors influence the usage of shared bikes?
  - How does land use impact the usage of shared bikes?
  - For what purpose are the system mostly used for?
- Are there any differences in use between different cities?
  - Are all cities impacted by the same variables?
  - Are all systems used the same way and for the same purposes?
- Is the usage of shared bikes influenced by global or local factors?
  - Are the systems more affected by short-term or long-term variables?

Chapter 2

State of the Art

This chapter examines the current state of the art in the topic of bike sharing analysis. Different studies have been conducted that examine which effects influence the usage of shared bikes. However, not all of the studies used the same methods or assessed the same impact variables. This chapter is divided into the following sections: "Ride Analysis," "Influence of Weather Events," "Influence of Infrastructure," and "Influence of Land Use."

2.1 Ride Analysis

This section considers the usage patterns that have been detected by different authors who have analysed various BSS around the globe. In the study conducted about Chicago's bike sharing system called Divvy, Zhou found a clear two peak characteristic on weekdays regarding the number of rides. The first peak was around 8:00, and the second was at 18:00. The peak at 18:00 was slightly higher. Furthermore, for weekends the usage was much less, which is a sign that the BSS was used for commuting purposes. Another interesting fact from that study is that during weekdays more subscribers used the service, whereas on the weekends more one-day casual users used the service. This finding leads to the assumption that subscribers use the system to commute, whereas one-day users avail of the bikes for recreational purposes (Zhou, 2015).

Caulfield et al. analysed bike sharing trends in Cork, Ireland. They found out that the majority (70%) of all rides took less than 9 minutes, and they also detected two major peaks in usage around

9:00 and and 18:00 as well as a minor peak around 13–14:00. Regarding the difference between weekend and weekday trips, 82% of all rides occurred during weekdays. The authors discovered that 18% of the users used the system on a daily basis, whereas 60% of the bike users used the BSS once or twice a week (Caulfield et al., 2017).

Faghih-Imani et al. conducted research in Montreal, Canada, and they found that the BSS called BIXI was used mostly during the afternoon. Additionally, they detected an increased usage on Friday and Saturday nights, which leads to the assumption that young people use the system to go out and then to return home. Moreover, they observed a reduced usage on weekends compared to weekdays (Faghih-Imani et al., 2014).

Sun et al. found a clear distinction between subscribers and casual users in Chicago, USA. First, the average trip duration of a subscriber was 12.1 minutes, and the average duration of a casual user's trip was 29.2 minutes. Annual riders, which accounted for 70% of all users, used the bikes predominantly in the afternoon peak hours (15–18:00) as well as in the evening (18-20:00) hours (Sun et al., 2017).

Mateo-Babiano et al. conducted a study in Brisbane, Australia and discovered two peaks for weekdays: one in the morning around 7–8:00 and one in the evening around 17:00. The peak in the afternoon was slightly higher than the one in the morning, and there was an additional smaller peak around noon. For the weekend, the usage increased gradually until approximately 11:00, from then on it was fairly stable until 17:00. After 17:00, the usage numbers decreased. While the travel time was longer on weekends than on weekdays, for both periods the travel speeds were faster in the morning and the evening and were slower during the day. Furthermore, on weekdays the trip distance was longer in the morning and evening compared to the afternoon. For the weekends, no clear pattern was visible (Mateo-Babiano et al., 2016).

El-Assi et al. examined the difference in usage between subscribers and casual users in Toronto, Canada. For the casual users the distribution between the weekend and weekdays was balanced, but subscribers used the bike more during weekdays. The authors assume that this is the consequence of many commuting rides during weekdays. Similar to other authors, they also detected a morning and an evening peak. However, they observed that the peak changes with seasonal variation. In the fall, three peaks occurred: morning, midday, and evening. In the winter the midday peak declined slightly, and a greater number of rides were taken during the morning and evening peaks. The authors found that a decline of casual users occurred in winter, whereas the commuter trips remained effectively the same (El-Assi et al., 2017).

Kim performed a study in Daejeon, South Korea. In contrast to the other studies, the usage between the weekday and weekend was not considerably different as both had the highest peak around 16–20:00. However, on weekdays a visible peak occurred during the morning rush hour around 8:00, but the peak was about half as large as the peak in the evening (Kim, 2018).

Reiss and Bogenberger analysed the ride characteristics of free-floating bike sharing in Munich, Germany. On average, each user used the bike sharing for 15 rides in the course of one year. However, 50% of all users used it less than five times a year. Alternatively, 80% of all the trips are taken by heavy users, who accounted for 20% of all users. The temporal analysis indicated results similar to other studies with a clear distinction between the weekend and weekdays. The weekday usage had the common peaks in the morning and evening as well as the smaller noon peak. For the weekend, the peak was around 15:00 (Reiss and Bogenberger, 2015).

In the study by Vogel et al. that was conducted in Vienna, Austria, they detected a visible peak around midnight. Furthermore, this peak slightly increased from Monday to Friday, which had the greatest midnight-peak followed by the one on Saturday. The only day without a midnight peak was Sunday, which only had one midday peak. In addition, the weekday usage was characterised by a large peak in the evening around 18:00 and a smaller peak in the morning between 6–8:00 (Vogel et al., 2011).

Nair et al. conducted a study in Paris about the BSS Vélib'. The ride analysis reflected the well-known two-peak characteristic with a morning and an evening peak. However, the peak in the morning was both shorter and smaller than the one in the afternoon. A small midday peak around noon was also evident. For the weekends the usage pattern was considerably different. The usage numbers started to increase around 9:00, and they peaked around 15:00 (Nair et al., 2012). Figure 2.1 offers an example of the common temporal distribution of bike sharing rides by presenting the riding patterns of Vélib'.

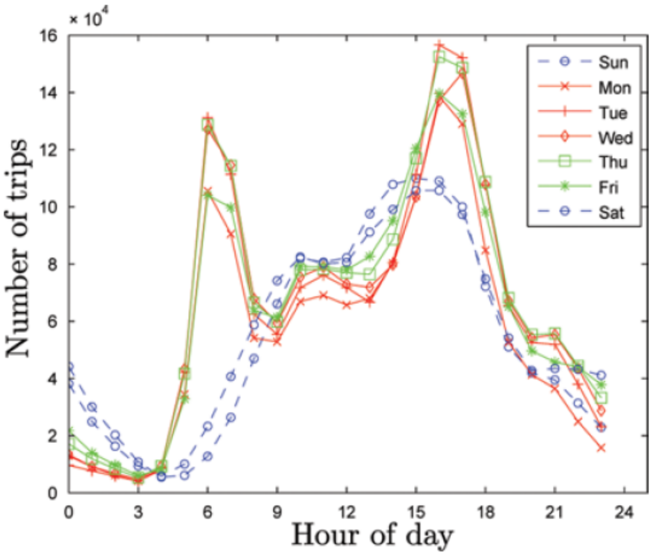


Figure 2.1 – Temporal ride characteristics for Paris.

Source: (Nair et al., 2012, p. 89)

2.2 Influence of Weather Events

Regarding weather, several authors have conducted studies to investigate how the usage of BSS is influenced by weather variables such as wind, humidity, temperature, and precipitation. However, not all the authors agree on the effects of the different variables on BSS usage.

Faghih-Imani et al. assessed the influence of the weather variables of temperature, rain, and humidity. Through their study conducted in Montreal, Canada they discovered a positive correlation between temperature and the number of bike sharing trips. While rain had no impact, humidity had a negative effect on the bike sharing usage (Faghih-Imani et al., 2014).

Caulfield et al. detected in their research in Cork, Ireland that good weather conditions influence the usage of bike sharing in two ways. Good weather results in both more rides as well as longer rides. To distinguish between good and bad weather they developed a threshold for all the variables that were used to categorize the data. The weather variables of temperature and hours of sunshine did not affect the usage of shared bikes if they were analysed separately (Caulfield et al., 2017).

Reiss and Bogenberger analysed the impact of weather events on the BSS in Munich. They distinguished between days with fair weather and days with bad weather based on a threshold that they developed. The results suggest that significantly less bikes are used during bad weather days than during fair weather days. The greatest outliers regarding bike sharing usage numbers, both positive and negative, are explained by good or bad weather. Additionally, they detected that people tend not to use bike sharing for a few hours after it has stopped raining (Reiss and Bogenberger, 2015).

Gebhart and Noland made an in-depth analysis regarding the influence of weather in their study in Washington, DC. They examined the influence that different weather variables have on both the usage numbers of shared bikes as well as the duration of the rides. The characteristic they chose for their investigations were; temperature, rainfall, snow, wind, fog, and humidity levels. They also looked how the impact of the variables differs between subscribers and casual users. Whereas one half of the subscribers tended to use the bike sharing even on rainy days (48% decrease), casual users were less likely to use bike sharing in rainy conditions (68% decrease). Moreover, the trip duration decreased more for the casual user (22.4% decrease) compared to the subscribers (10.1% decrease). In contrast to other authors' findings, high temperatures (32.2–37.2 C°) did not have an impact on the usage of shared bikes (Gebhart and Noland, 2014).

Figures 2.2 and 2.3 present the effect of rain on the riding behaviour of subscribers and casual users in Washington, DC. The difference of the temporal ride distribution between subscribers and casual users is evident, and the distribution has a similar pattern to the one described in Section 2.1.

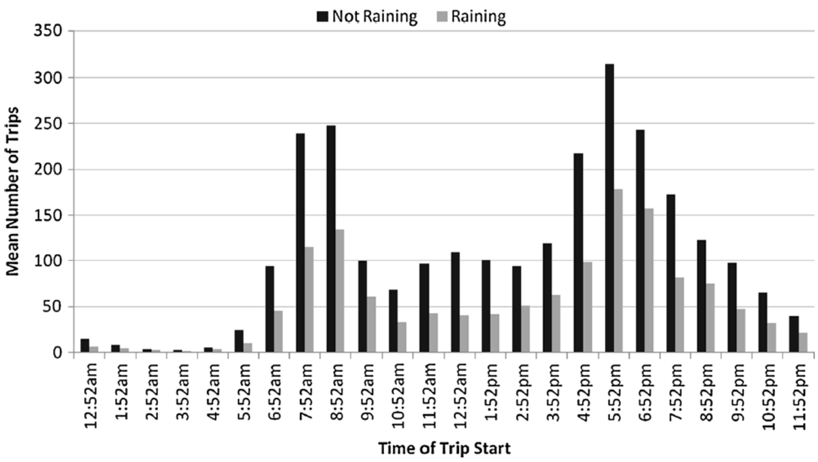


Figure 2.2 – Influence of rain on the riding behaviour of subscribers in Washington, DC.

Source: (Gebhart and Noland, 2014, p. 1211)

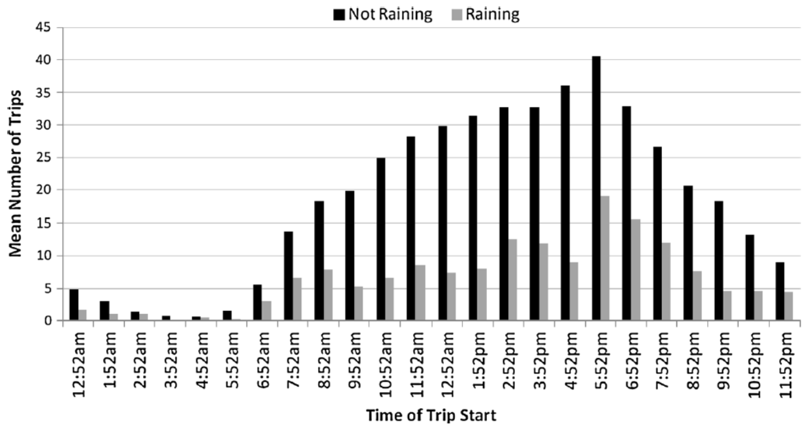


Figure 2.3 – Influence of rain on the riding behaviour of casual users in Washington, DC.

Source: (Gebhart and Noland, 2014, p. 1212)

Kim's study in Daejeon, South Korea indicated some contradictory results to the study of Gebhart and Noland. According to Kim, high temperatures (above 30C°) have a negative impact on bike sharing usage. In other respects, Kim agrees with Gebhart and Noland; for example, humidity, precipitation, and wind speed all have a negative effect on bike sharing usage. Temperature is positively correlated with bike sharing usage as long as the temperature is not more than 30C° (Kim, 2018).

El-Assi et al.'s study in Toronto, Canada, examined the influence of the factors of temperature, humidity, snow, and precipitation on bike sharing usage. They first calculated the perceived temperature by using the wind chill and a humidity index, and then they created five categories ranging from "very cold (below 0C°)" to "hot (30C° or more)." The results indicated that bike sharing usage is highest when the temperature ranges between 20–30C°. Although temperatures from 0–10C° and 10–20C° are correlated with a significant number of rides, the usage is less than that of the higher temperatures. Alternatively, precipitation, humidity, and snow are negatively correlated with the usage of bike sharing (El-Assi et al., 2017).

The study of Campbell et al. that was conducted in Beijing, China examined the factors of temperature, precipitation, and air quality. The BSS in Beijing uses both regular bikes as well as e-bikes, and this system allowed the authors to determine if the impact of weather variables differed from regular bikes to e-bikes. On the one hand, the results indicated that regular bikes were used less during days

with poor air quality and periods with high temperatures. On the other hand, low temperatures only exhibited a minor impact. Similar to other systems, precipitation had a negative impact on the usage of shared bikes. Regarding the e-bikes, the authors found that the usage was less sensitive to high temperatures and bad air quality than the regular bikes. The authors assume that this finding is due to the fact that the effort for riding an e-bike is less than a regular bike, and thus, these factors are less important (Campbell et al., 2016).
Corcoran et al. examined the impact of the weather variables of temperature, rainfall, and wind speed in Brisbane, Australia. The results did not agree with the majority of the authors aforementioned regarding temperature as it was not found to be a significant independent factor. The authors assume that this phenomenon is caused by the minor temperature variations in Brisbane compared to other cities. Rainfall had a negative effect on the usage of shared bikes, which is consistent with other studies. One additional finding of this study was that strong wind (daily average over 5km/h) was found to have a negative influence on the usage of shared bikes (Corcoran et al., 2014).

2.3 Influence of Infrastructure

The influence of infrastructure on the usage of bike sharing is categorised based on various variables. In the literature, not all the authors have considered the same infrastructural features. The most commonly studied factors have been road types and public transport. Another factor assessed in the literature review is altitude, which has exhibited promising results in several studies. This section is subdivided based on these three infrastructure features.

2.3.1 Road Types

Faghih-Imani et al. investigated the impact that different street types, such as bike paths, minor roads, or major roads, have on the usage of bike sharing in Montreal, Canada. They first created a buffer around each bike station and then calculated the total street length of each road type inside each buffer. Then, the road characteristics were compared with the number of rides starting and ending at each station to determine the trends. The results indicated that cycling facilities like bike paths and bike lanes as well as minor roads had a positive impact on the usage of bike sharing.

Alternatively, major roads had a negative impact (Faghih-Imani et al., 2014).

Faghih-Imani and Eluru also studied this topic in Chicago, USA. The results did not differ much from their previous study as bike paths near the bike sharing stations had a positive effect on the usage of bike sharing. However, major roads hindered the use of bike sharing systems as the researchers recorded less traffic at bike stations with many major roads in its surrounding area (Faghih-Imani and Eluru, 2015).

Mateo-Babiano et al. used the same method in their study to calculate the length of different types of bike infrastructure inside a certain area around each bike station. However, they did not include regular streets like minor or major roads, and they focused only on bike-specific road types. The results indicated that off-road bike paths had the most positive effect on the usage of bike sharing. This type of cycle lane is completely segregated from the other traffic. Furthermore, bike-friendly infrastructure, such as shared pathways, tends to increase the number of rides, albeit not as much. The authors assume that the effect was not greater, because the cycling infrastructure in Brisbane, Australia, where the study was conducted, lacks satisfying solutions at junctions regarding connectedness and directness (Mateo-Babiano et al., 2016).

Sun et al. calculated the total length of roads and bike lanes in the surrounding area of each bike station to then compute the road density. They distinguished between the arrivals and departures of bikes at stations to determine if a difference existed. The results indicated that road density had no impact on neither the arrivals nor the departures. However, bike lanes had a positive impact on both parameters (Sun et al., 2017).

Zhang et al. assessed different road types in the catchment area of bike stations. They studied bike lanes, main roads, secondary roads, and branch roads. Of these types of roads, bike lanes had the most positive impact as rides increased when more bike lanes were present. Branch roads also had a positive impact whereas the two larger road types, main and secondary, did not influence usage (Zhang et al., 2017).

Noland et al. investigated the specific impact of bike lanes, and they distinguished between subscribers and casual users. The results indicated that bike lanes have a greater impact on casual users, who tend to use bike sharing more when more bike lanes are present. Alternatively, subscribers are not affected by the amount of bike lanes (Noland et al., 2016).

El-Assi et al. researched what type of bike infrastructure affects bike sharing usage and found that shared bikes are used more when more bike lanes are present. Furthermore, shared bikes are also used more when less junctions with major roads exist. The authors claim that this effect is because intersections are a place of high risk for cyclists if no marked bike paths are present (El-Assi et al., 2017).

2.3.2 Public Transport

To assess the impact of public transport on the usage of shared bikes, the approach of the different authors was similar to each other. The most common approach was to count the number of public transport stations or calculate the length of public transport lines in the catchment area of each bike sharing station. This approach allowed the researchers to find correlations between the public transport and the rides for each bike sharing station. However, the results differ, because first, not every city has the same types of public transport available, and second, not all the authors assessed the same types of public transport.
In New York City, USA, Noland et al. compared the number of rides for each bike station with the subway ridership in the catchment area of the bike station. The analysis resulted in a positive and significant impact of subway ridership on the usage of bike sharing. Thus, a BSS station near a busy subway station has an increased number of rides. Other public transport types such as buses were not considered in this particular study (Noland et al., 2016).
O'Neill and Caulfield calculated a public transport score for Dublin, Ireland that depended on the number of railway and bus stops, and they compared the score with bike sharing station activity. In the comparison they included the 20 busiest and the 20 least busy bike sharing stations. The results indicated that the busiest stations had a higher public transport score than the least busy stations. Therefore, the authors argued that public transport favours the usage of bike sharing (O'Neill and Caulfield, 2012).
Tran et al. separately assessed the public transport types of the railway, metro, and tram in Lyon, France. The results indicated that only railway stations had an impact on the usage of bike sharing, and the metro and tram were not significant factors. The authors assumed that bike sharing was often used by people coming from outside the city who then take a shared bike from a dock near

the railway station. They also supposed that since the city is relatively small, the users from within the city do not need to combine the different transport modes as they can travel completely by bike (Tran et al., 2015).

Faghih-Imani and Eluru discovered similar results in their study in Chicago, USA as the regional metro stations were highly used destinations for subscribers, which highlights public transport's impact on bike sharing. However, casual users did not use public transport stations as bike sharing trip destinations. The authors assumed that casual users tend to use bike sharing instead of classic public transport options such as trains and buses (Faghih-Imani and Eluru, 2015).

Sun et al. discovered in their study, which was also conducted in Chicago, that the impact on bike sharing usage differs between bus- and metro-access. Whereas bus-access is positively correlated with bike sharing usage, metro-access is negatively associated (Sun et al., 2017).

The study by Zhang et al. in Zhongshan, China found that public transport had no effect. There, using the BSS is cheaper than using the public transport as the bike sharing is free for the first hour and public transport tickets are relatively expensive. This dynamic leads to the residents preferring not to connect the two types of transport as they would rather complete their entire trip by bike (Zhang et al., 2017).

2.3.3 Altitude

This sub-section discusses how the altitude of the different bike sharing stations has an impact on their usage. The difference between arrivals and departures is noteworthy to observe as one might expect that bike sharing is primarily used for descending rides than for ascending rides. Mateo-Babiano et al. calculated the hire/return ratio for each bike station to determine correlations between the ratio and the altitude of the bike stations. The results indicated that a strong correlation exists as stations with higher altitudes had more hires than returns. However, when considering other infrastructure variables the impact decreased. In this case, many of the stations with a higher altitude had less suitable cycling infrastructure, which led to less rides. Despite this finding, a slight tendency still existed for bike sharing users to travel in a downhill direction rather than in an uphill direction (Mateo-Babiano et al., 2016).

Faghih-Imani et al. discovered a correlation between altitude and the number of rides in Barcelona and Seville, Spain. Although they detected a decreased demand for bike sharing at stations with a higher altitude, no difference was found between arrivals and departures. In this case, the altitude affected both arrivals and departures negatively (Faghih-Imani et al., 2017).

In the study of Tran et al., altitude was negatively significant in all the models, both for arrivals and departures, which indicates that altitude is a hinderance for bike sharing trips (Tran et al., 2015).

2.4 Influence of Land Use

This section reviews how land use may affect the usage of bike sharing. The focus is on how the temporal ride characteristics for land use categories such as residential or commercial areas change over the course of the day. The aim is to determine if the usage behaviour differs depending on the land use categories.

Sun et al. calculated a land use mix score for each bike sharing station to illustrate the level of homogeneity of the land use in the area of each bike station. If only one land use type was in the catchment area of a bike station, the score would be 0, and if each land use type was represented equally, the score would be 1. The scores were then compared with the rides per station to determine if the land use mix influenced the usage of bike sharing. The results indicated that no effect occurred, meaning that the homogeneity of the land use around a bike station did not matter (Sun et al., 2017).

Faghih-Imani and Eluru investigated the impact that a park near a bike sharing station has on the way the station is used by subscribers and casual users. The results suggested that casual users used bike stations near parks more often than subscribers, especially on weekdays. On weekends, subscribers also registered more rides near parks. The authors state that these findings evidence that casual users use bike sharing for recreational activities, whereas subscribers use it mainly for commuting during weekdays, and on the weekends, they use it for recreational activities as well. Another factor that underlines the commuting assumption is that subscribers tend to use bike sharing in the morning to get into the city centre, and in the evening to get out of the city centre (Faghih-Imani and Eluru, 2015).

Noland et al. researched if usage patterns exist between the different land use categories and the subscribers and casual users. They calculated the percentage of each land use category in the catchment area of each bike sharing station, and these numbers could then be compared with the rides. The results indicated that residential land use was associated with subscribers but not with casual users. Commercial land use was correlated to weekday usage but not to weekend or holiday usage. Furthermore, both subscribers and casual users were positively affected by commercial land use. Alternatively, recreational land use did not impact the usage of bike sharing (Noland et al., 2016).

Mateo-Babiano et al. assigned each bike sharing station a land use type according to the land use present in the catchment area of the station. Then, they considered how cycle flows change between different origin-destination land use pairs. The results indicated that most of the rides from a residential land use area to a commercial land use area occur in the morning. As expected, in the evening, the flows go in the opposite direction, from commercial to residential. The stations in residential land use areas were used more often as origins on weekdays than on weekends. The authors assume that this underlines the fact that bike sharing is often used for commuting purposes. Moreover, trips originating from parks occurred more often on weekends, which is a sign that bike sharing is used for recreational purposes on the weekend. Trips to and from commercial land use types occurred both on weekdays as well as on weekends (Mateo-Babiano et al., 2016). Figure 2.4 illustrates the results from Mateo-Babiano et al.'s research regarding weekdays. The most significant origin-destination land use pairs are highlighted.

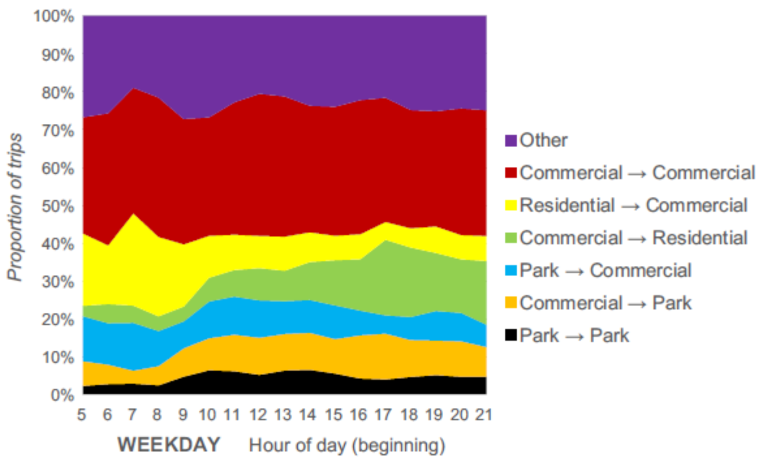


Figure 2.4 – Example of a land use analysis.

Source: (Mateo-Babiano et al., 2016, p. 304)

2.5 State of The Art Recap

Regarding the temporal ride distribution, the most common pattern was the two peaks on weekdays in the morning and the afternoon. For the weekends, most of the systems had a single peak. The difference between casual users and subscribers was obvious as subscribers tended to use bike sharing for commuting purposes, which leads to the morning and afternoon peaks, whereas causal users exhibited a less consistent usage pattern.

Regarding weather, rain was the factor that influenced BSS the most. Humidity and wind were significant influencing factors in some studies but not in all. The temperature was the characteristic with the most discrepancies. Some studies demonstrated a correlation between very high temperatures and decreased bike sharing usage while others did not observe such an impact.

In the studies about infrastructure, bike lanes favoured bike sharing rides whereas major roads hindered usage. Regarding public transport, not all the authors agreed with each other. However, a major part of the literature states that public transport stations near bike sharing stations increased the usage frequency of the bike station. However, not all the authors agree on the type of public transport that favours bike sharing rides. Some studies indicated a correlation with regional trains whereas other found a correlation with buses or tubes. The authors also do not agree with each other about altitude; some observed a correlation between the usage of bike sharing stations and their respective altitude while others do not.

The studies about land use indicated that depending on the type of land use, the bike sharing stations were used in a different way or at a different time. Moreover, the researchers have indicated that a difference exists between subscribers and casual users regarding which type of land use they favour for bike sharing rides.

As presented in this chapter, the studies of many researchers have investigated how different factors influence the usage of bike sharing. However, the results differ from city to city and amongst different research. The varying results may be due to factors such as different statistical approaches, data collection, the geographic location of a city and its climatic conditions, or the unique behaviour of people in different cities. Since most of the researchers only studied one city, these differences were not tackled. Thus, it is not examined if different BSS in different cities react the same way on given impact variables. This thesis should fill this gap by comparing cities by using the same methods for all cities to determine if differences exist between the cities.

Chapter 3

Methods

3.1 Data Collection

3.1.1 Bike Sharing System Data

After the topic was defined, the first step of the thesis was to search for available and usable data. As the bike sharing data was the most important to find, open access bike sharing data was searched for through the website bikesharemap.com[1] by Oliver O'Brien, which lists almost every BSS worldwide. Additionally, the website notes the size of the system as well as the system's webpage. By using this resource, the systems that have freely available data published on their webpage were identified. The initial intention of this study was to compare large European cities, such as London, Paris, Barcelona, Madrid, and Milan. However, most cities with the largest bike sharing systems in Europe do not provide open access to their data. The only exception was London. As a consequence, the study was adapted to include smaller cities with smaller bike sharing systems. Some cities only provided older data from 2016 and earlier, and since more recent data was preferred, some cities were eliminated as options. Ultimately, four cities in addition to London were chosen: Edinburgh, Oslo, Bergen, and Trondheim. Given the scarcity of data in Europe, the research was expanded to North America as most of the continent's bike sharing systems provide their data freely on their respective webpages. Ultimately, four additional US-based BSS were chosen to be included in the study from New York,

[1] https://bikesharemap.com

Boston, Chicago, and Washington, DC. Although adding more systems from North America would have been easier as many provide their data for free, a balanced number between European and US cities was determined to be preferred. The final dataset consisted of nine cities. Table 3.1 1 presents the providers of the bike sharing systems data for each city.

Table 3.1 – BSS Data Source

City	**Source**
New York	Citi Bike[2]
Boston	Bluebikes[3]
Chicago	Divvy[4]
Washington, DC	Capital Bikeshare[5]
London	Transport for London[6]
Edinburgh	Just Eat Cycles[7]
Oslo	Oslo City Bike[8]
Bergen	Bergen City Bike[9]
Trondheim	Trondheim City Bike[10]

After finding the data, the next step was to determine which period of time the data should span. The aim was to assess a single year in the same time span to compare the riding habits in the different seasons. This time span still allowed the data to be assessed in a smaller temporal resolution since all the data is trip-based, which is further discussed in Section 3.2. Since the most current data for the majority of the cities was from May 2019, the original goal was to download data from June 2018–May 2019. However, this was not possible in all cases, because the data of certain BSS for the full time span was not available to the public. However, since the BSS with less historical data had the most recent data, spanning an entire year was possible. Additionally, the BSS from Oslo and Trondheim are not in service during December, January, February, and March, and therefore,

[2] https://www.citibikenyc.com/
[3] https://www.bluebikes.com/
[4] https://www.divvybikes.com/
[5] https://www.capitalbikeshare.com/
[6] https://tfl.gov.uk/
[7] https://edinburghcyclehire.com/
[8] https://oslobysykkel.no/en
[9] https://bergenbysykkel.no/en
[10] https://trondheimbysykkel.no/en

data was not available for these months. Furthermore, Oslo's data system underwent a formatting change in 2019, and comparable data that spanned at least eight months (April–November 2019) was not available until January 2020. Finally, for six of the nine BSS the data was from June 2018–May 2019. Table 3.2 lists the time spans of the downloaded datasets. The New York dataset was later complemented with the months of March to May 2020, which were affected by the COVID 19 pandemic.

Table 3.2 – BSS Data Timespans

City	**Timespan**
New York	June 2018–May 2019/March–May 2020
Boston	June 2018–May 2019
Chicago	June 2018–May 2019
Washington, DC	June 2018–May 2019
London	June 2018–May 2019
Edinburgh	October 2018–September 2019
*Oslo**	April 2019–November 2019
Bergen	July 2018–June 2019
*Trondheim**	June 2018–May 2019

*The BSS in Oslo and Trondheim are in service only from April–November.

3.1.2 Weather Data

A first attempt to collect the weather data targeted the official weather websites of the respective states. However, historical weather data was not available for free. After an in-depth search, the open data website meteostat.net[11] was found, which collects weather data and provides it for free. The data were collected from various sources, such as the National Oceanic and Atmospheric Administration (NOAA), Environment Canada, and the National Weather Service of Germany (DWD). According to Meteostat, most of their data meets the requirements of the World Meteorological Organization (WMO), which has formulated standards to assure the quality of data regarding the location of the weather stations as well as the way observations are made (Meteostat.net, 2020). However,

[11] https://meteostat.net/de

finding comparable data was difficult as not all the weather stations collect the same data; for example, whereas some provide daily averages, others provide hourly averages. Additionally, not all the stations measure every weather variable. Some do not record precipitation, while others do not record wind. The preferred variables for the study were temperature (hourly and aggregates), wind, and precipitation, but it was difficult to find weather stations that covered all of these factors. As a consequence, for some cities data was combined from different weather stations to have a complete dataset. Furthermore, some of the weather stations are also relatively far from the respective city (up to 50 km), but this option was the only possibility to collect data for all the cities. Table 3.3 details which weather stations were used for each city. The names of the stations in the table are those listed on meteostat.net. For the daily weather data, data from the same time period as the BSS was downloaded (see Table 3.2), whereas for the hourly weather data only the data of one month was downloaded. May 2019 was chosen, because all BSS observed an increased number of trips, which led to a larger sample. Additionally, the weather was expected to be more heterogeneous in May than in the summer months of July or August.

Table 3.3 – Weather Stations

City	**Daily Data**	**Hourly Data**
New York	John F. Kennedy Airport	Newark Airport
Boston	Boston Logan International	Boston Logan International
Chicago	Chicago Midway Airport	Chicago Midway Airport
Washington, DC	Washington National Airport	Washington National Airport
London	London Heathrow Airport	London Weather Centre
Edinburgh	Edinburgh Airport*	Edinburgh Airport*
Oslo	Oslo-Blindern	Oslo-Blindern
Bergen	Bergen / Florida	Bergen / Florida
Trondheim	Trondheim / Vaernes	Trondheim / Høiset

*The precipitation data for Edinburgh was from the weather station Leuchars

3.1.3 Roads and Public Transport Data

The data for roads and public transport are from OpenStreetMap (OSM), which is a worldwide, open dataset that is primarily user generated and provides geographic information about various topics (OpenStreetMap, 2020a).

Since OSM only allows individuals to download a certain amount of data at a time, which was too small for the purposes of this study, a third-party website called geofabrik.de[12] was used to download the whole dataset for the cities. Geofabrik.de provides up-to-date OSM data, which can be downloaded for countries or states, depending on the scale, and then sub-sections can be isolated to download the required data. Table 3.4 presents the smallest sub-section for each city that was downloaded. For Washington, DC, three different files were downloaded, because the BSS of Washington, DC extends over the city's borders and has stations in Virginia as well as in Maryland.

Table 3.4 – OSM Datasets from Geofabrik

City	**Geofabrik - Dataset**
New York	New York
Boston	Massachusetts
Chicago	Illinois
Washington, DC	District of Columbia, Virginia, Maryland
London	Greater London
Edinburgh	Scotland
Oslo	Norway
Bergen	Norway
Trondheim	Norway

The table shows the downloaded sub sections. Only the smallest sub section is shown.

[12] https://www.geofabrik.de/de/index.html

3.1.4 Land Use Data

To collect land use data, OSM was first considered as a possible source, because it also provides land use data. However, the data quality varies greatly from city to city. Therefore, official land use maps were preferred to be used. In the case of Europe, the best solution was to use CORINE Land Cover, which is from the Coperinicus Land Monitoring Service. Copernicus is a European program that monitors the earth. Various geographical information is stored and provided for free, including land cover and land use (Land.copernicus.eu, 2020). Therefore, Corine Land Cover provided the land use data for all European cities. The newest dataset from 2018 was used.

Regarding the US cities, official land use data was collected from the individual states, except for Chicago, because an appropriate land use map was not found as the link for the official land use map was not functioning. Furthermore, since the BSS of Washington, DC crosses the borders of the city into the adjacent states, it was not possible to find compatible land use data for Maryland and Virginia. To cover all of the BSS's locations, many different land use maps with different land use categories and different spatial resolutions had to be combined into one map. As a consequence, only the land use data from the District of Columbia was used. The BSS was therefore limited to that area for the land use research in order to be comparable. Table 3.5 lists the sources of the different land use datasets.

Table 3.5 – Land Use Data Source

City	Land Use Data Source	Year
New York	NYC Planning[13]	2019
Boston	Massachusetts Document Repository[14]	2016
Chicago	*No data*	-
Washington, DC	Open Data DC[15]	2019
London	CORINE Land Cover[16]	2018
Edinburgh	CORINE Land Cover	2018
Oslo	CORINE Land Cover	2018
Bergen	CORINE Land Cover	2018
Trondheim	CORINE Land Cover	2018

[13] https://www.nyc.gov/

3.2 Data Preparation

This section explains how the data was prepared before beginning the analysis. The different data, namely BSS data, weather, infrastructure, and land use are explored in detail. The preparation of the different data was mainly achieved with R Studio and ArcMap, and Microsoft Excel was used for some minor issues.

3.2.1 Bike Sharing System Data Preparation

All the data of the bike sharing systems was trip-based and provided as CSV files, and for every trip one entry was created. Seven of the nine systems provided files that contained all trips from one calendar month. The exceptions were Chicago, which provided quarterly files, and London, which provided files for five to seven days. Thus, the first step was to edit the files of London and Chicago so that they covered a calendar month, like the other cities. This process facilitated use of the same programming code (in R Studio) for all the different datasets. Depending on the BSS, different data was stored for each trip. However, the main trip information of the start and end time as well as the start and end stations was available in all the datasets. Table 3.6 presents which type of information was stored in each dataset. For Oslo, not all the files contained the same information. The fields marked with an *(X)* in the table indicates that the information was not available in every file, whereas the fields with an *X* indicate the information was in every file. Since the station IDs did not change, completing the missing information regarding stations (location and name) was easy. The calculation of the trip duration was also possible as the start and end times for each trip was available. Finally, all the files from Oslo were updated to provide the same information as the other cities.

Regarding time-zones, some adaptations were made. The data from New York, Boston, Chicago, Washington, DC, and London were all stored in the local time, so no further actions were required. Since the data for Edinburgh, Oslo, Bergen, and Trondheim was stored in UTC (Universal Time Coordinated), the time had to be transformed to the corresponding local time-zone to have correct

[14] https://docs.digital.mass.gov/
[15] https://opendata.dc.gov/
[16] https://land.copernicus.eu/pan-european/corine-land-cover

and comparable data. In the end, all data were stored in their local time-zone.

Table 3.6 – BSS Trip Data

City	Start Time	Stop Time	Trip Duration	Start Station ID	Stop Station ID	Station Names	Station Locations	Station Descriptions	Bike ID	User type	Birthyear	Gender
New York	X	X	X	X	X	X	X		X	X	X	X
Boston	X	X	X	X	X	X	X		X	X	X	X
Chicago	X	X	X	X	X	X			X	X	X	X
Washington, DC	X	X	X	X	X	X			X	X		
London	X	X	X	X	X	X			X			
Edinburgh	X	X	X	X	X	X	X	X				
Oslo	X	X	(X)	X	X	(X)	(X)	(X)				
Bergen	X	X	X	X	X	X	X	X				
Trondheim	X	X	X	X	X	X	X	X				

The categories *Station Names, Station Locations*, and *Station Description* refer to both the start as well as the stop stations

For Chicago, Washington, and London, the stations' locations were not included, and external sources were used to complete the dataset. The sources for the stations' locations are listed in Table 3.7. While the official list of stations for Chicago and Washington could be downloaded, the official list from London was not functioning. Therefore, the list was pulled from a third-party website and adapted into a CSV file. The connection to the BSS trip dataset was simple as all BSS have unique station IDs, which were present in both files.

Table 3.7 – BSS Station Location Sources

City	Station Location Source
Chicago	Chicago Data Portal[17]
Washington, DC	Open Data DC[18]
London	bikeshare-research.org[19]

In May 2019, 600 of the 604 stations in Chicago from the trip data file could be matched (99.34%). For Washington, DC, 554 of the 556 stations matched (99.64%), and for London, which is the BSS with the most stations of the three, 780 out of 791 stations matched (98.61%). The reason for the differences might be that some stations were closed since May 2019. However, since the station's location was only used for the infrastructure analysis, the unmatching stations were retained for the other analyses such as weather and riding behaviour. After this step, all the trip datasets contained the following information:

[17] https://data.cityofchicago.org/
[18] https://opendata.dc.gov/

- Start Time
- Stop Time
- Trip Duration
- Start Station ID
- Start Station Name
- Start Station Latitude
- Start Station Longitude
- Stop Station ID
- Stop Station Name
- Stop Station Latitude
- Stop Station Longitude

The next step was to filter the data. Following Caulfield et al. (2017); Gebhart and Noland (2014); Zhang et al. (2017); Zhou (2015), all rides with less than a minute duration were removed from the dataset as these might be rides in which users pick up a bike and immediately return it. Furthermore, following Gebhart and Noland (2014) all rides with a duration of more than 24 hours were also removed. Since some data providers had used the same parameters to filter the data before providing it to the public, nothing was changed by the researcher for these cases. Finally, the trips with the same start and end location were omitted as well, following Gebhart and Noland (2014); Zhang et al. (2017); Zhou (2015).
After the data filtering, the preparation of the data began. As this study's analysis was based on counts, for each city, monthly, daily, and hourly counts were created for each month. As a result, the data included the number of rides per hour, per day, and per month for a whole year. For all these subsets, the average trip duration was calculated as well. To distinguish between weekdays and the weekend, each entry was coded with 0 for weekdays or 1 for the weekend, and the day of the week was added to the entry. Table 3.8 presents a sample of the hourly data from New York.

For each month, the average number of rides per hour and the average trip duration per hour were calculated both for weekdays and weekends, which was the foundation of further research.

As one part of the research was to distinguish between subscribers and casual users, the same data preparation was executed. However, the data was filtered beforehand based on the user type so that two separate datasets existed for subscribers and casual users. Then, both datasets were prepared as described. Since not all the BSS provided information about the user type, this step was only completed for New York, Boston, Chicago, and Washington, DC.

[19] https://bikeshare-research.org/

Table 3.8 – Hourly Data Example (New York)

By Hour	n	day	weekend	mean_trip_dur(s)
01.06.2018 20:00	3,026	Friday	0	938
01.06.2018 21:00	2,027	Friday	0	970
01.06.2018 22:00	1,683	Friday	0	993
01.06.2018 23:00	1,324	Friday	0	979
02.06.2018 00:00	561	Saturday	1	885
02.06.2018 01:00	396	Saturday	1	802
02.06.2018 02:00	315	Saturday	1	932
02.06.2018 03:00	213	Saturday	1	904
02.06.2018 04:00	133	Saturday	1	1,047
02.06.2018 05:00	147	Saturday	1	833

By Hour = Date and Time; n = number of rides; day = day of the week; weekend = 0 if weekday, 1 if weekend; mean trip dur(s) = average trip duration in seconds

3.2.2 Weather Data Preparation

As previously mentioned, two kinds of weather data were downloaded. The first was the daily averages, and the second was the hourly averages. The daily averages covered a whole year of data, having one value for each day for every variable. The hourly averages only covered one month of data (May 2019) and had one value for every hour. Since the daily averages could be downloaded as one file, no further action was needed. The hourly files first had to be merged as it was not possible to download all of them at once. The next step was to delete the categories that were not of interest. Finally, the daily file contained the following information:

- Avg. Temperature
- Max. Temperature
- Min. Temperature
- Precipitation
- Windspeed

The hourly file only had one value for temperature and did not have minimum and maximum values. To draw a connection to the bike rides, the weather data was merged with the BSS data by date and time, for the daily and hourly files as both covered the same time span. Ultimately, the data was a combination of the information in Table 3.8 and the weather data listed above, and every day or hour had corresponding BSS usage numbers as well as the weather information.

3.2.3 Bike Stations

To assess the different influences of the infrastructural features on the usage of BSS, the first step was to extract all the bike stations from the trip data files. Additionally, the number of trips starting and ending at each station for one month was calculated. For each BSS a bike station file was created that contained the station ID, station location, and rides arriving and departing from the station. All these steps were executed by using R Studio. The bike station files were stored and were later extended with information about infrastructure and land use. The month used for the creation of the station file was May 2019 as it was available for all datasets.

Altitude

To have altitude values for the different bike sharing stations, the website GPS Visualizer[20] was used. It transforms a text file with latitude and longitude values into a spreadsheet file with the respective altitude for each location. The altitudes were derived from different digital elevation models (DEM). The best available source was used to calculate the height. The altitudes for the US cities were from the U.S Geological Survey's National Elevation Dataset (NED) and the altitudes for the European cities were from a DEM of Austria's OpenDataPortal (ODP), which hosts various high-quality DEM from Western Europe. The NED has a horizontal resolution of 1 arc-second. The resolution of the ODP DEM is also 1 arc-second (GPSVisualizer, 2020).
The created station file was used as input as it contained the required information (latitude and longitude) for each station. As a result, the altitude of each bike station was added to the file. At this point in the process, the file stored the following information: station ID, station name, station location, arrivals, departures, and altitude.

[20] https://www.gpsvisualizer.com/

Bike Station Area

The bike station file with all the BSS stations and the associated content was loaded into ArcMap, and a point was created for each BSS station according to the stations' latitude and longitude. All the points were then collected in a point layer. As all the BSS stations' coordinates were in the coordinate system WGS84, they were transformed to a more suitable coordinate system for the corresponding city. Table 3.9 identifies which coordinate systems were used for each city. The table presents one geographic coordinate system and a projected coordinate system. The projected coordinate system was used to display the data in a better manner and was developed from the geographic coordinate system that the spatial analysis was based on.

The next step was to calculate "bike station areas" to represent the catchment area of each station. To do so, a buffer of 300 m was created for each station, following Faghih-Imani and Eluru (2015); Médard de Chardon et al. (2017); Sun et al. (2017); Tran et al. (2015); Zhang et al. (2017). These areas were the basis for the following infrastructure investigations. The calculations are explained later. For a general understanding, Figure 3.1 presents all the bike station areas from New York with an underlying OpenStreetMap base map. Each bike station area had the same information stored as the bike stations, so no data was lost during this process, and each bike station was associated with a bike station area.

Table 3.9 – Coordinate Systems

City	Geographic Coordinate System	Projected Coordinate System
New York	GCS_North_American_1983	NAD_1983_StatePlane_New_York_Long_Island_FIPS_3104
Boston	GCS_North_American_1983	NAD_1983_StatePlane_Massachusetts_Island_FIPS_2002
Chicago	GCS_North_American_1983	NAD_1983_StatePlane_Illinois_East_FIPS_1201
Washington	GCS_NAD_1983_2011	NAD_1983_2011_StatePlane_Maryland_FIPS_1900
London	GCS_OSGB_1936	British_National_Grid
Edinburgh	GCS_OSGB_1936	British_National_Grid
Oslo	GCS_ETRS_1989	ETRS_1989_NTM_Zone_30
Bergen	GCS_ETRS_1989	ETRS_1989_NTM_Zone_30
Trondheim	GCS_ETRS_1989	ETRS_1989_NTM_Zone_30

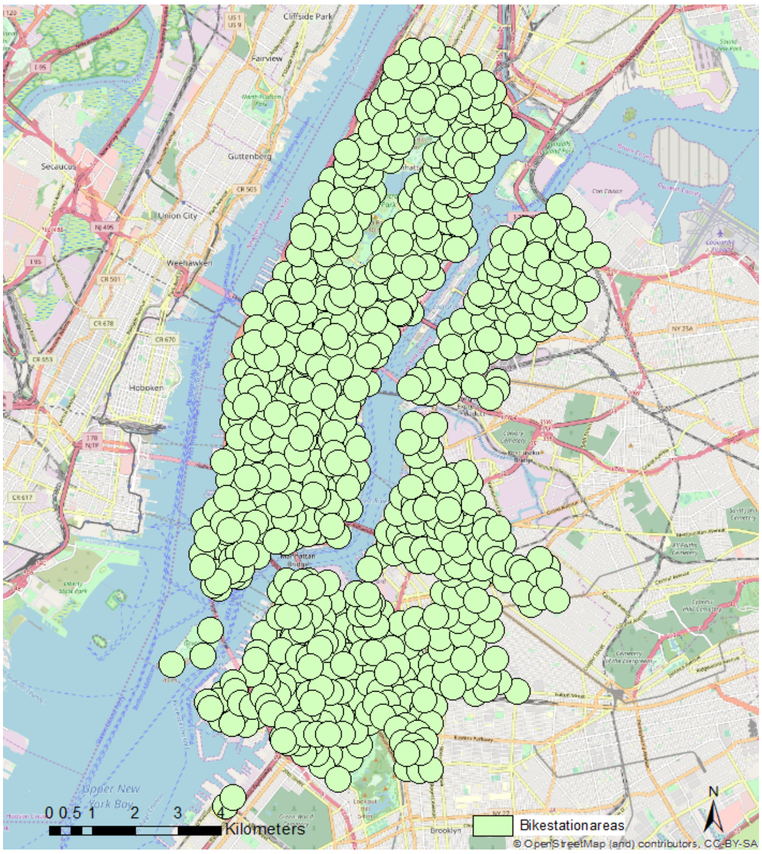


Figure 3.1 – Bike station areas of New York.

3.2.4 OpenStreetMap(OSM) Data

The downloaded OSM data contained many features that were not relevant for the research, which were deleted to keep the data volume low. The features that were retained were the public transport stations as well as the road types. Next, the data was transformed into the matching coordinate system (see Table 3.9) for each city as the OSM data was stored in WGS84.

Public Transport Stations

The public transport station point layer was loaded into the existing ArcMap project with the bike station areas. Since each station was attributed to a public transport type, such as bus, train, ferry, tram, taxi, or airport, the stations were split into these categories. The taxi stations as well as the airports were deleted, as these types of public transport stations were not relevant for the research.

Furthermore, some bus and train stations were labelled as larger stations, for instance bus hubs or railway stations compared to tube stations. However, as these were few, and for certain cities only one, they were aggregated to the other bus and train stations.
The next step was to intersect the bike station areas with the public transport stations and count the number of public transport stations in each bike station area. Consequently, a count of the different public transport stations was calculated for each bike station within the 300m radius, similar to the study by Faghih-Imani and Eluru (2015).

Roads

For the roads, a similar procedure was executed. First, the converted road layer was loaded into ArcMap. One difference to the public transport analysis was that none of the features were aggregated or deleted at first, which led to 26 road types. Each road type was used, because they can be aggregated or deleted later during the statistical analysis. The roads were also intersected with the bike station areas so that every bike station area had its assigned road segments. Although the segments were not counted like the public transport stations, the total street length within the station area was calculated for each road type similar to the studies of Faghih-Imani et al. (2014); Mateo-Babiano et al. (2016), as every street segment had its length as an attribute.

3.2.5 Land Use Data

As the land use data was from different sources, it had to be pre-processed to be compared with each other. Since every dataset had its own land use categories with its own code, a new code was developed which was used for all the datasets. The basis for the new code was the land use dataset from CORINE Landcover as it was the source of all the European datasets. The new land use code contained five major categories, which are listed in Table 3.10. The complete datasets and the process of adapting the new code to the existing datasets is presented in Table 3.11.

First, the land use layer was transformed into the respective coordinate system. Then, the new labels were assigned for each land use category. Similar to the OSM data, the land use layer was intersected with the bike station areas so that for each bike station the dominant land use category

Table 3.10 – New Land Use Code

New Code	Description
1	Continuous Urban Fabric
2	Discontinuous Urban Fabric
3	Commercial or Industrial
4	Transport Facilities
5	Residential/Forest
999	Mixed

could be calculated. This process resulted in a file that contained the dominant land use category for each bike station. This file was later merged with the trip data file so that each trip was assigned to a start and stop land use category, which was necessary to filter the trip data according to the land use. This filtering step had to be performed for both the start- and the stop-stations, which resulted in two files for each land use category that contained all the trips starting or ending at a station from the particular category. Next, the same steps as the regular trip data (see Section 3.2.1) were executed to prepare the land use data.

Table 3.11 – Original Land Use Categories

Europe			New York			Boston			Washington, DC		
Land Use Code	*Description*	*New Code*	*Land Use Code*	*Description*	*New Code*	*Land Use Code*	*Description*	*New Code*	*Land Use Code*	*Description*	*New Code*
111	Continuous urban fabric	1	1	One & Two Family Buildings	2	0	Unknown	999	LDR	Low Densit Residential	2
112	Discontinuous urban fabric	2	2	Multi-Family Walk-Up Buildings	1	2	Open land	5	LMDR	Low_Medium Density Residential	2
121	Industrial or commercial units	3	3	Multi-Family Elevator Buildings	1	3	Commercial	3	MDR	Medium Density Residential	1
122	Road and rail networks and associated land	4	4	Mixed Residential & Commercial Buildings	999	4	Industrial	3	HDR	High Density Residential	1
123	Port areas	4	5	Commercial & Office Buildings	3	6	Forest	5	GC;NC;C;O	Commercial	3
124	Airports	4	6	Industrial & Manufacturing	3	7	Agriculture	999	TCU	Transport, Communication, Utilities	4
131	Mineral extraction sites	999	7	Transportation & Utility	4	8	Recreation	5	HI; I; LI	Industrial	3
132	Dump sites	999	8	Public Facilities & Institutions	3	9	Tax exempt	999	MU	Mixed Use	999
133	Construction sites	999	9	Open Space & Outdoor Recreation	5	10	Mixed use, primarily residential	2	S	Institutional	3
141	Green urban areas	5	10	Parking Facilities	4	11	Residential_single_family	2	FP	Federal Public	3
142	Sport and leisure facilities	5	11	Vacant Land	999	12	Residential_multi_family	1	LP	Local Police	3
211	Non-irrigated arable land	999				13	Residential_other	2	P!P:I	Public, Quasi_Public, Insitutional	3
212	Permanently irrigated land	999				20	Mixed use_other	999	R	Parks and Open Spaces	5
213	Rice fields	999				30	Mixed use, primarily commercial	3	PARKING	Parking	4
221	Vineyards	999				55	Right_of_way	4	ROADS; ALLEYS; MEDIAN	Roads; Alleys; Median	4
222	Fruit trees and berry plantations	999				88	Water	999	TROW	Transporation Right of Way	4
223	Olive groves	999							Null	Undetermined	999
231	Pastures	999							CANAL; RIVER; LAKE	Water	999
241	Annual crops associated with permanent crops	999									
242	Complex cultivation patterns	999									
243	Land principally occupied by agriculture, with significant areas of natural vegetation	999									
244	Agro-forestry areas	5									
311	Broad-leaved forest	5									
312	Coniferous forest	5									
313	Mixed forest	5									
321	Natural grasslands	5									
322	Moors and heathland	5									
323	Sclerophyllous vegetation	5									
324	Transitional woodland-shrub	5									
331	Beaches, dunes, sands	5									
332	Bare rocks	999									
333	Sparsely vegetated areas	999									
334	Burnt areas	999									
335	Glaciers and perpetual snow	999									
411	Inland marshes	999									
412	Peat bogs	999									
421	Salt marshes	999									
422	Salines	999									
423	Intertidal flats	999									
511	Water courses	999									
512	Water bodies	999									
521	Coastal lagoons	999									
522	Estuaries	999									
523	Sea and ocean	999									
999	NODATA	999									
990	UNCLASSIFIED LAND SURFACE	999									
995	UNCLASSIFIED WATER BODIES	999									

For each city, *Land Use Code* represents the original land use code from the downloaded land use files whereas *New Code* represents the newly developed land use code (see Table 3.11).

3.3 Statistics

To quantify the impact of the different variables on the usage of bike sharing, different statistical models were used. The determination of the model was based on different factors. The first was that all the data was not normally distributed. This feature was detected by using the Shapiro-Wilks test, which rejected the hypothesis that the data was distributed normally. The next step was to test the data for overdispersion, which means the variance exceeds the mean. Since this feature was true for all datasets, the appropriate model was the negative binomial regression, which is a generalization of the Poisson regression, but it assumes overdispersed data. The negative binomial regression is a model which also fits for the count data (Gebhart and Noland, 2014; Kim, 2018; Nair et al., 2012; Noland et al., 2016; Rudloff and Lackner, 2014).
The negative binomial regression was used for the trip-based and the station-based datasets, because both were overdispersed and not normally distributed.

For the trip-based data, two models were used, one for the daily data and one for the hourly data. To account for seasonal variability in the daily data a season variable was added to the model. The seasons were split as follows: spring (March, April, May), summer (June, July, August), fall (September, October, November), and winter (December, January, February). To determine if the model with the season variable was better than the one without, an Anova test was executed. This test compared the models with each other and suggested that the model with the seasonal variable was better in all cases. Additionally, the number of active stations in each month was added to account for changes in the number of stations. For the hourly data, instead of a seasonal variable (which was useless as the data was only from one month), a time of the day (ToD) variable was added. The variable split the day into six three-hour periods, starting from 00:00-03:00, to consider the hourly variability in the data. Then, the model with the ToD variable was compared to the model without ToD by executing an Anova test. The model with ToD was determined to be a better fit for all the cases. However, the season and ToD variables were not displayed in the results table, because they were only used to improve the data. Moreover, as they are not continuous variables, the resulting coefficient did not reflect the impact on the rides but only the differences within the variable itself. The only other explaining variable which is in the results, apart from the rides and the weather, is the weekend variable. Being a binary variable, and thus comparable, the impact on number of rides is displayed.

For the station analysis, no further normalization variables were added, because the data was only from one month that was summed up, and a seasonal or ToD variable was not needed. However, since the data for the infrastructure analysis contained many variables, some aggregations had to be completed prior to running the model. Two models were run, each for arrivals and for departures. For the first one, all the road types were aggregated to their parental category and all public transport stations were aggregated. In the second model, the most important road types were split into subcategories. Regarding public transport, trains and buses were considered separately.

For the distinction between subscribers and casual users, the same model as that of the general weather model was run. The only difference was that the model was run twice, once for the subscribers and once for the casual users. The generalization stayed the same with the seasonal variable for the daily and the ToD variables for the hourly model.

3.4 Local and Global Variables

As mentioned in the research questions one aim of this study was to find out if the sharing bike usage is affected more by global or by local variables. To be able to answer that question, the different variables had to be categorized into local and global.

Global variables are expected to be consistent. Additionally, they are expected to be the same no matter where you are. Furthermore, they cannot be planned or influenced.

Local variables differ from city to city. Additionally, they are controllable by the bike providers or the local government. Another characteristic of local variables is that they change slowly.

Examples for a global variable are season and weekend which are consistent, not controllable events that occur in every city. One variable which is hard to categorize is weather. One key characteristic for weather variables is that they change fast, especially precipitation. These kinds of variables are difficult to categorise into a local and global label. On the one hand, they are local; for example, rain is not regular, and it affects cities in different ways as differences occur in terms of intensity and frequency. Therefore, people from different cities are affected differently by rain. People from rainy regions might be affected less by rain than people from a city where it rarely rains. The way people manage it can also be described as the soft factors of rain. Alternatively, rain can be described as global as the main effect, or hard factors; for example, the fact that the streets and people get wet, is the same over the world. Furthermore, it is still a continuous variable and is comparable with other cities. The same can be applied for temperature. As a consequence, weather is grouped as global variable. The variables from infrastructure and land use are local factors, as every city has its own infrastructure, topography, public transport network, and unique land use areas. However, compared to the previous variables, infrastructure and land use areas are the only variables which are partially controllable by the bike providers or the local government. Bike providers would be able to plan their stations near subway stations, or at locations with low altitude. Another difference between infrastructure and the other variables, especially weather, is that it changes very slowly. Additionally, infrastructure and land use are the only variables which are not the same for all the cities. While rain and the weekend do not change at their core, infrastructure and especially cycling infrastructure change drastically from city to city. Some cities might have a very good bike path network or a public transport network which is well-elaborated, whereas other cities might have few bike paths and an insufficient public transport coverage.

Chapter 4

Results and Interpretation

This chapter begins by presenting a system comparison and descriptive statistics to offer an overview of the data. The seasonal distribution of bike sharing rides is then explored, followed by the daily temporal ride characteristics. Furthermore, the impact of weather events is described in addition to the influence of both infrastructure and land use. Finally, subscribers and casual users are compared, and the impact of the COVID-19 pandemic on the usage of bike sharing in New York is discussed.

4.1 System Comparison and Descriptive Statistics

This section examines the nine BSS that were used in this thesis, and they are compared with each other in terms of system size. This analysis includes the number of rides, the number of stations, and the size of the system area. The factor of trip duration is not discussed in this section, because it is not relevant to the system size. However, it is covered later in the chapter in a separate section. Table 4.1 evidences that the system in New York is the largest in terms of rides per day as well as in terms of number of stations; New York is followed by London, Chicago, and Washington, DC. Edinburgh is the smallest system, with only 257 rides per day on average and a maximum of 908 rides in one day. Figure 4.1 illustrates these differences by presenting the average rides per month. Regarding the number of bike stations (see Table 4.2), the differences between New York and the next largest systems are not considerably significant compared to the daily number of rides. For example, London and New York have almost the same number of stations, but New York had almost twice as many rides per day. Chicago which has 600 bike stations (compared to 793 in New York), had less than one fifth of the rides of New York. Alternatively, Oslo has few stations compared to its number of rides, which are near the numbers of Washington, DC and Chicago. However, Oslo had less than half as many stations as Washington and Chicago. Furthermore, Oslo's BSS is in service only from April to November, which means that the average did not include any winter months, which are typically the months with reduced usage that in turn decrease the daily average. The same pattern was found for Trondheim, which would presumably have less rides per day if it would be in service the whole year.

The presented systems are divided into two groups: the first is the five major cities (New York, Boston, Chicago, Washington, DC, and London) and the second is the smaller systems (Edinburgh, Bergen, and Trondheim). Oslo is in the middle of these groups since it is not as big as the major cities, but it is much larger than the smaller systems.

City	**Ø Rides per Day**	**Max.**	**Min.**	**SD**
New York	50,446	81,392	8,395	17,547
London	28,488	43,839	6,328	8,578
Chicago	9,521	19,240	143	5,587
Washington, DC	9,126	17,046	598	3,733
*Oslo**	8,678	17,512	10	4,037
Boston	5,190	9,650	142	2,439
Bergen	1,183	6,079	20	1,317
*Trondheim**	746	2,471	4	531
Edinburgh	257	908	41	166

* System only in service from April until the end of November

Table 4.1 – System Comparison: Number of Rides per Day

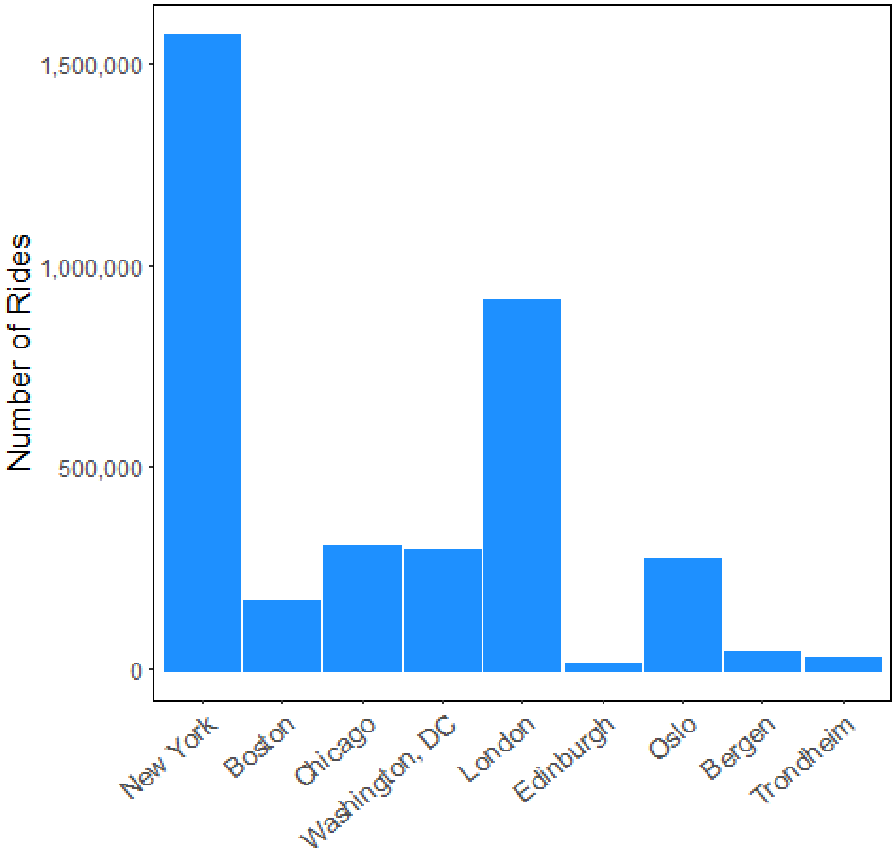


Figure 4.1 – System Comparison: Average Rides per Month

The "Rides per Station" ratio in Table 4.2 was determined by the "Ø Rides per Day" from Table 4.1 divided by "Bike Stations" in Table 4.2. New York had the most rides per station with a ratio of 63.6, followed by London and Oslo with 36.5 and 35.6, respectively. Boston, Washington, Chicago, Trondheim, and Bergen ranged between 10–20. Finally, Edinburgh had a small ride per station ratio of 3.2.

The service area was calculated using the convex hull. Washington, DC had the largest service area with 897.69 km^2. This large scale is because the system not only runs in Washington, DC but also in surrounding towns, which are kilometres away from the city. However, few stations exist between Washington, DC and its surrounding towns, leaving open spaces between Washington and its surrounding cities. This characteristic is demonstrated in Figure 4.2, which is a screenshot of Washington, DC's service area that was retrieved from the BSS provider's website. The smallest service area was in Bergen with 8.94 km^2. Based on the number of stations per square kilometre, Washington, DC had a low value with 0.6 stations per square kilometre. Bergen had the highest number of stations per square kilometre, followed by New York, Oslo, and London with 6.7, 6.4 and 6.1 Stations per km^2, respectively.

Table 4.2 – System Comparison: Bike-Stations

City	Bike Stations	Rides per Station	Area (km2)	Stations per km2
New York	793	63.6	118.86	6.7
London	780	36.5	126.93	6.1
Chicago	600	15.9	353.65	1.7
Washington, DC	553	16.5	897.69	0.6
Oslo	244	35.6	38.05	6.4
Boston	268	19.4	137.24	2.0
Bergen	80	12.3	8.94	8.9
Trondheim	54	14.8	11.22	4.8
Edinburgh	81	3.2	68.36	1.2

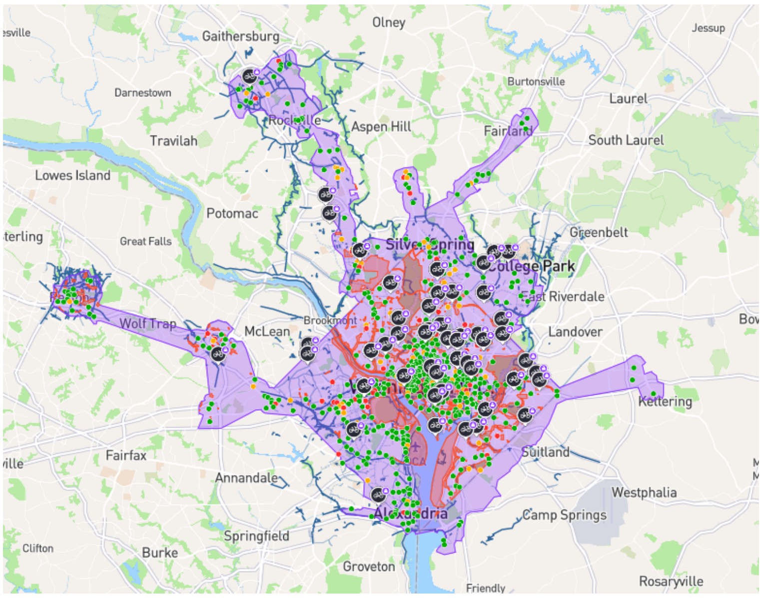


Figure 4.2 – Service area of Washington, DC.

Source: https://secure.capitalbikeshare.com/map/ (22.07.2020)

Table 4.3 presents the percentage of rides which fall into the initial free period of the corresponding system. All the systems offered their users a free initial period. The free period was between 30–60 minutes for all cities. If the ride takes longer than this initial free period, the user must pay an additional fee for the longer usage. Since not all systems had the same initial period and not all users had the same free period—for example, subscribers usually have a longer free period than casual users, and two-day pass users have a longer free period than one-day pass users—it was not possible to calculate the percentage for all systems. The five systems listed had the same initial free period for all their users, which allowed the percentage of the trips which fall into this free period to be calculated. The percentages were relatively high and all were over 90%. Oslo and Bergen had the highest percentages with 99.1% and 98.8% and an initial free period of 60 and 45 minutes, respectively. Washington, DC had the lowest percentage with 91.2%. However, it was also the system with the shortest free period of 30 minutes. Edinburgh had a free period of one hour, and its percentage was in between that of Washington, DC and of Oslo and Bergen.

The results demonstrate that bike sharing users mostly stay inside the free initial period. One reason might be that people do not want to pay extra for the usage. Alternatively, 30 minutes is a relatively long period, and a considerable distance can be covered. Since many rides fall into this period, the distance they cover takes less than 30 minutes.

Table 4.3 – Rides in Free Initial Period

City	**Free Initial Period**	**Rides in Free Initial Period (%)**
Washington, DC	30 min	91.2
Edinburgh	60 min	95.4
Oslo	60 min	99.1
Bergen	45 min	98.8
Trondheim	60 min	94.3

4.2 Seasonal Differences and Daily Ride Characteristics

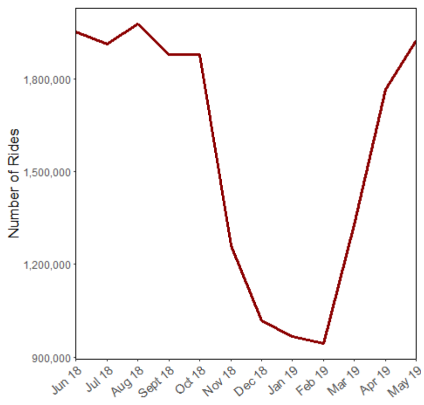
This section examines the seasonal variability regarding the number of rides as well as the duration of rides. The second part explores the daily temporal ride characteristics, including how the rides are distributed over one day and how the trip duration changes over the course of one day.

4.2.1 Seasonal Differences

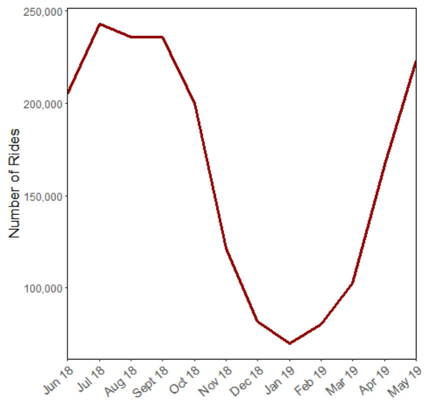
Number of Rides

Figure 4.3 illustrates the usage pattern for one year (except for Oslo and Trondheim, which are in service only for eight months). For the cities (a)–(e) a similar pattern is visible. The most obvious aspect is the strong decrease of rides in the winter months. The lowest ride counts occur between December and February. From March forward, the numbers start to increase, and the peak is in July or August. Oslo's graph (g) is similar to the first five aside from the fact the system does not operate during the winter months. Nevertheless, Oslo experienced a peak in August as well as a decline in fall, and the lowest counts were in November before the systems closes. For Bergen, the numbers significantly increase in March and remain high compared to the months before. This characteristic might be caused by a system enlargement or system improvement. Before the substantial increase, the system indicated a common pattern with a peak in August before declining in winter. Trondheim demonstrates the same characteristic as the system peaks in September, and the numbers decline until November. The system then closes and reopens in April with a much higher usage than before. In May, the system experiences a peak significantly larger than the one in September. The system from Edinburgh had a similar pattern to Bergen and Trondheim, but the increase is not as steep. Edinburgh also had an increase from March–August.

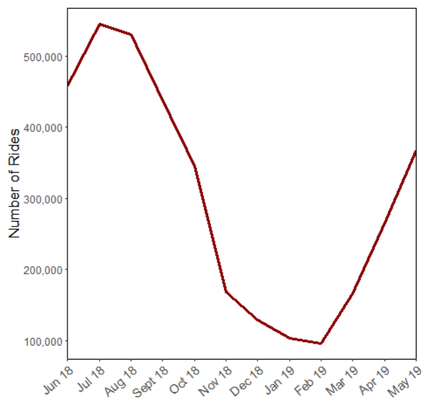
As such graphs reflect absolute numbers, they are prone to system changes. For example, if the system operators add several stations and bikes or increase the advertisement of the BSS, this information is not displayed in this graph, which influences the results. To exclude the factor of a possible addition of many new stations, the same graphs were created. Time was normalized by dividing the



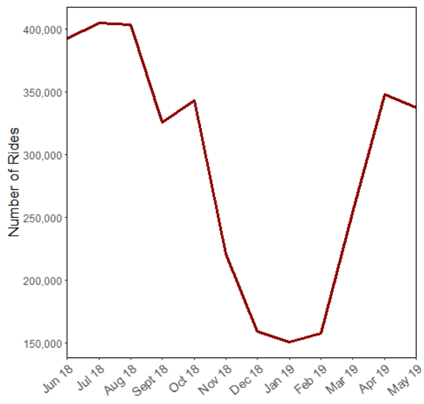
(a) New York



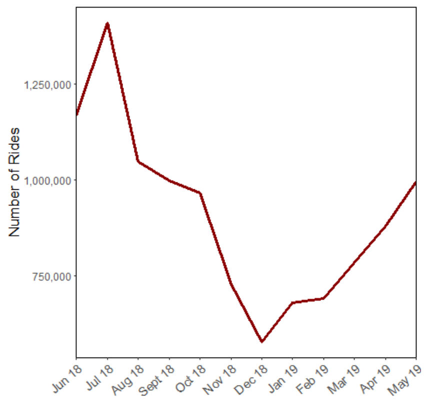
(b) Boston



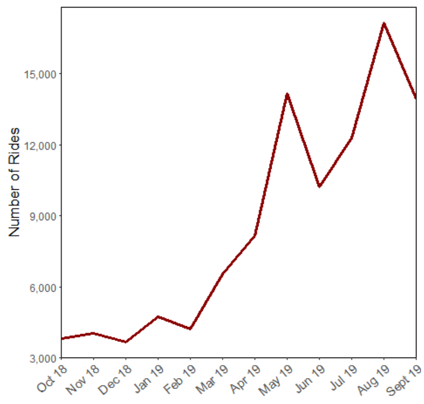
(c) Chicago



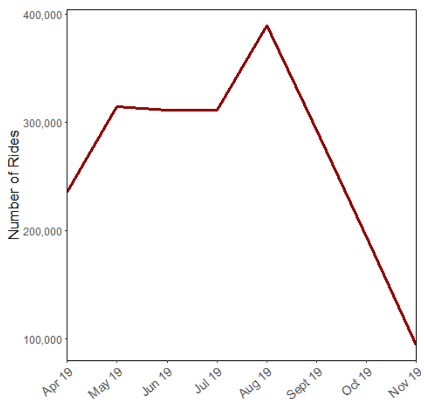
(d) Washington, DC



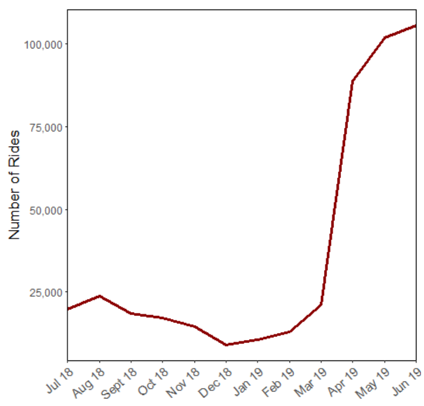
(e) London



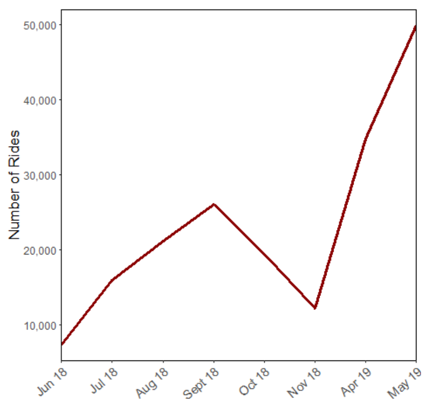
(f) Edinburgh



(g) Oslo*



(h) Bergen



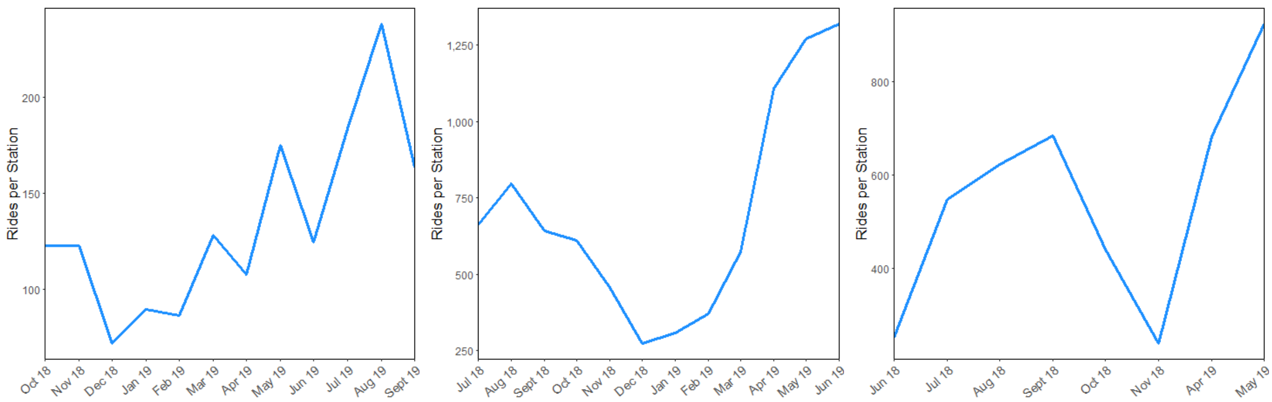
(i) Trondheim*

Figure 4.3 – Number of rides over one year.

* The systems in Oslo and Trondheim are in service from March–November

number of rides by the number of stations. In general, the graphs do not differ much, especially the five large cities (a)–(e) exhibit no visible difference. Oslo (g) is similar to these, indicating that for these six cities not many stations were added or removed. As a consequence only the graphs from Edinburgh, Bergen, and Trondheim are displayed in Figure 4.4 as these are the ones with visible changes.

The graph from Bergen is more familiar with a peak in August, a minimum in December, and an increase from March forward. Although a major increase is visible, it is not as significant as the one in Figure 4.3. The same pattern is observable for Trondheim. However, Edinburgh's graph is inconsistent and does not offer a better insight on how the number of stations might affect the total usage numbers. A reason for the increase might be that the operators added larger stations with more bikes compared to the older ones, which led to more rides per station. Another explanation could be that the existing stations were enlarged, rather than only building new ones, which would also lead to more rides per station. The third explanation might be that the system is more attractive, because it has more stations. More stations could lead to a tighter station network or an enlarged service area which might have a positive effect on all the stations and results in a higher usage and more rides per station.



(a) Edinburgh **(b)** Bergen **(c)** Trondheim*

Figure 4.4 – Ride per station ratio over one year.

* The system in Trondheim is in service from March–November.

Lastly, Table 4.4 compares the cities in terms of seasonal differences, and it lists the maximum deviations from the average number of rides per month. A smaller deviation means that the number of rides did not change significantly between the different months whereas a high value indicates that significant differences were experienced. The table presents the major increases in usage in Edinburgh, Trondheim, and Bergen that were previously discussed. For these three cities the maximum positive deviation ranges from 100–185%. These increases had an impact on the average, which led to higher negative deviations as well. Regarding the other cities, Chicago had both the highest maximum and minimum deviations, which suggests that the usage in Chicago varies strongly depending on the season. The reason for this might be that Chicago has heavier winter months with lower temperatures and more precipitation, which could lead to less usage of bike sharing. Alternatively, New York had the smallest deviations with 26.16% positive and 39.78% negative, which indicates a more balanced usage. Boston, Washington, DC, and Oslo are between the two extremes.

Table 4.4 – Maximum Positive and Negative Deviations from the Average Number of Rides per Month

City	**Max. Positive Deviation**	**Max. Negative Deviation**
New York	26.16%	-39.78%
Boston	48.23%	-57.36%
Chicago	80.94%	-68.05%
Washington, DC	38.85%	-48.28%
London	54.53%	-36.47%
Edinburgh	99.81%	-57.31%
Oslo	45.42%	-64.88%
Bergen	185.19%	-75.67%
Trondheim	113.81%	-68.67%

Trip Duration

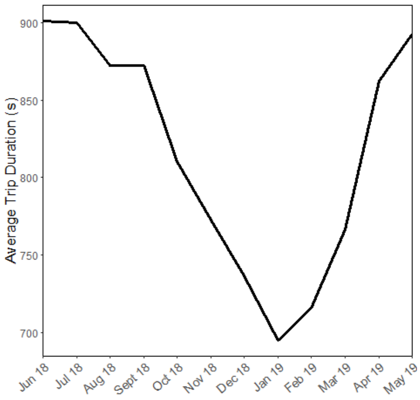
Table 4.5 reflects the average trip duration for each system for one year. To calculate this, the average trip duration for each month were used. Additionally, the maximum positive and negative deviations from the mean are displayed. Edinburgh had the longest trip duration average (23.6 min) which was almost six minutes more than London with the second longest duration (17.9 minutes). Compared to Oslo, Bergen, and Trondheim, Edinburgh's trip duration average was more than twice as long. Bergen had the lowest average with 11.4 minutes, which was similar to the other two Norwegian cities of Oslo (11.8 min) and Trondheim (11.6 min). Considering the deviations from the average, New York had relatively low percentages, meaning that the different months had similar average trip durations, which suggests a similar riding behaviour throughout the year. In contrast, Chicago had the most inconsistent riding behaviour with high deviations both in the positive and the negative direction. Edinburgh (17.52% and -10.64%) had similar values to New York. The other cities all had greater changes, with deviations up to almost 40%, which indicates that the monthly averages are dissimilar. The five major systems of New York, Boston, Chicago, Washington, DC, and London (as well as Oslo) had more consistent values than the smaller systems. The larger systems had similar deviations in the negative and positive directions, whereas the smaller systems tended to have higher values in one direction.

Table 4.5 – Descriptive Statistics: Trip Duration

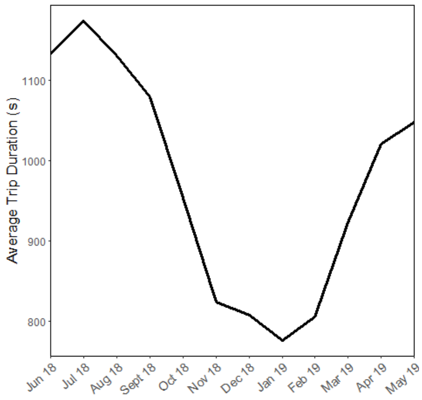
City	Mean Trip Dur. (s)	Mean Trip Dur. (min)	Max. Positive Deviation	Max. Negative Deviation
New York	816	13.6	10.36%	-14.91%
Boston	973	16.2	20.66%	-20.22%
Chicago	937	15.6	31.94%	-27.58%
Washington, DC	1,018	17.0	21.93%	-19.77%
London	1,071	17.9	14.13%	-15.07%
Edinburgh	1,417	23.6	17.52%	-10.64%
Oslo	693	11.6	19.12%	-22.02%
Bergen	682	11.4	39.56%	-18.12%
Trondheim	705	11.8	33.38%	-14.80%

Figure 4.5 presents the average monthly trip durations for the entire data period. The five major cities exhibit a similar pattern with longer trip durations in the summer months and shorter durations in the winter. January and February were the months with the shortest trip durations, and June and July were the months with the longest trip durations. London and Washington both experienced a minor peak in April before the average duration slightly decreased leading into May. Additionally, London had a small peak in December before the average duration dropped to a minimum in January. Edinburgh would look similar to the major five cities if it covered the same timeline; its trip duration decreased in the winter and had a minimum in February before it increased in the spring with a peak in May. However, the duration decreased after May and did not remain high like the previous examples. Oslo, Bergen, and Trondheim peaked in July before decreasing as the winter approached. All three of the Norwegian cities documented a decrease in May compared to April, which was similar to London and Washington. Regarding the difference between the longest and shortest trip duration, New York had a smaller difference than Chicago. New York's peak was around 900 seconds, and its minimum was around 700 seconds on average. Alternatively, Chicago peaked around 1,200 seconds, and its minimum was 800 seconds. These observations were congruent with the findings from Table 4.5.

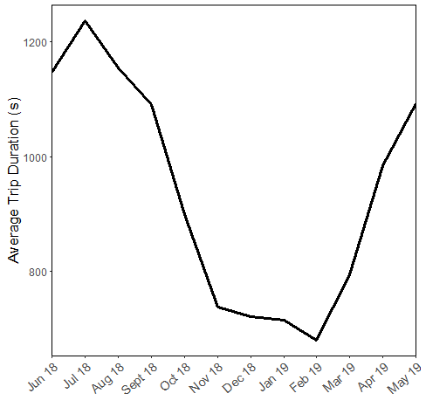
The reasons for the longer trip durations vary. Some explanations could be due to longer distances between the stations, different types of usage (commuting versus recreational), or the amount of tourists using the service (assuming that tourists take longer rides than commuters), which would also lead to longer trip durations in the high season for tourism. The trip durations could also be affected by the length of the initial free period. Cities with a shorter initial free period presumably tend to have shorter average trip durations, because users try not to exceed the time limit. Furthermore, the different types of subscriptions might have an impact. For instance, the system in Chicago offers unlimited three-hour rides by purchasing a one-day pass. This offer might also encourage people to take longer trips.



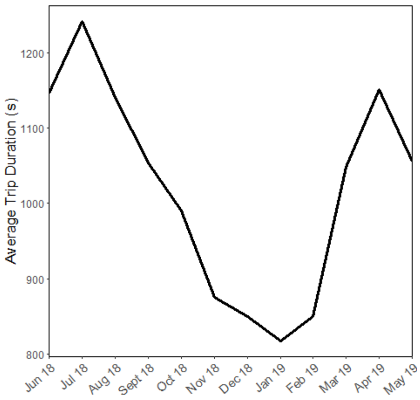
(a) New York



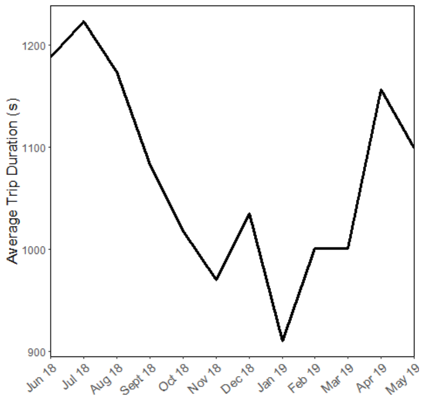
(b) Boston



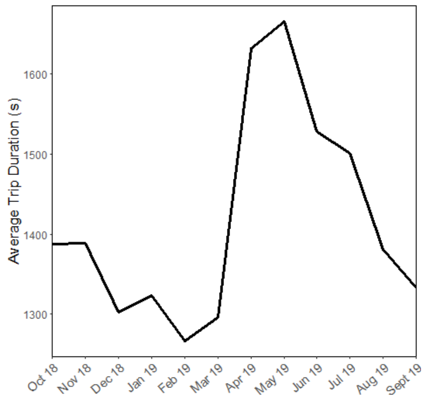
(c) Chicago



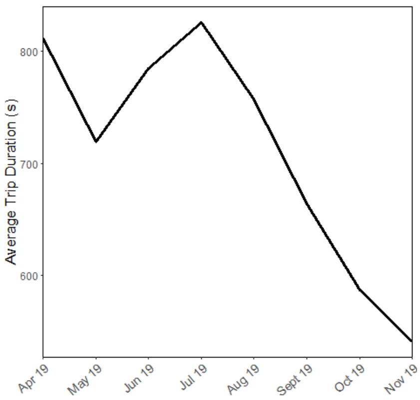
(d) Washington, DC



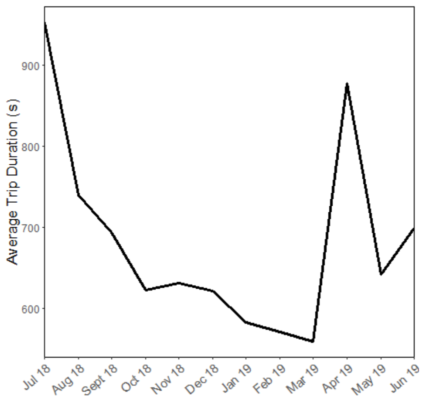
(e) London



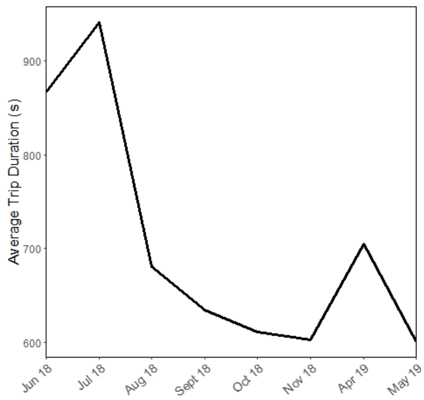
(f) Edinburgh



(g) Oslo*



(h) Bergen



(i) Trondheim*

Figure 4.5 – Trip durations over one year.

* The systems in Oslo and Trondheim are in service from March–November.

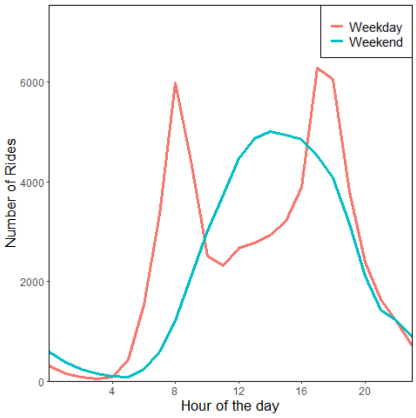
4.2.2 Temporal Ride Characteristics

Number of Rides

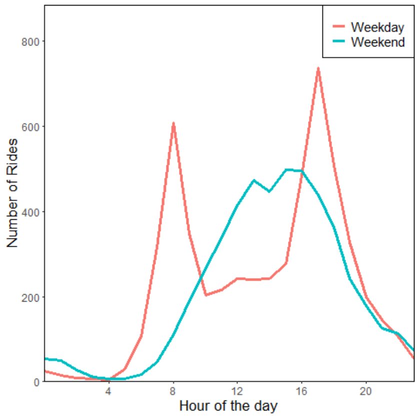
This sub-section focuses on the distribution of bike sharing rides over the course of one day. Particularly, the differences between the weekdays and weekends are considered, and different months are compared.

Figures (4.6, 4.7, and 4.8) present the distribution of the rides over weekdays and weekends. The x-axis represents the hours of the day, and the y-axis is the number of rides. The plots were created by using hourly counts, such as 7:00–7:59. Therefore, when a peak around 7:00 is mentioned, it means that the usage is highest between 7:00–7:59. The displayed value is the average of all the hourly counts over the course of one month. The y-axis is different in each graph so that the cities can be compared with each other in terms of daily usage characteristics. Additionally, the difference in size between the systems and the seasonal variability was described earlier. Figure 4.6 presents data from April 2019, because it was available in all the datasets. The systems in Bergen and Trondheim are not open 24 hours per day; they close at midnight and reopen at 6:00. The system in Oslo closes at 1:00 and reopens at 5:00; therefore, the straight line in Oslo's (g) graph is because a value exists for the midnight hour (00:00–00:59) but no values are registered until 5:00, which translates to a linear interpolation from the midnight hour to 5:00.

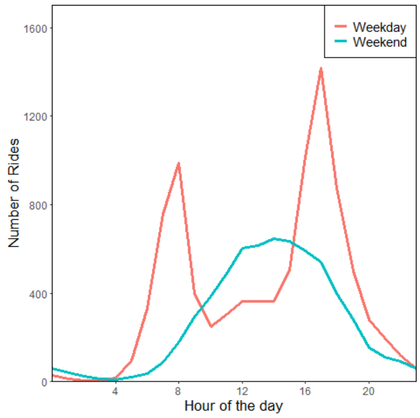
In general, a repeating pattern is evident for weekdays as each system had a considerable morning and afternoon peak. The morning peak for all the cities was at 8:00 and was smaller than the evening peak in all cases except London (e). Moreover, the evening peak lasted a bit longer than the morning peak, and in some cases it was not as sharp as the morning peak which was a result of higher numbers prior and after the main peak. The evening peaks were at 16:00 and 17:00. Furthermore, except for Oslo and Bergen, all the cities had a smaller midday peak at noon; however, the peak was not as significant as the morning and afternoon peaks. Overall, the usage was higher after midday, and it decreased in the evening. For the weekend, the pattern was similar for all the systems. There was one major peak in the afternoon, which is flat compared to the weekday peaks. The number of rides increased slightly over the day, peaked around 12–16:00 and then slightly decreased again. The weekend peak was smaller than the ones on weekdays.



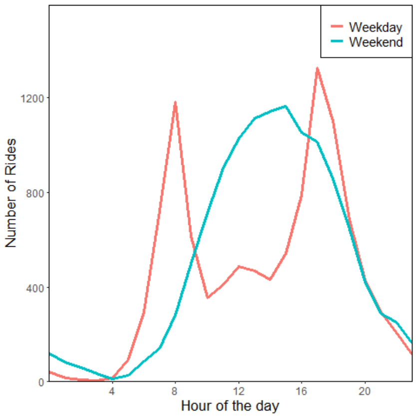
(a) New York



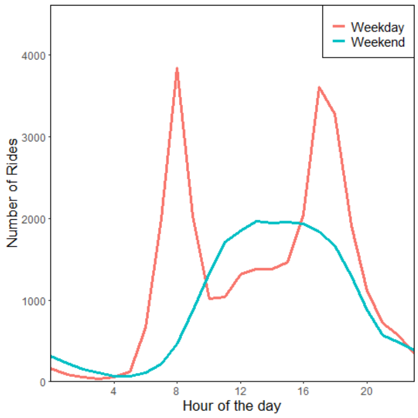
(b) Boston



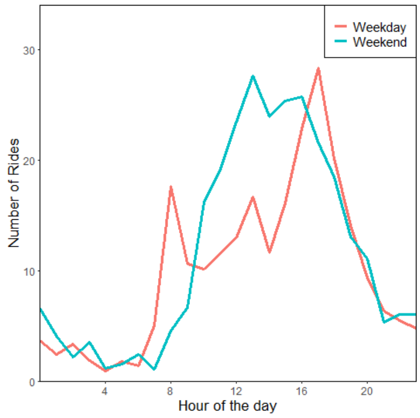
(c) Chicago



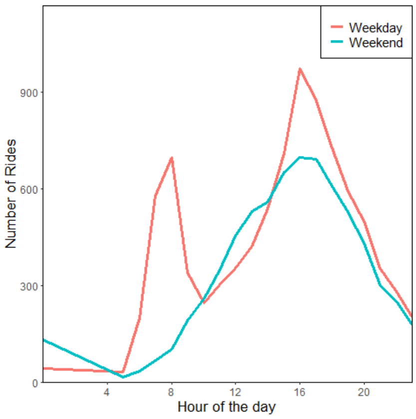
(d) Washington, DC



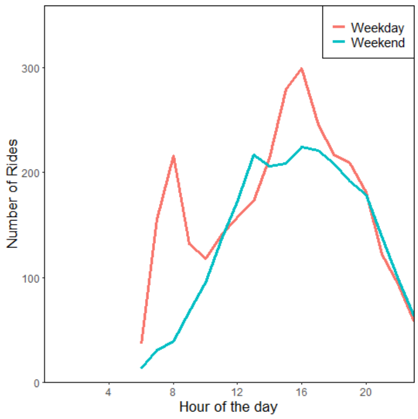
(e) London



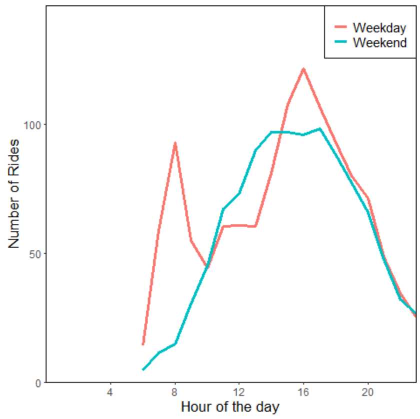
(f) Edinburgh



(g) Oslo**



(h) Bergen*

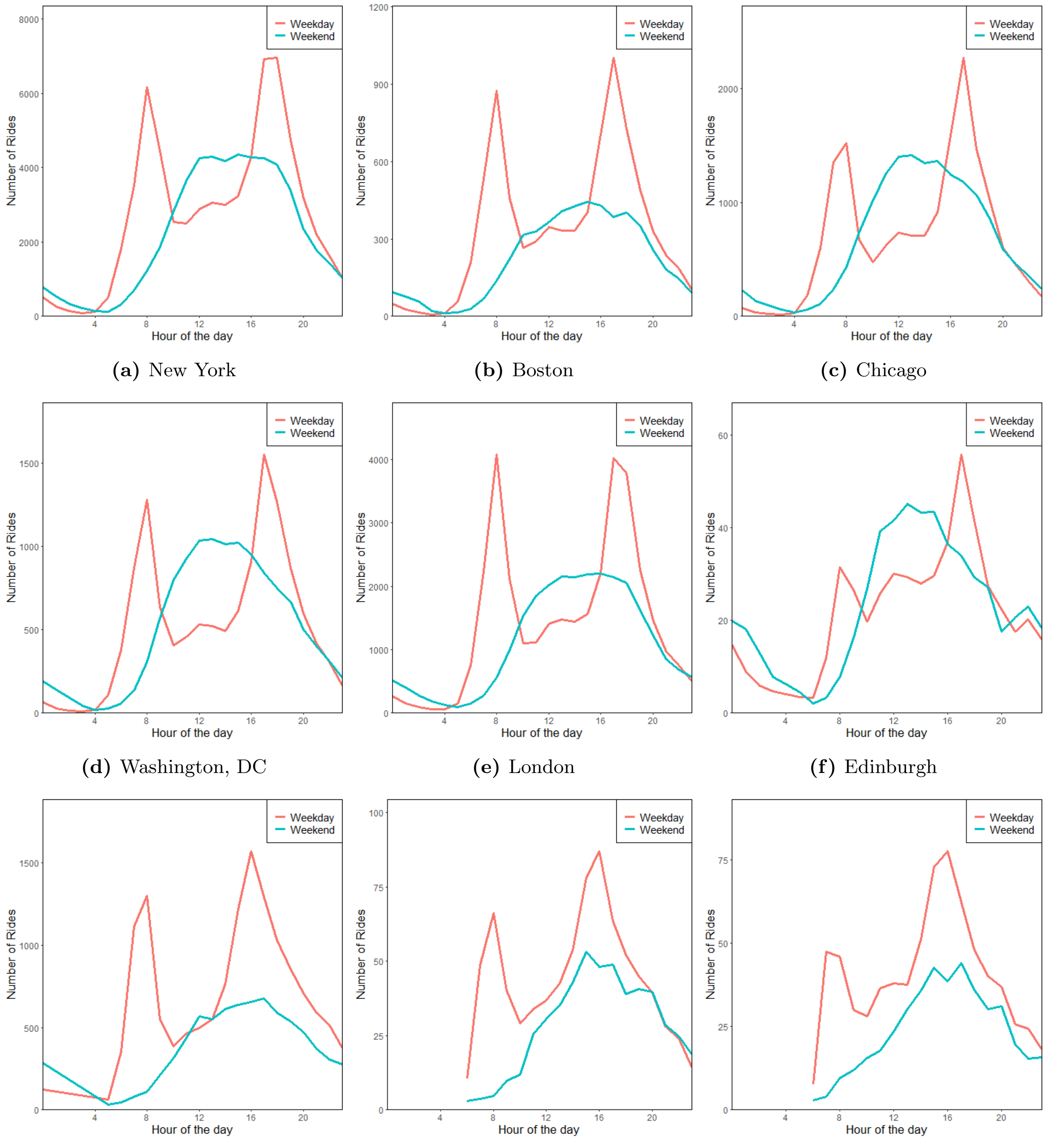


(i) Trondheim*

Figure 4.6 – Average hourly usage on weekdays and weekends in April.

* The systems in Bergen and Trondheim are in service from 6:00 until midnight.

** The system in Oslo is in service from 5:00 until 1:00.



(a) New York (b) Boston (c) Chicago

(d) Washington, DC (e) London (f) Edinburgh

(g) Oslo** (h) Bergen* (i) Trondheim*

Figure 4.7 – Average hourly usage on weekdays and weekends in August.

* The systems in Bergen and Trondheim are in service from 6:00 until midnight.

** The system in Oslo is in service from 5:00 until 1:00.

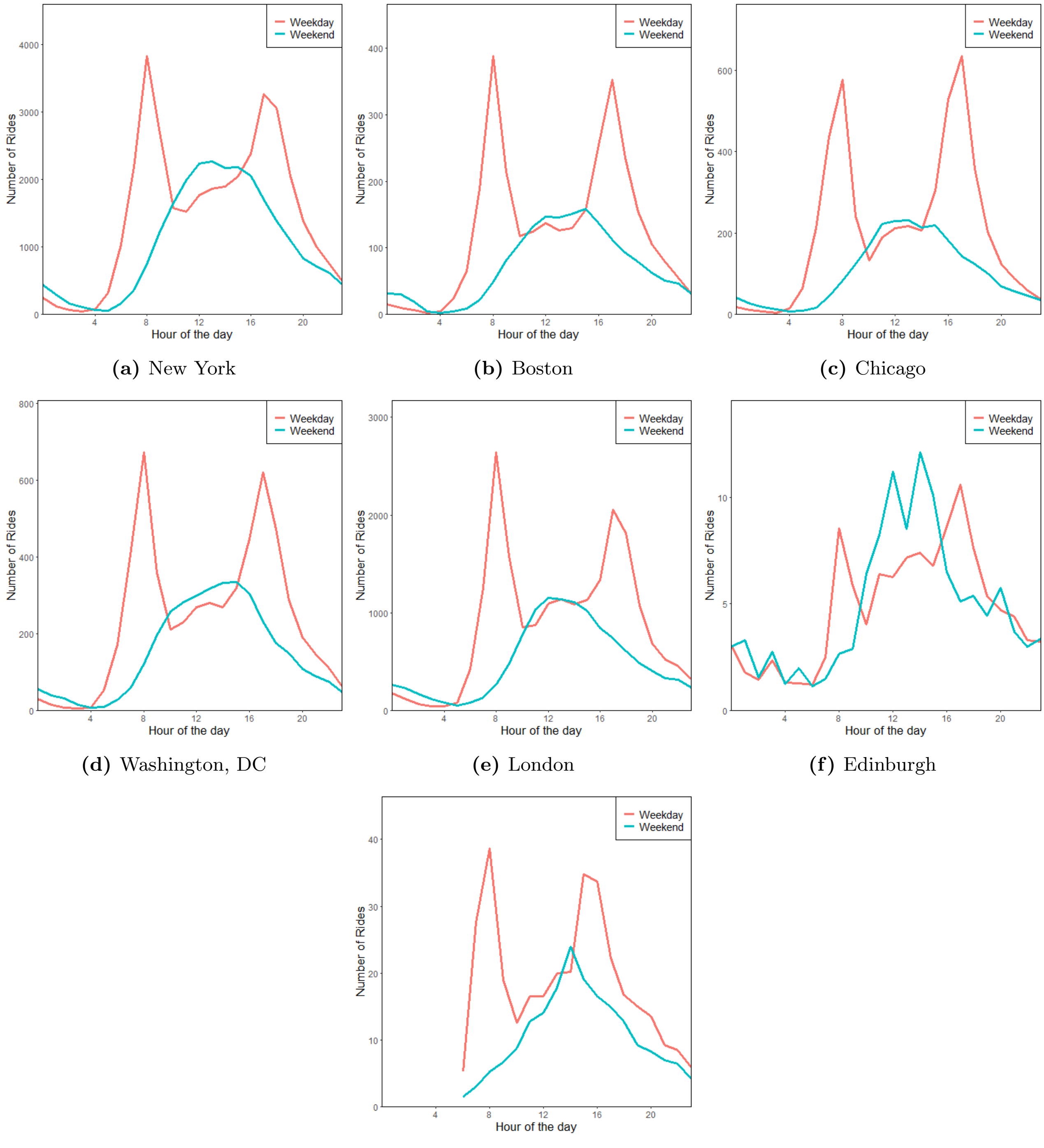
The hours with the least number of rides are from 2–4:00. However, one thing to note is the increased usage of bike sharing in the hours immediately after midnight on the weekend compared to weekdays.

The next month that was examined was August. The data is from August 2018, except that of Edinburgh and Oslo, which is from August 2019. In Figure 4.7 the daily patterns are evident, similar to the ones from April. Again, two major peaks occurred on weekdays with a small midday peak. The morning peak was at 8:00, except for Trondheim which was at 7:00. The afternoon peak was greater than the morning peak except for London. The afternoon peaks were similar to April around 16–18:00. The weekend pattern exhibited the familiar one peak characteristic, with a long peak from approximately noon through the evening. Again, the weekend graphs displayed increased usage in the hours after midnight.

The Figure 4.8 depicts the characteristics of all the systems that ran in December 2018. Since Oslo and Trondheim are not in service in the winter, only seven cities are included. The general pattern with the morning and afternoon peaks as well as the small midday peak for weekdays still existed. However, the morning peaks were mostly larger than the afternoon peaks, except in Chicago and Edinburgh. The time of the peaks did does not change in December, and all the cities had a morning peak at 8:00 and an afternoon peak between 16–18:00. For the weekend, the peak was smaller than the ones in April and August relative to the weekday peaks. Whereas in April and August the weekend peak was between the midday peak and the morning and afternoon peaks for weekdays, in December it was slightly greater than the midday peak. Similar to the previous two figures, the weekend exhibited an increased usage after midnight.

One outlier in all three months was Edinburgh (f), which had an increased weekend usage compared to the other cities and a weekend peak that exceeded the weekday peaks. Moreover, the morning peak was smaller than the afternoon peak and was slightly greater than the midday peak.

These graphs evidence how the systems are used differently on weekdays and on weekends. One possible explanation for the weekday pattern might be that bike sharing is used for commuting purposes in the morning and evening, which would explain the two major peaks. The fact that the afternoon peak is less sharp than the morning peak might be because in the evening some people might not



(a) New York **(b)** Boston **(c)** Chicago

(d) Washington, DC **(e)** London **(f)** Edinburgh

(g) Bergen*

Figure 4.8 – Average hourly usage on weekdays and weekends in December.

* The system in Bergen is in service from 6:00 until midnight.

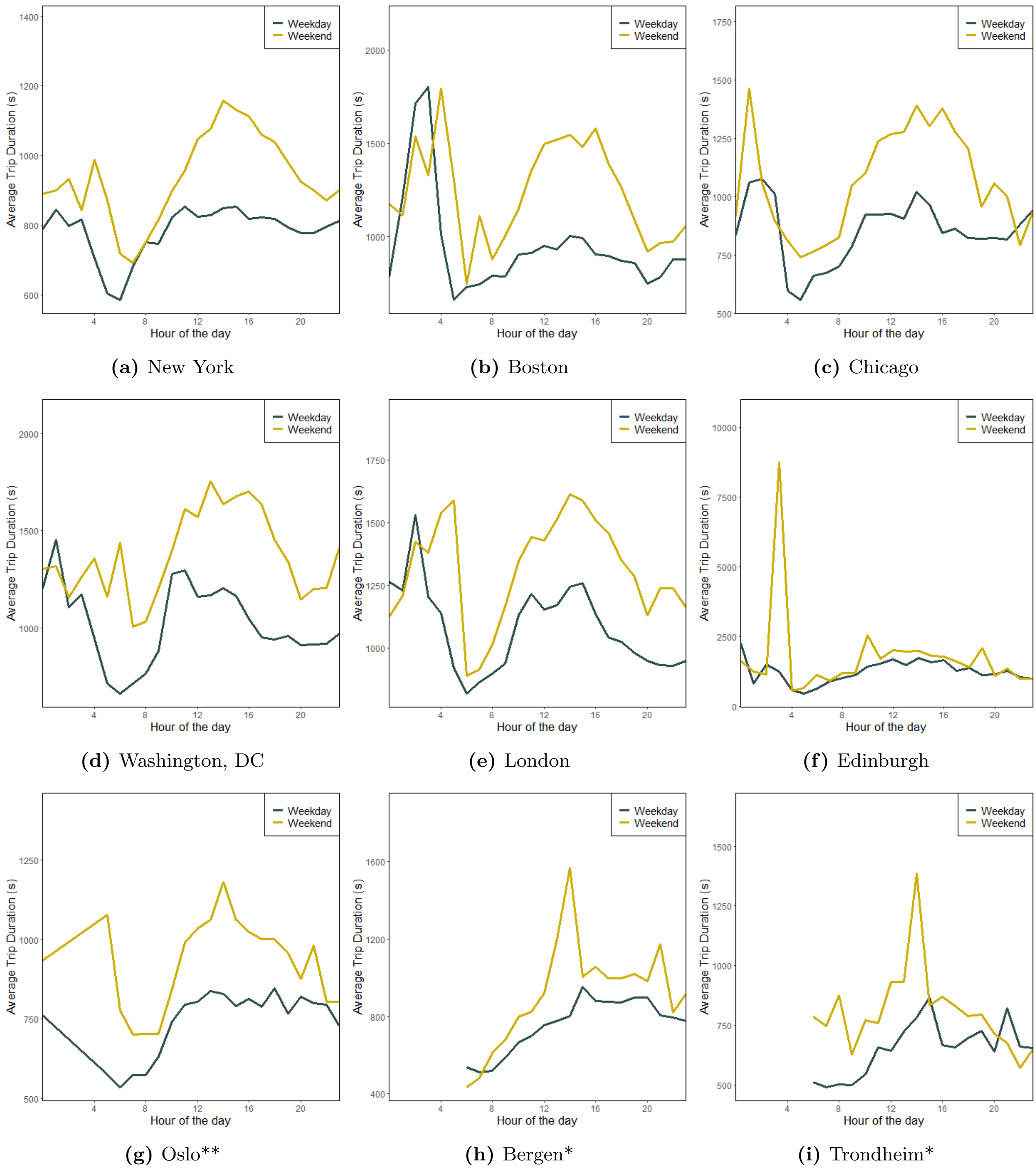
go home immediately but rather go to a pub or do their shopping before returning home by bike. This occurrence would explain the slightly longer peak. The fact that the evening peak was greater than the morning peak might be due to the fact that more casual users use the system in the evening than in the morning, or because people tend to take bikes to get home rather than to go to work. This trend could be the case, because people might estimate other ways of transport to be more reliable or faster. They might fear empty docking stations in the morning whereas in the evening there is less pressure of time, which would benefit the usage of bike sharing. The small midday peak is most likely caused by people using bikes during their lunch break.
For the weekend, the graph changes, because there are few commuters. One reason for the increase in trips after midnight could be that people use bike sharing to return home after going out or to go out after midnight. Furthermore, the weekend pattern is likely dominated by recreational activities as people tend to use shared bikes more in the afternoon. Moreover, the weekend usage decreased relative to the weekday usage in winter. This scenario could be because people tend to leave their houses less during colder months and stay home on weekends, or because they use other ways of transport that are less exposed to the cold climatic conditions. The smaller evening peak in December might also be due to similar reasons. On the one hand, there might be fewer casual users using bike sharing; on the other hand, as days are shorter in winter some people might take other ways of transport to get home. Individuals also probably do not want to ride in the dark as sunset is often before the typical working day ends, whereas during the morning peak between 8:00–8:59 the sun is already rising.

The ride characteristics of Edinburgh and some characteristics of Oslo, Bergen, and Trondheim differ from the major five systems. The ride characteristics of these cities is characterised with a smaller morning peak than the evening peak and a greater weekend peak than the weekday peaks. In the case of Edinburgh, the same size or greater than the evening peak. One possible explanation might be that these systems are more used by casual users, who presumably use bikes more during the afternoon than in the early morning, which is when the percentage of subscribers is expected to be the highest. The increased weekend usage supports this assumption as weekend usage is dominated in most of the cases by recreational rather than commuting purposes.

Trip Duration

After analysing the daily patterns regarding the number of rides, this section addresses the trip duration over the course of one day. The graphs in Figures 4.9, 4.10, and 4.11 were created by taking the average trip duration of all rides occurring in one particular hour (e.g. 7:00–7:59) over a one-month period. Since the result is an average, it is dependent on the number of data points. Therefore, the hours with less rides are more prone to outliers with extremely long trip durations as these have a greater impact on the average compared to a high frequency hour with many data-points. The smaller systems are therefore more prone to outliers as they have less rides than the major systems, which still have a reasonable number of rides during low-frequency periods. Additionally, over the course of one month the weekend will inherently have fewer data-points than weekdays, which partially explains the periodic large peaks in the early morning hours and on the weekend as this is when the fewest data-points were available.

Figure 4.9 presents the average trip duration for April 2019, and it indicates that the ride duration on the weekends is generally higher than on weekdays. Moreover, the systems (a)-(g) exhibit similar patterns with a decreased trip duration around 6:00. After this minimum, the trip duration was more or less stable. Before the low at 6:00, the systems demonstrate a higher average trip duration than during the rest of the day. Boston had an especially high peak in the early morning both on weekdays and weekends. Furthermore, these five systems all have one or two small peaks, both before and after noon. The graph for the weekend also has a decrease around 6:00, and then rises to a peak in the afternoon, before decreasing in the evening. Edinburgh had a similar pattern, but the graph is slightly distorted by the significant weekend peak around 3:00. On weekdays, the temporal characteristics experienced a longer trip duration in the early day hours, before decreasing in the morning and then rising again. It also had two small peaks before and after noon. Oslo had a similar pattern to the major cities with a long trip duration around midnight and a decreased duration around 6:00 before the level levelled off late in the evening. Bergen and Trondheim had a different pattern with no data until 6:00. The trip duration rose from approximately 8:00 until midday. In Bergen the graph levelled off, but in Trondheim the graph decreased again after midday and rose again around 21:00. In general, many systems exhibited long trip durations after midnight.



(a) New York **(b)** Boston **(c)** Chicago

(d) Washington, DC **(e)** London **(f)** Edinburgh

(g) Oslo** **(h)** Bergen* **(i)** Trondheim*

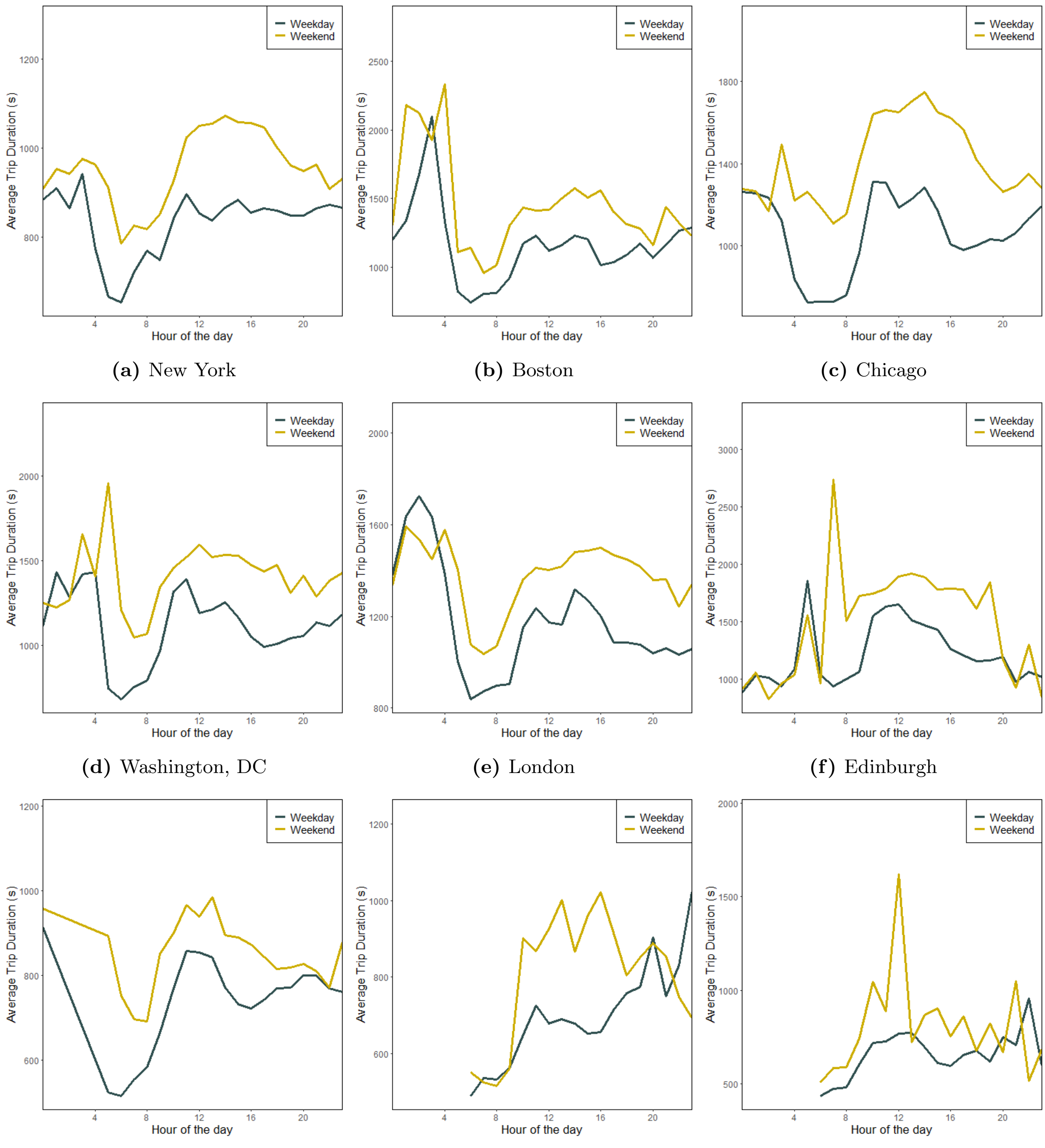
Figure 4.9 – Trip duration over one day in April.

* The systems in Bergen and Trondheim are in service from 6:00 until midnight.

** The system in Oslo is in service from 5:00 until 1:00.

Figure 4.10 depicts the trip durations for August. For Edinburgh and Oslo, data from August 2019 was used, and for the other cities data from August 2018 is reflected. The plots exhibit similar patterns to the ones from April. The high peak in the early morning hours for London, Boston, and Oslo is noteworthy. In August, the pattern of Edinburgh matches better with the ones from the major cities (a)–(e) as it decreased before 6:00 and also had a longer trip duration in the early morning. The plot from Bergen rises more or less from the start until the end. In general, April and August have a similar pattern in regard to trip duration.

Figure 4.11 presents the trip durations from December 2018, and it presents a completely different characteristic. The only city with a similar pattern to the two mentioned before is New York, and to a lesser extent London and Washington, DC. In New York, the graph decreases before 6:00; however, the trip duration in the early morning is shorter compared to the ones from April and August and to the trip duration of the day hours. The system peaked before noon and levelled off before 17:00, when a small peak occurs. Alternatively, in December London demonstrated an even greater peak in the early morning compared to April and August as well as the rest of the day. The trip duration decreases before 6:00 and rises before noon, before decreasing unevenly. Washington had a comparable pattern to London; however, it is less consistent. All seven cities except for Bergen peaked in the early morning. Except from that trend, there is not a visible pattern. The weekends followed a similar schema with a high peak in the early morning and a smaller peak in the afternoon. Furthermore, the trip duration for weekends was not much longer than the weekday trip duration, which is different compared to April and August when the trip duration on weekends was longer than that of weekdays.



(a) New York **(b)** Boston **(c)** Chicago

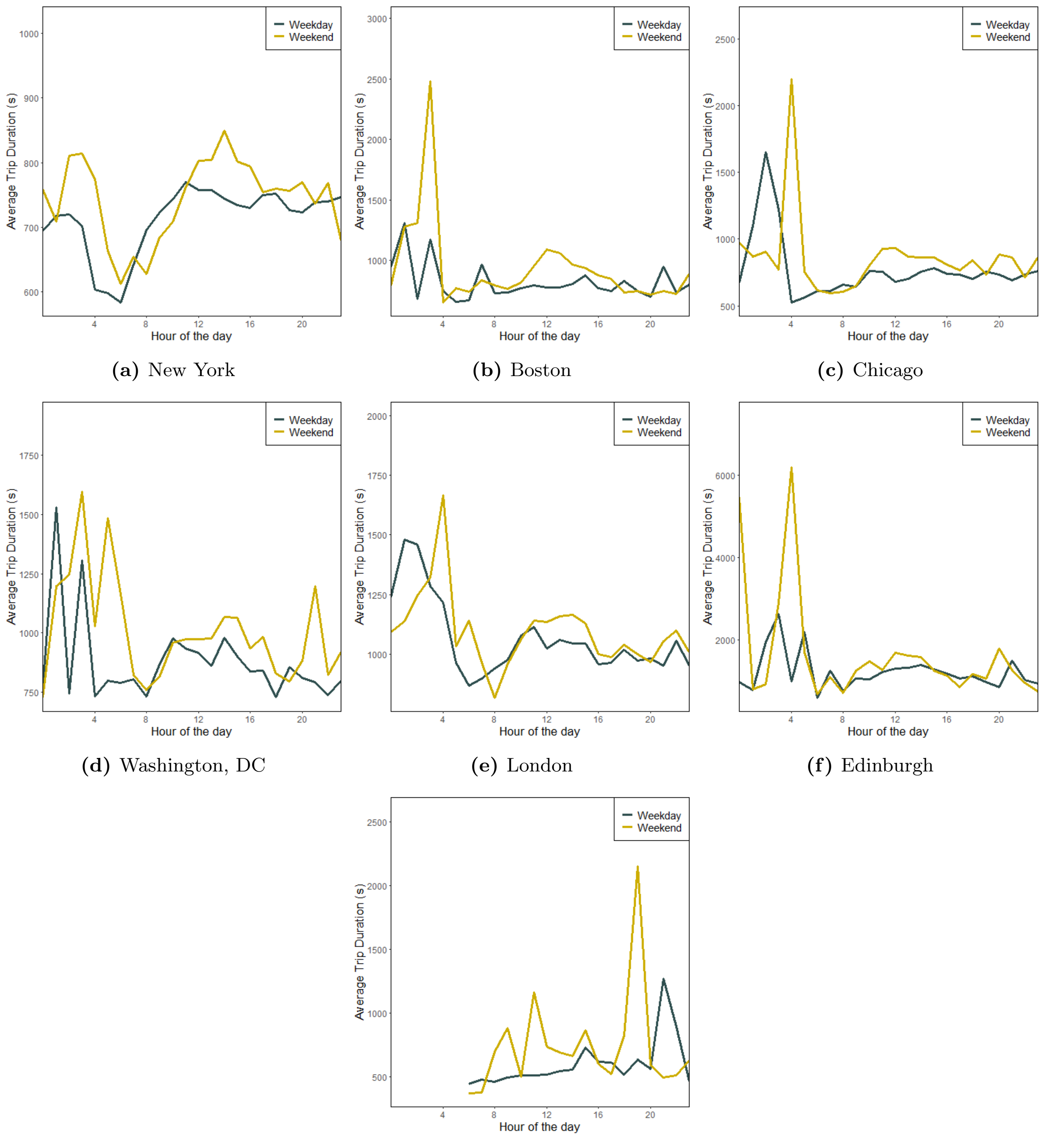
(d) Washington, DC **(e)** London **(f)** Edinburgh

(g) Oslo** **(h)** Bergen* **(i)** Trondheim*

Figure 4.10 – Trip duration over one day in August.

* The systems in Bergen and Trondheim are in service from 6:00 until midnight.

** The system in Oslo is in service from 5:00 until 1:00.



(a) New York

(b) Boston

(c) Chicago

(d) Washington, DC

(e) London

(f) Edinburgh

(g) Bergen*

Figure 4.11 – Trip duration over one day in December.

* The System in Bergen is in service from 6 am until midnight.

Overall, the most evident characteristic is the increased trip duration on weekends compared to weekdays. This trend might be due to the fact that on weekends people use bike sharing for recreational purposes rather than for commuting. This activity might entail that they take a journey with the bike, ride slower, or ride on quieter roads, whereas on weekdays people try to arrive as fast as possible to their destination. The decrease before 6:00 might also be a caused by the fact that people want to get to their destination quickly, but in the evening they have less time pressure and therefore ride slower or take other roads. The increased usage after midnight might be due to the decreased, or absent, public transport availability at that hour of the day. This scenario might prompt people to ride longer distances with the bikes as it might be the fastest or only way to get home without taking a taxi. For example, in Boston the trip duration increases after midnight. The subway in Boston is not open 24 hours and thus is closed between midnight and 1:00. In London, only a small number of tube lines run 24 hours. In New York and Chicago where the trip duration is not much different compared to the rest of the day, the subway runs 24 hours in New York, and the main subway lines run 24 hours in Chicago. The exception is Washington, DC. It did not experience an increase in trip duration despite the fact that its subway does not run 24 hours. However, all cities have a night schedule on weekends, which would not be congruent with this assumption, because some cities still have longer trip durations in the early morning hours on the weekend (Boston Discovery Guide, 2020; NYCGOycgo, 2020; Transit Chicago, 2020; Visitlondon.com, 2020; Washington.org, 2020).
The different trip durations on the weekend in December compared to weekdays might be due to cold weather conditions that impede people to take longer journeys. Additionally, journeys are probably mostly taken for transporting purposes which are typically shorter than recreational rides.

4.3 Influence of Weather Events

This section examines the influence of weather events on the usage of bike sharing. The factors considered are temperature (average, minimum, maximum), precipitation, and wind speed. Data with two different temporal resolutions was used for the analysis. The daily averages of these parameters is first analysed in correlation with the daily trip count number. The dataset for this part includes the entire data-period, which is mostly one year. However, the weather dataset was not complete for one year and contained missing values, which resulted in fewer data-points. In the second part, hourly data is analysed, and this dataset spans one month. A negative binomial regression model was used in both cases as it is an appropriate model for overdispersed count data. Each city was modelled separately. To account for seasonal variability in the model with daily aggregates, a season variable was added as well as the number of active stations. The latter accounts for changes in the number of stations that might have impacted the number of rides. For the model based on hourly data, a variable was created for every three hours (e.g. 3–6:00) to include hourly variability. A weekend variable was added to both models to account for the differences between weekday and weekends. The same models were used to calculate the impact of the different weather variables on the trip duration.

4.3.1 Number of Rides

Table 4.6 lists the model output for the daily data. At the bottom of the table is the number of observations. Not all datasets have the same number of observations as the different weather datasets differed in terms of completeness. The variable "Weekend" was significant for all nine cases; it had a negative coefficient, meaning that less rides were taken on weekends than on weekdays. Wind-speed and precipitation were significant in five of the cities and had a negative effect on the number of trips. The maximum temperature was the variable which explains the most out of the three temperature variables, and it was significant for four cities and had a positive impact. Minimum temperature was significant in two cases with a negative impact, and the average temperature was significant in three cases, which indicates that the number of rides increased with a higher average temperature.

Table 4.6 – Influence of Weather on the Daily Number of Rides

	Dependent variable: Number of Rides								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Avg. Temp.	0.022	0.017	0.064^{***}	0.026	0.070^{**}	0.036	−0.025	−0.020	0.073^{*}
	(0.015)	(0.018)	(0.023)	(0.016)	(0.028)	(0.056)	(0.051)	(0.059)	(0.042)
Min. Temp.	−0.014	0.001	−0.015	-0.027^{**}	-0.064^{***}	−0.038	0.002	−0.026	−0.001
	(0.009)	(0.012)	(0.012)	(0.011)	(0.018)	(0.028)	(0.027)	(0.035)	(0.021)
Max. Temp.	0.019^{**}	0.019^{**}	−0.003	0.032^{***}	0.023	0.040	0.068^{**}	0.073^{**}	−0.021
	(0.008)	(0.010)	(0.013)	(0.008)	(0.015)	(0.032)	(0.028)	(0.034)	(0.026)
Precipitation	-0.017^{***}	-0.016^{***}	-0.016^{***}	-0.014^{***}	−0.003	−0.001	0.0004	−0.0004	0.022^{***}
	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)	(0.007)	(0.006)	(0.003)	(0.008)
Wind speed	-0.007^{***}	-0.010^{***}	-0.017^{***}	-0.008^{**}	−0.005	−0.007	−0.015	−0.008	-0.018^{***}
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.005)	(0.011)	(0.007)	(0.006)
Weekend	-0.239^{***}	-0.457^{***}	-0.440^{***}	-0.195^{***}	-0.424^{***}	-0.148^{*}	-0.538^{***}	-0.608^{***}	-0.320^{***}
	(0.028)	(0.042)	(0.040)	(0.034)	(0.042)	(0.077)	(0.084)	(0.087)	(0.076)
Constant	9.088^{***}	7.316^{***}	4.018^{**}	10.333^{***}	43.772^{**}	3.718^{***}	−0.398	4.927^{***}	6.008^{***}
	(2.161)	(0.343)	(1.680)	(1.475)	(20.656)	(0.211)	(10.176)	(0.226)	(0.551)
Observations	324	217	293	308	129	149	203	162	195

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.7 provides the output for the model with the hourly data, and the number of observations are more similar to each other compared to the ones from the daily output. Moreover, the datasets for New York, Boston, Washington, DC, and London are complete with no missing values. The weekend value was significant in seven cities. In each city, the weekend usage was less than the weekday usage, which matches the previous findings.

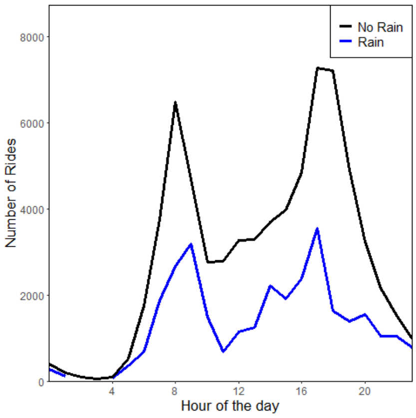
Table 4.7 – Influence of Weather on the Hourly Number of Rides

	Dependent variable: Number of Rides								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Temp.	0.045^{***}	0.034^{***}	0.049^{***}	0.039^{***}	0.052^{***}	0.004	0.021^{***}	0.040^{***}	0.031^{***}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.009)	(0.006)	(0.005)	(0.004)
Precipitation	-0.197^{***}	-0.456^{***}	-0.161^{***}	-0.086^{**}	-0.579^{***}	-0.830^{***}	-0.470^{***}	-0.273^{***}	-0.275^{***}
	(0.022)	(0.047)	(0.022)	(0.036)	(0.124)	(0.123)	(0.055)	(0.048)	(0.069)
Wind speed	-0.006^{**}	−0.002	−0.004	-0.011^{***}	-0.009^{*}	−0.005	−0.005	−0.002	0.005
	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)	(0.005)	(0.003)	(0.005)
Weekend	-0.222^{***}	-0.135^{***}	−0.071	-0.157^{***}	−0.012	0.086	-0.283^{***}	-0.415^{***}	-0.331^{***}
	(0.048)	(0.049)	(0.058)	(0.054)	(0.045)	(0.059)	(0.060)	(0.045)	(0.047)
Constant	7.409^{***}	5.446^{***}	5.661^{***}	5.609^{***}	6.779^{***}	3.336^{***}	6.150^{***}	4.849^{***}	4.092^{***}
	(0.110)	(0.099)	(0.128)	(0.158)	(0.136)	(0.138)	(0.129)	(0.101)	(0.087)
Observations	744	744	674	744	744	677	599	451	378

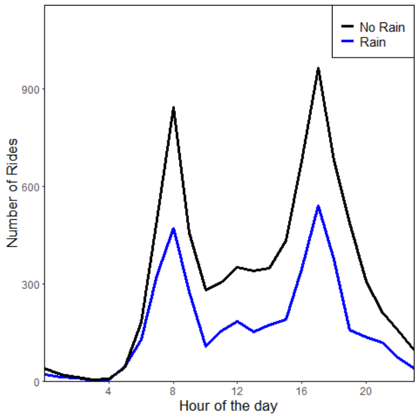
Note: *p<0.1; **p<0.05; ***p<0.01

Regarding the other values, wind speed was significant in only three cities, compared to five cities in the daily model. It still has a negative coefficient, which indicates less rides were taken with a higher wind speed. Precipitation was significant in all nine cities, compared to five in the daily model. It also had a negative coefficient, which indicates that rain negatively impacted the number of rides. For the hourly model, only the average temperature for one hour was available. Temperature was significant in eight cities with a positive coefficient, indicating that more rides were taken if the temperature was higher.

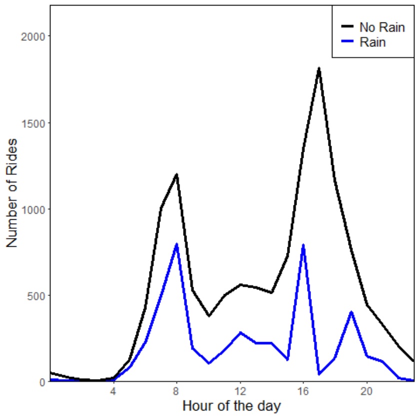
As precipitation in the hourly model was significant in all cases, Figure 4.12 demonstrates the difference in usage between rainy and dry hours in May 2019. The “No Rain” dataset contained all hours when there was no precipitation, the “Rain” dataset contained all entries with rain. The average number of rides was calculated for each hour, similar to the graphs in Section 4.2.2.Washington, DC’s graph has missing values, because for that particular hour no precipitation occurred for the whole month. The graphs indicate that the usage decreased if there was rain; however, the “Rain” graph still has the same general shape as the “No Rain” graph. In Trondheim, the difference in the morning was not significant, but it increased after 8:00. Edinburgh had similar graphs, which was unexpected, and it had a strong negative coefficient in the calculated model. The other cities had different graphs for “Rain” and “No Rain.” Overall, a difference is in the graphs is evident for the majority of the cities. While the shape of the rain graph primarily stays the same, it has less rides. The same plots were produced for weekends, but as the dataset for weekends is much smaller than the one for weekdays, the plots lack data and have many missing values, especially for the “Rain” graph. Additionally, Table 4.8 presents the average decrease per hour in rainy periods compared to dry periods. The average is 37.7%. All cities had a significant decrease in rides during rainy periods except for Edinburgh, which was only 3.53%. One reason might be that the precipitation data for Edinburgh was from a weather station in Leuchars which is approximately 50 km away. This distance might cause false precipitation values as precipitation in Edinburgh and Leuchars can be different. Another explanation could be that the rain in Edinburgh was less heavy than in the other cities, leading to a less hindering effect. A third explanation might be that people in Edinburgh are affected less by rain as they are more used to it. If this were true, a similar result would be expected for London; however, London had a decrease of almost 32%.



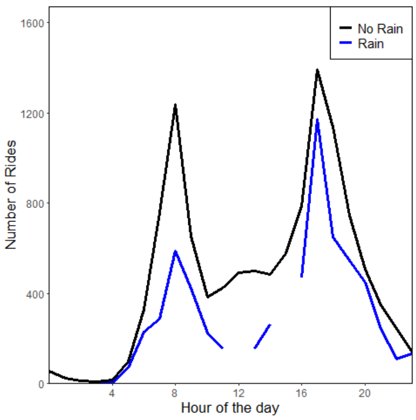
(a) New York



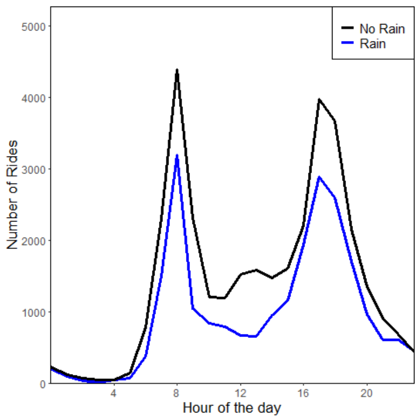
(b) Boston



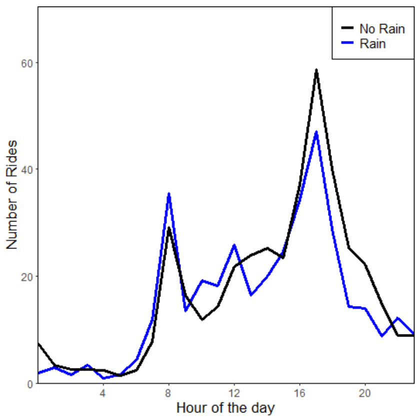
(c) Chicago



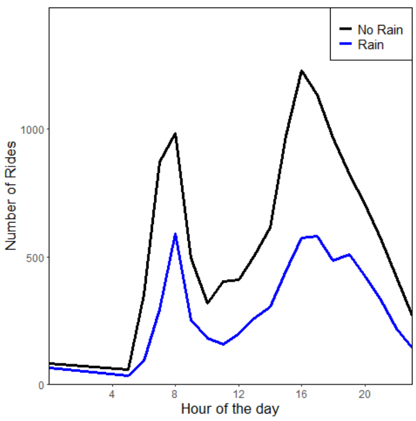
(d) Washington, DC



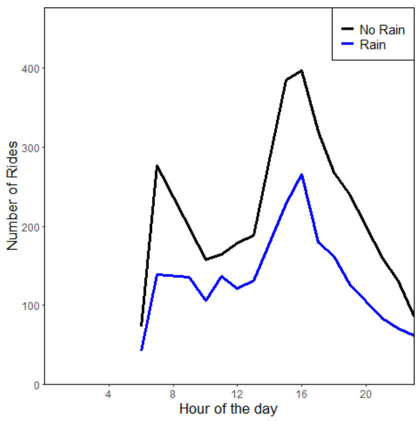
(e) London



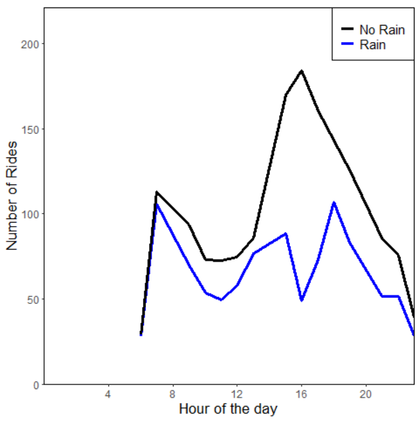
(f) Edinburgh



(g) Oslo**



(h) Bergen*



(i) Trondheim*

Figure 4.12 – Ride characteristics: rain versus no rain.

* The systems in Bergen and Trondheim are in service from 6:00 until midnight.

** The system in Oslo is in service from 5:00 until 1:00.

Table 4.8 – Percentage Decrease of Number of Rides During Rain

City	Avg. Decrease (%)
New York	48.72
Boston	43.11
Chicago	63.34
Washington, DC	32.37
London	31.98
Edinburgh	3.53
Oslo	48.30
Bergen	37.37
Trondheim	30.55

4.3.2 Trip Duration

Tables 4.9 and 4.10 present the effect of the different weather variables on the trip duration on daily and hourly basis. The daily analysis indicates an increased trip duration on weekends with a significant and positive coefficient in all nine cities, which agrees with the previous findings. Precipitation was significant in six cities and had a negative coefficient, indicating a negative impact of rain on the trip duration. Wind speed was significant in five cities, and it also had a negative coefficient. Considering the different temperature variables, the maximum temperature was the most meaningful variable, and it was significant in seven cities with a positive coefficient, indicating longer trip durations occur with higher temperatures. The average temperature was significant in three cities with a positive coefficient. The minimal temperature was significant in only four cities and had a negative coefficient. The hourly model was more consistent regarding significance. Similar to the daily model, the weekend was significant in all nine cities and had a positive coefficient. This result is congruent with the findings from Figures 4.9, 4.10, and 4.11, which indicate that the average trip on weekends was longer than the average trip on a weekday. Precipitation was significant in eight cities, and it had a negative impact on the trip duration for seven and a positive impact for one. Temperature was significant in each city with a positive coefficient in all cases, which leads to the assumption that higher temperatures lead to longer bike trips. Alternatively, wind speed was significant in only two cities, with one positive and one negative coefficient.

Table 4.9 – Influence of Weather on the Daily Average Trip Duration

	Dependent variable: Trip Duration								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Avg. Temp.	−0.004	0.012^{*}	0.007	0.013^{*}	0.016	−0.040	0.008	0.018	-0.030^{*}
	(0.005)	(0.007)	(0.010)	(0.007)	(0.017)	(0.029)	(0.008)	(0.019)	(0.017)
Min. Temp.	0.001	-0.013^{***}	−0.008	-0.011^{**}	−0.016	0.003	-0.008^{*}	−0.010	-0.015^{*}
	(0.003)	(0.005)	(0.005)	(0.005)	(0.011)	(0.014)	(0.004)	(0.011)	(0.009)
Max. Temp.	0.013^{***}	0.015^{***}	0.015^{***}	0.007^{**}	0.010	0.041^{**}	0.011^{***}	0.006	0.037^{***}
	(0.003)	(0.004)	(0.005)	(0.004)	(0.009)	(0.017)	(0.004)	(0.011)	(0.010)
Precipitation	-0.003^{***}	-0.003^{***}	-0.002^{**}	-0.003^{***}	−0.001	−0.001	-0.002^{**}	−0.001	-0.010^{***}
	(0.0005)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)
Wind speed	-0.001^{*}	−0.002	-0.005^{***}	-0.005^{***}	-0.005^{***}	−0.001	−0.002	-0.005^{**}	0.003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Weekend	0.157^{***}	0.266^{***}	0.309^{***}	0.276^{***}	0.250^{***}	0.288^{***}	0.219^{***}	0.174^{***}	0.234^{***}
	(0.009)	(0.017)	(0.016)	(0.015)	(0.025)	(0.040)	(0.013)	(0.028)	(0.030)
Constant	6.510^{***}	6.507^{***}	6.537^{***}	6.639^{***}	6.727^{***}	6.935^{***}	6.242^{***}	6.293^{***}	6.228^{***}
	(0.020)	(0.035)	(0.034)	(0.036)	(0.061)	(0.093)	(0.026)	(0.071)	(0.059)
Observations	324	217	293	308	129	149	203	162	195
Note:									*p<0.1; **p<0.05; ***p<0.01

Table 4.10 – Influence of Weather on the Hourly Average Trip Duration

	Dependent variable: Trip Duration								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Temp.	0.019^{***}	0.026^{***}	0.026^{***}	0.017^{***}	0.014^{***}	0.024^{***}	0.030^{***}	0.021^{***}	0.015^{***}
	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)	(0.007)	(0.004)	(0.003)	(0.002)
Precipitation	-0.032^{***}	-0.048^{*}	-0.030^{***}	-0.071^{***}	-0.123^{**}	0.041	0.058^{*}	-0.073^{***}	-0.053^{*}
	(0.006)	(0.026)	(0.010)	(0.019)	(0.052)	(0.088)	(0.031)	(0.025)	(0.032)
Wind speed	−0.001	−0.002	-0.005^{***}	−0.002	0.002	−0.002	0.005^{*}	0.001	−0.004
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Weekend	0.126^{***}	0.198^{***}	0.314^{***}	0.140^{***}	0.219^{***}	0.312^{***}	0.260^{***}	0.130^{***}	0.169^{***}
	(0.012)	(0.030)	(0.029)	(0.028)	(0.019)	(0.048)	(0.034)	(0.024)	(0.023)
Constant	6.431^{***}	6.574^{***}	6.627^{***}	6.671^{***}	6.809^{***}	7.143^{***}	6.127^{***}	6.224^{***}	6.266^{***}
	(0.027)	(0.061)	(0.063)	(0.084)	(0.057)	(0.115)	(0.073)	(0.055)	(0.043)
Observations	744	744	674	744	744	677	599	451	378
Note:									*p<0.1; **p<0.05; ***p<0.01

4.4 Influence of Infrastructure

This section focuses on the impact of infrastructure on the usage of bike sharing. The analysis is station-based and assesses the impact that the infrastructure has on the usage of the different stations within a city. Different parameters were calculated for each station as well as the total arrivals and departures over the course of one month. Each bike station was associated with its surrounding station area as the basis for the calculations, which included the length of different street types inside the station area and the number of public transport stations. The last variable was the altitude of the station. The analysis is divided into arrivals and departures to determine if a difference exists between the two. A negative binomial regression model was used, because it is a suitable model for overdispersed count data. For this analysis many factors were used, which could also be aggregated, such as road types or public transport station types. The different models are presented, and some have more aggregated data than others. Table 4.11 presents the descriptive statistics for the different cities.

Table 4.11 – Descriptive Statistics for May 2019: Stations

City	Ø Rides per Station	Max. Rides	Min. Rides	SD	Ø Altitude	Max. Altitude	Min. Altitude	SD
New York	4,742	29,107	3	4,472	12.5	48.1	0.3	8.95
Boston	1,595	11,052	9	1,611	8.8	50.4	0.9	8.96
Chicago	1,171	12,476	2	1,502	181.5	189.0	176.0	1.96
Washington, DC	1,169	11,965	3	1,468	51.4	150.7	1.6	42.56
London	2,436	15,657	120	1,690	19.1	52.2	-0.2	11.08
Edinburgh	297	966	8	240	60.4	101.6	5.0	26.48
Oslo	2,543	11,434	115	1,862	37.0	131.3	1.6	30.60
Bergen	2,552	6,993	323	1571	9.7	52.2	1.0	10.99
Trondheim	1,794	5,052	247	1,070	20.0	114.0	2.1	21.90

The table includes the descriptive statistics for the departures in May 2019. New York had the highest average rides per station in one month with almost 5,000. Edinburgh had the lowest with 297 rides. The smaller systems of Oslo, Bergen, and Trondheim had more rides per station than Boston and Washington. The maximum and minimum values indicate the stations with most and least rides. Many cities had stations with less than 10 rides per month. In New York, the most used station had over 29,000 departures in one month, and the least used station had three departures. The number of rides for the busiest stations in Boston, Chicago, Washington, London, and Oslo all had over 11,000 rides. Edinburgh was the smallest system with its busiest station only having 966 rides in one month. Based on altitude, the most striking values are the maximum, minimum, and standard deviation. Significant differences between the minimum and maximum are likely attributed to a hillier city with stations located at both high and low altitudes. Small differences indicate that the city is generally flat. One might expect that altitude more heavily influences cities with greater minimum and maximum differences compared to a flat city. The table demonstrates that Washington, DC, Oslo, Edinburgh, and Trondheim have the greatest fluctuations in altitude with 95–150 m difference between the highest and the lowest station. The other cities' difference ranges 20–50 m.

4.4.1 Influence on Trip Departures

For the model in Table 4.12, the different road types and public transport stations were aggregated to their parent category and then inserted into the model. The variable "Altitude" was significant in eight cities, having a negative impact seven times and a positive impact once. The variables of "Major Roads" and "Minor Roads" were both significant in five cities with a positive coefficient. "Highway Links" and "Very Small Roads" had a negative impact four times, and "Roads Unsuitable for Cars" had a positive impact five times. "Public Transport Stations" had a significant value three times, two were positive impacts and one was negative.

In the next model, the categories of "Major Roads," "Minor Roads," and "Unsuitable for Cars" were split as they were the most significant in the previous model. This process help determine if certain types of roads in these categories are more significant than others. For the category of "Cycleway," one might expect more rides at stations with cycling infrastructure like bike paths. The "Major Roads" category was also split, because the range from "Motorway" to "Tertiary Roads" is quite different; for example, bikes are prohibited on motorways and tertiary roads are roads with low to moderate traffic (OpenStreetMap, 2020b). "Very Small Roads" was also split into road types that favour cycling such as "Residential," "Living Street," and "Pedestrian." Finally, the public transport was split into train and bus as these are the two most common public transport stations, and the ferry and tram categories were eliminated. The categories that were not split were "Highway Links" and "Very Small Roads," because the latter is mostly road types that are unsuitable for cycling like

Table 4.12 – Infrastructure: Aggregated Model for Departures

	Dependent variable:								
	Departures								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Public Transport Stations	0.023***	0.009	0.005	0.015***	0.015	−0.009	0.001	−0.022	−0.021*
	(0.007)	(0.015)	(0.009)	(0.005)	(0.023)	(0.007)	(0.013)	(0.015)	(0.012)
Altitude	−0.016***	−0.046***	−0.019***	0.003*	0.001	−0.009***	−0.030***	−0.008**	−0.146***
	(0.004)	(0.007)	(0.001)	(0.002)	(0.004)	(0.001)	(0.006)	(0.003)	(0.027)
Major Roads	0.0003***	0.0002**	0.0003***	0.0001***	0.0003	−0.00002	−0.00001	0.0001	0.0003***
	(0.00003)	(0.0001)	(0.00005)	(0.00003)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Minor Roads	0.00004*	0.0001**	0.0001	0.00005***	0.0001	−0.00004	−0.00003	0.0001**	0.0002***
	(0.00002)	(0.0001)	(0.00005)	(0.00002)	(0.0001)	(0.00004)	(0.0001)	(0.00004)	(0.0001)
Highway Links	−0.0003**	−0.0001	−0.0005***	−0.001***	0.002	0.00001	−0.00003	−0.0004	−0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.014)	(0.0001)	(0.0001)	(0.0003)	(0.0001)
Very Small Roads	−0.0004***	0.0002**	−0.0001***	0.00002	−0.0001	−0.0002**	−0.00003	−0.0001	−0.00005
	(0.0001)	(0.0001)	(0.00004)	(0.00003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Unsuitable for Cars	0.00004***	0.0002***	−0.00002	0.0001***	0.00005	0.0001***	0.0001	0.0001	0.0002***
	(0.00001)	(0.00003)	(0.00002)	(0.00001)	(0.0001)	(0.00002)	(0.00004)	(0.00004)	(0.00002)
Constant	7.173***	5.533***	6.763***	6.201***	4.170***	7.388***	7.339***	6.569***	31.714***
	(0.139)	(0.315)	(0.227)	(0.100)	(0.471)	(0.256)	(0.292)	(0.361)	(4.929)
Observations	793	268	553	780	85	247	80	54	600
Note:									*p<0.1; **p<0.05; ***p<0.01

tracks or service roads.

Table 4.13 presents the results of the model that is described above. The most significant attribute for all the cities was "Altitude," which was significant in six. In all of these it had a negative coefficient and thus a negative impact on the number of departures. Furthermore, the major five cities had a positive correlation with "Primary," "Secondary," and "Tertiary" roads as 13 of the possible 15 observations were significant. The smaller cities had no significances out of the possible 12 observations. Similarly, the road types of "Unclassified," "Residential," "Living Street," and "Pedestrian" had 13 significances out of 20 observations for the major cities, and 12 were positive, and one was negative. The small cities had three significances out of 15 observations. Two of which were positive, and one was negative. Regarding "Cycleways," only four cities reported a significance, and one was negative. This finding was unexpected as a greater positive impact was anticipated. The other roads from the category "Unsuitable for Cars" had inconsistent significances and coefficients for the different cities. Most of them had a positive impact if they were significant. The most significant were "Footway" and "Bridleway." The latter was significant three out of four times with a positive coefficient. The variable of "Steps" was significant twice with a positive impact. This finding was noteworthy, because as one might expect steps to have a negative impact since they are typically a hindrance for bikes.

Regarding public transport, “Train Stations” had a greater impact on the departure rates than “Bus Stations.” However, five significances out of nine were found, and four were positive, and one was negative. All four positive significances were in the five major cities. The result might be explained by the fact that Oslo, Bergen, and Trondheim do not have a subway, which reduces the number of train stations. Of the six cities with a subway, only Chicago was not significant, which indicates that subways had an impact on the usage of bike sharing.

Table 4.13 – Infrastructure: Extended Model for Departures

	Dependent variable:								
	Departures								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Train Station	0.112***	0.201***	0.095	0.240***	0.093***	−1.769**	−0.094	−0.452	−0.093
	(0.033)	(0.075)	(0.080)	(0.075)	(0.034)	(0.776)	(0.092)	(0.283)	(0.247)
Bus Station	−0.007	−0.013	−0.031**	−0.016	0.014***	0.021	−0.018**	0.006	−0.035
	(0.008)	(0.016)	(0.013)	(0.010)	(0.005)	(0.026)	(0.009)	(0.017)	(0.023)
Altitude	−0.016***	−0.051***	−0.148***	−0.020***	0.0004	0.0002	−0.007***	−0.027***	−0.004
	(0.003)	(0.007)	(0.027)	(0.001)	(0.002)	(0.005)	(0.002)	(0.007)	(0.005)
Motorway	−0.0003***	−0.0002	−0.0002	−0.0001				0.048	
	(0.0001)	(0.0001)	(0.0002)	(0.0001)				(0.060)	
Trunk	0.001***	−0.001**	−0.0003	0.0002	0.0001*		−0.0001	−0.001***	0.0003
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.00005)		(0.0001)	(0.0002)	(0.0003)
Primary	0.0002***	0.0002*	0.00001	0.0003***	0.0002***	0.0003	−0.0001	−0.0003	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00004)	(0.0004)	(0.0001)	(0.0002)	(0.0002)
Secondary	0.0003***	0.0002*	0.0003***	0.0003***	0.0002***	0.0004	−0.00002	−0.00002	0.00001
	(0.00005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0005)	(0.0001)	(0.0002)	(0.0002)
Tertiary	0.00002	0.001***	0.0003***	0.0002***	0.0001***	0.0002	0.0001	0.00002	−0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00004)	(0.0002)	(0.0001)	(0.0002)	(0.0003)
Unclassified	0.001***	0.001*	0.001	−0.001*	0.0001*	0.00002	−0.0002***	−0.00001	−0.0001
	(0.0001)	(0.001)	(0.001)	(0.0003)	(0.00003)	(0.0003)	(0.0001)	(0.0001)	(0.0002)
Residential	−0.00000	0.0002***	0.0002**	0.00004	0.00003	−0.00002	0.00005	−0.0002*	0.0001
	(0.00004)	(0.0001)	(0.0001)	(0.00005)	(0.00002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Living Street	0.006***	0.0002	0.002	0.005***	0.001**		−0.0002	−0.0002	−0.0001
	(0.002)	(0.001)	(0.003)	(0.001)	(0.0003)		(0.0003)	(0.0004)	(0.007)
Pedestrian	0.0001***	0.0001	0.0005***	0.001**	0.0001**	0.0004	0.0002*	0.00001	0.0001
	(0.00003)	(0.0002)	(0.0002)	(0.0002)	(0.00004)	(0.0003)	(0.0001)	(0.0002)	(0.0001)
Highway Links	0.0001	0.0001	0.0001	−0.0002*	−0.001***	0.015	0.0002	0.0002	−0.0003
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.017)	(0.0002)	(0.0002)	(0.0003)
Very Small Roads	−0.0004***	0.0002***	−0.00001	−0.0001***	0.00002	−0.0001	−0.00002	−0.0001	−0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.00004)	(0.00003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Bridleway	0.001*			0.002**	0.001***		0.001		
	(0.001)			(0.001)	(0.0002)		(0.004)		
Cycleway	0.001***	0.0003**	0.001***	−0.0003**	0.0001	0.00001	−0.0001	0.0002	−0.0001
	(0.0001)	(0.0002)	(0.0003)	(0.0001)	(0.00004)	(0.0004)	(0.0001)	(0.0001)	(0.0001)
Footway	−0.00001	0.0001***	0.0001***	−0.00003*	0.00003	0.0001	0.0001***	0.0003***	0.0001*
	(0.00001)	(0.00003)	(0.00003)	(0.00002)	(0.00002)	(0.0001)	(0.00002)	(0.0001)	(0.0001)
Path	−0.0001	−0.0004	−0.0002	−0.0005**	0.0001	0.0003	−0.0002	−0.0001	0.0001
	(0.0003)	(0.0004)	(0.0001)	(0.0002)	(0.0001)	(0.001)	(0.0001)	(0.0004)	(0.0002)
Steps	0.0002	0.005***	−0.001	0.005***	0.0002	−0.001	0.0001	0.0003	0.0004
	(0.0004)	(0.002)	(0.001)	(0.001)	(0.0002)	(0.001)	(0.0004)	(0.001)	(0.001)
Unknown		−0.011			−0.029				
		(0.007)			(0.018)				
Constant	7.265***	5.113***	32.019***	6.775***	6.351***	4.482***	6.972***	7.059***	6.544***
	(0.189)	(0.348)	(4.946)	(0.232)	(0.122)	(0.558)	(0.329)	(0.375)	(0.617)
Observations	793	268	600	553	780	85	247	80	54
Note:									*p<0.1; **p<0.05; ***p<0.01

4.4.2 Influence on Trip Arrivals

The same models were used to assess arrivals, which was the updated dependent variable. An aggregated as well as an extended model were created. The goal was to determine if people favour different infrastructure for their arrival stations compared to their departure stations. Table 4.14 presents the results of the aggregated model. Similar to the departures model, the most significant attribute was "Altitude" with a significant and negative impact in eight cities. The only city where it was not significant was Washington, DC, which is the city with the greatest altitude difference between the lowest and the highest stations. "Roads Unsuitable for Cars" was significant for six cities, and it had a positive coefficient for each one. "Major Roads" and "Minor Roads" were both significant five times with a positive impact. "Highway Links" and "Very Small Roads" were negatively associated three times each. "Public Transport Stations" was significant for only two cities. The major five cities had more significant variables than the smaller cities.

Table 4.14 – Infrastructure: Aggregated Model for Arrivals

	Dependent variable:								
	Arrivals								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Public Transport Stations	0.022***	0.012	0.009	0.015***	0.010	−0.009	0.001	−0.022	−0.018
	(0.007)	(0.015)	(0.008)	(0.005)	(0.022)	(0.007)	(0.013)	(0.014)	(0.012)
Altitude	−0.019***	−0.055***	−0.022***	0.001	−0.008**	−0.017***	−0.045***	−0.024***	−0.148***
	(0.004)	(0.007)	(0.001)	(0.002)	(0.004)	(0.002)	(0.006)	(0.003)	(0.027)
Major Roads	0.0003***	0.0002**	0.0003***	0.0002***	0.0002	0.00002	0.00001	0.00000	0.0003***
	(0.00003)	(0.0001)	(0.00005)	(0.00003)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Minor Roads	0.00004**	0.0001**	0.00001	0.0001***	0.0001	0.00001	−0.00003	0.0001**	0.0002***
	(0.00002)	(0.0001)	(0.00004)	(0.00002)	(0.0001)	(0.00004)	(0.0001)	(0.00004)	(0.0001)
Highway Links	−0.0003**	−0.0001	−0.0004***	−0.001***	0.005	0.0001	−0.00004	−0.0001	−0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.013)	(0.0002)	(0.0001)	(0.0003)	(0.0001)
Very Small Roads	−0.0004***	0.0002**	−0.0001***	0.00001	−0.0001	−0.0001*	−0.0001	−0.0001	−0.00004
	(0.0001)	(0.0001)	(0.00003)	(0.00003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Unsuitable for Cars	0.00004***	0.0002***	−0.00001	0.0001***	0.00003	0.0001***	0.00003	0.0001***	0.0002***
	(0.00001)	(0.00003)	(0.00002)	(0.00001)	(0.0001)	(0.00002)	(0.00004)	(0.00004)	(0.00002)
Constant	7.189***	5.626***	6.768***	6.090***	4.813***	7.225***	7.532***	6.552***	32.079***
	(0.138)	(0.310)	(0.217)	(0.104)	(0.460)	(0.267)	(0.281)	(0.339)	(4.919)
Observations	793	268	553	780	85	247	80	54	600
Note:									*p<0.1; **p<0.05; ***p<0.01

The results of the extended model in Table 4.13 sindicate that "Altitude" was again the factor that was significant in most cases. In this model, it was significant for Washington, DC but not for London. However, it was again significant eight times with a negative impact. Furthermore, the results resemble the ones from the departures as "Primary," "Secondary," and "Tertiary Roads" were the most favoured road types for cycling as they had between three and five positive and significant

coefficients. “Pedestrian” and “Living Street” were significant four and three times, respectively, with positive coefficients. “Cycleways” were significant four times and had a positive impact on the station usage. The remaining street types did not have a clear result and had few significant variables that often contradicted with each other. The unexpected positive correlation with “Steps” and “Bridleways” occurred again, which are difficult to explain. For this model, the major five cities again had more significant values than the smaller four.
Regarding public transport the pattern was the same as that in the model for the departures. “Train Stations” were more significant than “Bus Stations” as five cities reported a significance with train stations and only three with bus stations. Of the five cities that had a correlation with train stations, four were positive, and the negative one was Edinburgh. All of the five cities have a subway. Of the remaining four cities for which train stations were not significant, only Chicago has a subway. This finding indicates that subway stations benefit the usage of bike sharing.

In general, altitude had the greatest impact on the usage of bike sharing, and it had a negative impact on both departures and arrivals in most of the cities. This finding is understandable for the arrivals as one can imagine that people do not want to ride uphill to a bike station. Most BSS usually use heavier bikes with limited gears. The fact that departures were also affected negatively by altitude was not expected as altitude should not be a hinderance if the slope is negative, unless the brakes of the bikes are in poor condition. Furthermore, subways have an impact on the bike sharing usage too as it was significant for five of the six cities that have one. The positive correlation is understandable as people might use bike sharing paired with the subway. The negative correlation for Edinburgh is harder to explain. One reason might be that people use bike sharing as a result of a lack of public transport connectivity; therefore, a subway station would have a negative impact on the bike sharing station since it is in an area with good public transport connectivity.
Regarding road types, bike sharing users seem to prefer larger roads for their journeys. Instead, one might expect them to use cycle paths and smaller, quieter roads. One possible explanation might be because the most used stations are most likely in the city centre where there are more major roads than smaller, quieter roads. Additionally, it might be that the major roads are the fastest option, and as bike sharing is mostly used for commuting purposes, this tendency might favour major roads, compared to smaller roads.

Table 4.15 – Infrastructure: Extended Model for Arrivals

	Dependent variable:								
	Arrivals								
	New York	Boston	Chicago	Washington, DC	London	Edinburgh	Oslo	Bergen	Trondheim
Train Station	0.115***	0.228***	0.074	0.227***	0.111***	−1.768**	−0.118	−0.450	−0.049
	(0.033)	(0.073)	(0.080)	(0.072)	(0.035)	(0.756)	(0.097)	(0.276)	(0.233)
Bus Stations	−0.009	−0.011	−0.028**	−0.012	0.014**	0.026	−0.018**	0.009	−0.021
	(0.008)	(0.016)	(0.013)	(0.009)	(0.006)	(0.025)	(0.009)	(0.017)	(0.022)
Altitude	−0.019***	−0.061***	−0.146***	−0.023***	−0.002	−0.010**	−0.016***	−0.041***	−0.023***
	(0.003)	(0.007)	(0.027)	(0.001)	(0.002)	(0.005)	(0.002)	(0.007)	(0.005)
Motorway	−0.0003***	−0.0002*	−0.0003	−0.0001				0.046	
	(0.0001)	(0.0001)	(0.0002)	(0.0001)				(0.059)	
Trunk	0.001***	−0.001***	−0.0004	0.0002	0.0001		−0.00001	−0.001**	0.0001
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.00005)		(0.0001)	(0.0002)	(0.0003)
Primary	0.0002***	0.0002	0.00001	0.0003***	0.0002***	0.0003	−0.00005	−0.0003	0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00004)	(0.0004)	(0.0001)	(0.0002)	(0.0002)
Secondary	0.0003***	0.0002*	0.0004***	0.0003***	0.0003***	0.0002	−0.0001	0.0001	−0.00003
	(0.00005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0005)	(0.0001)	(0.0002)	(0.0001)
Tertiary	0.00002	0.0005***	0.0003***	0.0002***	0.0001***	0.0002	0.0001	−0.0001	−0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00005)	(0.0002)	(0.0001)	(0.0002)	(0.0003)
Unclassified	0.001***	0.001	0.001	−0.0005*	0.0001**	−0.0002	−0.0001	−0.0001	−0.00001
	(0.0001)	(0.001)	(0.001)	(0.0003)	(0.00003)	(0.0003)	(0.0001)	(0.0001)	(0.0002)
Residential	−0.00000	0.0002***	0.0002***	0.00000	0.00003	0.00001	0.0001	−0.0002	0.0001
	(0.00004)	(0.0001)	(0.0001)	(0.00005)	(0.00002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Living Street	0.006***	0.0001	0.002	0.005***	0.001***		−0.00003	−0.0001	−0.004
	(0.002)	(0.001)	(0.003)	(0.001)	(0.0004)		(0.0003)	(0.0004)	(0.006)
Pedestrian	0.0001***	0.0001	0.001***	0.001**	0.0001***	0.0005	0.0002	−0.00003	0.0001
	(0.00003)	(0.0002)	(0.0002)	(0.0002)	(0.00004)	(0.0003)	(0.0001)	(0.0002)	(0.0001)
Highway Links	0.0001	0.0001	0.0001	−0.0001	−0.001***	0.017	0.0002	0.0001	−0.0001
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.016)	(0.0002)	(0.0002)	(0.0003)
Very Small Roads	−0.0004***	0.0002***	0.00001	−0.0001***	0.00001	−0.0001	0.00001	−0.0001	0.00003
	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.00003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Bridleway	0.001*			0.002**	0.001***		−0.0003		
	(0.001)			(0.001)	(0.0002)		(0.004)		
Cycleway	0.001***	0.0003**	0.001***	−0.0002	0.0001	−0.0001	0.00001	0.0003*	0.0001
	(0.0001)	(0.0002)	(0.0003)	(0.0001)	(0.00004)	(0.0004)	(0.0001)	(0.0001)	(0.0001)
Footway	−0.00001	0.0001***	0.0001***	−0.00003*	0.00003	0.0001	0.0001***	0.0002***	0.0001
	(0.00001)	(0.00003)	(0.00003)	(0.00002)	(0.00002)	(0.0001)	(0.00002)	(0.0001)	(0.0001)
Path	−0.00004	−0.0004	−0.0002	−0.0004*	0.0001	0.0002	−0.0002	0.00001	0.0001
	(0.0003)	(0.0004)	(0.0001)	(0.0002)	(0.0001)	(0.001)	(0.0001)	(0.0004)	(0.0002)
Steps	0.0003	0.005***	−0.001	0.004***	0.0001	−0.001	0.0002	0.0003	0.001
	(0.0004)	(0.002)	(0.001)	(0.001)	(0.0002)	(0.001)	(0.0004)	(0.001)	(0.001)
Unknown		−0.012*			−0.025				
		(0.007)			(0.018)				
Constant	7.290***	5.250***	31.541***	6.780***	6.329***	5.084***	6.857***	7.354***	6.278***
	(0.188)	(0.341)	(4.920)	(0.221)	(0.127)	(0.545)	(0.347)	(0.365)	(0.584)
Observations	793	268	600	553	780	85	247	80	54
Note:									*p<0.1; **p<0.05; ***p<0.01

4.5 Influence of Land Use

This section explores how the arrivals and departures rates differ between the land use categories. After each bike station was assigned to a certain land use category (see Sections 3.2.5 and 3.1.4), each ride log was coded with a land use category for both the start and the stop station. With this information, the average number of rides per hour for every land use type could be calculated over the course of one month, which was May 2019. As mentioned in Section 3.1.4, land use data for Chicago was not able to be retrieved. Therefore, only eight cities are displayed.

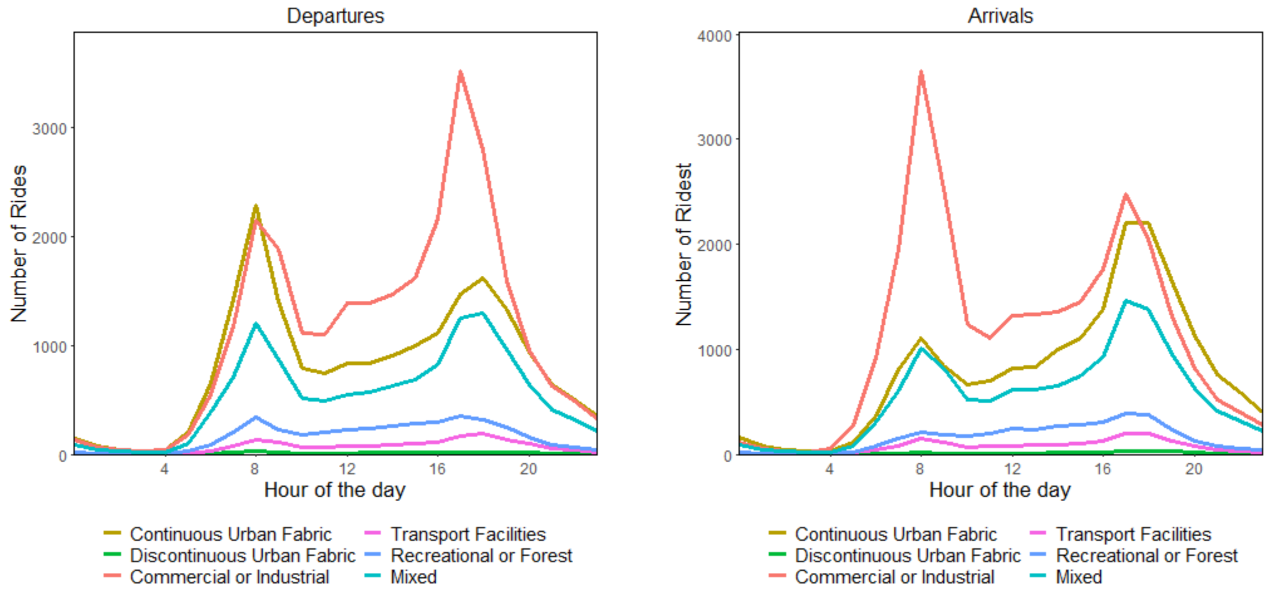
Figures 4.13 - 4.16 present the departure rates on the left and the arrival rates on the right for one day. The first noticeable fact is that not all the systems have the same dominant land use categories. The three US cities particularly differ in this respect. This finding might be because the base land use categories were from the European dataset CORINE whereas the datasets and categories from the US cities had to be transformed to the European categories (see Table 3.11). Furthermore, some cities do not have stations in all the categories. In Edinburgh some categories have missing values, which means that in that particular hour no departures or arrivals occurred in that land use category during the whole month. For the five European cities the most dominant land use types was “Continuous Urban Fabric” followed by “Discontinuous Urban Fabric.” For the US cities it was “Commercial or Industrial” for New York, “Mixed” for Boston, and ”Transport Facilities” for Washington, DC. Only the graphs for the weekdays are displayed as they are more meaningful in terms of differences between departures and arrivals. All graphs for the weekend had few differences between departures and arrivals. The weekend graphs were effectively the same for all cities with the number of rides slightly increasing over the day and a long peak in the afternoon before decreasing before midnight. All the land use categories shared this characteristic, but with varying numbers of rides. One difference between weekdays and the weekend was that for New York, London, and Washington, the land use category “Recreational or Forest” was greater relative to the other categories. This occurrence might be because people tend to use bike sharing more for recreational purposes on the weekends, and therefore more rides depart or end at stations belonging to the recreational land use category.

New York

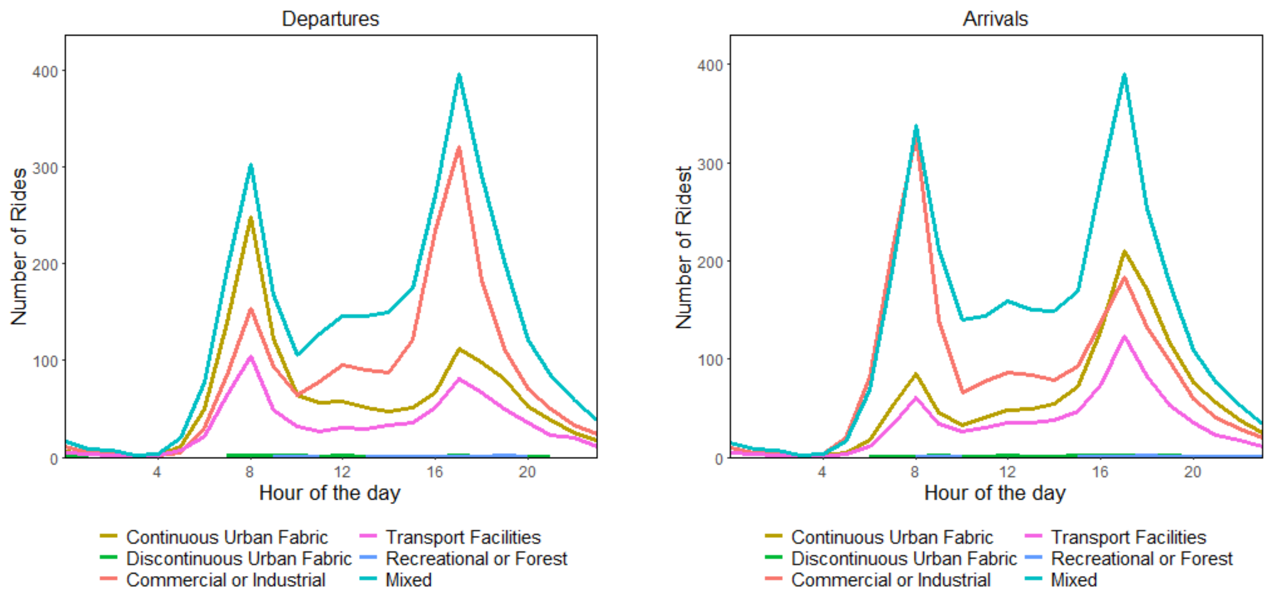
Some unique patterns are evident for New York. The most obvious is the difference between departures and arrivals for the land use category "Commercial or Industrial" as well as "Continuous Urban Fabric." The former has its departure peak in the late afternoon around 17:00, and the arrival peak is in the morning at 8:00. The characteristic of "Continuous Urban Fabric" is the opposite as the departures peak around 8:00, and the arrivals peak in the afternoon between 16–18:00. One possible explanation is that in the morning people pick up a bike close to their home and ride to work or towards the city where they return the bike. This occurrence would explain both the high number of departures in the morning for "Continuous Urban Fabric" as well as the high number of arrivals in the morning for "Commercial or Industrial." In the afternoon when people return from work, the flow goes in the other direction which also fits with the graphs. The other land use categories do not exhibit a striking difference between departures and arrivals.

Boston

As mentioned earlier, Boston's most frequent land use category was "Mixed," which does not allow any conclusions. However, "Commercial or Industrial" and "Continuous Urban Fabric" had a reasonable number of rides both for departures as well as for arrivals. The characteristics of both is similar to the ones from New York. In the morning, there were many departures linked to "Continuous Urban Fabric" and many arrivals for "Commercial or Industrial," and the opposite happened in the evening. The graphs are more distinct than those for New York. The category "Transport Facilities" also indicates a slightly different pattern between departures and arrivals. Whereas the greatest number of people depart from this kind of land use category in the morning, the most arrivals are in the evening. This category also demonstrates an opposite characteristic for departures and arrivals. A possible explanation might be that people use public transport or cars to drive inside the city and then use bike sharing for the last part of the commute. In the evening, the flow would be in the other direction as they return to their car or public transport line by bike.



(a) New York



(b) Boston

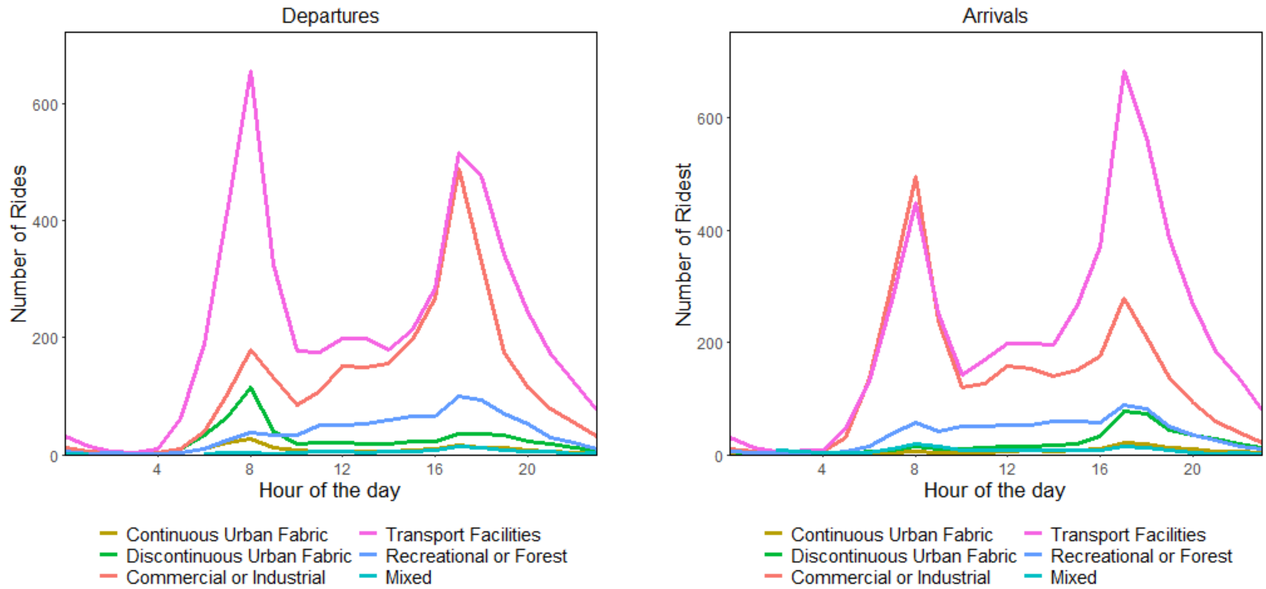
Figure 4.13 – Land use: New York and Boston.

Washington, DC

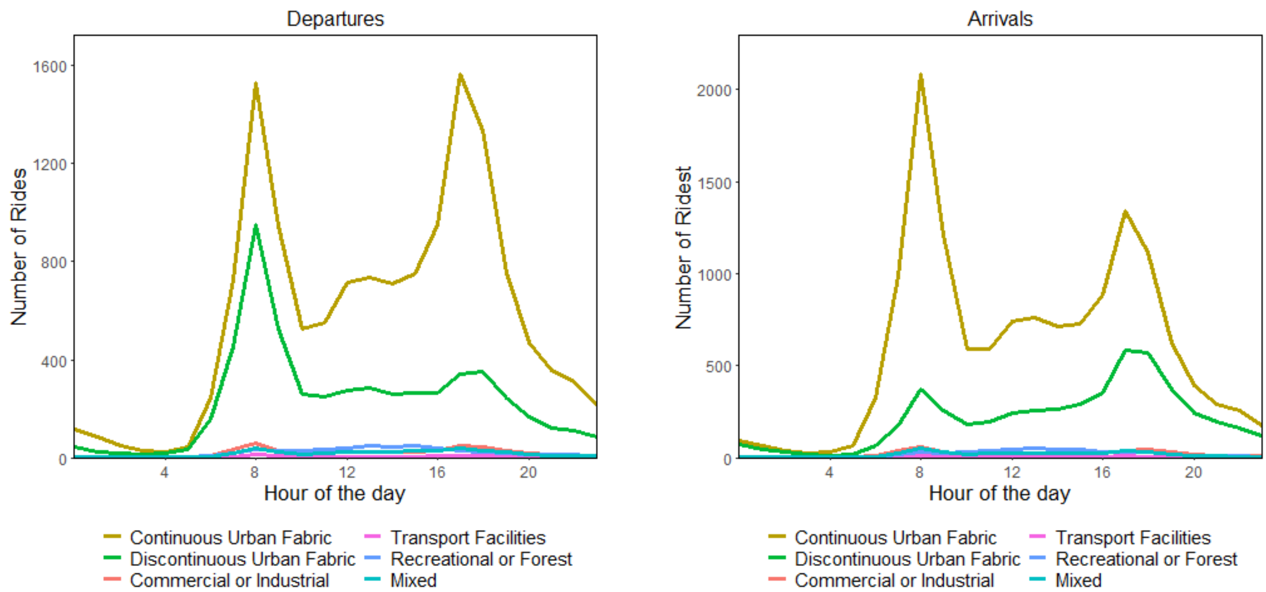
In Washington, the category with the most rides was “Transport Facilities,” both for departures and arrivals, and it was followed by “Commercial or Industrial.” The pattern for “Transport Facilities” had two peaks for both departures and arrivals, one in the morning and one in the afternoon. For the departures, the morning peak was larger, and for the arrivals, the afternoon peak was larger. “Commercial or Industrial” had the common pattern of a departure peak in the afternoon and an arrival peak in the morning. Another category with one peak for both departures and arrivals was “Discontinuous Urban Fabric.” Its departure peak was in the morning, and the arrival peak was in the evening. “Recreational or Forest” had a small afternoon peak for departures and a smaller peak for arrivals. Compared to the previously discussed cities, “Continuous Urban Fabric” had considerably few rides. The reason for this characteristic might be the same as the ones for the previous cities. The increase in the land use category “Recreational or Forest” in the evening might be caused by people using bikes to reach recreational activities after work.

London

London’s two dominant categories are “Continuous Urban Fabric” and “Discontinuous Urban Fabric.” The other four categories had few rides, and no pattern was visible. Regarding departures, “Continuous Urban Fabric” had two large peaks, one in the morning and one in the afternoon. “Discontinuous Urban Fabric” had a major peak in the morning whereas in the afternoon the peak was marginal. For the arrivals, “Continuous Urban Fabric” had a major peak in the morning and a smaller peak in the evening. The arrival peaks of “Discontinuous Urban Fabric” were smaller than the ones from the departures, but they were mostly the same size. This type of characteristic is difficult to interpret. One possible explanation might be that almost all stations are in these two land use categories. However, this does not explain the differences between the departures and the arrivals. One further explanation might be that the station areas have more mixed land use than the previous cities that were discussed, but by detailing only the dominant category, this was not displayed.



(a) Washington, DC



(b) London

Figure 4.14 – Land use: Washington and London.

Edinburgh

Edinburgh has one major category, "Continuous Urban Fabric." For the departures it had a major peak around 17:00 and two smaller peaks at 8:00 and around noon. The arrivals had two peaks, one at 8:00 and one around 17:00. Moreover, it had a midday peak from approximately noon to 14:00. The only other category which had some visible pattern was "Discontinuous Urban Fabric" with two small peaks for the departures (8:00 and 17:00) and one larger peak for the arrivals at 17:00.

Oslo

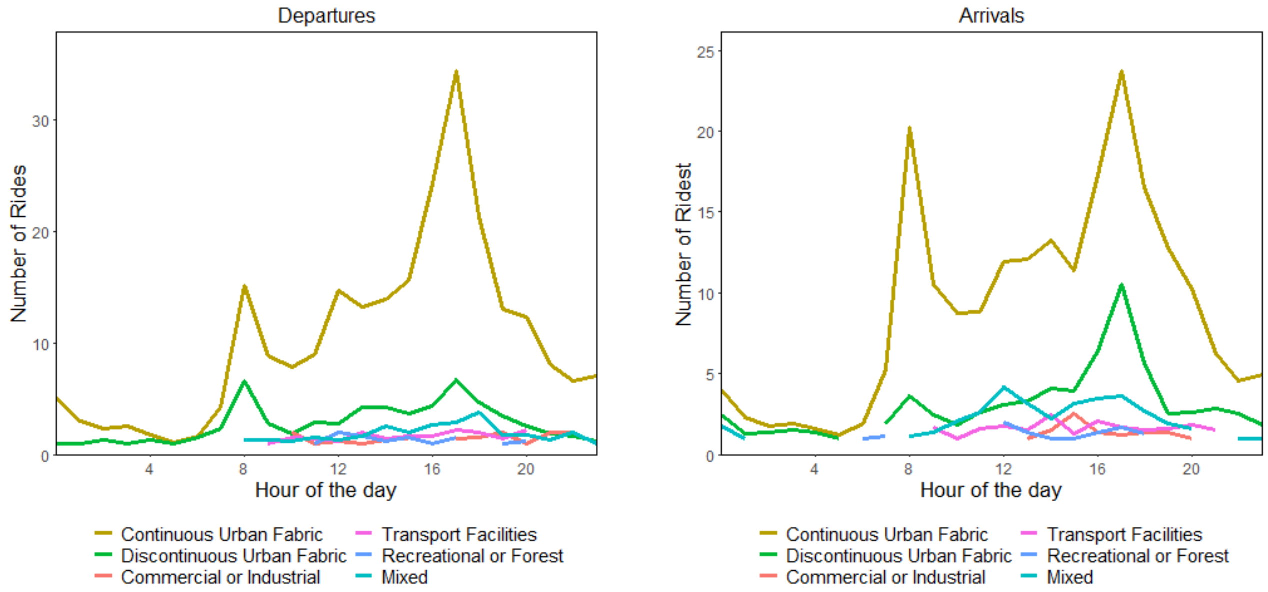
In Oslo, the different categories are distributed in a more balanced way. The most significant categories are "Continuous Urban Fabric" and "Discontinuous Urban Fabric," followed by "Transport Facilities." However, no real pattern was visible. One difference is the peak of "Transport Facilities" for arrivals at 8:00 when it exceeds the number of rides connected to "Discontinuous Urban Fabric."

Bergen

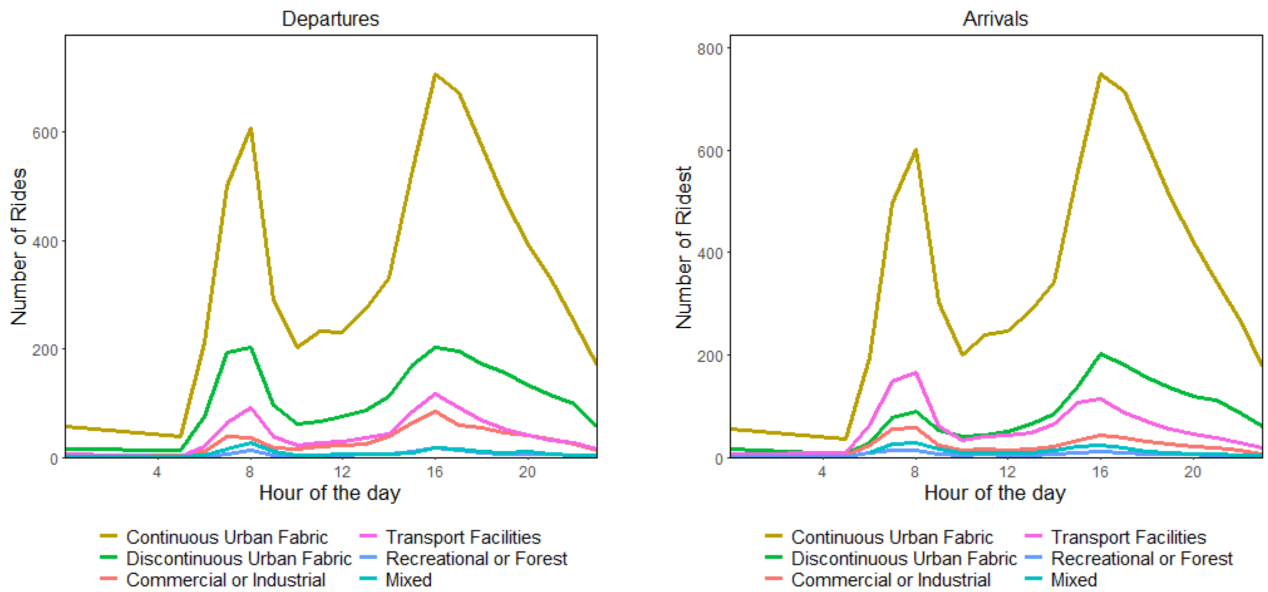
Bergen's most dominant land use category is "Continuous Urban Fabric," followed by "Discontinuous Urban Fabric," "Transport Facilities," and "Commercial or Industrial." "Discontinuous Urban Fabric" was the category with the greatest difference between departures and arrivals, and it had a major peak both in the morning for departures and in the afternoon around 17:00 for arrivals.

Trondheim

Trondheim has stations in only four land use categories, and two of which, "Continuous Urban Fabric" and "Discontinuous Urban Fabric," are significantly larger than the others. These two did not have differences between the departures and arrivals. "Transport Facilities" peaked in the afternoon around 17:00 for the departures and in the morning at 8:00 for the arrivals. However, with only approximately 10 rides per hour in this category, describing a true pattern is difficult as the lines are noisy.

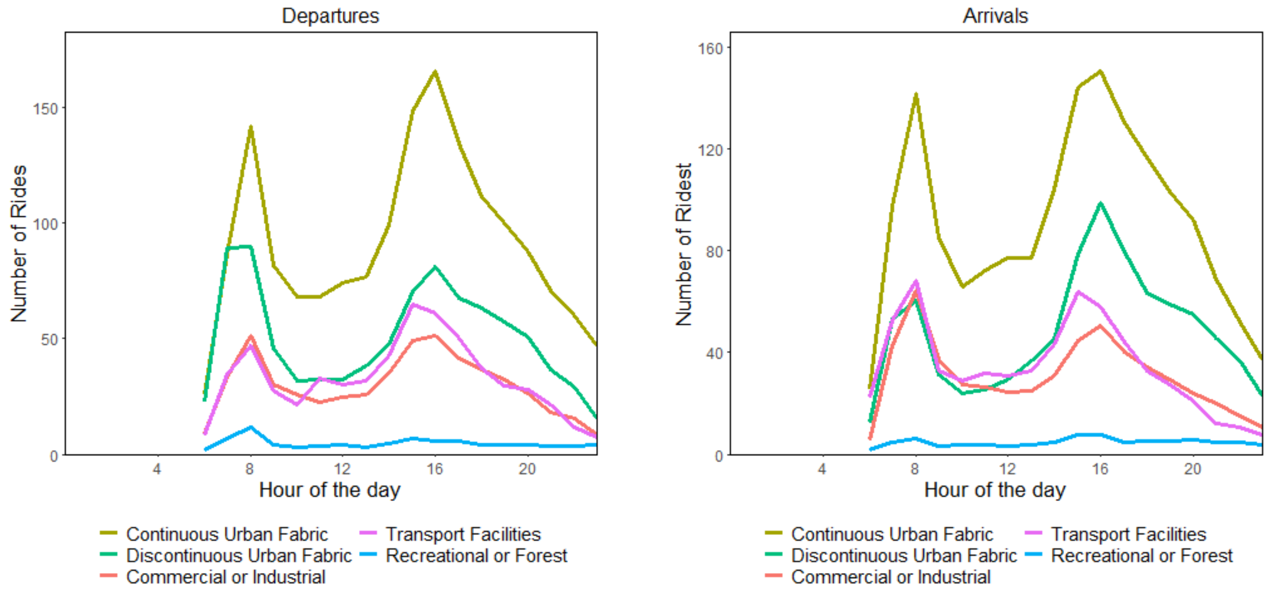


(a) Edinburgh

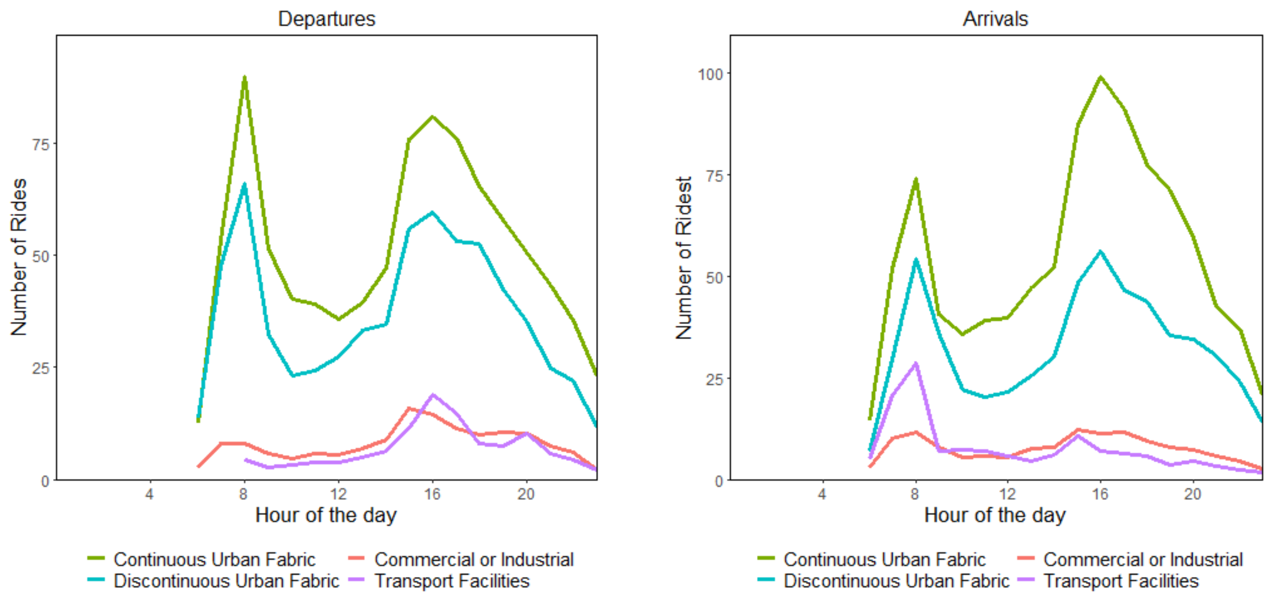


(b) Oslo

Figure 4.15 – Land use: Edinburgh and Oslo.



(a) Bergen



(b) Trondheim

Figure 4.16 – Land use: Bergen and Trondheim.

4.6 Casual Users versus Subscribers

This section examines the differences in usage between subscribers, people who pay a monthly or yearly fee for the service, and casual users, people who pay for one single ride. Since not all the datasets had this information stored, this analysis was only able to be executed with the four US cities, New York, Boston, Chicago, and Washington, DC. The section first explores how the different user types use the service in terms of trip duration, including the number of trips within the cost-free initial period. The behaviour of different users in different season is then discussed as well as the different ride characteristics for one day. The impact of weather is discussed afterwards, before the section concludes with the impact of land use on the two groups.

Table 4.16 presents the differences between subscribers and casual users in terms of trip duration and percentage of rides in the free initial period over the course of one year.

Table 4.16 – Subscriber versus Casual Users: Trip Duration and Free Initial Period

City	Avg. Trip Dur. Sub.	Avg. Trip Dur. Cust.	Free Initial Period (min.)	Sub. in Free Period	Cust. In Free Period
New York	11.1	26.9	S: 45 C:30	99.21	78.96
Boston	11.7	34.2	S: 45 C: 30/120	99.08	
Chicago	11.2	36.3	S: 45 C: 30/180	99.21	
Washington, DC	12.3	36.7	S: 30 C: 30	95.60	66.41

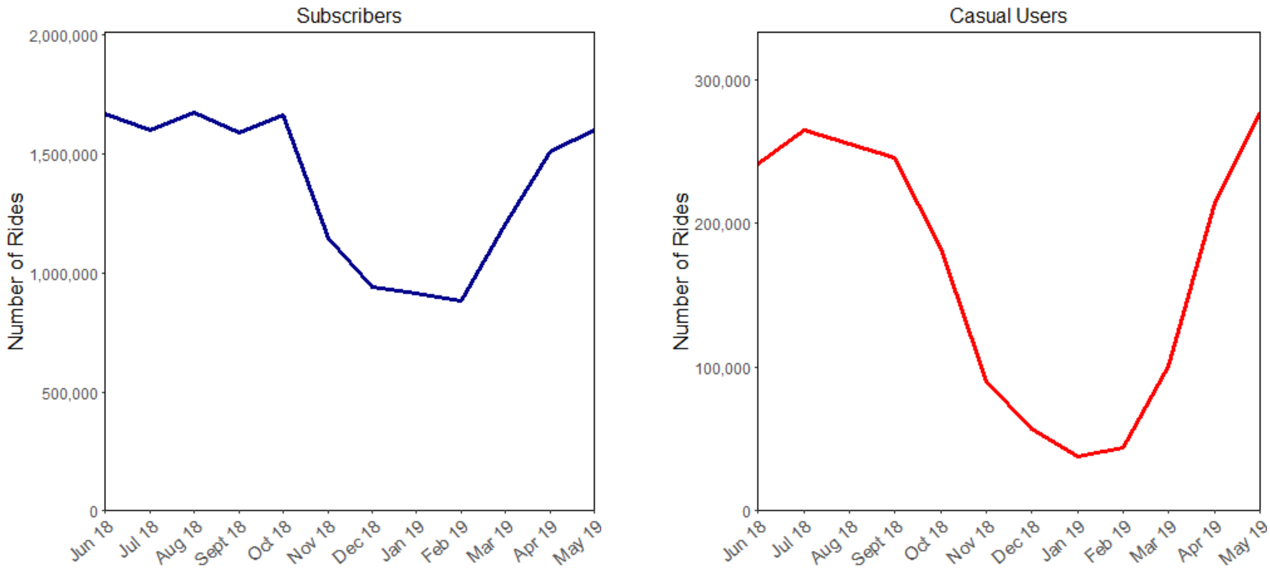
In all four cities, the subscribers' average trip duration was significantly shorter than that of the casual users. Whereas the trip duration for subscribers ranges from 11.1–12.3 minutes, the duration for the casual users ranges from 26.9–36.7 minutes. In the third column, the duration of the free initial period for subscribers (S) and casual users (C) is displayed. In all the cities apart from Washington, DC, the subscriber gets more free minutes than the casual user. Boston and Chicago's both offer a daily pass which allows users to ride for free for 120 respectively 180 minutes. However, these two separate customer groups are not labelled individually in the dataset as both are labelled as casual users. Determining how many day passes are sold might be a good predictor for tourist usage as this offer seems most attractive for people who want to explore a city by bike. The missing information is the reason the percentage of casual users within the free initial period could not be calculated for Boston and Chicago.

Almost all the subscribers stayed within the free initial period. The values for New York, Boston, and

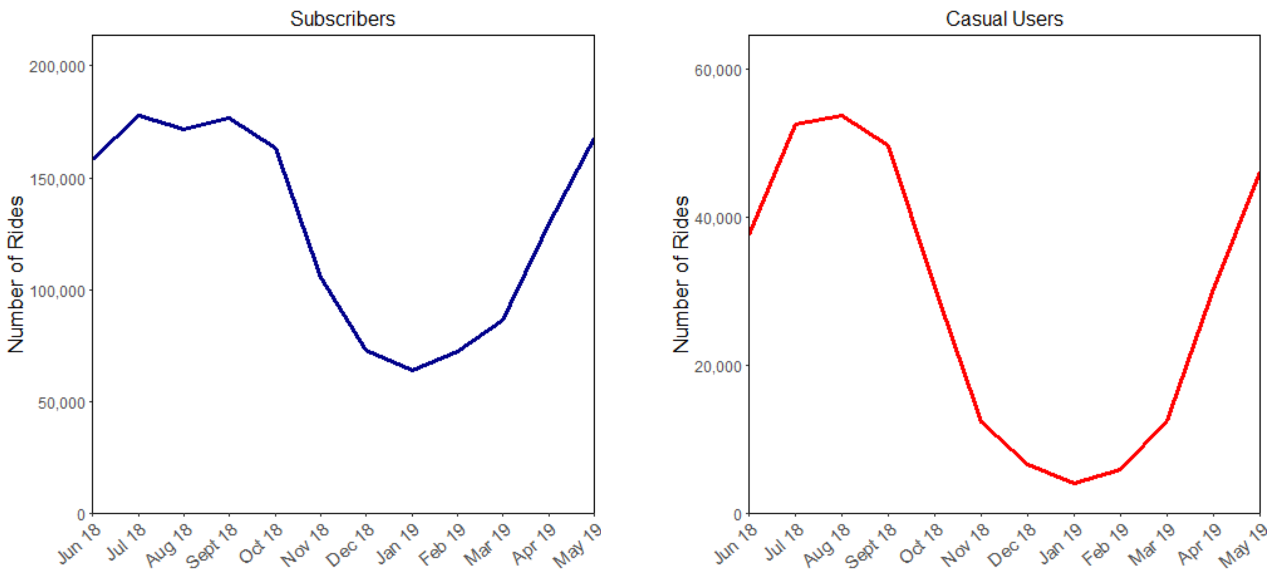
Chicago were all over 99%, and in Washington, DC the percentage was 95%. However, Washington, DC was the only city of the four which allowed their subscribers only 30 minutes of free riding compared to 45 minutes in the other three cities. In the last column, the percentage of the casual users is lower with 66.41% for Washington, DC and 78.96% for New York. One possible explanation for this might be that subscribers use the shared bikes mostly for transporting purposes such as commuting when they want to arrive at their destination as fast as possible. Alternatively, the casual users are most likely people who use the bikes more for recreational purposes. For example, tourists typically use the bikes to travel around the city for sightseeing, which takes more time than a regular commute.

4.6.1 Seasonal Differences

Figures 4.17 and 4.18 illustrate the number of rides over one year both for subscribers and casual users. All four systems had a decrease in usage during the winter months. However, for the casual users the usage decrease was greater. For Chicago and Washington, DC, the casual user usage significantly decreased with considerably few rides during winter. The number of rides for the subscribers also decreased, but the number remained relatively high compared to the casual users. One exception was Washington, DC experiencing a significant drop in December 2018. The reason might have been a severe bad weather period with cold temperatures and heavy snow. In general, the casual user usage decreases more than the subscriber usage. A reason might be that many of the subscribers use the bikes for commuting purposes, which explains why there are also rides in the winter months. However, some people might use other transport methods like the car or public transport, because they do not want to ride a bike in winter weather conditions. The decrease in ridership for the casual users might be also due to winter weather conditions. However, since casual users use the bikes for recreational activities rather than commuting purposes, the number decreases more. Another factor for the decrease might be that in winter there are less tourists.

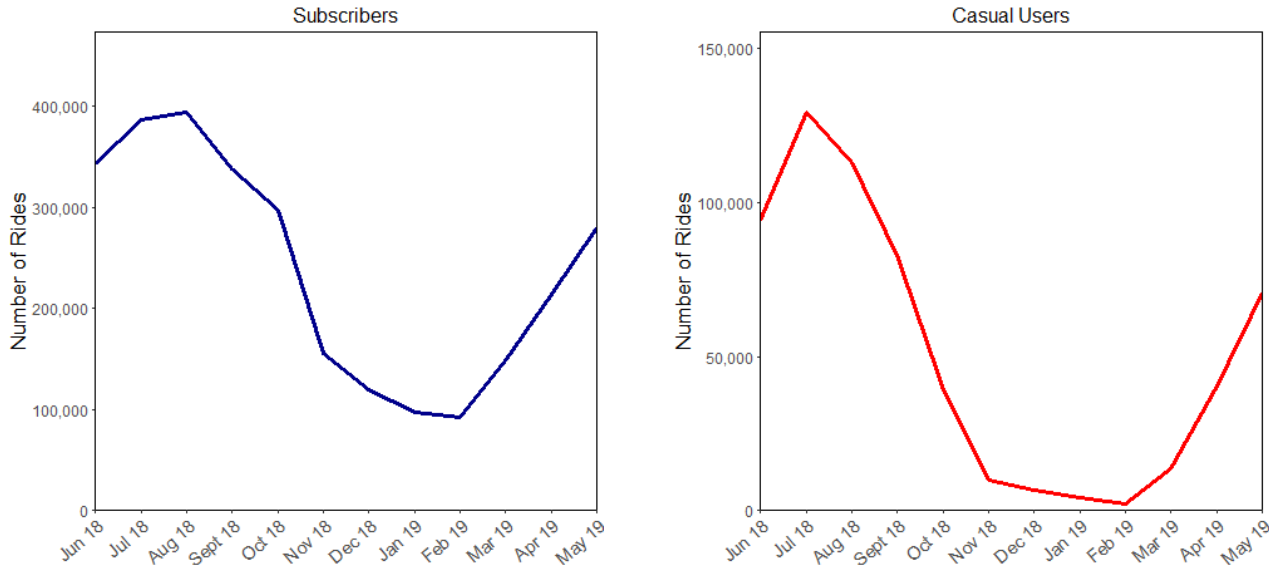


(a) New York

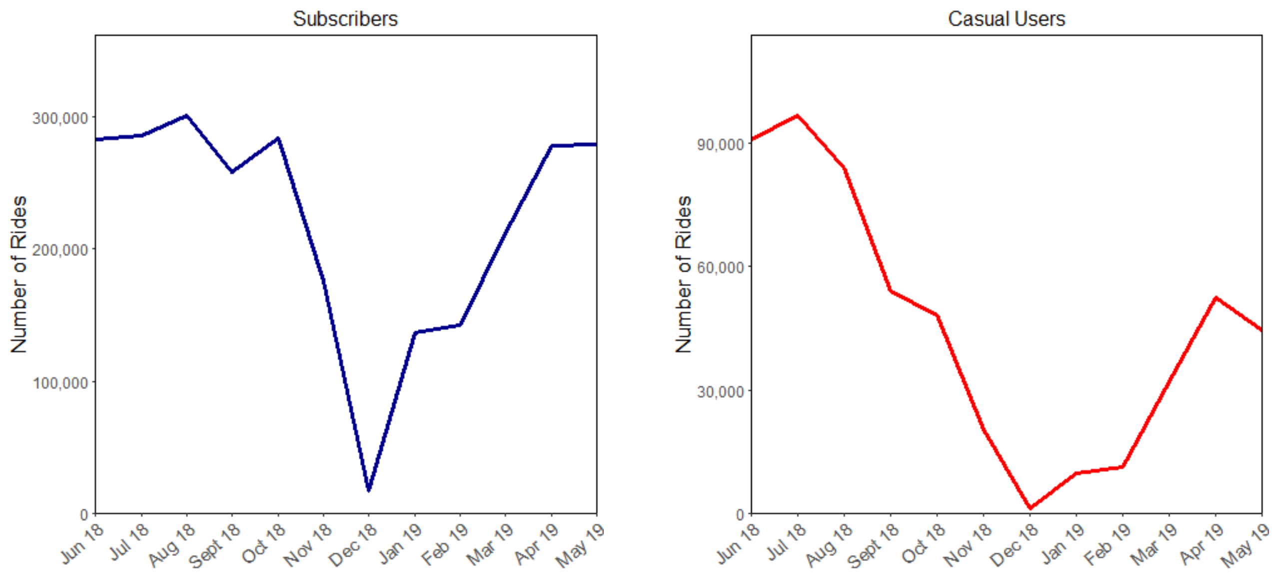


(b) Boston

Figure 4.17 – Subscriber versus casual user: New York and Boston number of rides over one year.



(a) Chicago



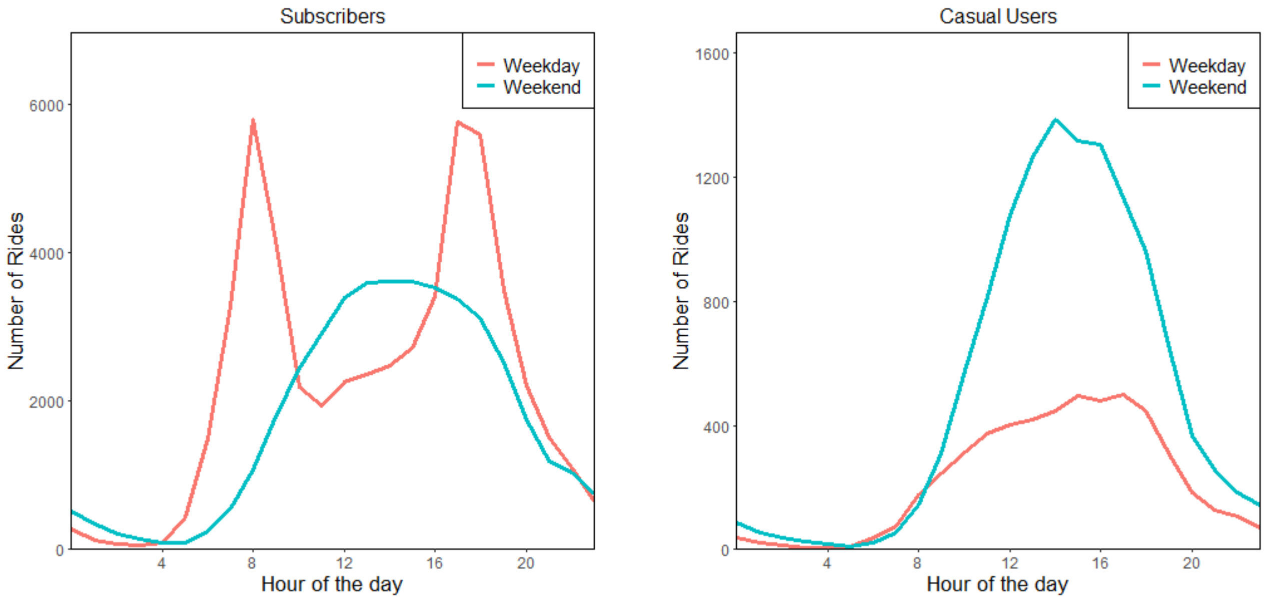
(b) Washington, DC

Figure 4.18 – Subscriber versus casual user: Chicago and Washington number of rides over one year.

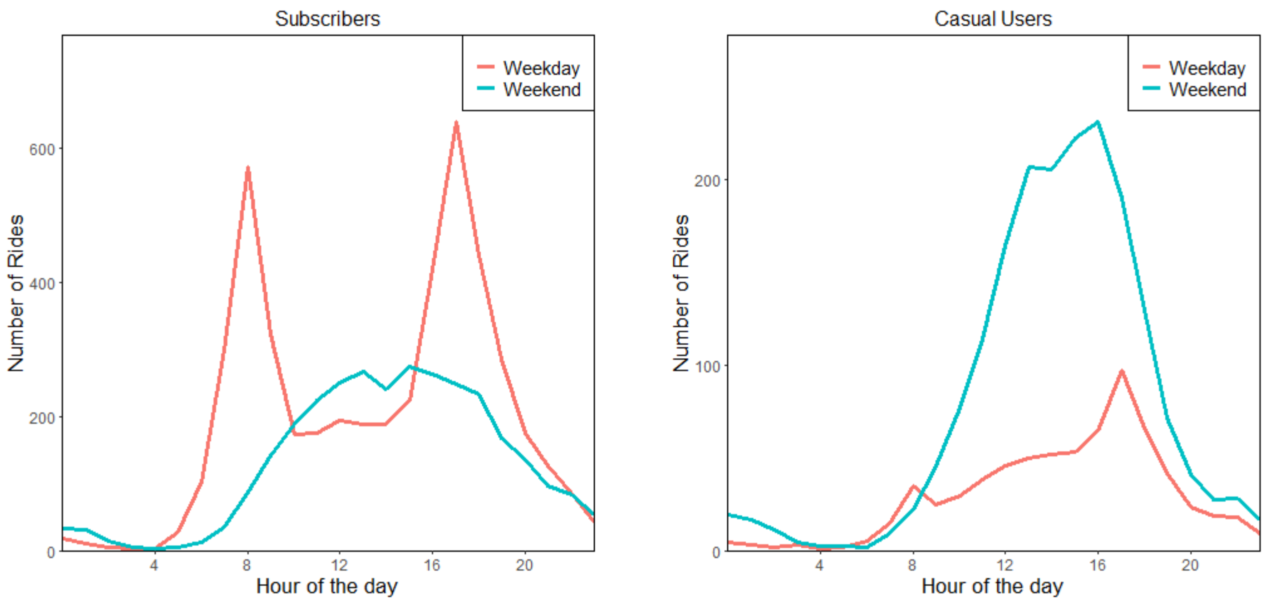
4.6.2 Daily Ride Characteristics

In Figures 4.19 and 4.20 the daily patterns for subscribers and casual users are displayed. Each city has two figures. The left-hand side details the subscribers, and the right-hand side presents the casual users. Both figures include the weekday and weekend usage. The graphs were calculated in the same manner as the ones from Section 4.2.2, and they display the average number of rides per hour for one month.

The ride characteristics of the two groups differ significantly. While the ride characteristics of the subscribers reflects the general two-peak characteristic on weekdays and the one-peak characteristic on the weekends, the casual users' ride characteristic is considerably different. First, the weekend usage is more than twice as high as the weekday usage, which is the opposite of the subscribers. Moreover, the graph for the weekdays is similar in shape to the one from the weekends, but smaller and more inclined towards the evening with a peak around 18:00. The only system with a small morning peak for the casual users is Boston, and the casual users' characteristics are more similar to the subscribers of the other cities. In all the four cities the casual users took less trips than the subscribers, which explains why the general ride characteristic was similar to the subscriber characteristic as it is the dominant user type.



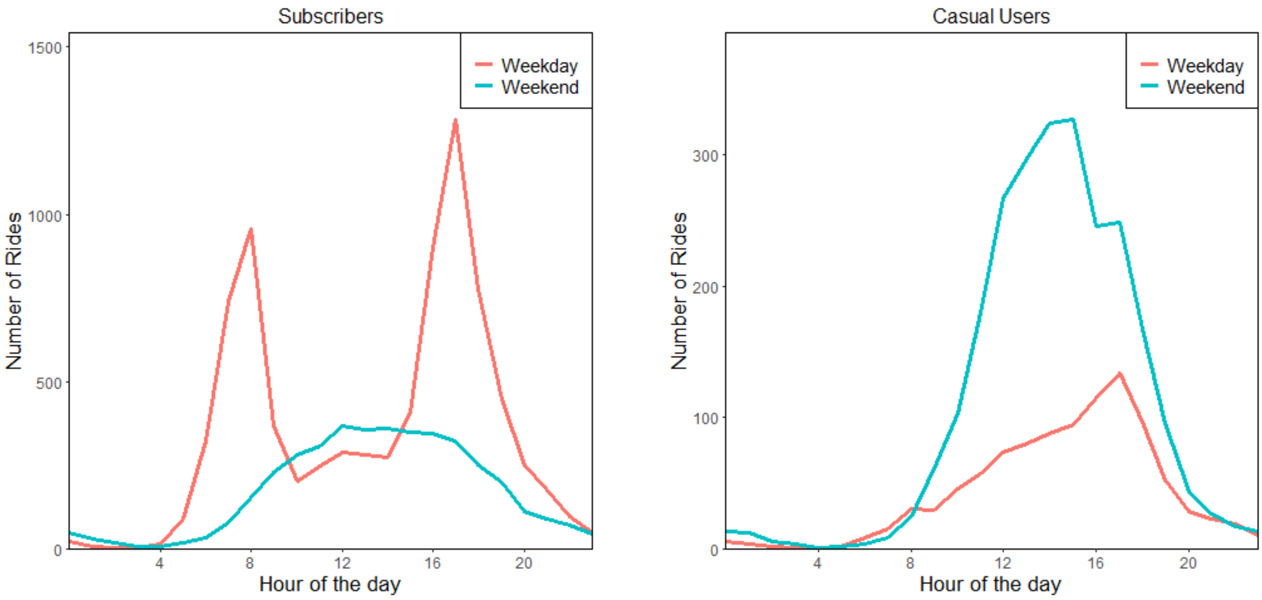
(a) New York



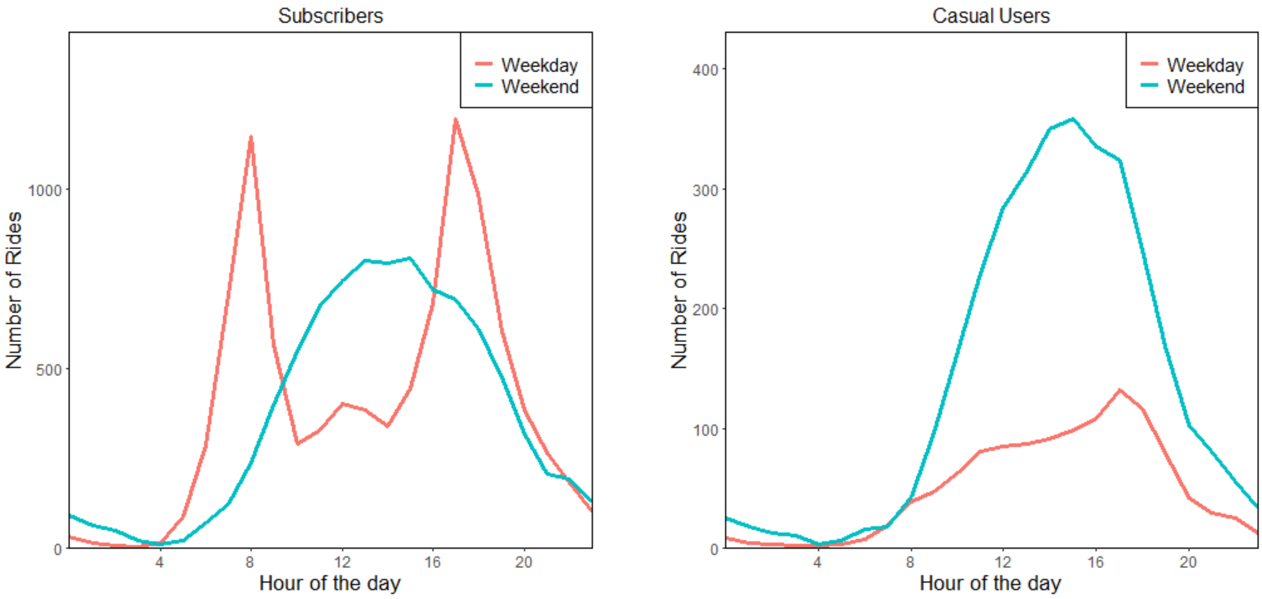
(b) Boston

The graphs show the average number of rides per hour for one month.

Figure 4.19 – Subscriber versus casual user: New York and Boston in April 2019.



(a) Chicago



(b) Washington

The graphs show the average number of rides per hour for one month.

Figure 4.20 – Subscriber versus casual user: Chicago and Washington, DC in April 2019.

4.6.3 Influence of Weather Events

This section focuses on how the weather impacts both user groups in terms of the number of rides (Table 4.17) and duration of the rides (Table 4.18). A negative binomial regression model was used for both user types like the one in Section 4.3. The weather data is the same for both cases as both datasets cover the same time span. The data is from May 2019 and consists of hourly values. The differing number of observations between subscribers and casual users was because there are no casual user rides for some hours. Furthermore, the subscribers' observations are complete only twice, for New York and Washington, DC. The other two cities have some missing values in the weather dataset.

Number of Rides

In Table 4.17, the most obvious difference is that subscribers had a negative coefficient for the weekend variable whereas casual users had a positive coefficient. This is congruent with previous results which indicate that casual users ride more on the weekends and subscribers ride more on weekdays. Apart from that trend, not many differences are evident. For all the cities both for subscribers and casual users temperature was positively correlated, and precipitation was negatively correlated. One small difference is noted for wind speed, which was significant twice for the subscribers but four times for the casual users. In all the cities, it has a negative impact. In general, the impact of weather was effectively the same for all users during May 2019.

Trip Duration

Regarding the impact of weather on the trip duration, not much difference is evident. The weekend variable was positive and significant in all four cities for subscribers and in three cities for casual users. One difference is that the coefficient was higher for the casual users, meaning that the trip duration increase was greater compared to that of the subscribers. Temperature also had a positive effect and was significant four times for the subscribers and three times for the casual users. Precipitation was negatively correlated and was significant four times for subscribers and two tines for casual users.

Table 4.17 – Influence of Weather on the Hourly Number of Rides: Subscriber versus Casual Users

	Dependent variable: Number of Rides							
	Subscriber				Casual Users			
	New York	Boston	Chicago	Washington, DC	New York	Boston	Chicago	Washington, DC
Temp.	0.036***	0.019***	0.036***	0.033***	0.107***	0.088***	0.097***	0.072***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.005)	(0.005)
Precipitation	−0.188***	−0.438***	−0.168***	−0.078**	−0.282***	−0.567***	−0.111***	−0.147***
	(0.023)	(0.048)	(0.022)	(0.038)	(0.023)	(0.062)	(0.025)	(0.034)
Wind speed	−0.006**	0.001	−0.002	−0.011***	−0.006**	−0.009***	−0.013***	−0.012***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Weekend	−0.349***	−0.339***	−0.293***	−0.255***	0.453***	0.404***	0.633***	0.370***
	(0.049)	(0.050)	(0.058)	(0.056)	(0.047)	(0.052)	(0.057)	(0.048)
Constant	7.376***	5.345***	5.549***	5.544***	4.363***	3.270***	3.467***	3.037***
	(0.112)	(0.100)	(0.129)	(0.164)	(0.106)	(0.105)	(0.123)	(0.142)
Observations	744	743	674	744	744	726	641	733

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.18 – Influence of Weather on Ride Duration: Casual Users versus Subscriber

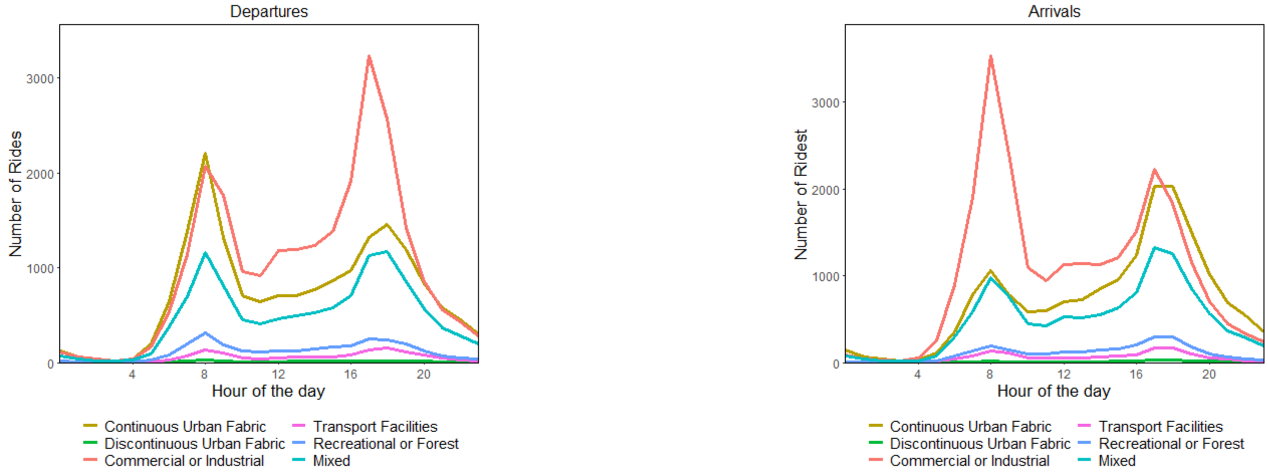
	Dependent variable: Trip Duration							
	Subscriber				Casual Users			
	New York	Boston	Chicago	Washington, DC	New York	Boston	Chicago	Washington, DC
Temp.	0.014***	0.014***	0.018***	0.013***	0.003	0.012***	0.017***	0.011**
	(0.001)	(0.002)	(0.001)	(0.003)	(0.003)	(0.005)	(0.003)	(0.004)
Precipitation	−0.021***	−0.054***	−0.025***	−0.043**	−0.030**	0.025	−0.012	−0.105***
	(0.004)	(0.016)	(0.007)	(0.020)	(0.013)	(0.041)	(0.018)	(0.029)
Wind speed	−0.00004	−0.002	−0.001	0.001	−0.001	0.002	−0.004	−0.003
	(0.0005)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Weekend	0.046***	0.091***	0.133***	0.089***	0.119***	0.113**	0.158***	0.049
	(0.009)	(0.018)	(0.018)	(0.030)	(0.028)	(0.044)	(0.043)	(0.043)
Constant	6.353***	6.430***	6.284***	6.460***	7.289***	7.362***	7.512***	7.566***
	(0.020)	(0.037)	(0.041)	(0.088)	(0.064)	(0.090)	(0.095)	(0.129)
Observations	744	743	674	744	744	726	641	733

Note: *p<0.1; **p<0.05; ***p<0.01

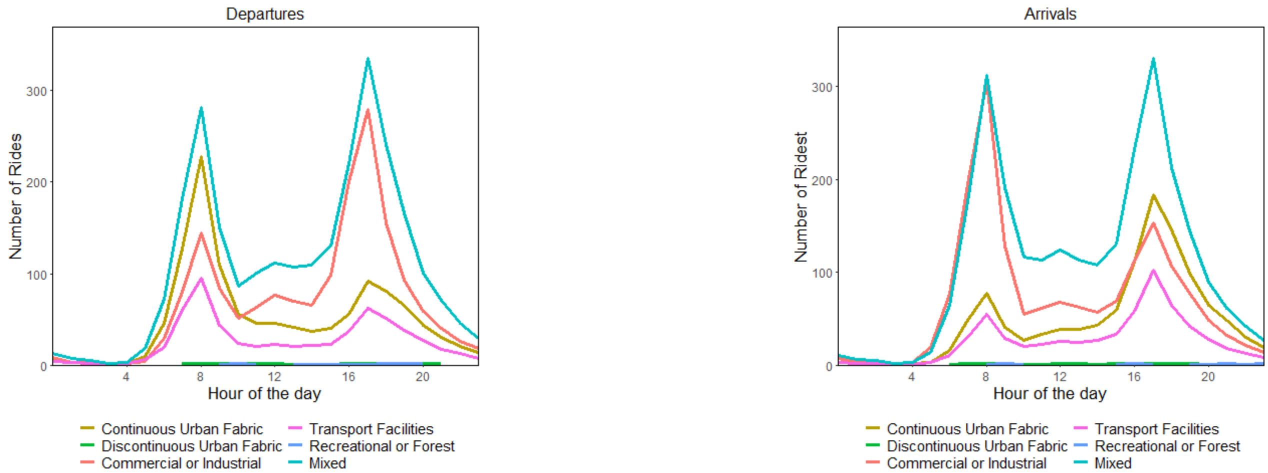
4.6.4 Land Use

This section assesses if subscribers and casual users had similar or different land use patterns. The temporal ride characteristics for weekdays, distinguished by each land use category is presented in Figure 4.21 for the subscribers and in Figure 4.22 for the casual users. Since the temporal characteristics reflect those that are discussed in Section 4.6.2, this section focuses on the different land use categories and how they influence both user groups. The characteristics for the subscribers is similar to the general land use characteristics in Section 4.5, because subscribers are the majority of all users. Therefore, these characteristics are not described again in this section. Instead, the casual users' characteristics are focused on and are compared with the subscribers.

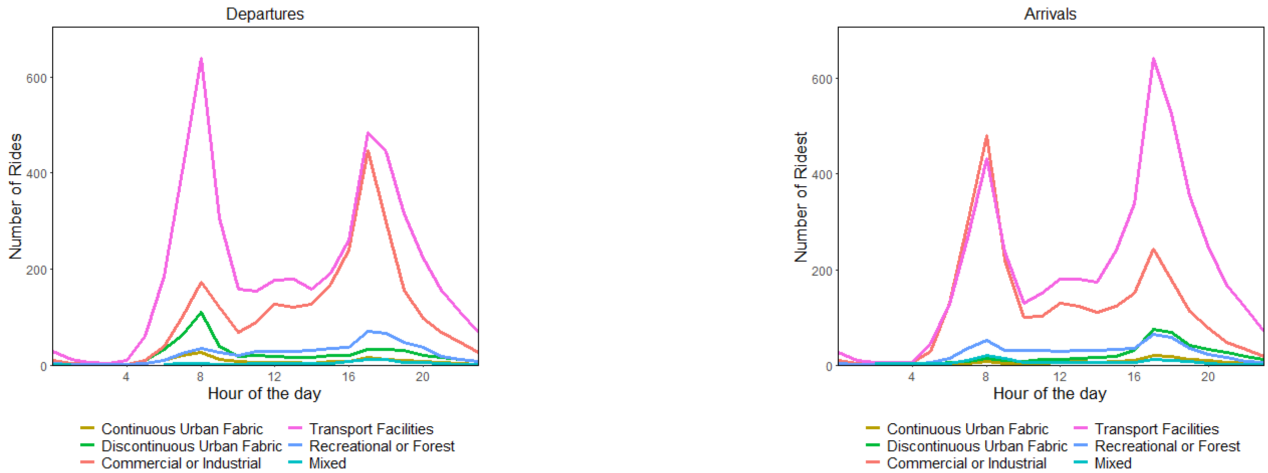
One interesting difference was the casual users' considerably higher usage of stations in the land use category "Recreational or Forest" in New York and Washington for both departures and arrivals. This trend might be due to the fact that casual users use bike sharing for recreational purposes rather than commuting. Furthermore, the shape of the recreational category is different to the others having an earlier peak. The peak was around 15:00, and the other categories peaked at 19:00. Before and after the peak, the number of rides increases and decreases gradually. The other categories' number of rides increase until 19:00 before decreasing rapidly. One possible explanation might be that during the day casual users use bikes to go to recreational places or to return from them. Then in the evening they use the bikes for activities which do not take place at recreational places, such as shopping or dining at a restaurant. The stations in the residential areas of "Continuous Urban Fabric" and "Discontinuous Urban Fabric" had a less high usage number relative to the other categories compared to the subscribers. This finding is expected as stations in residential areas might be used more often as commuting start and end points. Furthermore, the "Commercial or Industrial" category had the highest usage numbers in all three cities (New York, Boston, and Washington, DC). The characteristic of Boston suggests that some casual users also use bike sharing for commuting, because "Continuous Urban Fabric" had a departure peak in the morning and an arrival peak in the evening. Moreover, an arrival peak in the morning and an increased departure peak in the evening is visible for the land use category "Commercial or Industrial." Another fact is that the differences between arrivals and departures are much smaller for the casual users than for the subscribers.



(a) New York

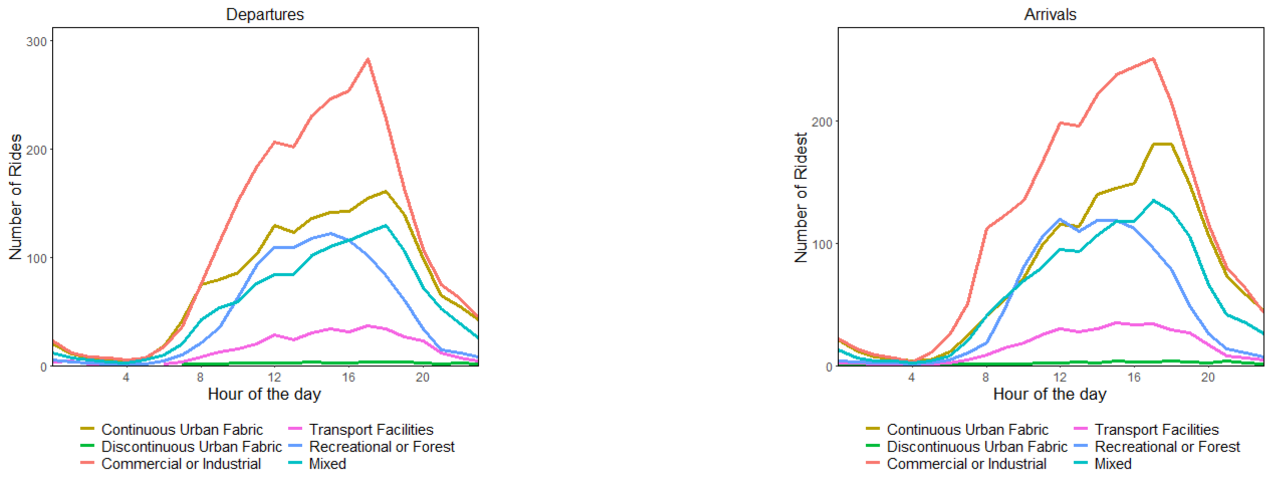


(b) Boston

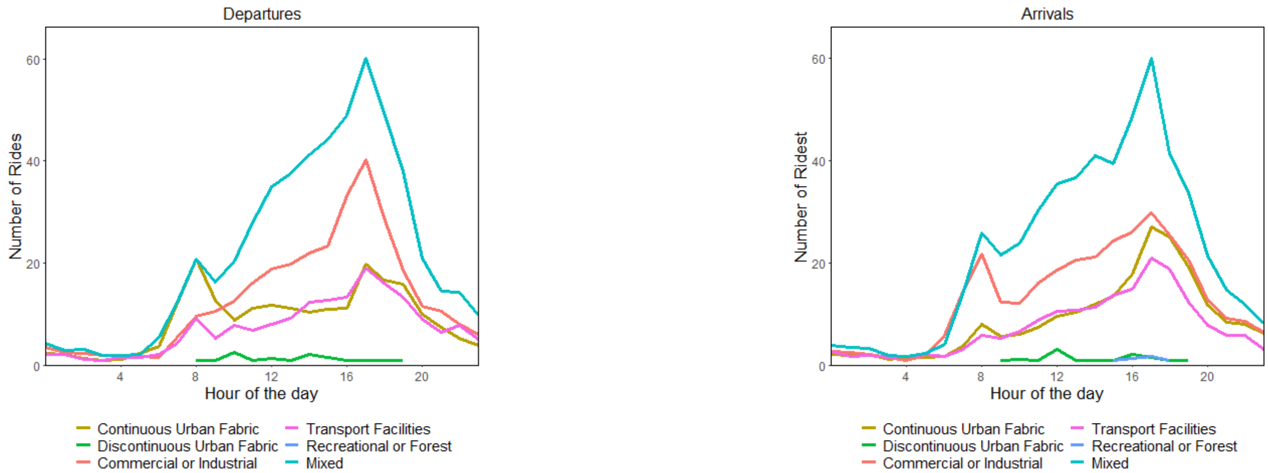


(c) Washington, DC

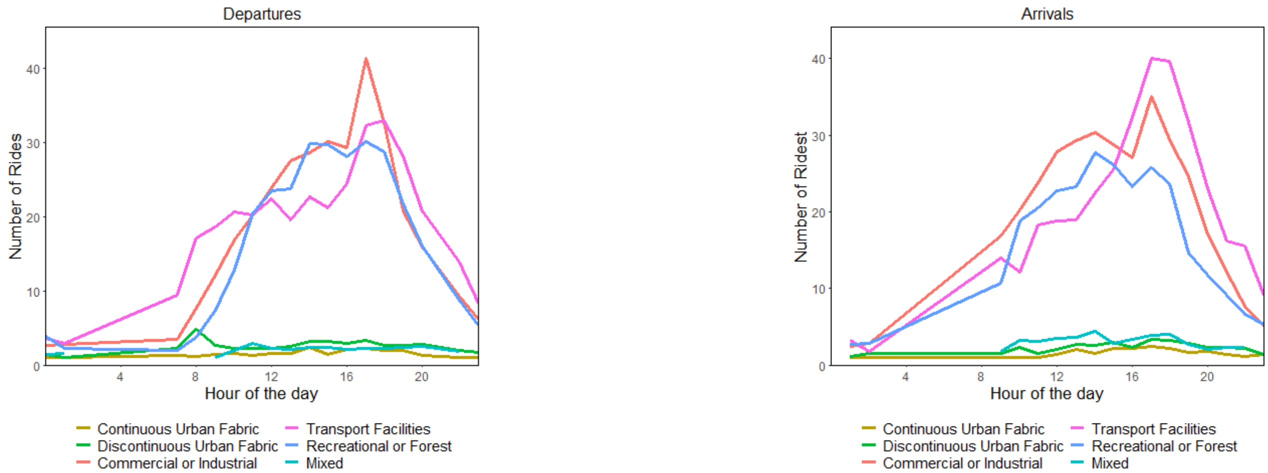
Figure 4.21 – Land Use: Subscribers.



(a) New York



(b) Boston



(c) Washington, DC

Figure 4.22 – Land Use: Casual Users.

4.7 Covid-19

During the process of writing this master's thesis, the Coronavirus Pandemic, or Covid-19, developed and has had serious impact on the daily life of many people. As people have had to stay at home or work from home, movement patterns have changed drastically. Since a major part of bike sharing users are commuters, the impact of the Coronavirus pandemic on the usage of bike sharing is notable to research. To assess the changes, March, April, and May 2020 were compared with the same months from 2019 to determine if a visible difference in usage had occurred. The analysis used the data from New York, because it is the most stable system in terms of size as it does not have major changes in ridership or station numbers.

Figure 4.23 illustrates the ridership numbers from 2019 and 2020. On the left-hand side are the weekdays, and on the right-hand side are the weekends. In March, there was not a considerable difference between 2019 and 2020. While there are less rides in 2020 for weekdays and weekends, the shapes of the curves are similar, which could possibly be due to worse weather conditions. In April, the situation is different as the ridership collapsed to a minimum for 2020, and the pattern is no longer comparable with 2019. In 2019, the peak hours had approximately 6,000 rides, and in 2020 there was a maximum of approximately 2,000 rides per hour. Given these changes, the major part of the commuting flow disappeared as the main peak was in the evening around 18:00. For the weekend the shape of the curve remains the same. The ridership significantly decreased but not as much as for weekdays. The increased ridership after midnight also decreased strongly, indicating that many people stayed at home and did not go out. In May, the shape of the curve stayed the same as in April 2020, but the ridership slightly increased. For the weekend graph the number of rides in 2020 was considerably above the number of rides from 2019. This occurrence might be due to the cycling boom that developed during the Coronavirus pandemic, because many recreational activities could not be performed due to the closings and restrictions. Moreover, since many people had to stay in their homes during the week, this may have caused them to leave the house on the weekend. Additionally, the risk of being infected while cycling is relatively small. In general, based on the findings from New York, the impact of the Coronavirus pandemic on bike sharing behaviour is evident.

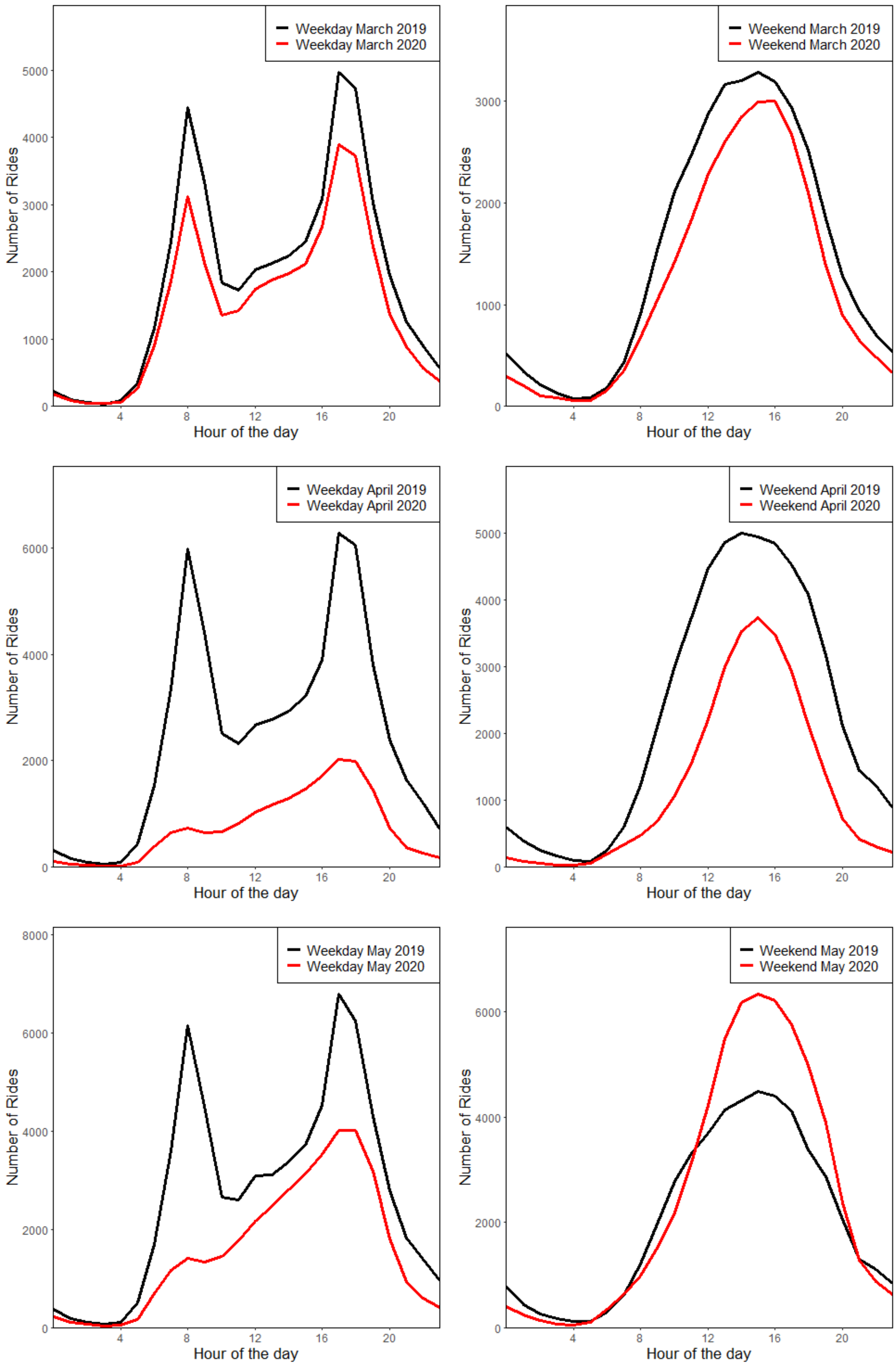


Figure 4.23 – Covid-19 impact in New York: ridership comparison of 2019 versus 2020.

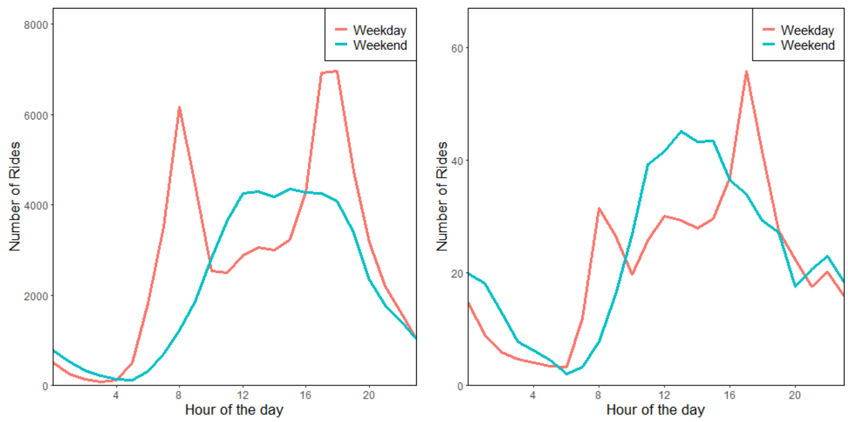
Chapter 5

Discussion

This chapter discusses the results from the analysis and compares it with the findings from the literature. This discussion focuses on the temporal ride characteristics, the influence of weather events, the impact of infrastructure, and the impact of land use. Additionally, this chapter reflects on how suitable the different data were for such an analysis. Finally, the differences between the cities are discussed and the research questions are answered. Furthermore, the approach for this thesis is critically examined.

5.1 Temporal Ride Characteristics

In Section 4.2, the seasonal variability as well as the daily ride characteristic of the different cities were analysed. For the major five cities (New York, Boston, Chicago, Washington, DC, and London) the weekday ride characteristics were dominated by two major peaks, one in the morning and one in the evening, as well as a small peak around midday. The weekend ride characteristics were dominated by one long, major peak. O'Brien et al. generated a qualitative classification for BSS based on temporal characteristics in which they describe different usage patterns and link them to a predicted demographic. As a result, a table was produced which lists the characteristics for each demographic with a list of cities that are a good example of each type. For the previously described ride characteristics, the predicted demographic type is "Commuters and Weekend Leisure User" which is congruent with this study's assumptions. Moreover, the systems from Boston, London, and Washington, DC are in O'Brien et al.'s list of example systems for this demographic type. For Edinburgh, Oslo, Bergen, and Trondheim where the evening peak is greater and the usage between the morning and the evening peak higher, the closest demographic type is "Commuters with Some Utility Users" or "Utility Users With Some Commuters" (O'Brien et al., 2014). Figure 5.1 visualises the two different demographic types.



(a) New York **(b)** Edinburgh

Figure 5.1 – Average hourly usage on weekdays and weekends in August in New York and Edinburgh.

The evening usage is generally higher than the morning usage, except for December. As mentioned in the chapter four "Results and Interpretation," this finding might be due to colder temperatures and an earlier sunset. According to Faghih-Imani et al. and Kim, this trend might be because people are in a rush in the morning which hinders them from using bike sharing (Faghih-Imani et al., 2014; Kim, 2018). In the case of Edinburgh in which the weekend usage was higher than the weekday usage, Kim suggests that the main purpose of a system with such a characteristic is not commuting (Kim, 2018). The higher usage after midnight on weekends is also described by different authors who have discovered the same pattern in their respective studies, and they suggest that this is caused by people going out on the weekends (Faghih-Imani et al., 2014; Reiss and Bogenberger, 2015; Vogel et al., 2011).

Regarding trip duration, the most obvious findings were that almost all the rides fell into the initial free period which is mostly between 30 and 60 minutes. Except for Edinburgh which had an average trip duration of 23.6 minutes, all the cities had an average trip duration of less than 20 minutes.

These findings are in line with those from Nair et al. who stated that the systems are designed for short trips. In their study in Paris, 92% of all trips were less than 30 minutes, and 98% were shorter than 45 minutes (Nair et al., 2012). For London, 30 minutes tends to be the threshold in terms of trip duration (Wood et al., 2011). Zhao et al.'s findings were similar as 95% of all trips in Nanjing, China were less than 30 minutes (Zhao et al., 2015). Regarding the different trip durations during the course of one day, no connection to the literature could be made as no study has included this factor.

The analysis of the difference between subscribers and casual users regarding BSS usage indicated that the two user groups have different usage patterns. O'Brien et al.'s system classification is referenced again to compare the results. The subscribers fit into the demographic of "Commuters and Weekend Leisure Users," whereas the casual users' demographic is "Leisure Users." Since the weekend usage was higher than the weekday usage, these labels fit for all the cases, and the study's assumptions are confirmed in that subscribers tend to use the bike sharing for commuting purposes, and the casual users use it for leisure purposes (O'Brien et al., 2014). Several authors have found similar results by concluding that subscribers tend to use bike sharing for commuting purposes whereas casual users use it more for leisure and various other activities (El-Assi et al., 2017; Sun et al., 2017; Tran et al., 2015; Zhou, 2015).

5.2 Covid-19

This sub-section addresses the results from the Covid-19 analysis that compared the rides in New York from March, April, and May 2020 with the same months in 2019. Figure 5.2 shows again the differences in weekday ridership of March, April, and May 2020 versus 2019 in New York. As the results indicated, the ridership in March was slightly lower in 2019 than in 2020. In April, the usage numbers decreased drastically, especially on weekdays. The reasons for the decrease might be the stay-at-home order and more people working from home. In May the numbers started to rise again, but with different characteristics as there were many rides in the afternoon and more rides than 2019 on the weekend. Regarding this topic, some research has already been conducted, including the analysis of New York by Teixeira and Lopes. Another finding was that the ridership in New York at the beginning of the pandemic increased by 33% in the second week of March compared to the week prior. The citizens of New York probably avoided public transport to prevent infection with the virus. After the emergency status declaration on March 12, the rides began to drop significantly. Then, after the stay-at-home order was issued from the state of New York, the ridership decreased even more. At the end of the month, there was a decrease of almost 70% compared to the beginning of the month. The average trip duration also increased from 13 minutes in the first week of March to 19 minutes in the last week, indicating that people used bikes more for longer distances. Moreover, a shift from using the subway to using BSS was detected as many people avoided public transport (Teixeira and Lopes, 2020). These findings are interesting, but they do not go beyond March 2020. The webpage MOBV[21] visualises the change in mobility during Covid-19. The authors conducted a longer analysis of the ridership anomalies in New York and confirmed the findings from Teixeira and Lopes. They also reported an increased usage in the second week of March before having a strong decrease after the lockdown was declared. The negative anomalies continued until the beginning of May when the numbers recovered slowly. Moreover, a strong positive anomaly for weekends was detected starting at the end of April. This occurrence is congruent with the findings from this study as a strong increase in weekend usage was detected. Furthermore, the visualization from the MOBV webpage continued until mid-June and reported a significant increase in ridership from mid-May through mid-June.

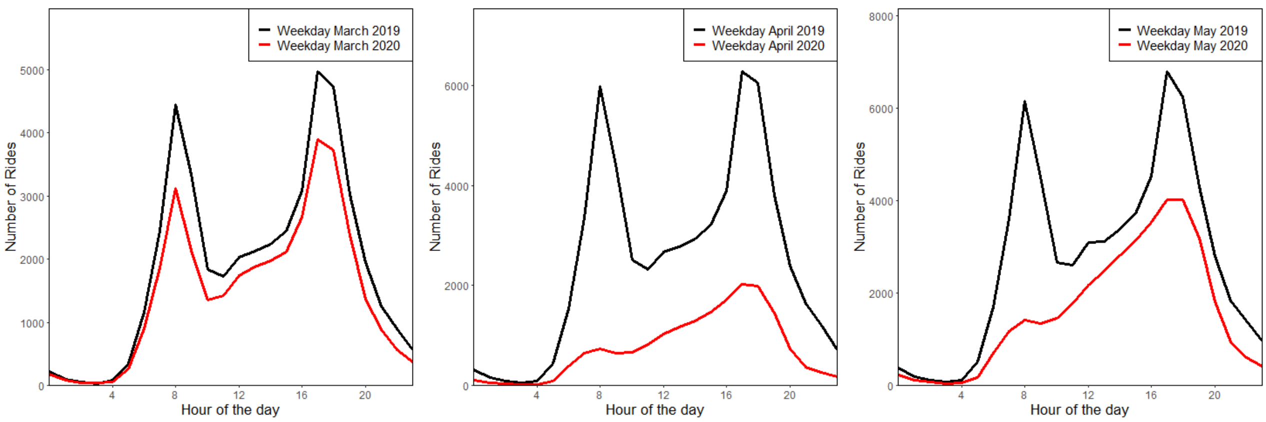


Figure 5.2 – Covid-19 impact in New York: weekday ridership comparison of 2019 versus 2020.

[21] https://jwolondon.github.io/mobv/docs/newyork/ by Purves et al. (2020)

5.3 Influence of Weather Variables

5.3.1 Number of Rides

The results from the weather impact analysis suggest that the most important weather variables are precipitation and temperature. The results indicated an average decrease in ridership of 37.7% during rainy periods in all nine cities (see 5.3). Higher temperatures also encourage people to use bike sharing more. Moreover, high wind speeds throughout the day seem to be a hindrance in New York, Boston, Chicago, Washington, DC, London, and Trondheim. The negative impact of precipitation is described in various previous studies (Campbell et al., 2016; Caulfield et al., 2017; Corcoran et al., 2014; El-Assi et al., 2017; Gebhart and Noland, 2014; Kim, 2018). Furthermore, Figure 4.12 that depicts the differences in usage between rain and no rain is similar to Figure 2.2 from Gebhart and Noland (2014), which indicates a lower riding volume in the case of rain.

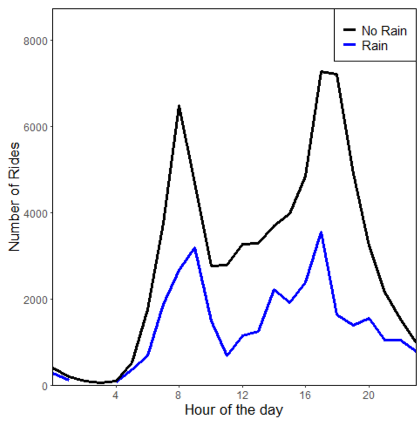


Figure 5.3 – Ride characteristics: rain versus no rain in New York

Regarding the impact of temperature, the understanding is not as evident as for precipitation, and it also partially contradicts with the findings from the literature. In this study's results, the temperature variable is more significant in the hourly analysis than in the daily analysis. The hourly analysis is congruent with the findings from Faghih-Imani et al., who also detected a positive correlation with higher temperatures (Faghih-Imani et al., 2014). For the daily analysis, the maximum daily temperature was the most meaningful variable, indicating that the maximal temperature is the best explaining variable temperature wise. Kim and Campbell et al., found a decrease in usage with high temperatures (Campbell et al., 2016; Kim, 2018). This result cannot be confirmed or denied by this study as the temperature was analysed as one and not split into different temperature ranges (e.g. 10-20°C, 20-30°) which would be a prerequisite for such an analysis. Caulfield et al. and Corcoran et al. did not detect a significant impact of temperature at all, neither positive nor negative (Caulfield et al., 2017; Corcoran et al., 2014). However, since the number of rides drops in almost all the cases towards winter (see Figure 4.1) a correlation is undeniable. The usage of a BSS might be impacted by significant temperature changes, such as the difference between summer and winter, and less impacted by smaller day to day changes. Another approach may need to be used for this variable, similar to the one used by El-Assi et al. which looked at the impact of different temperature ranges and determined that the most favourable temperature for BSS usage is 20–30C° (El-Assi et al., 2017). Likewise, elaborating a temperature threshold might be more meaningful, both for low and high temperatures, as the results from the literature report that both are hindering factors.

The factor wind speed was more significant in the daily analysis rather than in the hourly analysis, which indicates that people tend to use bike sharing less if the daily wind speed is high, compared to situations when there is only a short period of strong winds. However, as the variable was significant in only five cities, this was not determined to be a general trend. In the literature, three authors included wind speed in their analyses, and all three reported a negative impact of strong winds on the number of rides (Corcoran et al., 2014; Gebhart and Noland, 2014; Kim, 2018). Furthermore, wind, and especially strong and short term winds, can vary across a city. Some bike stations might also be sheltered better than others through high buildings or denser areas. Gebhart and Noland's results are particularly applicable to this thesis as their study was also executed in Washington, DC. This thesis' research results for Washington, DC are congruent with those of Gebhart and Noland; for Washington, DC, wind speed was a significant factor for both the daily and the hourly analysis. The difference between subscribers and casual users in the context of weather is that casual users seem to be impacted more by the different variables. For both subscribers and casual users temperature and precipitation are significant in all cases. However, for both variables the coefficient was higher for the casual users, indicating that casual users are affected more than subscribers by weather events. Moreover, casual users are affected more by wind, as wind speed was significant in four cases for the casual users versus two cases for the subscribers. Not much research has been conducted comparing the two user groups in regard to weather conditions. Therefore, this study only has one reference to the impact of rain on the two groups. The findings of Gebhart and Noland are in line with those from this study as they found a decrease for both groups during rainy periods. However, the decrease of casual users was 20% greater than the decrease of subscribers (Gebhart and Noland, 2014).

5.3.2 Trip Duration

Unlike the number of rides in which much research has been done to analyse the impact of weather variables, the impact of rain on bike sharing trip duration has been studied less. Only Gebhart and Noland have included it in their research. To have more reference literature regarding this topic, this study's findings are compared with studies from Gao et al. and Böcker and Thorsson who assessed the impact of weather events on the trip duration of regular, non-bike sharing rides. Their results suggest that precipitation and temperature are the most significant variables for the daily and the hourly analysis. For the daily analysis, the maximum temperature was the most significant factor. Wind speed was less significant in both results, and it had a greater impact on the daily usage. The positive impact of the temperature and the negative impact of the precipitation is in line with the findings of Gao et al., who reported shorter general bike rides during rainy periods and longer rides with higher temperatures (Gao et al., 2018). Böcker and Thorsson had similar results that indicated that precipitation has a negative effect. Regarding temperature, they stated that rides are impacted negatively by temperatures that are both too hot and too cold (Böcker and Thorsson, 2014). Regarding wind Böcker and Thorsson and Gao et al. agree with each other, but their findings are not aligned with the results of this study. They both record a negative impact of wind on trip duration (Böcker and Thorsson, 2014; Gao et al., 2018) which is not the case for the results of this study. Regarding the difference between subscribers and casual users, this study's findings were not as clear for the number of rides. For the subscribers, both precipitation and temperature were significant in all cases, and higher temperatures had a positive effect while precipitation had a negative effect. For the casual users, the temperature was significant three times and precipitation was significant twice. For the two cases in which precipitation was significant it had a higher coefficient, meaning it had a greater impact compared to the subscribers. For the temperature, no such trend was visible. Thus, this study's findings were partially congruent with the findings of Gebhart and Noland who report a stronger decrease for the casual users during rainy periods (Gebhart and Noland, 2014).

5.4 Impact of Infrastructure

Regarding the infrastructure, a pattern is less clear than for the weather. The most significant factor in this analysis was altitude as the variable with the most significances in all models, both for arrivals and departures. The results suggest that altitude is a hindering factor for bike sharing usage as it influences the arrivals and departures negatively as higher stations have less departures and less arrivals. These findings are in line with those of Faghih-Imani et al. and Tran et al. who also detected a negative impact on arrivals and departures (Faghih-Imani et al., 2017; Tran et al., 2015). Alternatively, Mateo-Babiano et al. reported a difference between arrivals and departures as stations with a higher altitude recorded significantly less arrivals than departures (Mateo-Babiano et al., 2016).

The results for the impact of road types were not evident. As mentioned in the results section, there was no clear understanding that suggested a direct impact. Moreover, some results contradicted some findings from the literature. In this study's results, bike lanes were significant in less than half the cities, and one negative impact on the usage was found. In contrast, several authors detected a positive correlation between rides per station and proper cycling infrastructure, including cycleways, bike paths, bike lanes, and bike roads (El-Assi et al., 2017; Faghih-Imani and Eluru, 2015; Faghih-Imani et al., 2014; Sun et al., 2017; Zhang et al., 2017). Moreover, this study's results suggest a positive correlation with major roads for half of the cities, both for departures and arrivals. These findings also strongly contradict with recent literature, which suggests a negative impact of major roads on the usage of bike sharing (El-Assi et al., 2017; Faghih-Imani and Eluru, 2015; Faghih-Imani et al., 2014). The only street type that had congruent results with the literature was minor roads, which in this study's results was positively significant five times, for both arrivals and departures. This finding is in line with those of Faghih-Imani et al. as well as Faghih-Imani and Eluru who also found a positive impact of minor roads on the station usage (Faghih-Imani and Eluru, 2015; Faghih-Imani et al., 2014). The results of this study do not offer a distinct pattern, which suggests that one cannot assume a clear relationship between road types and station usage. This is in line with the findings of Tran et al. who did not detect a significance, meaning that infrastructure has no impact (Tran et al., 2015). One possible explanation for these results might be that the location of a bike station is more important than the road infrastructure around it. Moreover, it might be more important for people to get close to the city centre, to a train station, or to a point of interest. In these cases, the start and stop points are not affected by the type of infrastructure near the bike station but rather by the demand of the users. In terms of public transport, recent studies did not agree with each other completely. The results of this study demonstrate that in general the public transport has no significant impact on the bike station usage. This finding is congruent with Zhang et al. who found no impact on the bike sharing usage in Zhongshang, China (Zhang et al., 2017). However, this finding contradicts with O'Neill and Caulfield who found public transport in general to be positively correlated to usage (O'Neill and Caulfield, 2012).

Taking a closer look at the public transport, by splitting it into the different types of public transport, the results indicated that it is likely that subways and tubes have an impact on the usage of bike sharing. Five of six cities with a subway resp. tube have significant values for the public transport mode train. This is the case for departures as well as arrivals. These findings are congruent with Noland et al., which detected a positive impact of subway stations on BSS ridership in Washington, DC (Noland et al., 2016). On the other hand it is contradicting with Sun et al. and Tran et al., where the first found a negative correlation between public transport and bike sharing, the latter found no correlation, neither positive nor negative. However, he found a correlation between railway and bike sharing which could be congruent with the findings of this study as in this model railway and subway are combined one category (Sun et al., 2017; Tran et al., 2015).

5.5 Impact of Land Use

Noland et al., who focused on the differences between subscribers and casual users, considered how these two user groups are affected by different land use categories. The most meaningful findings which are comparable to this study's results were that recreational land use has no impact on the usage and that subscribers tend to ride more in residential land use areas (Noland et al., 2016).

These findings are partially congruent with this study's results that suggest that casual users are likely to ride in recreational areas whereas subscribers are not. The fact that residential land use is linked with subscribers is confirmed by this study's results as subscribers took relatively more rides in residential areas than casual users. Combining the findings from different results suggests that stations in recreational land use areas are affected the most by rain, because these stations are mostly used by casual users, who are most affected by rain. Moreover, stations in recreational land areas probably have the highest usage volume on weekends as the weekend usage is higher for casual users. Additionally, some subscribers take recreational trips at the weekend too. Mateo-Babiano et al. investigated how bike sharing flows change over the course of one day, in respect to land use areas. One of the key findings of the study was that the most residential-to-commercial rides take place in the morning, and in the evening the most commercial-to-residential rides occur (Mateo-Babiano et al., 2016). These findings cannot be equally compared to the findings of this study, because it considered the departures and arrivals rather than the flows. However, one can still draw connections. For example, New York and Boston exhibited a pattern which fits these assumptions. The departures from "Continuous Urban Fabric" areas as well as the arrivals in "Commercial" areas peaked in the morning which suggests that many rides were taken between these two. In the evening, the same pattern occurs in the other direction, with many departures from "Commercial" areas and many arrivals in "Continuous Urban Fabric" areas. Nonetheless, not all the cities demonstrated these characteristics, and therefore, this assumption cannot be fully confirmed.

Figure 5.5 depicts a simplified visualisation of the results from the weather analysis as well as the infrastructure analysis. The weekend variable is added as well. The different colours indicate if the different impact variables have a positive (green), negative (red), or neutral impact on the dependent variables. For the weather analysis the dependent variables were "Number of Rides" and "Duration." For the infrastructure analysis the dependent variables were "Departures" and "Arrivals." The figure demonstrates the described effects all at once. Temperature, precipitation, and the weekend variable are the most consistent. Furthermore, the differences between subscribers and casual users in terms of weekend usage are visible. The figure also reveals that the results of the infrastructure analysis are not the same for every city, and making a general statement about it is impossible. Moreover, more variables had no impact on the European cities compared to the US cities.

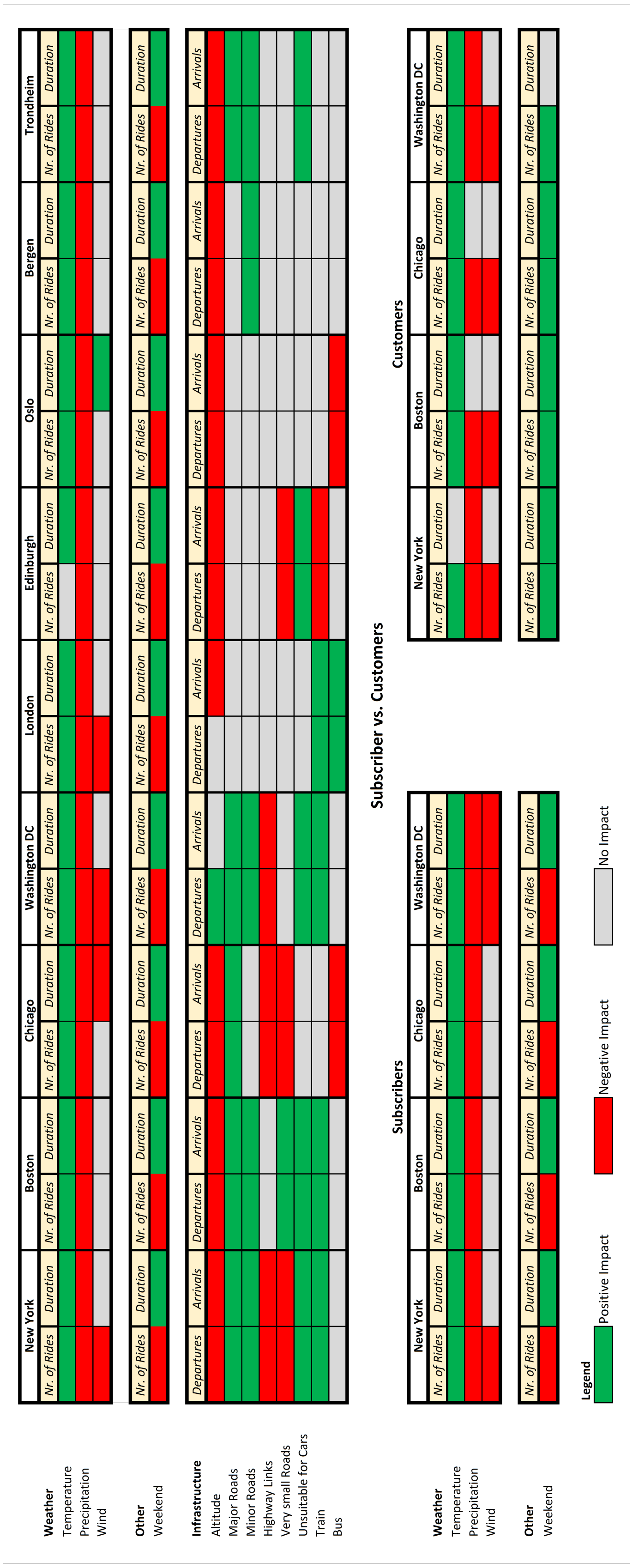


Figure 5.4 – Generalized Results for Weather and Infrastructure

5.6 Data Assessment

This section assesses the quality of the data and discusses how suitable the data was for this analysis. The main data for this analysis was the BSS data, or trip data. The most obvious aspect of the data was the different dimensions. Some datasets contained almost 2 million rides per month, and others barely had 10,000. This difference was evidenced through the ride characteristics graph in Chapter 4. The cities with more rides have smoother graphs. Moreover, the averages of various occasions are less prone to outliers if there is more data, and in this context more rides. In general based on all the results including the graphs and statistics, the systems with more rides offer better results. The graphs with smoother lines and the statistics with more rides have more significances, especially for the infrastructure analysis. This might also be influenced by the fact that systems with fewer rides generally also have fewer stations. This relationship leads to a smaller sample and less statistically representative results. However, the systems do not need 2 million rides per month to offer comparable data. The results for Boston, which has approximately 200,000 rides per month, are similar to the those from New York, with 2 million rides, which suggests that the threshold for a reasonable number could be even less than 200,000 rides per month. However, other factors might play a role as well. New York and Boston might be more similar to each other in many ways than New York and Edinburgh. The way cities are built and people behave as well as the attitude towards bike sharing may differ. Nonetheless, the data from the other cities is not useless. It is still possible to identify patterns of the ride characteristics as well as usage trends in the weather analysis. Although all the weather data originated from the same source, the completeness of the data varied greatly. Some datasets contained hardly any missing values, whereas others had recordings for only half of the data points. Moreover, not every weather station was in the city in question, and some weather stations were far away from the city. The weather station used for the precipitation in Edinburgh was Leuchars, which is about 50 km away from the centre of Edinburgh. Since some of the nearer weather stations offered insufficient data, this was the only possibility to get consistent data. The farther weather station, and thus inaccurate information, might explain why Edinburgh had the smallest decrease during rainy periods. The lack of consistent data was also the reason factors like humidity and peak gust were not included into the analysis as they had too many missing values. When comparing daily and hourly data, the hourly dataset was more complete than the daily dataset. Moreover, the results suggest better results with the hourly dataset which would lead to the conclusion that hourly data are more fitting for this purpose, especially since weather conditions can change significantly during the day. These findings also suggest that bike sharing users are prone to fast weather changes and try to use slots with good weather for their trips. However, collecting hourly data for a whole year would be much more time- and storage consuming.

Regarding the infrastructure analysis, the results offered few visible trends for all the cities. However, the smaller cities mostly had no significances, and this occurrence may be due to the fewer bike stations. Some findings were peculiar, for example the negative impact of cycleways. This finding calls into question whether the OSM dataset was feasible for such an analysis. The positive impact of major roads was initially unexpected, but as mentioned in Section 4.4 this might be because the most frequently used stations are in the city centre and thus in an area with many major roads. Additionally, by using a circle as station area, many road segments which do not affect bike sharing usage are included. Furthermore, some infrastructure that might have an impact could be more than 300 m away from the station. The weighting of the different factors might vary too. A very busy junction could be more hindering than a separated bike path is attracting, or vice versa.

The land use data can be divided into two parts: the European data from CORINE Landcover and the US data from different governmental sources. In retrospect, the US data seemed to be more meaningful than the European dataset. The different categories indicate more variability within each other, and the stations are spread among more categories. This outcome was not expected as the categories from the US datasets were transformed to match similar categories like the European dataset. Alternatively, the European dataset indicated much less variability, and almost all the rides were in two categories. One reason for that might be that the US datasets have a smaller spatial resolution and focus more on land use, whereas the European datasets also consider land cover, which generally requires less spatial resolution than land use in a city.

5.7 Summary

In summary, given the different variables which were analysed, the ride characteristics, the weather analysis, and the land use analysis agree with the results in the literature. Although the infrastructure analysis offered the least promising results, it provided some meaningful results such as the impact of altitude. The following section will discuss the research questions from Section 1.2. Some variables play an important role in more than one question. Therefore, the questions are answered together and not separately.

The three main research questions are the following:

- Which factors influence the usage of shared bikes? (RQ1)
- Are there any differences in use between different cities? (RQ2)
- Is the usage of bike sharing influenced by global or local factors? (RQ3)

RQ1/RQ2: The most obvious variable that influences the usage of bike sharing strongly for all nine cities is the season. All cities recorded less rides during the winter months than during the summer months. Moreover, all cities exhibit a different usage characteristic for weekdays and the weekend. Furthermore, the time of the day is also an important factor that influences the usage of bike sharing as certain hours have more or less rides. These three factors can be called global factors as they are the same for all cities. All cities have seasons, weekends, and rush hours, and these factors cannot be controlled by the bike sharing providers or the city government. The same rule applies for the weather variables that significantly impact the usage of bike sharing, and precipitation and temperature in particular seem to affect the usage of bike sharing for all cities. Wind speed cannot be stated as a relevant variable for all cities.
RQ2/RQ3: As described in Section 3.4, season, weekend and weather are categorised as global variables, whereas infrastructure and land use are categorised as local variables. The results of this study showed that global variables are affecting the usage of shared bikes more than local variables. Additionally, the impact of the global variables is very similar from city to city, whereas the impact of the local variables change from city to city. This dynamic might be the reason altitude was the

most significant factor for infrastructure, because altitude is the only factor in the infrastructure analysis which does not change from city to city. If one has to ride uphill, it does not matter which city, because it will be harder than riding on flat ground. In retrospect, it could have been classified as global variable. Alternatively, riding a bike around a city might be a completely different experience depending on one's location, as some cities might be bike friendly, and others are not. These described factors are assumed to be the reason for the heterogeneous results in the infrastructure analysis, as infrastructure is not the same in all cities. Thus, infrastructural variables do not have the same effect on all cities. To conclude, BSS all react the same way to global factors which are not controllable whereas the impact of local variables changes from city to city.

5.8 Limitations

The weaknesses of this thesis are now briefly discussed in addition to what could have been changed. The first thing to change would be to choose more systems with higher usage numbers and not choose systems with low usage numbers. As previously described, the smaller systems do not have enough rides and/or stations to deliver satisfying results, especially for the infrastructure analysis. As it is hard to gather data from big European cities, it would lead to a research based mainly on non-European cities. In this study a compromise was reached between European cities and cities with larger systems. The results of the infrastructure analysis are not completely satisfying, and the reasons for this might be various. One alternative approach would be to obtain governmental data to ensure that the data is accurate. With OSM data, one is never sure of the quality of the data. For example, the dataset for New York may be of a much higher quality than the dataset of Bergen. The two could also have the same quality data or have less data, meaning that the data has less riders and less variation, and thus it could have less significance. This scenario would lead to different results. The same applies for the land use data. Although all data are from trustworthy sources, the results differ. Especially for Europe the used dataset does not seem to be suitable for such an analysis compared to the US datasets. Therefore, collecting a genuine land use dataset would be ideal in comparison to a mixed dataset for land cover and land use like the one from CORINE

Landcover. This particular dataset seemed to be less detailed. Regarding weather data, a completer dataset with more variables like humidity and peak gust would be beneficial. Moreover, gathering more hourly data rather than daily averages would be preferred.

However, many of these changes entail costs, and if a researcher wants to work with open data only, many of these improvements would not be possible. Regarding the chosen methods and models, elaborating an infrastructure model with less variables would be ideal as it seems to be a slightly overloaded, which leads to a confusing output. Furthermore, many more variables should be included to have a complete dataset for the infrastructure analysis. Station capacity, population density, distance to city centre, points of interest, and distance to other stations are a few examples of other explanatory variables which might impact the distribution of rides over the different stations. Moreover, splitting the train variables into a normal train and subway would be preferred to determine if a difference exists. Furthermore, distinguishing between large train stations, normal train stations, and small train stations would be advantageous, because the impact of a large main station is assumed to be greater than the impact of a small tube station.

Chapter 6

Conclusion

The aim of this thesis was to determine if the usage of different bike sharing systems in Europe (London, Edinburgh, Oslo, Bergen, and Trondheim) and the United States (New York, Boston, Chicago, and Washington, DC) are affected by the same impact variables. The variables which were analysed were the bike sharing trips, the weather, infrastructure (including altitude), and land use areas. A reproducible approach and freely available data were used. First, the general ride characteristics were analysed. For the yearly analysis the results indicated a decreased usage during the winter months. For the daily analysis a distinction is made between weekdays and the weekend. For the weekdays a two-peak characteristic was recorded with one peak in the morning around 8:00 and one peak in the afternoon around 17:00. This two-peak characteristic suggests that bike sharing is used for commuting purposes. For the weekend there was one peak characteristic indicating that weekend usage is dominated by leisure users. Furthermore, the difference between subscribers, people with an annual or monthly subscription, and casual users who pay for one ride at a time, was analysed. The results indicated that subscribers use bike sharing for commuting purposes whereas casual users use it for leisure activities. Trip duration was affected by the same parameters as the number of rides. The trip duration decreased towards winter and differed between the weekend and weekdays. Trips with bike sharing were longer on the weekends than on weekdays. Additionally, the average trip duration of casual users (33.5 minutes) was considerably longer than the average trip duration of subscribers (11.6 minutes). For both user-groups the majority of all rides are within the free initial period. However, the percentages differ. For the casual users it was between 66% and 79%, and for the subscribers it was between 95% and almost 100%.

The impact of weather variables was visible in all the cities. The most influencing variables were precipitation and temperature. Precipitation had a negative effect on the bike sharing usage whereas people tended to ride more with higher temperatures. Both precipitation and temperature affect both subscribers and casual users. However, casual users are affected more by both variables. The impact of the weather variables on trip duration is less clear. Nonetheless, the results suggest that both have an impact: precipitation had a negative impact and temperature had a positive impact. This time, the trip duration of casual users seems to be less affected by the weather variables than the trip duration of subscribers. Wind speed was not significant for any of the cities, so it cannot be stated as a variable that impacts all the cities.

Regarding infrastructure, the results are less promising. The only variable affecting the usage for all cities was altitude. Fewer arrivals and departures were recorded at bike sharing stations with higher altitude. Regarding the impact of road types, none of them had an impact on all nine bike sharing systems. The impact of public transport was also not obvious. However, a closer analysis suggests that subways and tubes had a positive impact on the usage of bike sharing.

The land use analysis for the US cities indicated promising results. A pattern was evident which demonstrates that many rides in the morning start at a station in a residential area, and many arrivals in the morning are recorded in commercial areas. In the evening the characteristic is reversed with many starts in commercial areas and many arrivals in residential areas. For the European cities no such trend was visible as almost all the rides started and ended in residential areas. The reason behind this occurrence might be due to the less detailed land use data. The COVID-19 discussion indicated that the ridership in New York during the pandemic deceased drastically in April, before strongly increasing from mid-May forward. This pattern suggests that people used bike sharing as an alternative travel mode during the pandemic.

6.1 Future Work

The thesis provides some suggestions for future investigations about this topic. All of the investigations could be expanded by adding more variables. For example, one could analyse if a difference exists in usage between the genders and ages. One could also deepen the analysis regarding the differences between casual users and subscribers as well as consider more weather variables like humidity and peak gust. A flow analysis would also be interesting as one could detect the most used connections, and thus extract the city's most used cycling routes. Another opportunity for a deeper approach would be the infrastructure analysis, as many more variables could be included like population density, neighbouring effects, points of interest, and station capacity. However, assessing only one city would make more sense as infrastructure varies greatly between the different cities. The land use analysis could also be expanded by examining flows rather than only arrival and departure rates. Lastly, the analysis of the ridership during the Covid-19 pandemic indicated that this is a topic which would be interesting to explore. One could determine if the cycling boom remains or if people resume their old movement behaviour. In general, as bike sharing systems are still growing and evolving all around the world, this topic provides many approaches. Furthermore, with the emergence of e-bikes, free floating bikes, and bike sharing with GPS new possibilities for future research are particularly offered.

Bibliography

BFS (2018). Mobilität und Verkehr - Statistischer Bericht. Technical report.

Böcker, L. and Thorsson, S. (2014). Integrated weather effects on cycling shares, frequencies, and durations in Rotterdam, the Netherlands. *Weather, Climate, and Society*, 6(4):468–481.

Boston Discovery Guide (2020). Boston Subway - The "T" - Boston Public Transportation. https://www.boston-discovery-guide.com/boston-subway.html.

Campbell, A. A., Cherry, C. R., Ryerson, M. S., and Yang, X. (2016). Factors influencing the choice of shared bicycles and shared electric bikes in Beijing. *Transportation Research Part C: Emerging Technologies*, 67:399–414.

Caulfield, B., O'Mahony, M., Brazil, W., and Weldon, P. (2017). Examining usage patterns of a bike-sharing scheme in a medium sized city. *Transportation Research Part A: Policy and Practice*, 100:152–161.

Corcoran, J., Li, T., Rohde, D., Charles-Edwards, E., and Mateo-Babiano, D. (2014). Spatio-temporal patterns of a Public Bicycle Sharing Program: The effect of weather and calendar events. *Journal of Transport Geography*, 41:292–305.

DeMaio, P. (2009). Bike-sharing: History, Impacts, Models of Provision, and Future. *Journal of Public Transportation*, 12(4):41–56.

El-Assi, W., Salah Mahmoud, M., and Nurul Habib, K. (2017). Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. *Transportation*, 44(3):589–613.

Faghih-Imani, A. and Eluru, N. (2015). Analysing bicycle-sharing system user destination choice preferences: Chicago's Divvy system. *Journal of Transport Geography*, 44:53–64.

Faghih-Imani, A., Eluru, N., El-Geneidy, A. M., Rabbat, M., and Haq, U. (2014). How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal. *Journal of Transport Geography*, 41:306–314.

Faghih-Imani, A., Hampshire, R., Marla, L., and Eluru, N. (2017). An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville. *Transportation Research Part A: Policy and Practice*, 97:177–191.

Frade, I. and Ribeiro, A. (2014). Bicycle Sharing Systems Demand. *Procedia - Social and Behavioral Sciences*, 111(1999):518–527.

Gao, J., Kamphuis, C. B., Dijst, M., and Helbich, M. (2018). The role of the natural and built environment in cycling duration in the Netherlands. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1):1–16.

Gebhart, K. and Noland, R. B. (2014). The impact of weather conditions on bikeshare trips in Washington, DC. *Transportation*, 41(6):1205–1225.

GPSVisualizer (2020). GPS Visualizer: Assign elevation data to coordinates. https://www.gpsvisualizer.com/elevation.

Kim, K. (2018). Investigation on the effects of weather and calendar events on bike-sharing according to the trip patterns of bike rentals of stations. *Journal of Transport Geography*, 66:309–320.

Land.copernicus.eu (2020). About Copernicus Land Monitoring Service. https://land.copernicus.eu/about.

Mateo-Babiano, I., Bean, R., Corcoran, J., and Pojani, D. (2016). How does our natural and built environment affect the use of bicycle sharing? *Transportation Research Part A: Policy and Practice*, 94:295–307.

Médard de Chardon, C., Caruso, G., and Thomas, I. (2017). Bicycle sharing system 'success' determinants. *Transportation Research Part A: Policy and Practice*, 100:202–214.

Meteostat.net (2020). Our Data. https://meteostat.net/en/data.

Nair, R., Miller-Hooks, E., Hampshire, R. C., and Bušić, A. (2012). Large-Scale Vehicle Sharing Systems: Analysis of Vélib'. *International Journal of Sustainable Transportation*, 7(1):85–106.

National Association of City Transportation Officials (2017). Bike Share in the United States: 2017. Technical report.

Noland, R. B., Smart, M. J., and Guo, Z. (2016). Bikeshare trip generation in New York City. *Transportation Research Part A: Policy and Practice*, 94:164–181.

NYCGOycgo (2020). NYC Subway Secrets and Insider Tips. https://www.nycgo.com/articles/15-secret-subway-tips.

O'Brien, O., Cheshire, J., and Batty, M. (2014). Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography*, 34:262–273.

O'Neill, P. and Caulfield, B. (2012). Examining user behaviour on a shared bike scheme: the case of Dublin Bikes. *The 13th International Conference on Travel Behaviour Research*, (July):1–13.

OpenStreetMap (2020a). About OpenStreetMap - OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/About_OpenStreetMap.

OpenStreetMap (2020b). OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Main_Page.

Parkes, S. D., Marsden, G., Shaheen, S. A., and Cohen, A. P. (2013). Understanding the diffusion of public bikesharing systems: Evidence from Europe and North America. *Journal of Transport Geography*, 31:94–103.

PubliBike (2020). How it works. https://www.publibike.ch/en/publibike/how-it-works/.

Purves, R., Straumann, R., and Wood, J. (2020). Active Travel in Covid-19 World. MOBV. https://jwolondon.github.io/mobv/docs/newyork/.

Reiss, S. and Bogenberger, K. (2015). GPS-Data Analysis of Munich's Free-Floating Bike Sharing System and Application of an Operator-based Relocation Strategy. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, volume 2015-Octob, pages 584–589. Institute of Electrical and Electronics Engineers Inc.

Rudloff, C. and Lackner, B. (2014). Modeling Demand for Bikesharing Systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2430(1):1–11.

Shaheen, S., Guzman, S., and Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, Present, and Future.

Sun, Y., Mobasheri, A., Hu, X., and Wang, W. (2017). Investigating Impacts of Environmental Factors on the Cycling Behavior of Bicycle-Sharing Users. *Sustainability*, 9(6):1060.

Teixeira, J. F. and Lopes, M. (2020). The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York's Citi Bike. *Transportation Research Interdisciplinary Perspectives*, 6:100166.

Tran, T. D., Ovtracht, N., and D'Arcier, B. F. (2015). Modeling Bike Sharing System using Built Environment Factors. *Procedia CIRP*, 30:293–298.

Transit Chicago (2020). Schedules & Route Info - CTA. https://www.transitchicago.com/schedules/.

Visitlondon.com (2020). Lodnon Underground — Official Visitor Guide. https://www.visitlondon.com/traveller-information/getting-around-london/london-tube?ref=header.

Vogel, P., Greiser, T., and Mattfeld, D. C. (2011). Understanding Bike-Sharing Systems using Data Mining: Exploring Activity Patterns. *Procedia - Social and Behavioral Sciences*, 20:514–523.

Washington.org (2020). Navigating Washington, DC's Metro System — Metro Map & More. https://washington.org/navigating-dc-metro#Metro-Hours.

Wood, J., Slingsby, A., and Dykes, J. (2011). Visualizing the dynamics of London's bicycle-hire scheme. *Cartographica*, 46(4):239–251.

Zhang, Y., Thomas, T., Brussel, M., and van Maarseveen, M. (2017). Exploring the impact of built environment factors on the use of public bikes at bike stations: Case study in Zhongshan, China. *Journal of Transport Geography*, 58:59–70.

Zhao, J., Wang, J., and Deng, W. (2015). Exploring bikesharing travel time and trip chain by gender and day of the week. *Transportation Research Part C: Emerging Technologies*, 58:251–264.

Zhou, X. (2015). Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in Chicago. *Plos One*, 10(10).

Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Date:

10.09.2020

Signed: