Norwegian School of Economics,

Bergen Autumn 2022

Impact of COVID-19 on bicycle sharing system in Oslo

Annel Meirkhanova

Supervisor: Julio Cesar Goez

Master thesis, Master of Science in Economics and Business Administration, Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

The impact of COVID-19 on Oslo bicycle sharing system was analyzed in this thesis. Weather data in the form of temperature, precipitation and wind speed, and calendar event factor (whether the day, when the ride was taken, was workday or weekend) were used as explanatory variables to assess shared bikes usage. Independent variables were chosen using LASSO shrinkage method. Response variables were daily rides and average trip duration. Predictions of sharing system usage, if no COVID happened, were made using generalized additive model. Afterwards, predictions and actual values were compared to estimate the effect of COVID-19 on Oslo system and suggestions for improvements of Oslo City Bike were given.

Weather data and bicycle usage data were obtained from the open sources managed by the official providers of this information such as Norwegian Meteorological Institute and Oslo City Bike respectively. R was used to analyze this pool of data. Data were divided for train (pre pandemic period) and test (pandemic period).

Based on conducted analysis, it was concluded that overall daily rides on shared bikes in Oslo decreased and average trip duration increased due to COVID-19. Number of rides also could have been affected by the appearance of massive amount of electric scooters on the streets of Oslo in 2019-2021.

Oslo bicycle sharing system was able to satisfy the decreased aggregate demand for rides, but in terms of satisfying the demand for longer trips the supply side was limited to Ring 3 (road #150) of Oslo. This could be improved by rearranging placements of existing stations at a lower total cost or by expanding the bicycle sharing system's coverage area with new bikes and new stations at a higher total cost.

Acknowledgments

This thesis was written as part of master's degree at the Norwegian School of Economics in Autumn 2022. Writing this thesis was challenging and interesting as it helped to practice applying knowledge acquired during studying at NHH to a project of a bigger scope. I started studying at NHH with zero experience in programming and now while writing this thesis I was able to do the analysis of millions of data using software.

I would like to thank my supervisor, Julio Cesar Goez, and my family for giving me useful points for thought and needed support during writing this thesis.

Norwegian School of Economics

Bergen, December 2022

Annel Meirkhanova

Table of contents

List of figures

List of tables

1. Introduction

COVID-19 pandemic has influenced many aspects of everyday life for everyone all over the world. National and local rules and restrictions affected daily operations of businesses, educational institutions, governmental bodies, public transport, etc. As a part of public transport Oslo City Bike (bicycle sharing services in Oslo, Norway provided by UIP Oslo Bysykkel AS) should have been affected by the pandemic also. Common sense suggests that people (specifically those who live in the zones of operation of shared bikes) would prefer riding bikes in the open air to using closely confined buses, trams and metro carriages. Also, duration of the trip should be longer as one would like to go further distances than before to avoid crowds in other types of public transport.

But at the same time one can argue that going online (work from home and online education) could influence public bicycle usage negatively. In addition, temporal closure of sport and culture facilities, non-essential shops, cancellation of sport and cultural events reduced level of mobility everywhere and Oslo is no exception.

There could be simultaneously an increase in demand for bikes due to the travelers' shift from other (riskier in terms of spread of disease) modes of transport and a decrease due to increased home offices / online schooling / lockdowns. Overall change in shared bikes usage in Oslo is to be estimated in this thesis.

Total number of daily rides and daily average ride duration were used as indicators to assess bike sharing system usage. Based on previous studies in various cities, the pandemic effect and the method to measure it were different in each location. As for Oslo, in order to estimate the impact of COVID the author followed two step approach. First, pre-pandemic data were used to build a correct model. Data was gathered from open data sources and analyzed using R. Secondly, a comparison was made between predicted by this model values (as if there was no pandemic) with actual pandemic bicycle usage values.

The author expects total number of daily rides to decrease (as the work and study from home and governmental restrictions decreased mobility of people and that impact was greater than shift from other types of public transport out of fear of getting a disease) and average trip duration to increase (because commuters are willing to travel for longer distances to avoid catching the disease, for example, if earlier pre-pandemic someone would just travel from bike station 1 to bike station 2 which is closer to metro, nor the same user could travel from to the final desired destination by bike instead of metro).

Thus, thesis hypotheses could be stated as follows:

Hypothesis 1: Number of daily rides on Oslo's shared bikes decreased during COVID-19, which could be explained by decreased overall mobility caused by the pandemic and by uncontrollable inflow of electric scooters.

Hypothesis 2: Average trip duration on Oslo's shared bikes increased during COVID-19, meaning that commuters were willing to travel for longer time and distances and replaced their previous preferred public transport choice by shared bikes.

With a fleet of 3 000 bicycles Oslo City Bike should have been able to meet the demands of the users as total trips decreased, but users' willingness to travel for longer distances could have been limited by Ring 3 (limit of operation of Oslo City Bike, Ring 3 is not covering the whole city of Oslo). That means if the author's hypothesis is confirmed, Oslo City Bike should rearrange existing bikes' stock and expand area of operation beyond Ring 3. This could also help achieve other goals such as creation of additional mode of public transportation for Oslo residents outside Ring 3 and increasing preparedness of Oslo City Bike for possible new pandemics.

In the next chapter overview of bicycle sharing system in Oslo and timeline of COVID in the Norway will be given. Next, previous findings in the field (general factors influencing public bicycle usage and studies of pandemic effect on bicycle sharing systems in various cities) will be discussed. After that in the third part, overview of the data and model will be given. In Chapter 4 results will be presented and discussed. Finally, Chapter 5 will conclude the thesis and the author will discuss ideas for future work.

2. Literature review

In this section, first, brief history and description of bike sharing situation in Oslo as well as timeline of the pandemic in Norway and COVID statistics for Oslo will be provided. Then, previous studies on factors influencing shared bicycles usage (specifically weather factors) and previous studies on pandemic impacts on bicycle sharing system in various cities will be discussed.

2.1. Bicycle sharing system in Oslo

Bicycle sharing system (hereafter BSS) is defined as a "network of public use bicycles distributed around the city for use at a low cost" (New York City Department of City Planning, 2009). Shaheen et al. (2010) defines bike-sharing as a "short-term bicycle access which provides its users with an environmentally friendly form of public transportation", and which allows individuals to "use bicycles on "as-needed" basis without the costs and responsibilities of bike ownership".

First BSS appeared in Amsterdam, Netherlands in the mid-1960s and now in 2022 bike sharing is available in over 3000 cities worldwide (O'Sullivan, 2022).

As for Oslo, information about BSS prior to 2016 is almost nonexistent. In April 2015 Oslo City Council adopted Bicycle Strategy 2015-2025. According to this strategy in 2013 8% of commuters used bikes as their mode of transport for everyday travel and one of the main goals of this strategy is to double that number to 16% by 2025. One of the measures to fulfill this strategy is measure 1H: "expanding the bike sharing system". So, there was definitely a BSS running in Oslo prior to 2016 which has to be expanded. Then in April 2016 updated BSS was launched with 72 stations, 770 bicycles and 1400 docks with a new mobile application and website (Knudsen, 2016). By November 2022 Oslo's BSS (Oslo Bysykkel) increased to 251 stations, 6000 bikes and 6000 locks (UIP, 2022).

According to Global Bicycle Cities Index 2022 (Luko, 2022) Oslo's BSS is on 55th place. Ranking is based on six categories: weather, percentage of bicycle users, crime & safety (fatalities and accidents per 100 000 cyclists, bicycle theft score), infrastructure (specialized bicycle roads and road quality, number of bicycle shops per 100 000 cyclists, investment and infrastructure quality), sharing (number of sharing and rental stations and number of shared bicycles divided by 100 000) and events (no car day, critical mass score). From Table 1 it can be observed that Oslo outperforms number one city in the ranking (Utrecht) in crime& safety score and sharing score (sharing here includes not only BSS, but also rental shops), in all other categories Oslo's scores are lower than Utrecht's.

Table 1. Extract from Global Bicycle Cities Index 2022

Ranking	City	Country	Weather score	Bicycle usage	Crime and Safety Score	Infrastructure score	Sharing score	Events score	Total score
1	Utrecht	Netherlands	63.83	51%	82.46	57.51	17	279.88	77.84
55	Oslo	Norway	47.45	7%	90.79	30.31	45	226.4	31.31

Also, it is worth mentioning that weather score was derived using "total annual hours of sunshine, average annual precipitation in millimeters and number of days with temperature below 0°C and above 30°C (Luko, 2022). Key point to emphasize here is that weather was used in estimating the city ranking. In this thesis the weather data will also be used to analyze bike sharing numbers.

Another mode of shared (similar to BSS) transport, which is quite popular in Oslo, is electric scooters. E-scooters appeared during years 2018-2019. In 2019 there were 5 000 units on the streets of Oslo, by summer of 2021 total number of shared e-scooters rose to around 20 000 (Myhre & Uglum, 2021). Due to the problems associated with use of e-scooters (increased traffic safety concerns for scooters' drivers and other road users, reduced accessibility for disabled and elderly), since September 2021 local authorities had intervened in the operation of sharing-scooters program and had established a limit of 8 000 units to be available for shared renting (Oslo commune, 2021a).

There are a lot of similarities between shared bikes and shared scooters: both are types of shared public transport, both have easy-to-entry systems for users (simple registration in user-friendly mobile application), both are modes of transportation free of CO2 emissions. One major difference is that shared bikes are dock-based (meaning that users have to take and to return bikes to specific bike stations), whereas scooters are stationless and could be dropped off at almost any location.

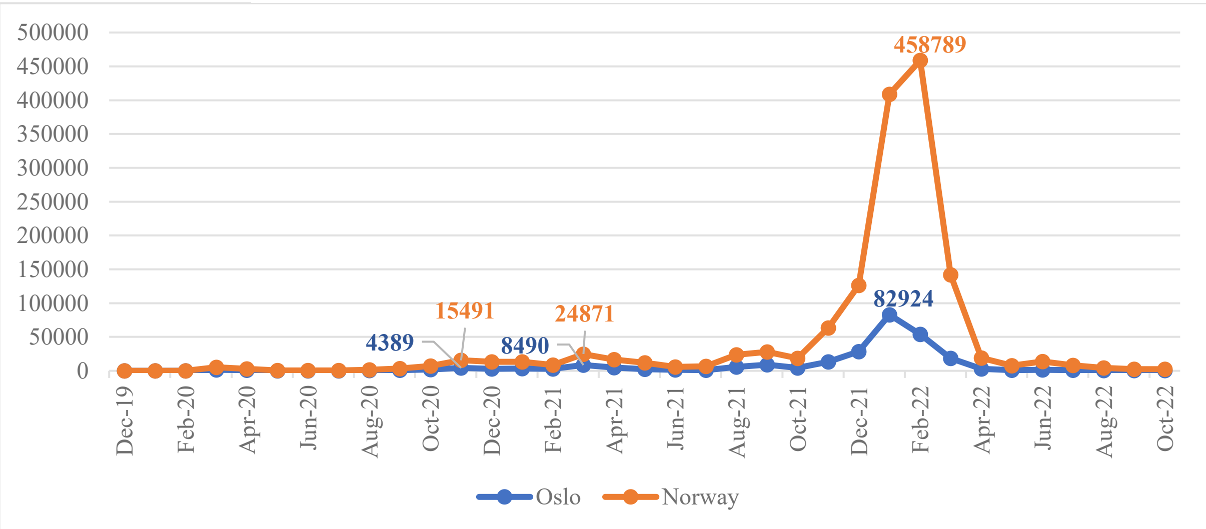
Even without the pandemic, it was expected that scooters would impact usage of Oslo shared bikes and would lure away some portion of Oslo BSS users because these two services are almost perfect substitutes, except for the prices and the level of physical effort needed to move the transport. Based on the information provided on the official website and applications of the service providers, annual pass for Oslo City Bike costs 549 NOK (around 46 NOK per month), monthly – 79 NOK, both with unlimited unlocks and free rides up to 1 hour, whereas fares for using electric scooters are higher: monthly pass costs 299 NOK (provider – Voi) or 389 NOK (provider – Bolt) just for 300 minutes a month (around 10 minutes a day).

Peak of scooters popularity among users and caused by that uncontrolled inflow of them on the streets happened exactly during analyzed pandemic years of 2020-2021. At the maximum there were 3 000 bikes and 20 000 (Nikel (2021) states that number was around 30 000 units) electric scooters in Summer 2021. This means that definitely part of decrease in demand for shared bikes during pandemics could be explained by appearance and rising popularity of e-scooters.

2.2. COVID-19 in Oslo

In the end of 2019 COVID pandemic began. On 11 March 2020 World Health Organization (2020) declared COVID outbreak a pandemic.

Figure 1. COVID cases overtime in Oslo and Norway



In February 2020 first cases were registered in Norway (Folkehelseinstituttet 2022a) and in Oslo (Folkehelseinstituttet 2022b). Figure 1 (source – https ://statistikk.fhi.no) presents monthly number of cases in Oslo and Norway.

Here is a short timeline of major governmental actions to fight the pandemic in Norway:

On 12 March 2020 Norwegian government introduced for the first time restrictions on the national level (Lokkevik et al., 2020). Schools and universities were closed. Cultural and sports events were prohibited. Restaurants and bars were not operating. Restrictions (at times eased, at times tightened) continued throughout 2020 and beginning of 2021.

In April 2021 gradual reopening plan was presented (The Office of the Prime-Minister, 2021, April 10). Since 16 April 2021 first step of reopening plan had started to be implemented.

On 25 September 2021 Norway moved to “normal everyday life with increased emergency preparedness” (The Office of the Prime-Minister, 2021, September 25).

However, in November-December 2021 some restrictions were brought back due to spread of omicron variant.

Finally, on 12 February 2022 all infection control measures (including face mask requirements, one meter distance, etc.) were removed (The Office of the Prime-Minister, 2022, February 12).

In this thesis there are two major time periods: pre-covid and covid. Pre-covid period is before 11 March 2020 (last day before national measures were announced). Covid period is from 12 March 2020 (date when first national measures against COVID were imposed) to 12 February 2022 (date when all infection measures were dropped).

2.3. Literature review on dependent and independent variables

A lot of factors could influence bicycle usage in general and specifically shared bicycle usage. Undoubtfully, weather conditions should play a significant role in individuals' decision whether to use a bike. No one would like to ride a bike in heavy rain or during windy day. Below previous research studying relationship between weather conditions / calendar events and BSS usage are discussed. Analyzing this research will help the author choose candidates for independent variables, then after running tests selecting the final ones.

Corcoran et al. (2014) studied the effect of weather and calendar events on public bicycle sharing program on the example of Brisbane's CityCycle. Effect of each factor was estimated separately (using two-sampled Poisson test) and simultaneously (multivariate regression). To estimate the simultaneous effect, multivariate Poisson regression was used. Dependent variable was log of average number of daily trips per station. Independent variables were three weather variable (temperature, rainfall and wind speed) and three dummy calendar events variables (weekend, school holiday, public holiday). Regression showed that public holiday, weekend, and warmer temperatures had significant positive effect, whereas rainfall and higher wind speed had significant negative effect on number of trips. Also, the authors found that "temperature was the only independent variable found to possess non-linear properties in its relationship to the dependent variable" (Corcoran et al., 2014).

Gebhart and Noland (2014) estimated the impact of weather conditions on bike sharing in Washington. Contrary to other research, the authors used two different models for different dependent variables: a negative binomial model for number of trips and an ordinary least squares regression for average trip time. Weather data (temperature, humidity, wind speed, fog, rain, thunderstorm, snow, darkness), dummy variables for month/weekend/federal holiday/peak travel time, and number of stations in BSS were used as independent variables. Gebhart and Noland (2014) found that "very cold temperatures, rain, high humidity and increased wind speed" decreased number of trips and shortened average trip duration. Those authors also believe that relationship between temperature and bicycling behavior is not linear.

Eren & Uz (2020) reviewed different studies on station-based bike sharing to derive common factors which influence trip demand. Those factors were divided into six main categories: weather, built environment, public transport, socio-demographic attributes, temporal factors, and safety. In the scope of this thesis weather category is of the most of interest. The authors highlighted that temperature is "one of the most investigated factors" to affect bike-sharing demand and that there is a positive correlation between temperature and demand. The authors concluded that adverse weather conditions (such as rain, strong wind, high temperature (above the maximum at the BSS location) and humidity) decrease trip demand.

Flynn et al. (2012) estimated the effect of weather on individual decision to ride a bike to work. Study was conducted in Vermont, USA, on 183 adults. Dependent variable was binary variable with 1 if the person biked and 0 if the person didn’t bike that specific day. Independent variables were mean temperature, mean wind velocity, precipitation, snow depth, daylight, distance, age and gender. The results have shown that precipitation and temperature increased (consistent with other research) the odds of bicycle commuting, whereas wind speed (contrary to the other research analyzed by the authors) decreased the probability of using a bicycle modestly.

The purpose of looking at all this research is to understand what could influence BSS usage and how it was estimated, which models were used. Analysis of applicability of these approaches to this thesis was done in Section 3.5.

Thus, based on the analysis of previous studies, weather conditions are playing key role in estimating bike sharing usage. In this thesis, weather parameters such as average daily temperature in °C, precipitation in millimeters, wind speed in meter per second and calendar event (weekend or workday) are candidates to be used as explanatory variables to analyze bicycle-sharing numbers in Oslo. A test will be run in Section 3.3. to select which variables should be used in the model.

As for measure of BSS usage, there will be two different dependent variables: daily total number of trips and daily average trip duration. Most of the previous studies used those variables as dependent and even without this fact it is logical to assess BSS by number of trips and trip duration.

2.4. Literature review on the effect of COVID-19 on bicycle sharing systems across the world

Worth mentioning is another analogous research. There have been several studies on measuring effect of COVID-19 pandemic on bicycle sharing systems in various cities across the world.

For example, Padmanadhan et al. (2021) measured the impact of pandemic on shared biking in New York, Boston and Chicago. The authors used correlation coefficients analysis to understand relation between number of COVID cases and trip frequency and average trip duration during different stages of pandemic (pre covid, uphill and downhill in cases). They found that there was a negative correlation between COVID cases and number of trips and a positive correlation between cases and average trip duration in all three cities.

Another study is based on London bicycle sharing system. Heydari et al. (2021) estimated the effect of pandemic over March-December 2020 in London using second order random walk time series model and found that number of rides decreased in March and April 2020, then increased in May 2020 to expected (without pandemic) level which showed London BSS resiliency. Average hire time was substantially higher than predicted in April, May and June 2020 and afterwards followed a predicted pattern.

Berezvai (2022) estimated short-term and long-term effects of COVID on Budapest BSS (MOL Bubi). In the fixed effect panel regression, the author used change between 2019 and 2020 in daily number of trips and average duration of trips per specific station as dependent variables and change in government restrictions (2019 – no restrictions, 2020 – measured as the Government Stringency Index based on Oxford COVID19 Government response tracker dataset), change in weather (daily temperature, daily precipitation, daily wind speed) and change in traffic as independent variables. Workday and weekend data was analyzed separately. Berezvai (2022) found that in the short run government restrictions in the beginning of pandemic increased bicycle usage substantially, however in the long run when some of the restrictions were lifted, bicycle sharing dropped (despite the fact that fares decreased).

Based on those studies, it can be observed that COVID-19 influenced BSS systems across the world in different ways. For some cities there was an increase in usage, for some – a decrease. The author does not have any knowledge of analogous study on Oslo BSS, and the effect of pandemic on bike sharing in Oslo will be analyzed in this thesis.

3. Data, method and model

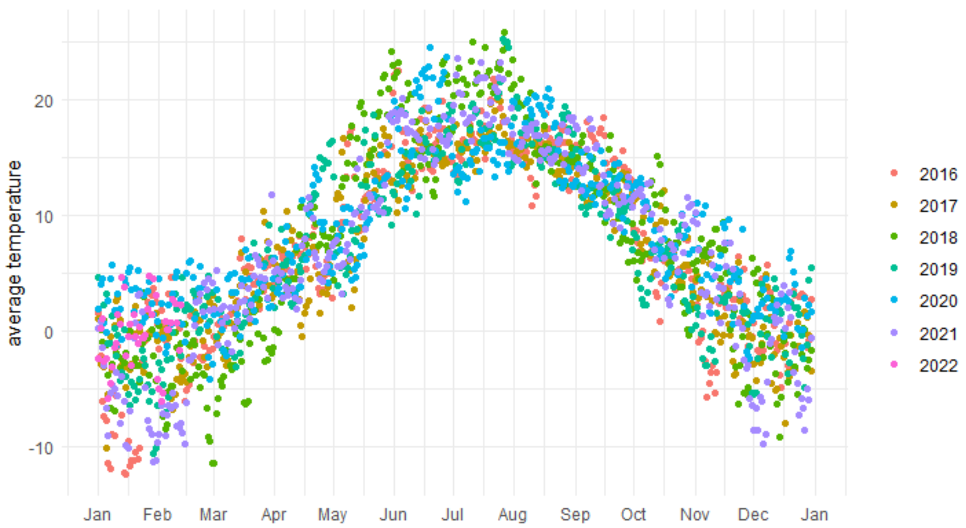
3.1. Data and data analysis

3.1.1. Weather data and calendar events

Weather data were scrapped from yr.no website (official website of Norwegian Meteorological Institute) for the period of 7 years (January 2016 – February 2022) on daily basis for Oslo. Chosen meteorological station Oslo (Blindern) is located at the following coordinates 59°56’32.3”N and 10°43’12.0”E and is exactly inside the coverage area of Oslo BSS.

Average temperature in °C, precipitation in mm and wind speed in m/s for each day were gathered in one weather dataset.

Figure 2. Average temperature in Oslo over Jan 2016-Feb 2022



From Figure 2 it can be observed that daily average temperature in Oslo follows almost the same pattern over the years. In those years mean average temperature was always in the range of 7.00-7.30 °C, except hotter 2018 (7.81°C) and 2020 (8.73°C).

Table 2. Descriptive statistics of daily weather data (explanatory variables)

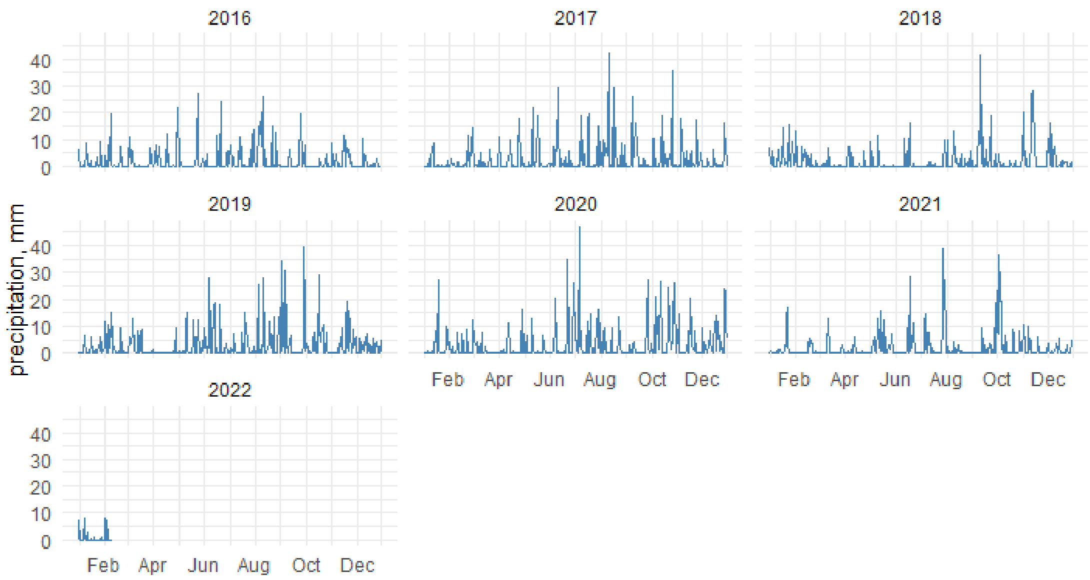
		2016	2017	2018	2019	2020	2021	Jan-Feb 2022
Average temperature, °C	mean	7.29	7.00	7.81	7.32	8.73	7.31	-0.31
	st dev	8.01	7.24	9.36	7.54	6.45	8.75	2.80
	max	22.60	20.40	25.90	25.30	24.60	23.60	4.80
	min	-12.40	-10.10	-11.40	-10.60	-4.20	-11.30	-6.00

Table 2 (cont.)

		2016	2017	2018	2019	2020	2021	Jan-Feb 2022
Precipitation, mm	mean	1.98	2.59	1.80	2.87	2.95	1.91	0.93
	st dev	4.23	5.55	4.40	5.61	6.03	4.75	2.26
	max	27.00	42.20	41.50	39.60	47.00	39.00	8.20
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wind speed, m/s	mean	2.72	2.64	2.74	2.76	2.85	2.68	2.52
	st dev	1.22	1.17	1.14	1.13	1.17	1.10	1.19
	max	6.90	8.00	8.10	7.10	8.40	7.70	5.20
	min	0.70	0.70	0.70	0.80	0.70	0.80	1.00

Controrary to the average temperature, daily precipitation varies over the years with different peaks and lows each year. Minimum mean daily precipitaion at 1.80 mm with standard deviation of 4.40 mm was observed in 2018, whereas maximum was at 2.95 mm with standard deviation of 6.03 mm in 2020 (Figure 3).

Figure 3. Daily precipitation (mm) in Oslo over Jan 2016-Feb 2022



As for the wind speed, 2.64-2.85 m/s was average daily wind speed in Oslo throughout January 2016-December 2021. Standard deviation was also stable over the years and in the range of 1.10-1.22 m/s. Minimum wind speed was always around 0.70-0.80 m/s throughout 2016-2021, and maximum was in the range of 6.90-8.40 m/s (see graph in Appendix 1).

Calendar event variable was created as a binary variable with value 0 for working days and 1 for weekend days. In the whole data there were 1591 working days and 638 weekend days.

3.1.2. Bicycle sharing system data

Data on Oslo BSS is publicly available and was obtained from service provider's official website (Oslo City Bike, 2022). Original BSS data was divided by the provider in two periods (not connected to COVID): data before April 2019 and data after April 2019. Data before April 2019 contained the following information about individual trips: 1) start time, 2) end time, 3) start station id, 4) end station id. Data after April 2019 contained the same information as above plus 5) duration, 6) start station name, 7) end station name, 8) start station description, 9) end station description, 10) start station longitude and latitude, 11) end station longitude and latitude. Provider states that all recorded trips were minimum one minute long, but this was not always the case. During tidying up the data, trips with duration of less than one minute were removed.

Pre-covid data frame was composed by combining data from before April 2019 (2016-April 2019 7.75 million rides) and after April 2019 (April 2019 - March 2020 2.25 million rides), a total of almost 10 million unique trips. During the process of cleaning data there were found trips with negative values for duration (possibly due to technical mistakes) and with duration longer than 24 hours (not relevant to analysis as weather changes every day). Those trips (212 793 trips) were removed from the final datasets. Final pre-pandemic dataset consists of data on 9 780 271 unique trips.

Covid data frame covered a period from 12 March 2020 (first national anti-COVID measures were imposed) till 12 February 2022 (all infection control measures were lifted) and is composed of 3.14 million unique rides. In the same manner as with pre-pandemic dataset, 19 trips with duration longer than 24 hours were also removed. The author observed that quality of data on Oslo BSS definitely improved over the time, for example, in the later data (after April 2019) there were no negative values for duration and substantially less trips with duration longer than 1 day (which of course could be explained by commuters' personal choices, but could be also caused by fewer technical mistakes/bugs in bicycle sharing IT system).

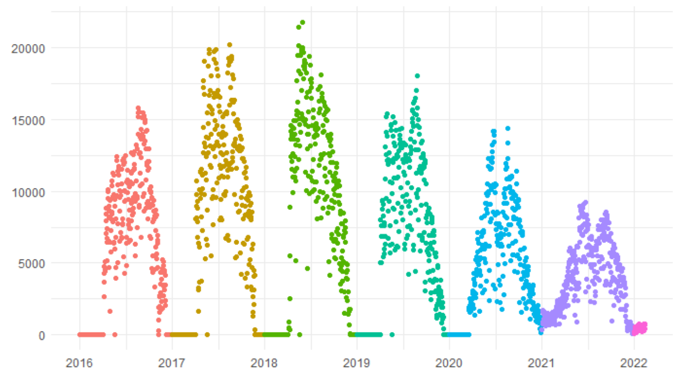
Both datasets were aggregated to daily basis, total number of rides per day and average trip duration during that day were calculated. Pre-pandemic and pandemic datasets were considered as train and test datasets respectively. Final pre-covid train dataset consists of 1528 days and covid test dataset of 701 days.

Table 3. Descriptive statistics of number of trips and trip duration (dependent variables)

	Year	Total number of rides	Daily number of rides		Average daily trip duration in seconds	
			mean	st dev	mean	st dev
Pre-COVID	2016	2 090 165	5 711	4 909	868	671
	2017	2 654 614	7 273	6 523	487	370
	2018	2 790 712	7 646	6 794	522	383
	2019	2 244 780	6 150	5 417	494	356
	Jan-11 Mar 2020	0	0	0	0	0
COVID	12 Mar-Dec 2020	1 696 039	5 749	3 528	766	181
	2021	1 419 280	3 888	2 462	745	114
	Jan-12 Feb 2022	21 778	506	187	690	89
	Total Pre-COVID	**9 780 271**	**6 384**	**6 033**	**565**	**495**
	Total COVID	**3 137 097**	**4 462**	**3 192**	**750**	**146**

There was a constant annual increase in trips on shared bikes since the introduction of updated version of BSS in Oslo in 2016 (Figure 4). Then, in 2019 there was a decline in total number of rides which could be explained by appearance and rising popularity of electrical scooters (discussed in Section 2.1.). Usually BSS was not operational during January-March of each year, but since 2021 there were 1 000 available bikes with spiked tires (UIP, 2022).

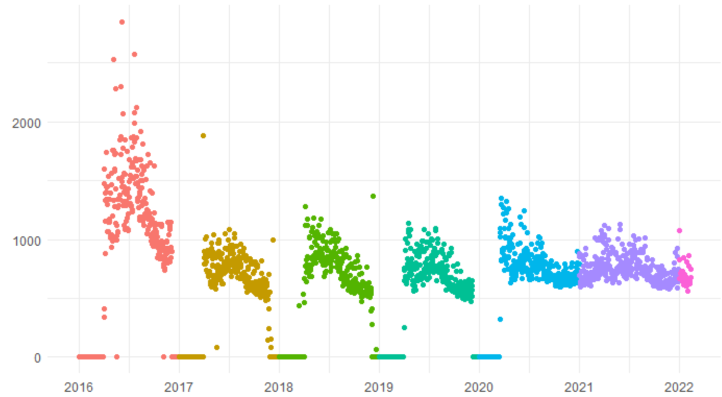
Figure 4. Daily number of rides on BSS over 2016-2022 in Oslo



In absolute numbers, total number of BSS rides dropped during pandemic: from 2.2 million rides in 2019 to 1.7 million rides in 2020 and even further to 1.4 million rides in 2021.

As for average trip duration, it was higher in 2016 with mean of 868 seconds per one trip (could be caused by introduction of new updated BSS with new user-friendly app and riders desire to try it), then it stabilized and followed the same pattern in years 2017, 2018 and 2019 with mean around 487-522 seconds per trip. In 2019 average trip duration was not affected to the same extent by the appearance of e-scooters as was total number of rides.

Figure 5. Average trip duration (in seconds) on BSS over 2016-2022



During COVID mean of average trip duration increased from 494 seconds to 766 seconds per one trip. This could be explained by BSS users' desire to travel longer distances to avoid other types of public transport (buses, trams, metros).

Additional candidate explanatory variable is binary variable with value 1 if the ride was taken during weekend (Saturday and Sunday), 0 if the ride was taken during working days (Monday – Friday). Historically mean daily number of rides was smaller on weekends than on working days (could be explained that Oslo residents used shared bikes mostly to travel to offices on workdays), whereas average trip duration was longer on weekends (possibly, bicycles were used longer for recreational use on weekends).

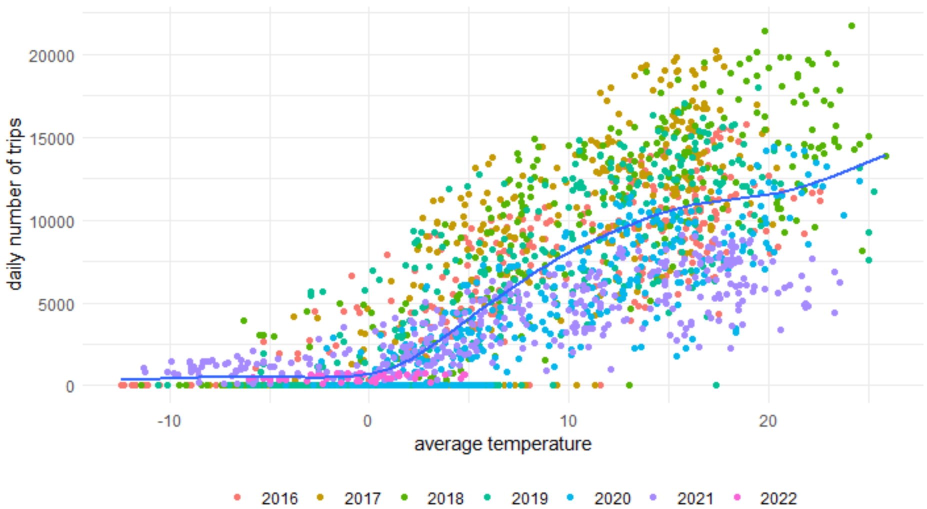
Table 4. Comparison between number of rides and average trip duration on workdays and weekend

Year	Daily number of trips (mean)		Average duration (mean)	
	Workdays	Weekend	Workdays	Weekend
2016	5 925	5 186	866	873
2017	7 313	7 174	475	516
2018	7 745	7 397	515	540
2019	6 366	5 609	483	522
2020	4 699	4 470	612	632
2021	3 925	3 797	735	768
Jan-Feb 2022	525	463	678	718

The same pattern as before is observed here for both type of days: for number of trips increase from 2016 to 2018, drop in 2019 and then decrease since start of pandemic in 2020; for average duration peak in 2016, stable 2017-2019, increase since start of pandemic (Table 4).

As for weather parameters, simple graphical analysis shows that there is positive (as expected) correlation between temperature and number of daily rides (Figure 6). Eren & Uz (2020) states that for the temperatures in the range of 0-20 °C, there is a positive correlation between number of rides and temperature; for temperatures in the range of 20-30 °C BSS demand is at maximum; for temperatures above 30 °C studies are different (for some citites there is an increase, for some decrease in number of rides). For Oslo maximum average temperature thorough 2016 – 2022 was never above +30°C , it was +25.9°C and minimum was –12.4°C (and usually BSS was closed during January-March, except 2021-2022).

Figure 6. Relationship between average temperature and number of daily rides over the years



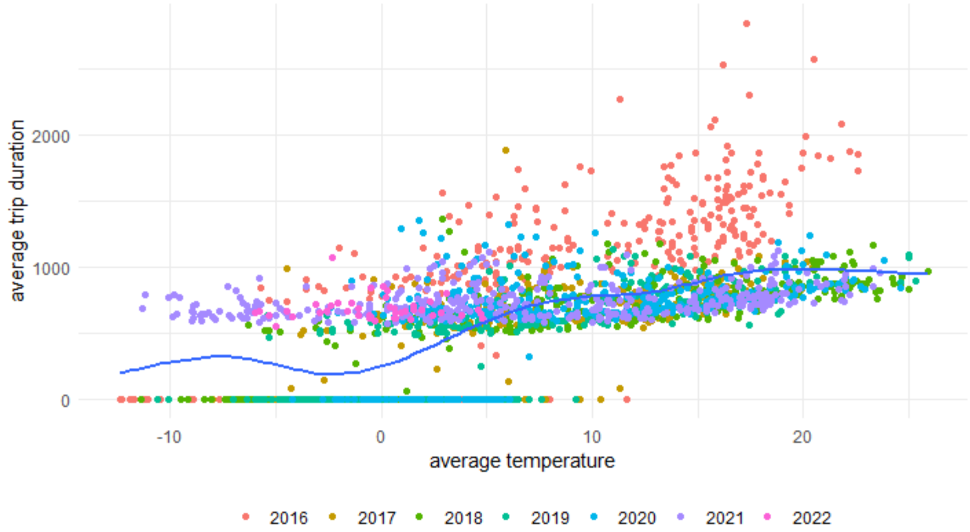
Each dot represents one specific day. Blue line on Figure 6 and Figure 7 is a smoothed regression line added using geom_smooth() function in R. Method for constructing this curve is LOESS for data containing less than 1000 observations and GAM for more than 1000 observations (which is the case for Figure 6 and 7) (RDocumentation, n.d.a).

From figure 6 the author observes that for temperatures below 0 °C number of daily rides does not change a lot due to change in temperature (keeping in mind that system was not operational during winter months in 2016-2020), whereas for temperature between 0 to 10 °C there is a sharp increase in rides with increase in temperature. For temperatures between 10 and 20 °C there is also an increase in rides, but to a lesser extent than before and even smaller increase for temperatures above 20 °C. From this graph (and regression smooth line), there is a clear non-linear relationship between temperature and daily rides.

As for relationship between average trip duration and average daily temperature (Figure 7), it is not as steep as with daily rides, but still positive, meaning the higher the temperature the longer the rides. The pattern for temperature ranges “below 0 °C”, “0 to

10°C", "10 to 20 °C" and "above 20°C" follows the above-described pattern for daily rides, with the exception that for range "above 20°C" it looks almost flat, meaning trip duration does not change with increase in temperature and for "below 0 °C" category blue regression line looks more curvy than the same one for daily rides. This also shows a non-linearity between temperature and trip duration.

Figure 7. Relationship between average temperature and trip duration over the years



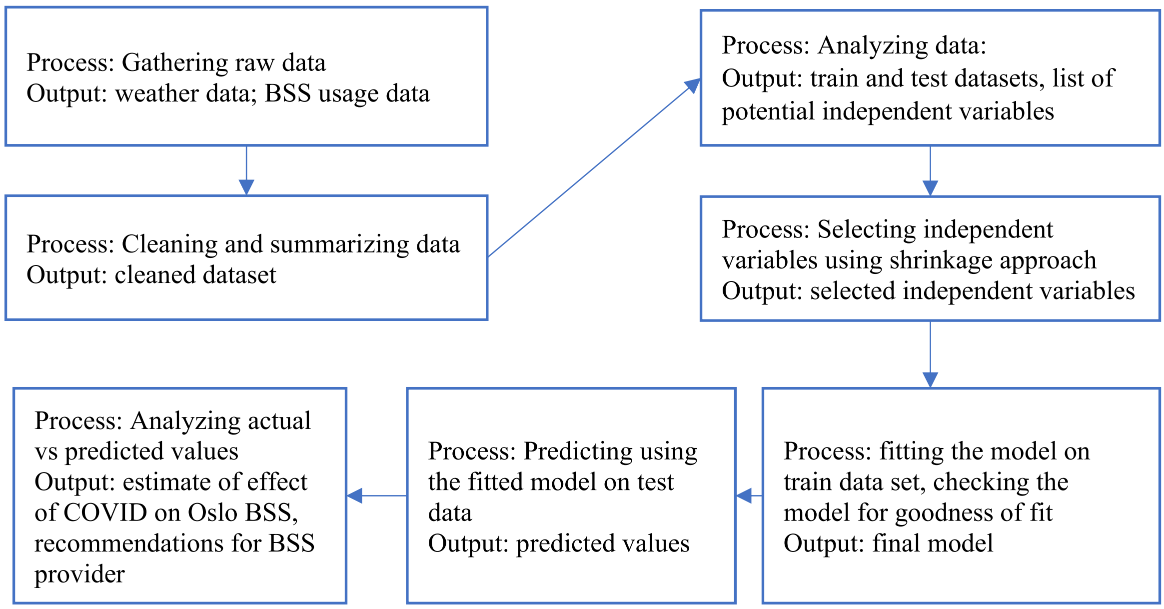
As for precipitation and wind speed, the relationships between those weather variables and dependent variables (number of trips and average trip duration) are quite ambiguous from the graphs (see Appendix 2). Ride duration seems to be unaffected by change in precipitation which is surprising. Generally, those blue regression lines gave the author just a preliminary understanding about the relationship between dependent and independent variables and significance of these variables in explaining Oslo BSS usage will be tested later in Sections 3.3. and 3.5.

3.2. Method

Weather conditions impact BSS usage (measured by number of trips and average trip duration) significantly. In this thesis, daily total number of trips and daily average trip duration were used as dependent variables. Weather data in the form of average daily temperature in °C, precipitation in mm and wind speed in m/s were candidate variables to be used as explanatory variables. In addition, calendar event (weekend or workday) also could play a role in explaining BSS usage, so binary variable "weekend" was also tested as candidate as independent variable in the analysis.

After analyzing data and defining a list of potential explanatory variables, selection of independent variables from this list was done using shrinkage method LASSO (will be described in detail in Section 3.3.).

Figure 8. Schematic view of the approach used in the thesis



Next, train pre-pandemic data were fitted to the model. The model was checked for goodness of fit. Then using the estimated model, predictions were made using the test pandemic dataset. Those predicted values were compared with actual values to estimate an impact of COVID on Oslo bike sharing system. Based on those results, recommendations to Oslo BSS provider were made.

RStudio version 2022.07.2 Build 576 was used to analyze data and for modelling.

3.3. Selection of independent variables

After analyzing previous research, it was found that weather data should play a key role in estimating BSS usage. But those studies were estimated not on Oslo. Thus, in order to choose correct explanatory variable a test was conducted on Oslo data.

The author chose average temperature, precipitation, wind speed and factor of weekend as candidates for independent variables.

Next, a shrinkage method LASSO was used to select variables. LASSO (stands for least absolute shrinkage and selection operator) is a technique that minimizes *RSS (residuals sum squared)* + $\lambda \sum_{j=1}^{p} |\beta_j|$ (where λ is a tuning parameter) by shrinking the coefficients estimates towards zero (James et al., 2021). Thus, this method works as variable selection method. Variables, which have coefficient estimates approaching zero, should be removed.

First, optimal λ was estimated using leave-one-out cross validation approach (which includes dividing the data into two parts: one single observation is used as a validation set and all the remaining observations are used as a training set, cross-validation errors for different values of λ are calculated, optimal λ is λ with the smallest error (James et al., 2021)), then a model was fitted using this optimal lambda.

Table 5. LASSO results for variable selection

Dependent variable	**Daily number of rides**	**Daily average trip duration**
(Intercept)	5360.38675	5543.4511
average	499.72393	507.40768
precipitation	-73.97429	-82.38957
wind speed	.	-37.4588
weekend	-705.31029	-815.5364

From LASSO shrinkage results (table 5), coefficient estimate on wind speed for dependent variable “daily rides” is zero, so this variable can be removed from further modelling. For response variable “average trip duration” all tested explanatory variables will be fitted in modelling.

Independent variables: multicollinearity

To check for potential multicollinearity between independent variables, correlation coefficients were calculated. From the correlation matrix it could be derived that none of the regressors is in linear dependency with another.

Table 6. Correlation matrix between independent variables

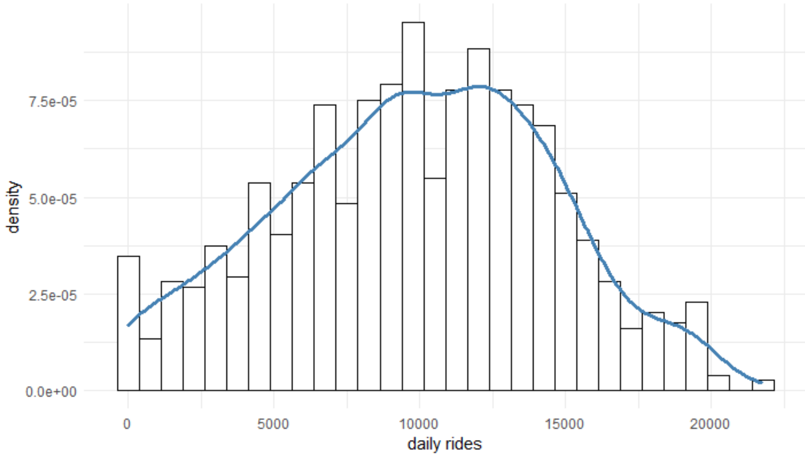
	average temperature	precipitation	wind speed	weekend
average temperature	1.000	0.006	-0.013	0.012
precipitation	0.006	1.000	0.056	0.0003
wind speed	-0.013	0.056	1.000	0.009
weekend	0.012	0.0003	0.009	1.000

3.4. Analysis of dependent variables

3.4.1. Dependent variable: daily rides

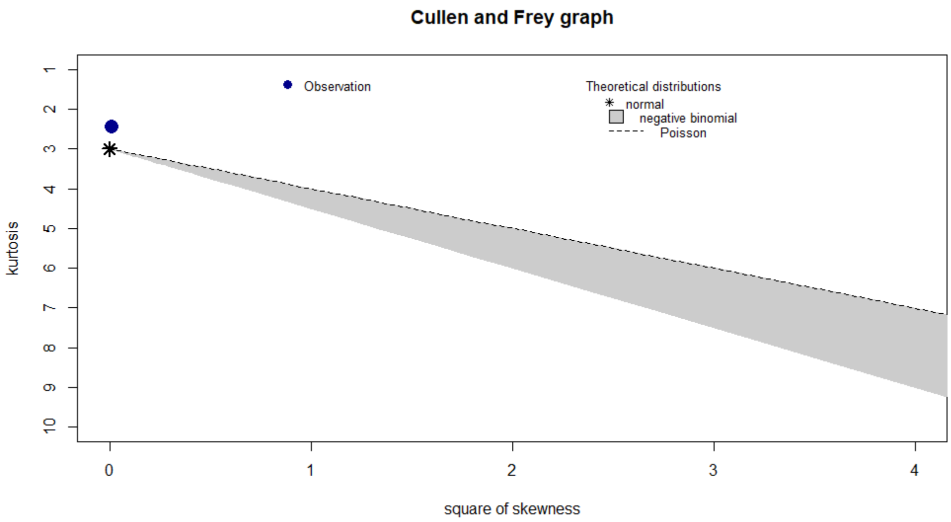
Since daily rides are positive and discrete numbers and represent trips per unit of time (day), one would expect a Poisson distribution for daily rides (figure 11). This was checked using the software.

Figure 9. Histogram of daily rides



Using descdist() function from fitdistrplus() package, distributions of variables could be found. Descdist() function "computes descriptive parameters of an empirical distribution for non-censored data and provides a skewness-kurtosis plot" (Rdocumentation, n.d.b). For discrete data Poisson, negative binomial and normal distributions, and for non-discrete data uniform, normal, logistic, lognormal, beta and gamma distributions are considered.

Figure 10. Fitting distributions to "daily rides" variable



Surprisingly, dependent variable "daily rides" does not have a Poisson distribution, in fact none of the tested distributions was of a good fit (Figure 12). Daily rides vary from minimum of 1 trip to a maximum of 21 thousand trips per day, with a mean of 9 849 and a standard deviation of 4 690.48 or variance of 22 000 600. Indeed, it is confirmed again that the distribution is not a Poisson since variance exceeds mean by a lot.

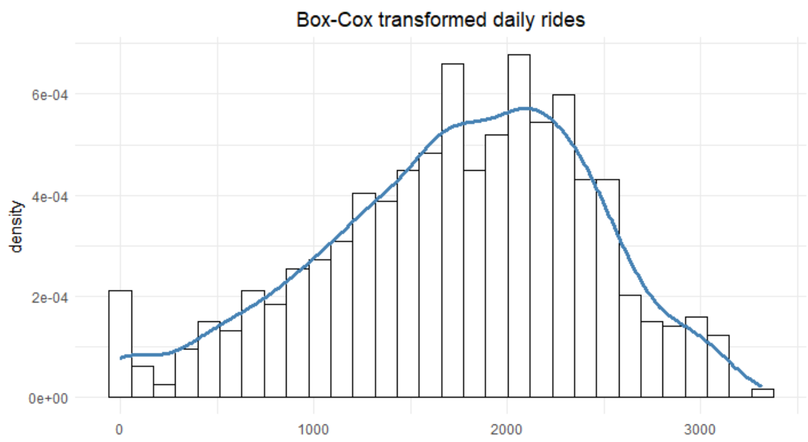
Since the test for the distribution gave no specific result and to increase the chance of normality of residuals, the author decided that a Box-Cox transformation of the daily rides need to be done.

A Box-Cox transformation is a transformation which depends on parameter λ and is defined as:

$w_t = \begin{cases} \log(y_t) \ if \ \lambda = 0, \\ \frac{y_t^{\lambda}-1}{\lambda} \ otherwise. \end{cases}$ where $y_1, y_2, \ldots y_T$ are original observations, $w_1, w_2, \ldots w_T$ are transformed observations (Hyndman & Athanasopoulos, 2018)

R can compute an optimal lambda for the data, and for daily rides it was 0.788. Histogram of transformed daily rides is presented in Figure 13. There is a slight change in right tail comparing with original histogram (figure 11).

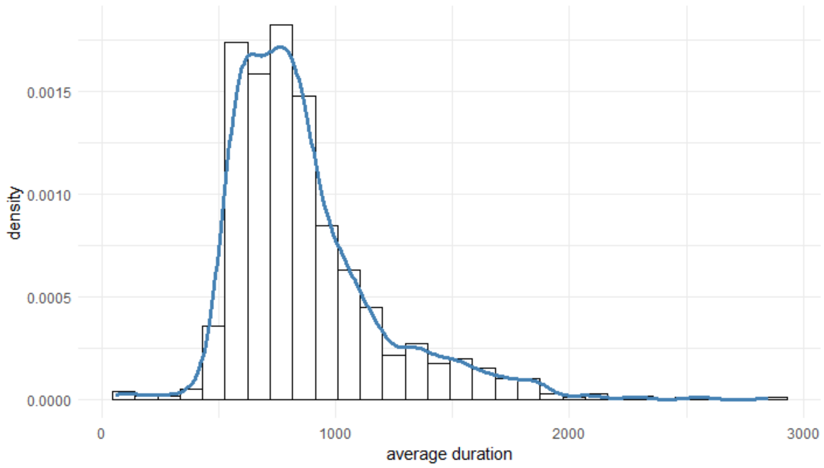
Figure 11. Histogram of Box-Cox transformed daily rides



3.4.2. Dependent variable: average trip duration

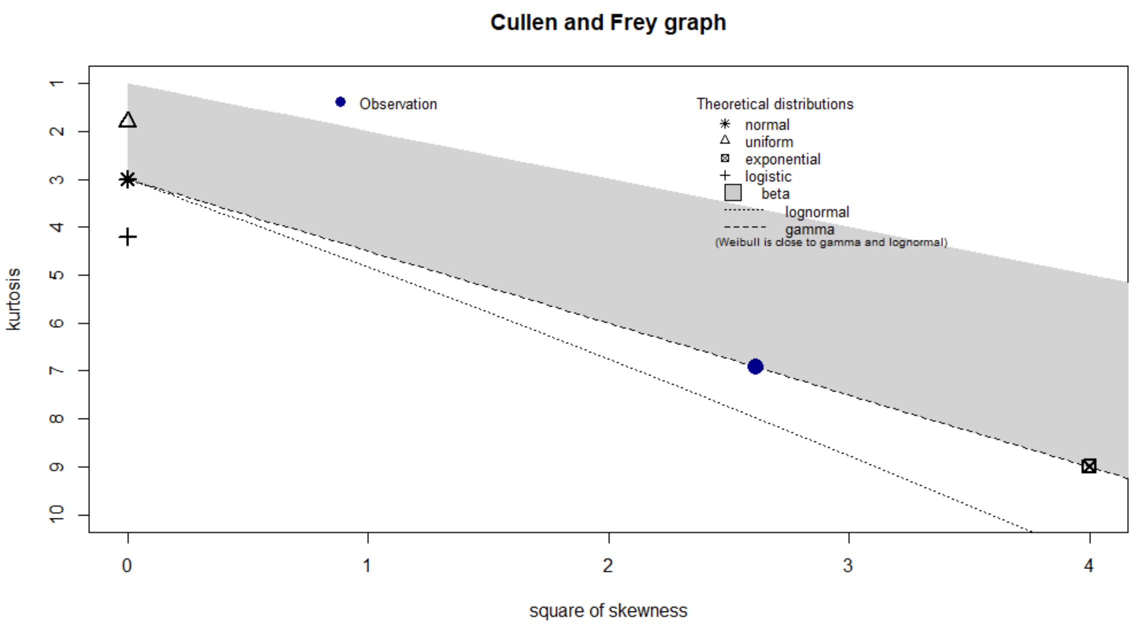
Average trip duration is a continuous variable. Following the same procedure as above with daily rides, first is the analysis of histogram (figure 12). There is a positive skew in values based on the plot, the mean is at 872.4 seconds and is greater than the median of 792.9 seconds.

Figure 12 . Histogram of average trip duration



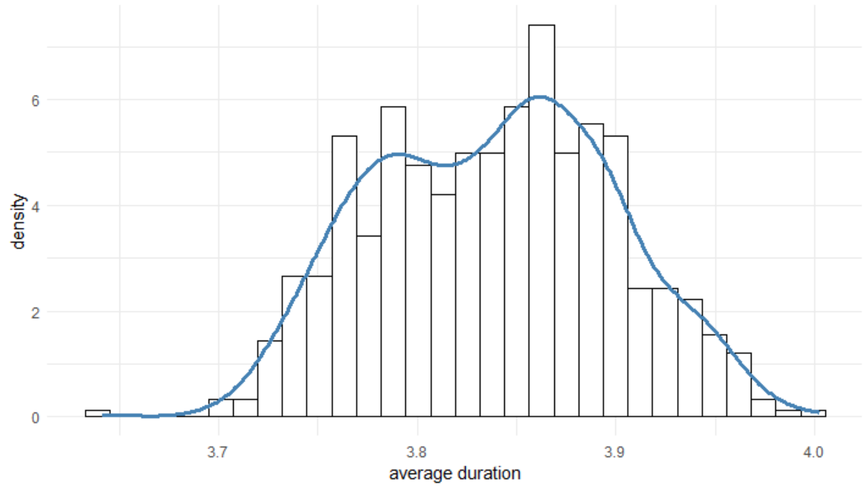
Using R allows the author to find that average trip duration values have a Gamma distribution. This assumption about the distribution will be used later in finding the right model. The dark blue dot (tested values) is on the gamma distribution line, which means that observations follow that specific type of distribution.

Figure 13. Fitting distributions to "average trip duration" variable



Even though distribution of values was determined, the author believes that a Box-Cox transformation for average trip duration is also needed to normalize the data and to increase the chances of getting normal residuals in the final model. Analogously, optimal lambda for a transformation was found to be –0.182. Histogram of Box-Cox transformed values (figure 14) has no positive skew as histogram of original raw values (figure 12). Both types of variables: original values following Gamma and Box-Cox transformed values following Gaussian distributions will be tested in the next section.

Figure 14. Histogram of Box-Cox transformed average trip duration



3.5. Model selection

Up to this point the following was assessed:

1) dependent variables were chosen as following:
    a. daily rides – Box-Cox transformed daily rides,
    b. average trip duration – original duration following Gamma distribution and Box-Cos transformed duration;
2) independent variables were selected:
    a. for daily rides – average temperature, precipitation, factor of weekend;
    b. for average trip duration – average temperature, precipitation, wind speed, factor of weekend

Next step is choosing appropriate model. All previous studies discussed in Section 2. Literature review used different approaches and models to analyze bicycle usage. Among those used models were multivariate Poisson regression (Corcoran et al., 2014), binomial model (Gebhart & Noland, 2014), ordinary least squares regression (Gebhart & Noland, 2014; Flynn et al., 2012), correlation coefficient analysis (Padmanabhan et al., 2021), fixed effect panel regression (Berezvai, 2022).

Padmanabhan er al. (2021) used correlation coefficient analysis between number of COVID cases and BSS usage indicators. Berezvai (2022) also incorporated COVID connected variables inside the model. In this thesis the author follows a different approach (COVID variables are not incorporated directly into the model, but predictions (if no COVID existed) were made and compared with actual values (which contained in themselves changes associated with COVID) during pandemic, meaning models used in mentioned studies are not suitable.

Main goal of Flynn et al. (2012) and their model was to predict whether an individual would ride a bike, dependent variable was binary, this is a classification problem which is also not relevant in the scope of this thesis. Most of weather variables were categorical variables in the models used by Gebhart & Noland (2014) unlike in this thesis in which weather data are values. Corcoran et al.'s (2014) approach cannot be used on Oslo data as it was found that dependent variable does not follow a Poisson distribution despite its nature.

To wrap up, none of these discussed models will be appropriate to use here to fulfill the intended purpose of this thesis. A different model is needed.

So far, a non-linear relationship has already been observed between each dependent and each independent variables separately (analysis in Section 3.1.). In other words, there are number of non-linear functions $f_i(x_i)$ which can describe connections between response and each of explanatory variables. Combining all these functions together in one model can be done using generalized additive models (GAM).

GAM extends linear model by allowing a smooth non-linear functions $f(x)$ define a non-linear relationships between response variable and explanatory variables (James et al., 2021). Model's form is as following:

$$y_i = \beta_0 + f_1(x_{i1}) + f_2(x_{i2}) + \cdots + f_p(x_{ip}) + \epsilon_i$$

In the context of this thesis $f_1(x_1)$ could explain relationship between average temperature and response variable, $f_2(x_2)$ could explain relationship between precipitation and response variable, and so on. $F_1(x_1)$ and $f_2(x_2)$ could be completely different functions. This is a one of the main advantages of GAMs. Various functions, even linear functions for binary variables (weekend variable in this thesis), could be easily used in one model.

One of GAM's key features is additivity. On the one hand, because the model is additive, it is easy to analyze the effect of each explanatory variable separately by holding other variables (and their respective functions) fixed. On the other hand, this additivity could omit possible interactions between independent variables. But in the scope of this thesis, this should not be a problem because pool of independent variables is not so large and collinearity between independent variables was checked, there are no important interactions between independent variables (for example, calendar event (factor of weekend) is not connected or not affecting average temperature in any manner).

By allowing inclusion of different functions GAM can be very flexible. This is advantageous because functions could describe the nature of relationship between response and explanatory variables closer to true form, and this in turn will increase the accuracy of predictions. On the other hand, this flexibility in choosing functions could lead to overfitting. In the case of Oslo BSS flexibility will allow adding closer to the true functions and overfitting will be controlled by limiting number of basis functions used in smoothing splines functions.

Thus, GAM should be a good choice for Oslo BSS usage data. Bias will be reduced by using non-linear functions, variance will be controlled by restricting some characteristics of these functions. Of course, choosing a perfect model which will reflect a true nature of the relationship between variables is a hard task, some approximations need to be done. With data used in this thesis (their nature and relationship between each other), GAM should be a suitable model to fit the model and afterwards to make predictions on test data.

3.6. Model results

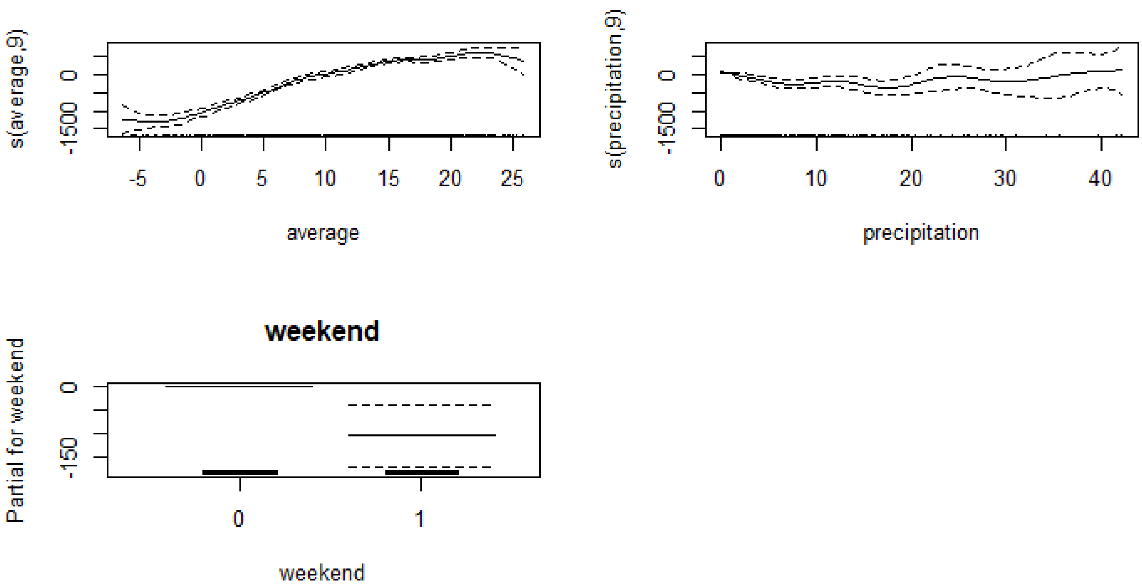
3.6.1. GAM model 1

In the model 1 Box-Cox transformed daily rides were used as a dependent variable. Based on selection made in Section 3.3. independent variables are average temperature, precipitation and factor of weekend. Formula for model 1 is as follows:

$$Box - Cox\ transformed\ daily\ rides_i = \beta_0 + f_1(average\ temperature_i) + f_2(precipitation_i) + weekend_i + \epsilon_i$$

As f_1 and f_2 functions smoothed regression splines are used. Regression splines divide values of the variable into several intervals, then fit a polynomial function to each interval with a constraint that there is a smooth joint at the ends of the intervals (James et al., 2021).

Figure 15. Relationship between "daily rides" and independent variables based on GAM model 1



Now after fitting a model, a true (or closer to the true) relationship between average temperature and daily rides could be observed (unlike in the data analysis done in the Section 3.1., there inferences were preliminary). From plots based on regression results, relationship between average temperature and daily rides is positive (as expected): overall as average temperature rises, so do daily rides up until 23°C, then there is a slight decrease in daily rides as temperature increases. The rise in daily rides is not the same for all temperatures. For example, for temperatures lower than –4°C Oslo riders are indifferent to riding a bike (almost a flat line), for temperatures in the range from – 4°C to + 10°C there is a steepest increase of all, meaning that any increase in temperature in that interval will result in substantial increase in daily rides. For temperatures from +10°C to +15°C increase in temperature will rise trips but smaller than in the previous range and even smaller for temperatures from +15°C to +23°C. After +23°C there is a decrease in daily rides, which could be explained as temperatures are getting a little bit uncomfortable for riders.

Precipitation's relationship with daily rides varies: when precipitation is from 0 to 20 mm per day, number of daily trips decreases. It is understandable: if there is no rain (0 mm), then when it starts to rain, some users may refrain from using bikes. Unexpectedly, for precipitation in the range from 20 to 40 mm per day, there is even an increase (just a slight one) in number of rides with increase in precipitation. This shows that Oslo BSS users are getting indifferent if precipitation is 30mm or 35 mm.

There were more trips made on weekdays than on weekends, which could mean that Oslo BSS users mostly ride bikes to the offices during Monday – Friday than ride them possibly for leisure on Saturday – Sunday.

From regression results, it can be concluded that factor of weekend (p-value of 0.00174), average temperature (p-value less than $2e^{-16}$) and precipitation (p-value less than $3.55e^{-9}$) are significant at 5% level of significance. All chosen independent variables are significant, which means that they were chosen correctly.

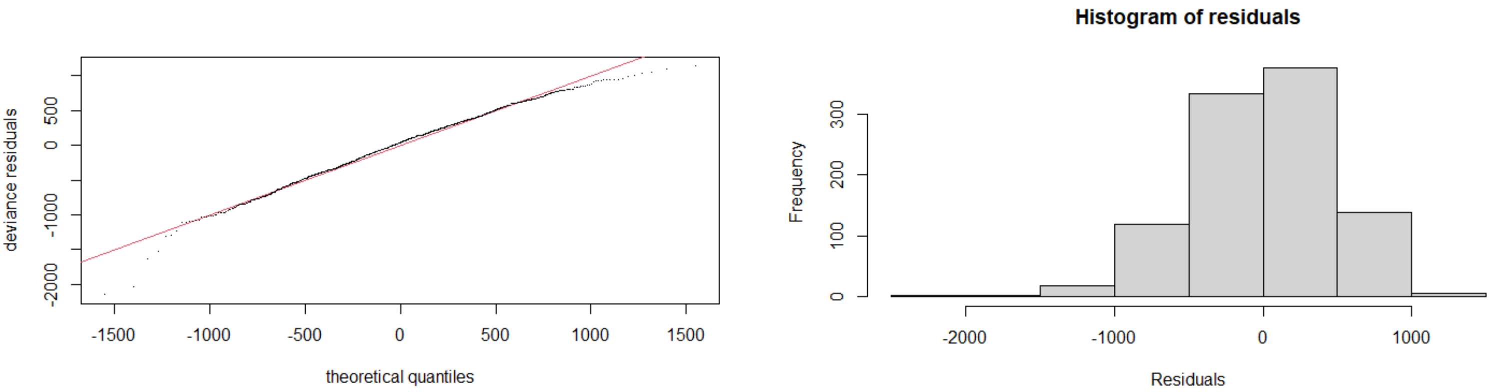
Table 7. Regression results GAM model 1

```
Family: gaussian
Link function: identity

Formula:
n_rides_bc ~ s(average, fx = T) + s(precipitation, fx = T) +
    weekend

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1763.43      17.74   99.41  < 2e-16 ***
weekend1     -104.73      33.35   -3.14  0.00174 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                 edf Ref.df       F  p-value
s(average)         9      9 125.652  < 2e-16 ***
s(precipitation)   9      9   6.592 3.55e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.548   Deviance explained = 55.7%
-REML = 7403.6  Scale est. = 2.2211e+05  n = 993
```

Model check

After fitting the model, model has to be checked for goodness of fit. First, residuals' behavior was checked from the graphs. From figure 16 histogram of the residuals resembles a normal bell-shaped curve and residuals distribution quantiles (black line) almost follows a normal distribution quantiles (red line).

Figure 16. Residuals of GAM model 1



AIC for the GAM model 1 is 15 064.56.

Next, in addition to graphical analysis, running a statistical test will help precisely decide a normality of residuals. During this test a ratio of an "estimate of the residual variance, based on differencing residuals that are near neighbors", divided by the residual variance is calculated. The more this ratio is away from 1, the bigger the chance "that there is missed pattern left in the residuals" (RDocumentation, n.d.c).

Test shows that residuals for the function of average temperature are randomly distributed (p-value is equal to 0.15 and is greater than 5% level of significance, so null hypothesis (that residuals are randomly distributed) is not rejected), whereas residuals for function of precipitation are not (p-values less than $2e^{-16}$). Re-fitting the model without precipitation variable results in AIC of 15105.33 (slightly higher than before, but one explanatory variable was removed) and all the tests for normality of residuals are passed. This refitted model is going to be used for final predictions of daily rides.

3.6.2. GAM model 2

GAM can also handle a situation where dependent variable has other than Gaussian distribution. For average trip duration the author is going to test two models: one with original value with gamma distribution and one with Box-Cox transformed duration, the best one will be used for further predictions.

In the model 2 original average trip duration following Gamma distribution was used as a dependent variable. Based on selection made in Section 3.3. independent variables are average temperature, precipitation, wind speed and factor of weekend. Formula for model 2 is as follows:

$$\begin{aligned} &average\ trip\ duration_i \\ &\qquad = \beta_0 + f_1(average\ temperature_i) + f_2(precipitation_i) \\ &\qquad + f_3(\text{wind speed}) + weekend_i + \epsilon_i \end{aligned}$$

From regression results (table 8), coefficients on weekend and wind speed are insignificant at 5% level of significance (p-value at 0.101 and 0.62174 respectively). AIC for GAM model 2 is 13938.08

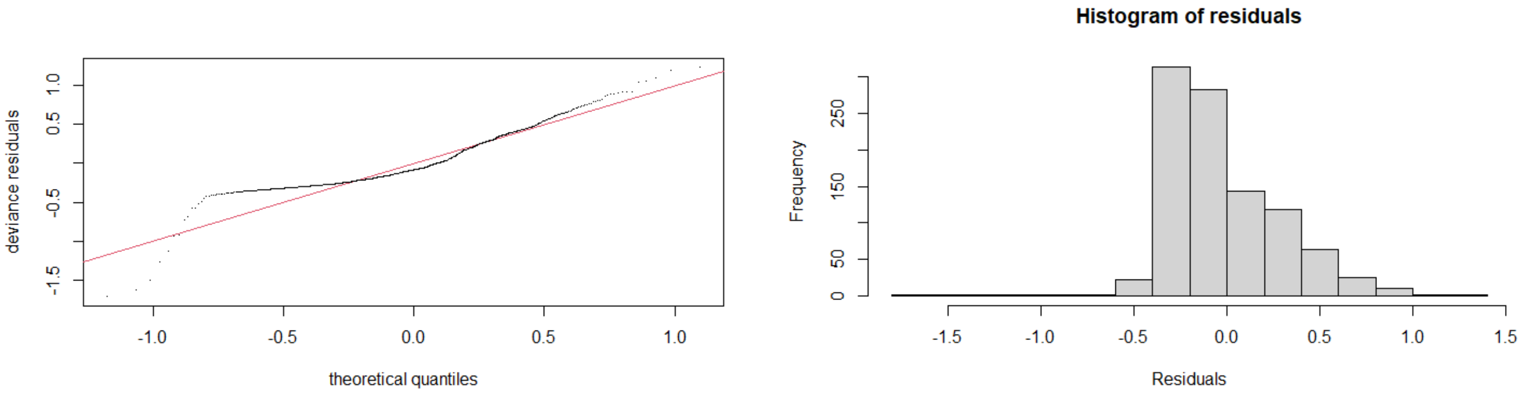
Table 8. Regression results GAM model 2

```
Family: Gamma
Link function: inverse

Formula:
aver_dur ~ s(average, fx = T) + s(precipitation, fx = T) + s(wind_speed,
    fx = T) + weekend

Parametric coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.192e-03  1.524e-05  78.202   <2e-16 ***
weekend1    -4.406e-05  2.683e-05  -1.642    0.101
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                 edf Ref.df      F p-value
s(average)         9      9 16.322 < 2e-16 ***
s(precipitation)   9      9  2.911 0.00209 **
s(wind_speed)      9      9  0.794 0.62174
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =   0.16   Deviance explained = 21.5%
-REML = 7188.4  Scale est. = 0.11577    n = 993
```

Model check

Analogously, as with previous model, residuals were checked for normality. It can be concluded that residuals are skewed and not normally distributed from histogram of residuals and Q-Q plot (figure 17).

Figure 17. Residuals of GAM model 2



Running a numerical test once again confirms that residuals are not normally distributed with p-values on all variables smaller than $2e^{-16}$ rejecting the null hypothesis of normally distributed residuals.

3.6.3. GAM model 3

The GAM model 2 has not passed the residuals test. Thus, in the final model 3 Box-Cox transformed average trip durations were used as a dependent variable instead of original values. Independent variables stayed the same as in the model2. Formula for the model 3 is as follows:

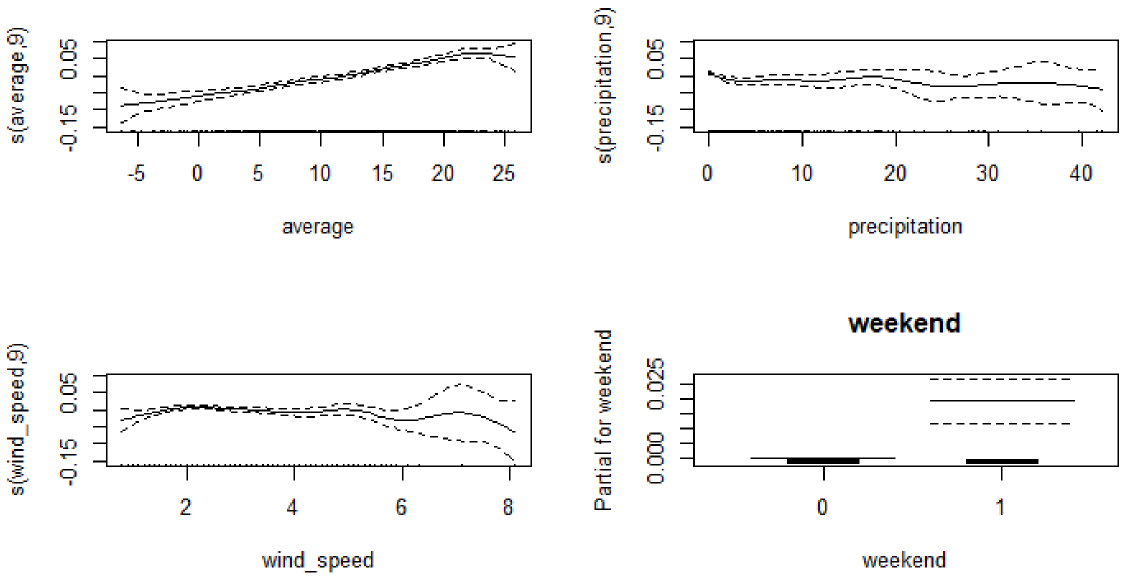
$$\begin{aligned} Box - Cox\ transformed\ average\ trip\ duration_i \\ &= \beta_0 + f_1(average\ temperature_i) + f_2(precipitation_i) \\ &+ f_3(wind\ speed) + weekend_i + \epsilon_i \end{aligned}$$

Plot of functions between response (duration) and each independent variables are shown in figure 18. Temperature and duration have a slightly different relationship than temperature and daily rides. An increase in duration caused by increase in temperature is smoother than previously observed for daily rides and almost the same up until temperature of +23°C. After +23°C the same as with the number of trips, average duration slightly decreases, which could be explained by temperatures becoming uncomfortable (too hot for Oslo) to ride bikes.

For precipitation between 0 to 18 mm, there is almost no change in duration, whereas with larger than 18 mm a day there is a slight decrease in duration, but not substantial, which could mean that overall Oslo riders are not so sensitive to the precipitation. The same is true for wind speed: up to 5 m/s trip duration is unchanged with increase in wind speed, for stronger winds there is slight decrease in riding time.

For the factor weekend, it can be concluded that Oslo BSS users prefer riding bikes for longer periods on weekends than for workdays (just the opposite of what was observed for daily rides).

Figure 18. Relationship between "average trip duration" and independent variables based on GAM model 3



From regression results (table 9), p-values for factor of weekend at 4.2.e^{-7}, for average temperature at smaller than 2e^{-16}, for precipitation at 0.000186 and for wind speed at 0.0284420 are smaller than 5% level of significance. All independent variables are significant, meaning that they were selected correctly.

Table 9. Regression results GAM model 3

```
Family: gaussian
Link function: identity

Formula:
aver_dur_bc ~ s(average, fx = T) + s(precipitation, fx = T) +
    s(wind_speed, fx = T) + weekend

Parametric coefficients:
            Estimate Std. Error  t value Pr(>|t|)
(Intercept) 3.832980   0.001990 1925.755  < 2e-16 ***
weekend1    0.019210   0.003761    5.108  4.2e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 9 (cont).

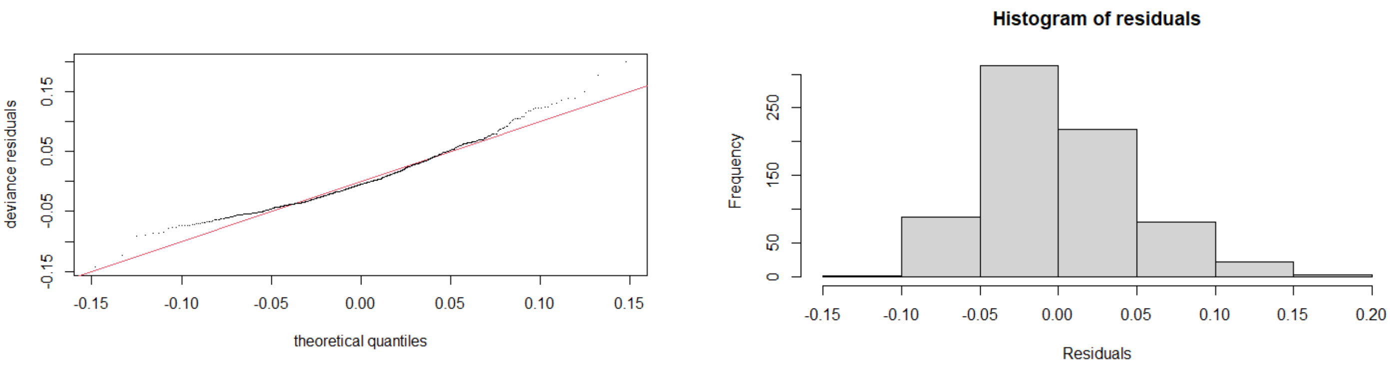
```
Approximate significance of smooth terms:
                 edf Ref.df      F  p-value    
s(average)         9      9 41.748  < 2e-16 ***
s(precipitation)   9      9  3.640 0.000186 ***
s(wind_speed)      9      9  2.089 0.028420 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.439   Deviance explained = 46.1%
-REML = -1126.8  Scale est. = 0.0020436  n = 728
```

AIC for the model 3 is – 2412.157, which is substantially lower than AIC for the model 2 (13398.08).

Model check

The residuals are better distributed for the model 3, no such skewness as was observed in the model 2 (figure 19).

Figure 19. Residuals of GAM model 3

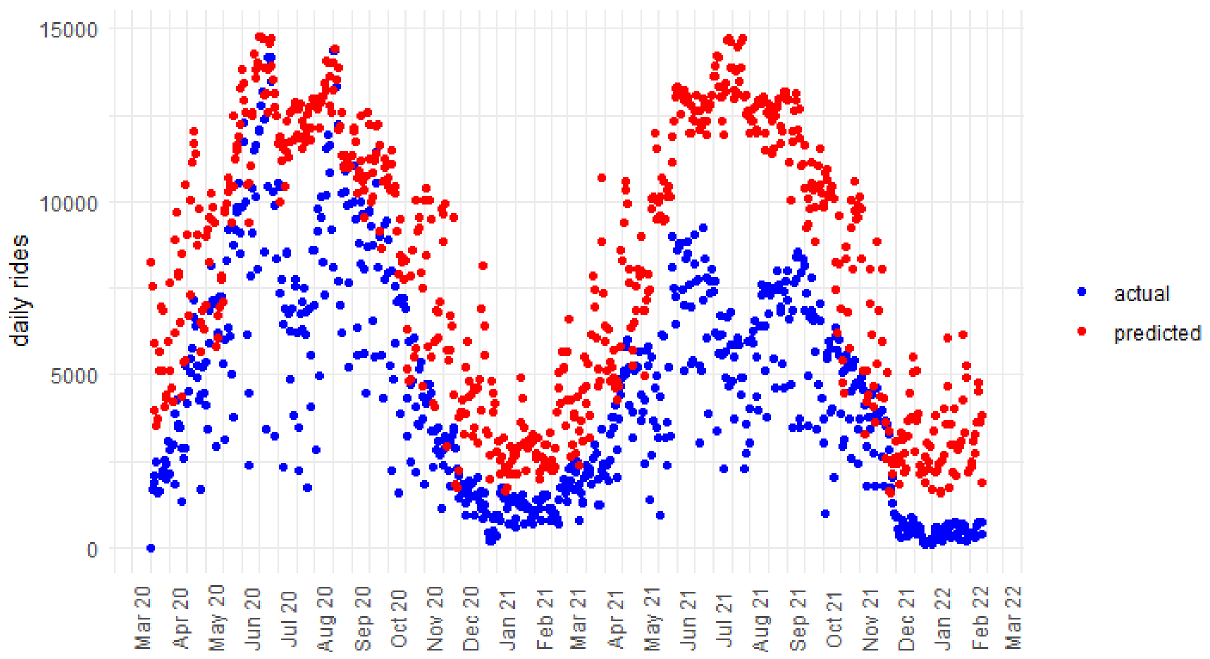


P-values in the normality of residuals test are 0.325 for function on average temperature, smaller than $2e^{-16}$ for spline on precipitation, 0.015 for function on wind speed. The model was re-fitted without the variables that did not pass the normality test (precipitation and wind speed). Refitted AIC is now – 2394.418, refitted p-value for normality is 0.43, meaning that null hypothesis of normality of residuals is not rejected. All tests are passed, this model will be used for further predictions of average trip duration.

4. Results and Discussion

Based on the model 1 predictions were made of daily rides during COVID period (figure 20), analogously based on the model 3 daily average trip duration was estimated on test pandemic period data (figure 21). Next, those predictions and actual values were compared to estimate the effect of COVID-19 on Oslo BSS usage.

Figure 20. Predicted vs actual daily number of rides during COVID period



Overall, it can be concluded that COVID-19 negatively affected daily total rides on Oslo BSS, but the effect was different over this almost two year long period. It can be concluded that total demand for shared bikes decreased for Oslo. Thus, the hypothesis #1 is confirmed.

In the beginning of pandemic (March – April 2020) there was around 50% decrease in actual rides compared to predicted without pandemic. This could be explained by the fact that for Oslo BSS a decrease in overall mobility was substantially greater than an increase from shift from other types of public transport in the start of pandemic.

After that in May – June 2020 drop in rides was smaller (around 27%), meaning that riders came back to use shared bikes for transporation. For some days it can be observed that the actual values with COVID match the predicted ones without COVID, which shows that demand for shared bikes increased at those days.

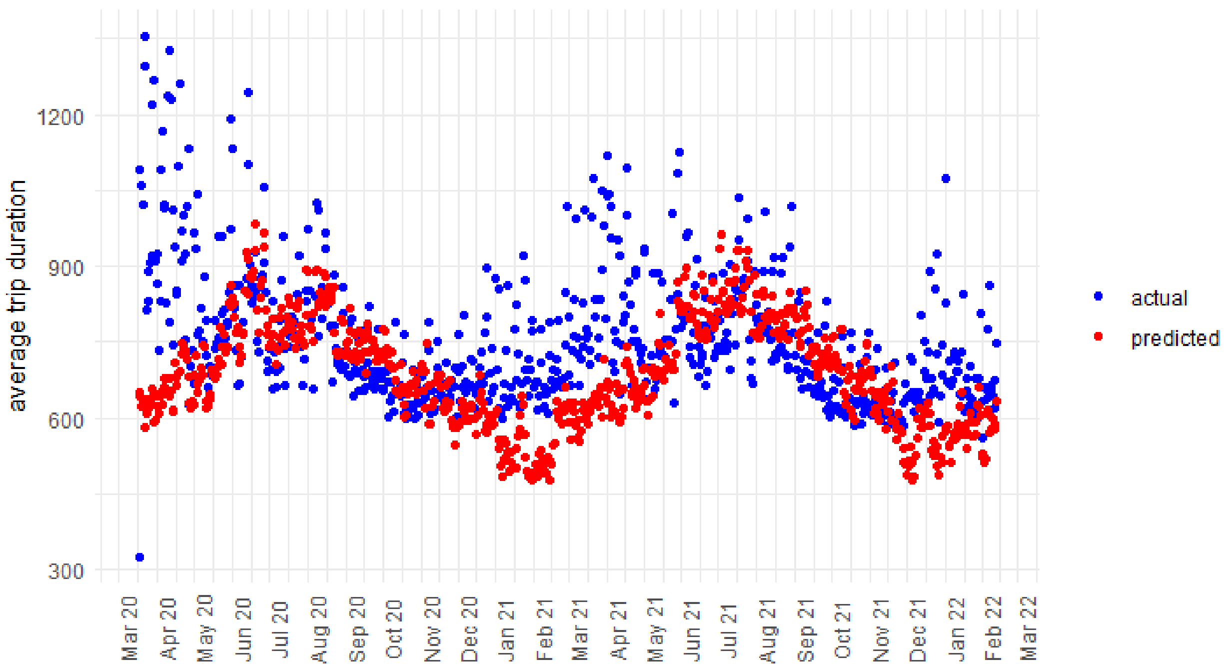
After that actual and predicted values go almost parallel with actuals being smaller (still staying at home and drop in the mobility dominates BSS usage). It can be noted that there were two significant (more than before) drops in rides: June – September 2021 and December 2021 – February 2022.

First drop could be explained by following facts: 1) continuation of reopening of society could have shifted back commuters to the other types of public transport (buses, metros, trams), meaning that it was again considered safe to travel in the closed and less ventilated vehicles; 2) there was an uncontrolled rise in quantity of shared e-scooters in Summer 2021, up to 20-30 000 units were on the streets of Oslo pushing local authorities to interfere and put a limit of 8 000 e-scooters allowed since Spetember 2021.

Second drop, which started in December 2021, could be explained by spread of more contagious omicron variant and by additional restrictions which were imposed to fight the spread of that variant and which dropped people's mobility.

Overall, because of almost constant decrease in rides, of course Oslo BSS met the aggregate demand. Oslo City Bike even made bikes available during winter months for the first time in COVID year 2021.

Figure 21. Predicted vs actual average duration during COVID period



With average trip duration, the impact of COVID-19 was completely different (figure 21). Overall, it was observed that average trip duration was even greater than predicted without COVID, which means that riders were willing to travel and travelled for longer distances. Thus, the hypothesis #2 is also confirmed.

In March – May 2020 increase in average trip duration (+34% on average from predicted values without COVID) was the highest throughout 2020-2022. It was the start of pandemic, less information about the disease was available, individuals responded quickly with travelling for longer time. Though the total number of daily trips decreased, trip duration increased substantially.

During June – November 2020 actual pandemic values were the same as the predicted non-pandemic values. That could mean that after the start of pandemic and initial drastic increase in hire time, users got used to the pandemic (initial fear has passed, more information about coronavirus became available) and used shared bikes as everything was normal.

In December 2020 – May 2021 again there was a significant increase (+25% than predicted on average) in trip durations. This could be explained by the appearance of new variants and imposed national restrictions associated with those variants. The riders reacted to that by taking longer trips.

As in previous year 2021, in June – November 2021 actuals and predicted were almost the same. Start of the opening of the society in that period could have affected the trip duration: commuters could prefer to ride for a shorter distance, for example, to the closest tram or metro stop instead of going by bike to the final desired destination.

As with daily rides, in December 2021 – February 2022 average trip duration was influenced by spread of omicron variant and related anti-pandemic measures. That resulted in 22% increase on average in trip duration.

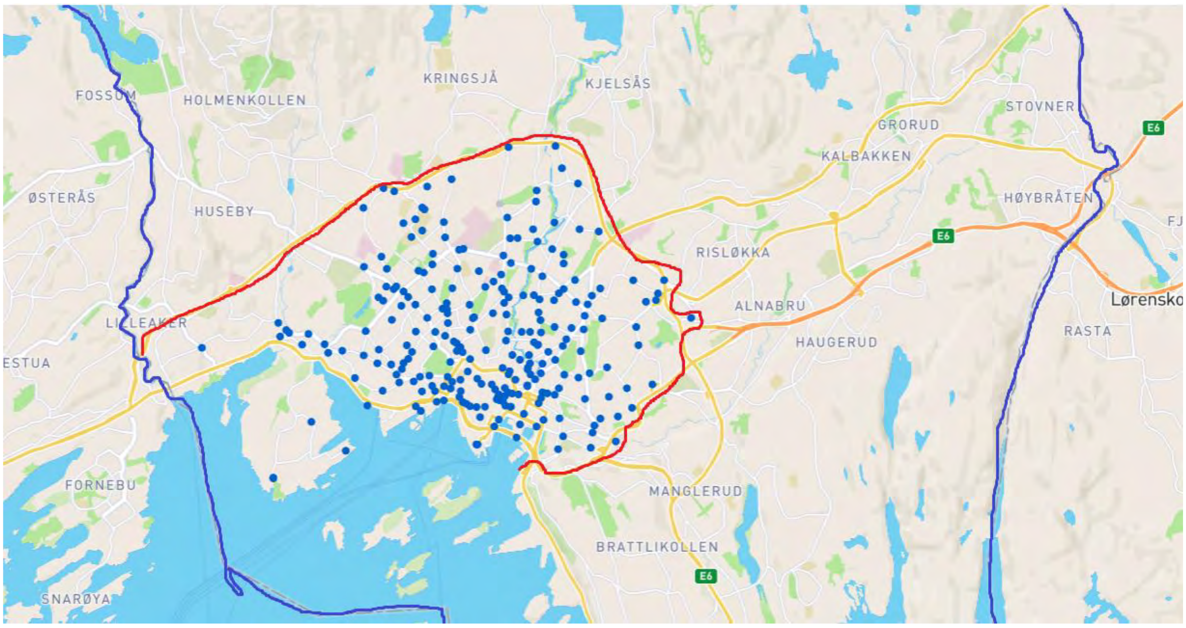
Overall, average bike hire time increased substantially over the pandemic period. Those increases were uneven, substantial rises happened when pandemic started, when new variants appeared, and corresponding governmental restrictions were imposed.

As for the tested calendar event factor, there were more rides made on working days than on weekend days, whereas Monday – Friday trips were on average shorter than Saturday – Sunday ones.

In this thesis it was confirmed that due to COVID-19 average trip duration increased. It means that users were willing to travel for longer distances to avoid catching a disease. It shows that people responded to pandemic by taking longer rides. This value is unaffected by the rivalry with the electric scooters' providers because it measures already taken rides using shared bikes. It is a pure COVID effect on the bicycle sharing system.

Even though users are ready to travel further, current Oslo BSS is limited to Ring 3, which is road #150 in Oslo. Road #150 is a red line on figure 22. Borders of city of Oslo are blue lines. Several boroughs of Oslo are not covered by Oslo City Bike stations. Just for example, population in three Oslo boroughs (which all are located beyond Ring 3) Alna, Stovner and Grorud exceeds 110 thousand persons (Oslo commune, 2022b).

Figure 22. Google Map of Oslo vs Map of Oslo City Bike coverage



There is a huge potential to expand existing Oslo's bicycle sharing system further. The analysis in this thesis has shown that users are willing to use shared bikes for longer time periods. The author recommends as the first step for Oslo City Bike to rearrange existing bike stations and to place them in boroughs adjacent to and beyond Ring 3. The costs of this step should not be substantial as it would not require acquiring new bikes or stations. Next step could be an even bigger expansion to cover all the other boroughs of Oslo. This will need larger investments than the first step as new bikes and stations are required. Partially those costs should be covered by the revenues gained from acquiring new users of the system.

To wrap up the discussion, detailed analysis of daily number of rides and average trip duration showed that the effect of COVID-19 on Oslo BSS was as follows: overall number of daily rides decreased, and overall average trip duration increased, there were periods when those changes were smaller or larger. The author's hypotheses were fully confirmed. Usually with appearance and spread of new variants and related governmental reactions Oslo's bicycle sharing system responded accordingly (substantial drop in daily rides and substantial increase in trip duration). With a fleet of 3 000 bikes and implementation of winter tires on bikes Oslo City Bike was able to meet the aggregate demand in terms of daily rides, but with limited coverage area some users' needs could not have been met. To satisfy the users' needs and to be prepared for the future possible analogous pandemic, the provider of shared bicycles service could rearrange existing bikes by moving existing stations to cover greater area of Oslo and if needed increase the quantity of bikes appropriately.

5. Conclusion

COVID-19 has changed many aspects of everyday lives. Some changes, that were caused by spread of pandemic, stayed, some aspects returned to pre-pandemic levels. In this thesis impact of COVID-19 on Oslo bicycle sharing system was estimated and based on the conducted analysis improvements to Oslo City Bike were suggested.

The raw data was in the form of information about 13 million of unique trips made on Oslo bicycle sharing system throughout the analyzed period and of daily weather data for more than 7 years. Using the software (specifically R) allowed the author to analyze this quite a big pull of data.

Various techniques were used throughout the thesis, including graphical and data analysis (to understand the patterns and preliminary relationships between various variables), shrinkage methods (to select independent variables), cross-validation methods (to select optimal tuning parameters for shrinkage methods,), and regression models (to make predictions).

There were some limitations to conducted research. Weather only in the form of temperature, precipitation and wind speed, and factor of weekend were used to explain shared bicycle usage. Potentially other weather factors, such as level of humidity, presence or absence of fog, or another type of calendar event (public holiday or school holiday) could influence the BSS usage. The author believes that chosen variables had a good predictive power.

The thesis hypothesis was that daily rides should decrease (as working and studying from home/closure of sport and cultural facilities, shops and restaurants/cancelation of events decreased people's mobility substantially) and average trip duration should increase (as users prefer to ride the bikes for longer distance to avoid other more contagious (in their opinion) modes of public transport). The hypotheses were confirmed. The impact of COVID-19 on daily rides was negative, there were substantial drops when new variants appeared, and also electric scooters took away part of BSS users. But now there is a sign of recovery of the demand for shared bikes. In the post-covid era since February 2022 number of daily rides keeps increasing: in 2021 mean of daily trips was 3 888, whereas in 2022 (February-November) it is already 4261. That shows that individuals are coming back to use Oslo BSS despite availability of e-scooters.

The effect of pandemic on average trip duration was quite the opposite. Hire time increased and in some periods rise was substantial. This indicates users' desire to use bikes for longer trips. It is interesting to mention that this pattern continues even today. Based on data from February 2022 to November 2022, mean of average trip duration was 741 seconds which is in the range of mean trip durations during COVID era in 2020 and 2021 (766 and 745 seconds respectively) and is greater than in pre-pandemic 2019 year, when mean duration was just 494 seconds. This means that commuters changed their behavior during COVID-19 by taking longer rides and kept doing so, even when pandemic was officially over in Oslo.

Based on these conclusions, it is recommended for Oslo City Bike to reconsider placements of the biking stations in a such manner that new mapping will cover areas of Oslo beyond Ring 3.

Oslo BSS reacted differently than bike sharing systems in other cities. For example, in London the same as in Oslo in the beginning of pandemics total daily rides decreased, but unlike Oslo, after initial drop, demand went up to pre-pandemic levels. Also like in Oslo average trip duration increased in London during covered period of analysis (Heydari et al., 2021).

On the contrary to Oslo and London, in Budapest bicycle sharing system usage increased drastically during the first wave of COVID and even more, there were more new users signed to use the system during that time than in pre-pandemic year 2019 (Berezvai, 2022).

Promoting the usage and expanding bicycle sharing system beyond Ring 3 have several advantages for the Oslo City Council and residents of Oslo in general. Health of BSS users could improve as riding a bike requires physical effort unlike riding an electric scooter. Since bikes produce no CO_2 emissions while in use, potentially greater usage of BSS could lead to a reduction in the city’s environmental pollution.

Future research could be made on how to expand the Oslo bicycle sharing system beyond Ring 3 with more efficient allocation of existing number of bikes and stations or with even greater expansion which will require increase in quantity of both bikes and stations. In addition, implementing this project of greater scope in real life could be beneficial for Oslo: if a new pandemic happens, then bicycle sharing system will be prepared for potential changes in usage patterns and will be available for use for everyone in Oslo; if there is no new pandemic, then still there are huge benefits to public health and environment caused by increased shared bikes usage.

References

New York City Department of City Planning. (2009). *Bike-share opportunities in New York City* New York City https://home3.nyc.gov/assets/planning/download/pdf/plans/transportation/bike_share_complete.pdf

Shaheen, S. A., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, Present and Future. *Transportation Research Record: Journal of the Transportation Research Board*, 2143(1), 159-167. https://doi.org/10.3141/2143-20

O'Sullivan, F. (2022). *The Radical roots of bike sharing*. Bloomberg. https://www.bloomberg.com/news/features/2022-02-26/the-dutch-anarchists-who-launched-a-bikesharing-revolution

Oslo City Council. (2015). *The City of Oslo Bicycle Strategy 2015-2025.* https://www.oslo.kommune.no/getfile.php/13415683-1630567227/Tjenester%20og%20tilbud/Gate%2C%20transport%20og%20parkering/Sykkel/Sykkelstrategier%20og%20dokumenter/The%20City%20of%20Oslo%E2%80%99s%20Bicycle%20Strategy%202015-2025.pdf

Knudsen, O. (2016). *Na kan du sykle pa de nye bysykklene.* Aftenposten https://www.aftenposten.no/oslo/i/rLAK3/naa-kan-du-sykle-paa-de-nye-bysyklene

UIP Drift AS. (2022). *Oslo City Bike.* https://urbaninfrastructure.no/oslo-bysykkel-as/?lang=en. Retrieved November 30, 2022

Luko. (2022). *Global Bicycle Cities Index 2022*. Luko. https://de.luko.eu/en/advice/guide/bike-index/

Myhre, S.M., & Uglum, M. (2021, June 6). *Slik skal de lose elsparkesykkel-kaoset i Oslo.* Aftenposten. https://www.aftenposten.no/oslo/i/WOwQld/slik-skal-de-loese-elsparkesykkel-kaoset-i-oslo

Oslo commune. (2021a). *Byradssak 1103/21 Forskrift om utleie av små elektriske kjøretøy på offentlig grunn.* Oslo Municipality. https://www.oslo.kommune.no/getfile.php/13412234-1627910141/Tjenester%20og%20tilbud/Politikk%20og%20administrasjon/Etater%2C%20foretak%20og%20ombud/Bymilj%C3%B8etaten/Kunngj%C3%B8ringer%20f ra%20Bymilj%C3%B8etaten/Byr%C3%A5dssak%20vedtatt%20-%20Forskrift%20om%20utleie%20av%20sm%C3%A5%20elektriske%20kj%C3%B8ret%C3%B8y%20p%C3%A5%20offentlig%20grunn.pdf

Nikel, D. (2021, 13 June). *City of Oslo cracks down on electric scooters rental*. Life in Norway. https://www.lifeinnorway.net/oslo-electric-scooter-rental-crackdown/

Folkehelseinstituttet. (2022a). *COVID-19 – dag*. https://statistikk.fhi.no/msis/sykdomshendelser?etter=diagnose&fordeltPaa=provedato&diagnose=713&fraDato=2020-02-21&tilDato=2022-02-12&diagramtype=tabell

Folkehelseinstituttet. (2022b). *Tilfeller etter sykdom, 2019-2022*. https://statistikk.fhi.no/msis/sykdomshendelser?etter=diagnose&fordeltPaa=aar&diagnose=713&diagramtype=tabell&maaned=1,2,3,4,5,6,7,8,9,10,11,12&geografi=fylkeKommune&kommune=0301&tidsrom=2019,2022

World Health Organization. (2020). *WHO-Director General's opening remarks at the media briefing on COVID-19 – 11 March 2020.* WHO. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Lokkevik, O., Ording, O., Holmes, M.C.S., Buggeland, S.A., & Johnsen, A.B. (2020, March 12) *Dette er tiltakene til Elna Solberg*: *- Vi star i en vanskelig tid*. VG. https://www.vg.no/nyheter/innenriks/i/2GeXaq/dette-er-tiltakene-til-erna-solberg-vi-staar-i-en-vanskelig-tid

The Office of the Prime-Minister (2021, April 10) *The Government's plan for gradual reopening: Out of the crisis together.* Norwegian Government. https://www.regjeringen.no/en/historical-archive/solbergs-government/Ministries/smk/Press-releases/2021/english/id2842670/

The Office of the Prime-Minister (2021, September 25) *Norway moves to normal everyday life with increased emergency preparedness.* Norwegian Government. https://www.regjeringen.no/en/historical-archive/solbergs-government/Ministries/smk/Press-releases/2021/norway-moves-to-normal-everyday-life-with-increased-emergency-preparedness/id2872539/

The Office of the Prime-Minister (2022, February 12) *The infection control measures are being removed on Saturday 12 February.* Norwegian Government. https://www.regjeringen.no/en/aktuelt/the-infection-control-measures-are-being-removed-on-saturday-12-february/id2900873/

Corcoran, J., Li, T., Rohde, D., Charles-Edwards, E., & 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. https://doi.org/10.1016/j.jtrangeo.2014.09.003

Gebhart, K., & Noland, R.B. (2014) The impact of weather conditions on bikeshare trips in Washington, DC. *Transportation,* 41 (2014), 1205–1225. https://doi.org/10.1007/s11116-014-9540-7

Eren, E., & Uz, V.E. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society*, 54 (2020) 101882. https://doi.org/10.1016/j.scs.2019.101882

Flynn, B.S., Dana, G.S., Sears, J., & Aultman-Hall, L. (2012). Weather factor impacts on commuting to work by bicycle. *Preventive Medicine*, 54 (2012), 122-124. https://doi.org/10.1016/j.ypmed.2011.11.002

Padmanabhan, V., Penmetsa, P., Li, X., Dhondia, F., Dhondia, S., & Parrish A. (2021). COVID-19 effects on shared-biking in New York, Boston and Chicago. *Transportation Research Interdisciplinary Perspectives*, 9 (2021) 10082. https://doi.org/10.1016/j.trip.2020.100282

Heydari S, Konstantinoudis G, & Behsoodi AW (2021) Effect of the COVID-19 pandemic on bike-sharing demand and hire time: Evidence from Santander Cycles in London. *PLoS ONE* 16(12): e0260969. https://doi.org/10.1371/journal.pone.0260969

Berezvai, Z. (2022) Short- and long-term effects of COVID19 on bicycle sharing usage. *Transportation Research Interdisciplinary Perspectives*, 15 (2022) 100674. https://doi.org/10.1016/j.trip.2022.100674

RDocumentation. (n.d.a) *geom_smooth: Smoothed conditional means*. RDocumentation. Retrieved December 08, 2022, from https://www.rdocumentation.org/packages/ggplot2/versions/3.4.0/topics/geom_smooth

Oslo City Bike (2022). *Historical data*. https://oslobysykkel.no/en/open-data/historical

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice* (2nd ed.). OTexts. https://otexts.com/fpp2/

Stock, J.H., & Watson, M.W. (2020) *Introduction to Econometrics* (4th ed., global ed.). Pearson Education

RDocumentation. (n.d.b) *descdist: Description of an empirical distribution for non-censored data*. RDocumentation. Retrieved December 06, 2022, from https://www.rdocumentation.org/packages/fitdistrplus/versions/1.1-8/topics/descdist

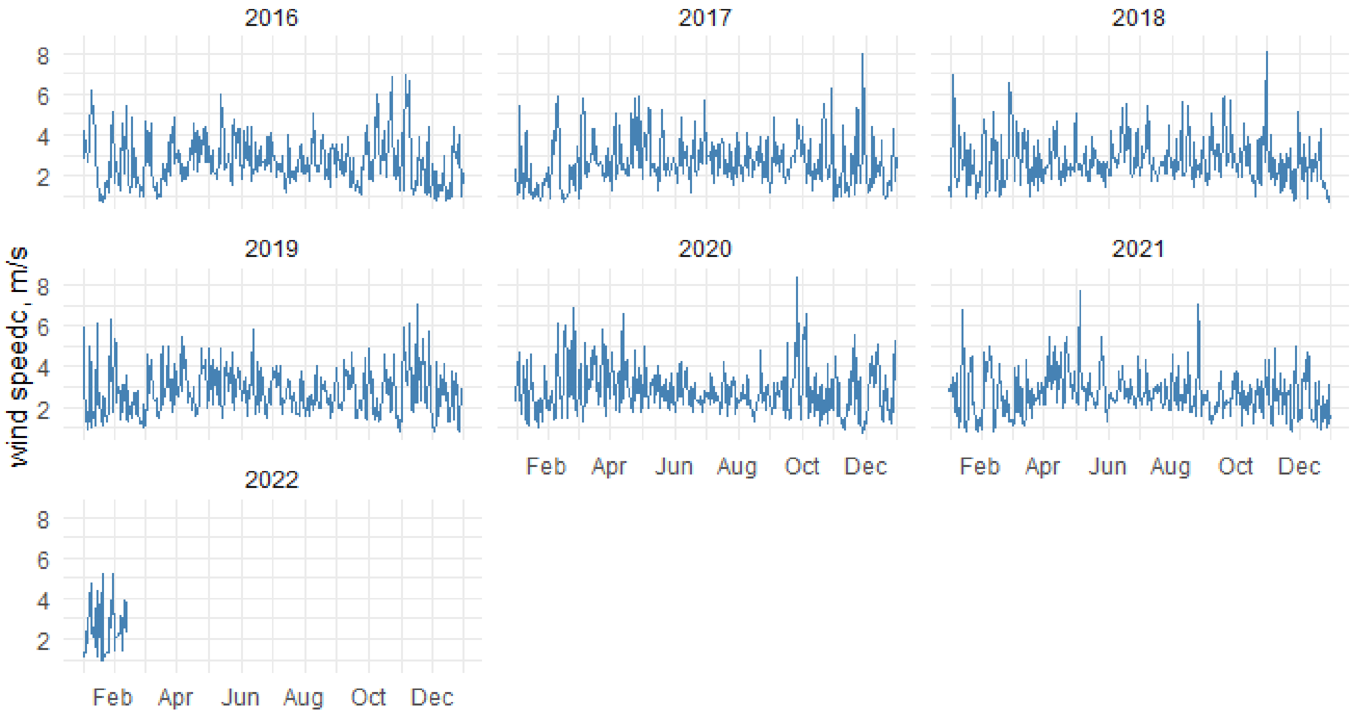
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021) *An Introduction to Statistical Learning with applications in R* (2nd ed.). Springer. https://hastie.su.domains/ISLR2/ISLRv2_website.pdf

RDocumentation. (n.d.c) *gam.check: Some diagnostics for a fitted gam model*. RDocumentation. Retrieved December 20, 2022 from https://www.rdocumentation.org/packages/mgcv/versions/1.8-41/topics/gam.check

Oslo commune. (2022b) *Folkemengden etter kjonn og alder (B).* Oslo commune. Retrieve December 17, 2022, from https://statistikkbanken.oslo.kommune.no/webview/velocity

Appendix 1.

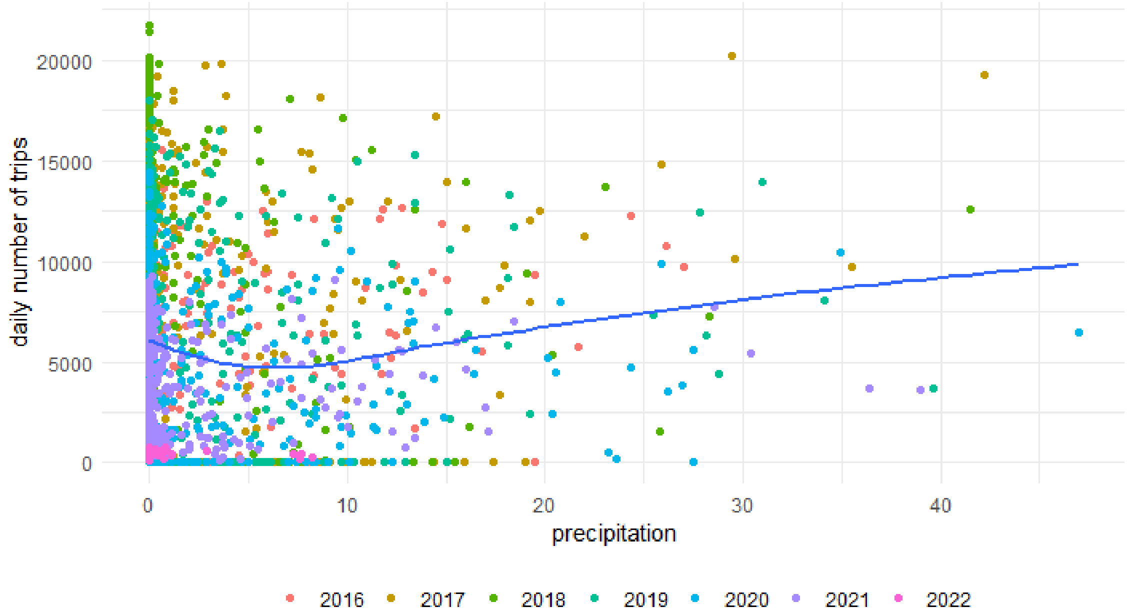
Wind speed in m/s in Oslo over Jan 2016- Feb 2022



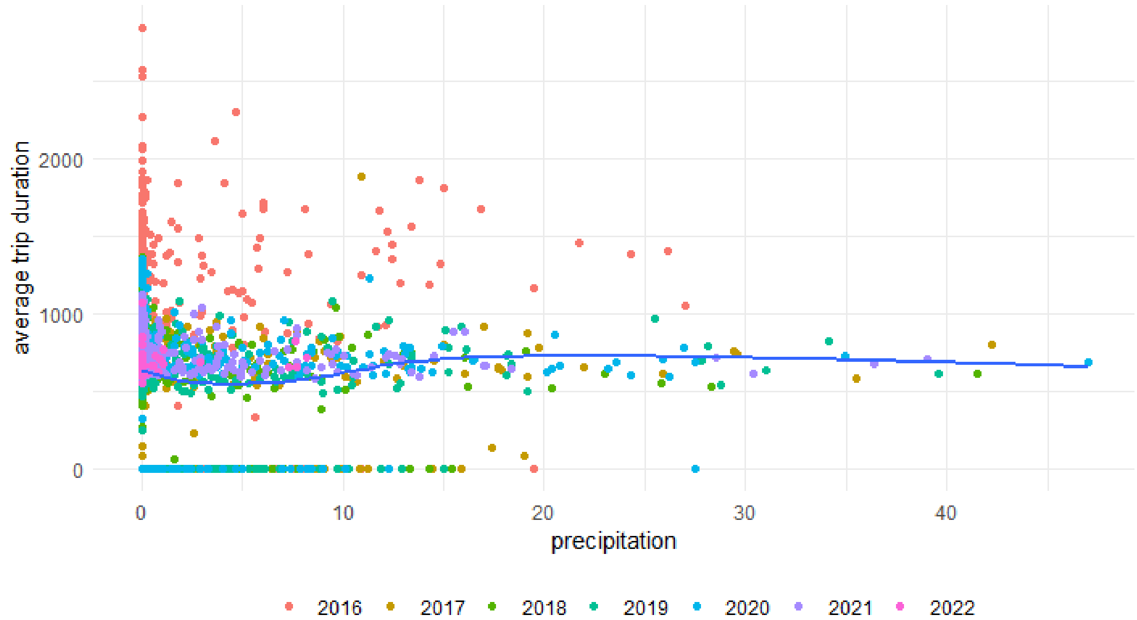
Appendix 2.

Graphs showing relationship between weather data (precipitation and wind speed) and dependent variables (daily number of trips and average trip duration) over time.

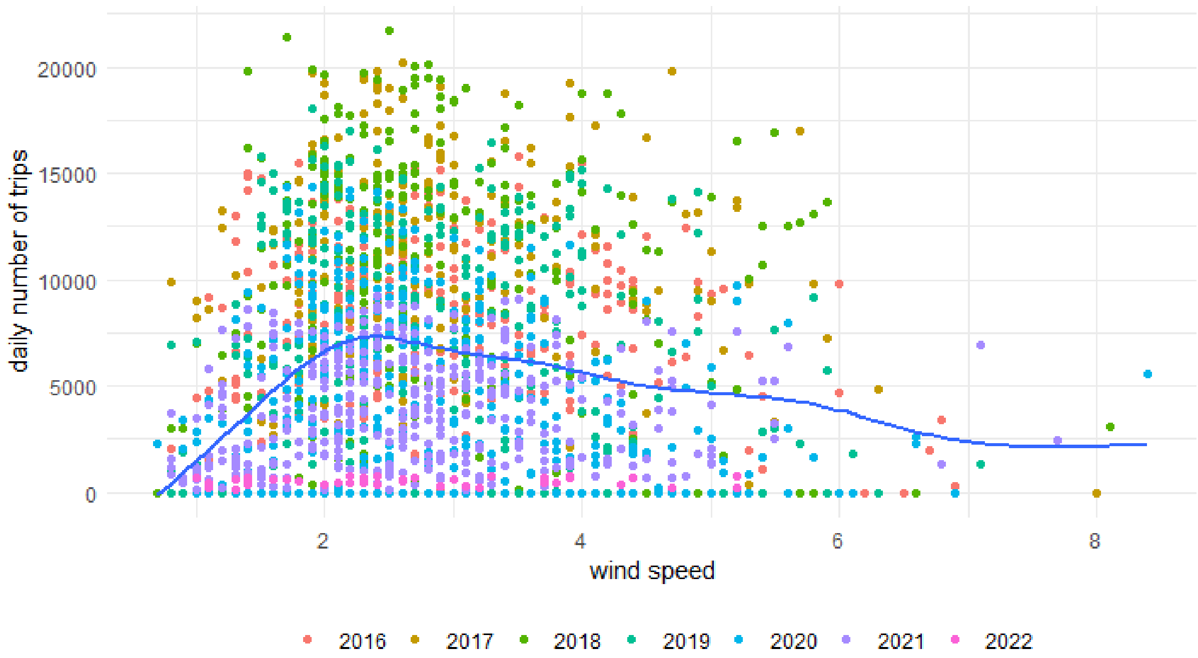
Daily rides vs daily precipitation in mm over the years 2016 – Feb 2022



Daily average trip duration (in seconds) vs precipitation (in mm) over the years 2016- Feb 2022



Daily rides vs wind speed (in m/s) over the years 2016 – Feb 2022



Daily average trip duration (in seconds) vs wind speed (in m/s) over the years 2016- Feb 2022