Analysing the Effect of the Covid-19 Pandemic on Just Eat Cycle Trip Behaviour

James Keane

5th April 2021

1 Overview

The coronavirus pandemic has had an impact on all types of business including the hire-bike industry. With changes to how we socially interact due to lockdown restrictions people have are limited to socialising outdoors, and with this an increased time spent partaking in outdoor activities. This study aims to see how the introduction of lockdown in march 2020 has had an impact on hire-bike usage in Edinburgh. In particular the change in behaviour of the bike usage, such as the peak usage hours, the change in route locations, the frequency of trips taken and the overall number of trips per day. hire-bike refers to the Just Eat Cycle Bikes that are available for hire for trips. They are shared bikes that anyone can use as long as they fit the required criteria.

Using 18 months of hire-bike data, 6 months before the first lockdown and 12 months after, that was provided by Just Eat Cycles and is open to the public. Using inferential statistics, Python and other python packages such as Pandas and Numpy to create insights and visualisations to provide insights into the data. The analysis of the data indicated an increased hire-bike usage during the pandemic. The results indicate a change in the type of trips taken from usage for commuting to recreational usage. Moreover, an increase in frequency of trips through a reduction of the time between trips during the lockdown in comparison to the months before the pandemic and a higher probability of higher trips per day.

2 Introduction

The Coronavirus Pandemic has had a major impact on almost all parts of society. One of these impacts is how we spend our recreational time, with the introduction of furlough and lockdown, people are spending more time outside socialising. The first lockdown in Scotland was announced in March of 2020 and since then Edinburgh has seen a constant level of harsh restrictions which has changed the way people interact and socialise. Because of this, outdoor activities have become the predominant form of social interaction. One such activity is cycling, Just Eat offers easy to access bicycles for rent. Each bicycle is equipped with a GPS tracker and its trip data is recorded, Just Eat makes the tracking data open for Edinburgh and the surrounding areas.

Currently there are several websites that provide both real time and historical visualisation for Edinburgh's Just Eat Cycle. One website is visualisation.bike, this website offers in-depth visualisations for the historical data for bike sharing in multiple cities throughout the world. Another website, bikesharemap offers real time statistics and visualisation worldwide. Other work includes an article on bikesharp.com by Oliver O'Brien which investigates hire-bike usage during Edinburgh's fringe festival. There is a blog post by David McArthur from Urban Big Data Centre which investigates the hire bike usage as lockdown was being introduced. From this post McArthur found that the hire bikes were being used more for leisure as a means of providing exercise for users. He also found that the duration of trips had increased significantly which indicated that the bikes where being used for tours instead of as a means for transport, this also suggested the change in usage to recreational use.

The aim of this study is to analyse the effect the Coronavirus pandemic has had on the usage of the Just Eat Cycle Bicycles. In General, it is expected that the overall usage of the hire-bikes to have increased since the first pandemic. In particular the change in frequency of trips, the peak usage times and the change in where the trips are starting and finishing. It would be expected that the increase in recreational use would cause the most popular stations to change to areas of recreational interest such as parks, beaches, or places of significance. Moreover, it can be expected the peak usage times would become more centred around mid to late afternoon as opposed to early morning, this would convey the change in usage from transport to leisurely cycles.

3 Data

Data provenance The data sets used in the study are historical data provided by Just Eat Cycles. The data is maintained and updated daily by Just Eat. All the historical data is available to download in the form of CSV or JSON files. All of the data is published under the Open Government License which allows for copying, publishing, adapting and transmitting the information, which is the data's purpose for this study.

Data description From the data available the study has used all records between the months of September 2019 and February 2021 for a total of 18 Months of hire-bike usage. The data is recorded on a month-to-month basis, in total there are 18 .CSV files. Of the data used in this study there is a total number of 280,376 recorded trips. There are 13 columns in the data set(see figure 1)

Table 1: Data Set Column Variable Description `https://edinburghcyclehire.com/open-data/historical`.

Variable	Format	Description
started_at	Timestamp	Timestamp of when the trip started
ended_at	Timestamp	Timestamp of when the trip ended
duration	Integer	Duration of trip in seconds
start_station_id	String	Unique ID for start station
start_station_name	String	Name of start station
start_station_description	String	Description of where start station is located
start_station_latitude	Decimal degrees in WGS84	Latitude of start station
start_station_longitude	Decimal degrees in WGS84	Longitude of start station
end_station_id	String	Unique ID for end station
end_station_name	String	Name of end station
end_station_description	String	Description of where end station is located
end_station_latitude	Decimal degrees in WGS84	Latitude of end station
end_station_longitude	Decimal degrees in WGS84	Longitude of end station

Data processing All of the .CSV files for each month were concatenated into one table using Pandas. The index number for each record was reset such that every record in the data set was serial. Any records with null values were dropped from the data set. In some visualisations timestamp formats were rounded to the nearest hour or day (for example Figure 2)

	started_at	ended_at	duration	start_station_id	start_station_name	start_station_description	start_station_latitude	start_station_longitude	end_station_id	end_station_name	end_station_description	end_station_latitude	end_station_longitude
0	2020-04-01 01:41:11.814000+00:00	2020-04-01 02:28:08.128000+00:00	46.933333	1808	Gorgie Road	Corner with McLeod Street	55.938741	-3.229909	1808	Gorgie Road	Corner with McLeod Street	55.938741	-3.229909
1	2020-04-01 04:52:10.586000+00:00	2020-04-01 05:45:32.129000+00:00	53.350000	1752	IGMM - Western General	The Institute of Genetics and Molecular Medicine	55.962642	-3.231916	1752	IGMM - Western General	The Institute of Genetics and Molecular Medicine	55.962642	-3.231916
2	2020-04-01 05:45:52.745000+00:00	2020-04-01 06:24:32.266000+00:00	38.650000	1752	IGMM - Western General	The Institute of Genetics and Molecular Medicine	55.962642	-3.231916	1753	Waitrose Comely Bank	On Fettes Avenue next to Waitrose	55.959536	-3.223434
3	2020-04-01 06:24:58.055000+00:00	2020-04-01 07:13:07.359000+00:00	48.150000	1753	Waitrose Comely Bank	On Fettes Avenue next to Waitrose	55.959536	-3.223434	1752	IGMM - Western General	The Institute of Genetics and Molecular Medicine	55.962642	-3.231916
4	2020-04-01 06:50:15.745000+00:00	2020-04-01 07:05:20.319000+00:00	15.066667	1769	Brunswick Place	Corner of Elm Row/Brunswick Street	55.960852	-3.180986	1752	IGMM - Western General	The Institute of Genetics and Molecular Medicine	55.962642	-3.231916
...	...	...	...	...	...	...	...	...	...	...	...	...	...
280371	2019-11-29 23:19:36.946000+00:00	2019-11-29 23:27:43.990000+00:00	8.116667	247	Charlotte Square	North Corner of Charlotte Square	55.952335	-3.207101	1720	Dundas Street	On corner of Henderson Row	55.960762	-3.201278
280372	2019-11-29 23:22:43.246000+00:00	2019-11-29 23:29:29.738000+00:00	6.766667	259	St Andrew Square	North East corner	55.954906	-3.192444	1743	Logie Green Road	Outside Lidl	55.964058	-3.195700
280373	2019-11-29 23:25:03.496000+00:00	2019-11-29 23:43:20.440000+00:00	18.266667	1097	Gladstone Terrace	Corner of Gladstone Terrace and Sciennes Road	55.937963	-3.185021	1091	Holyrood Road	Opposite St Leonards Land	55.949560	-3.180413
280374	2019-11-29 23:29:57.437000+00:00	2019-11-29 23:47:18.314000+00:00	17.333333	250	Victoria Quay	Entrance to Scottish Government Office	55.977617	-3.174126	1728	Portobello - Kings Road	Foot of Kings Road next to the promenade	55.957915	-3.118332
280375	2019-11-29 23:39:15.440000+00:00	2019-11-29 23:45:55.315000+00:00	6.650000	249	Fountainbridge	Fountainbridge by Gardner's Crescent	55.943357	-3.209248	248	Bristo Square	Bristo Square, near Potterrow	55.945834	-3.189053

280376 rows × 13 columns

Figure 1: Pandas Table Representation of Data Set

4 Exploration and analysis

The data will be analysed in 3 portions, the first is the 6 months leading up to the first lockdown, the second is the 6 months following the first lockdown and the third is the following 6 months after the previous portion. We see that in general the average number of trips taken per day is higher post during the pandemic, however we can assume that there will be a general increase in the summer months. By making a comparison between prepandemic and months 6 - 12 pandemic we see a sizable increase in the average number of trips taken per day.(see figure 2)

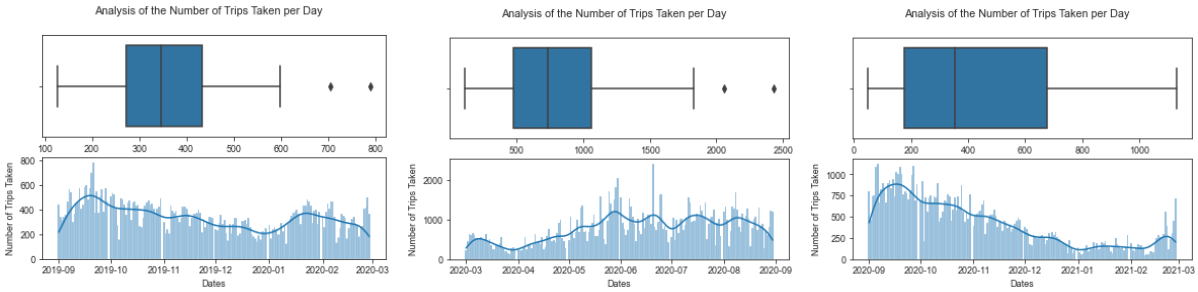


Figure 2: Visualisation of the frequency of trips before pandemic.

When analysing the total number of trips taken for each hour we see a change from peak times at 0700 - 0900 and 1600 - 1800 to a longer peak interval from 1200 - 1600. Along with the flatter peak in the 6 - 12 month interval these visualisations convey a change from the type of trips users are taking. With the prepandemic interval we see that peak times are early mornings and evenings which indicate that users are commuting to and from work or school. Comparatively we see that both other graphs show heavy usage late morning until early evening which suggests that trips are predominantly recreational (see Figure 3).

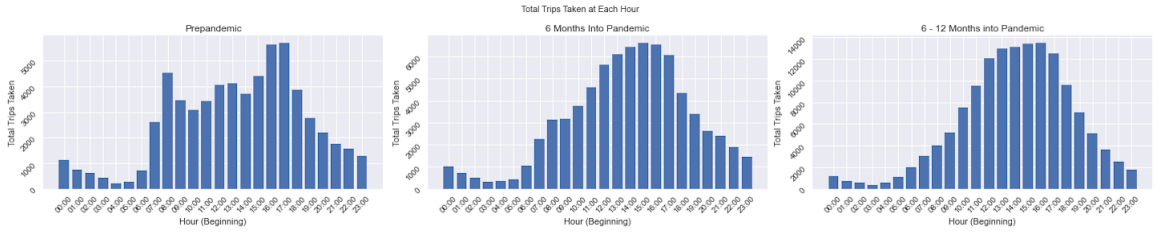
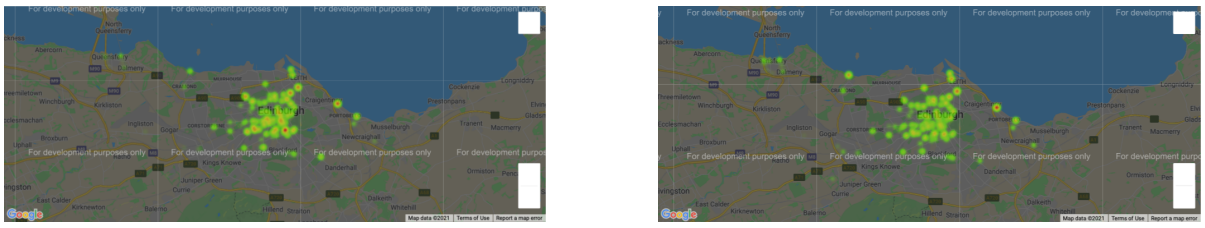


Figure 3: Visualisation of the total number of trips taken for each hour.

Using the GPS data for each trip and the Google Map API, we see a difference in proportions of traffic at each stop. Before the first lockdown we see that the stations with the heaviest traffic were in the city, this contrasts when we look at the first 6 months of lockdown as stations that receieve the most

traffic are stations closest to beaches e.g Cramond and Portobello. This indicates a change of the bike usage from commute to recreational use (see figure 4).



(a) Prepandemic (b) 6 Months into Pandemic

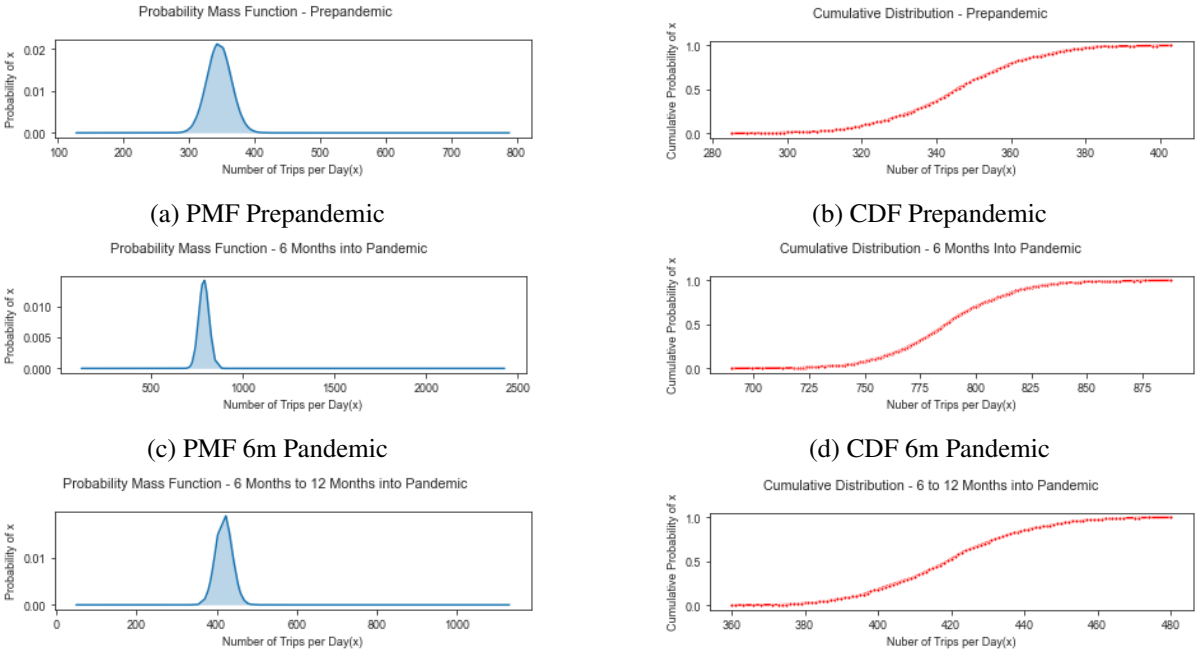
Figure 4: Heat Map of Station Traffic

Using the Poisson distribution we can model the frequency of trips per day. Using the formula (1) we derived a probability mass function for each part of the data set with our discrete random variable being the total number of trips per day (x) and lambda being being the total number of trips divided by the total number of days in the time-span used (see figure 5).

$$\bar{x} = P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \tag{1}$$

From the PMFs we can clearly see that both the intervals 0-6 months and 6-12 months into the pandemic have a higher mean number of trips at around 800 and 420 respectively. This is compared to the prepandemic mean at around 350 trips per day, however it does have a higher probability for reaching this mean which indicates more consistent daily trips (see figure 5).

From the PMFs we can construct cumulative distribution functions this gives us an indication of how likely an interval is to surpass a certain number of trips. An important observation is that Prepandemic there is a 80% chance on a given day that there will be more 329 trips in comparison to months 6-12 of the pandemic there is an 80% change on a given day that there will be more than 418 trips. With both time spans covering the same months, the effect of weather on usage is minimal.



(a) PMF Prepandemic (b) CDF Prepandemic

(c) PMF 6m Pandemic (d) CDF 6m Pandemic

(e) PMF 6-12m Pandemic (f) CDF 6-12m Pandemic

Figure 5: Poisson Process for Probability of Number of Trips per Day

Using the exponential distribution we can model the time elapsed before a trip occurs. For each interval we calculated the time between each trip to the next chronologically and plotted it alongside an exponential sample using the mean time difference and a sample population of 40,000 (2). The exponential function modelled the total population quite well for each of our 6 month time spans. (see Figure 6)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

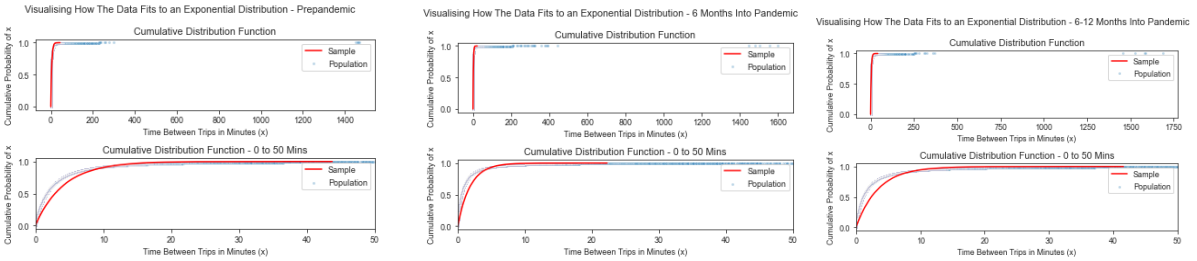


Figure 6: Modelling Exponential Distribution

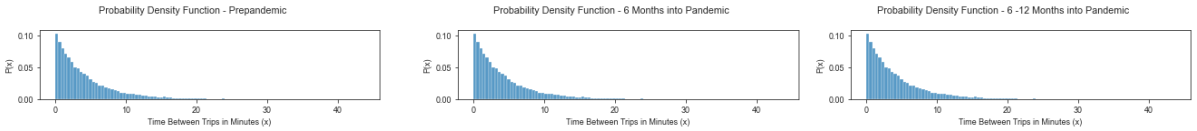


Figure 7: Probability Density Functions

Using bootstrap sampling with our cumulative exponential distributions we can recreate each of the time spans multiple times using sampling with replacement. From this we can use confidence intervals to determine where the true mean of time between trips falls. Using this method we are able to create a mean time between each trip for 10,000 sample populations for each time span. We can see from Table 2 that the mean for the time between each trip prepandemic is roughly 4.1 minutes. The mean time between trips of months 0-6 of the pandemic is roughly 1.8 minutes, this is a very sizable difference and indicates a lower frequency over the time-frame in comparison. A reduction in the mean time between trips can also be seen for months 6-12 of the pandemic with a mean roughly around 3.4 minutes. This follows the trend of overall usage increase over time.

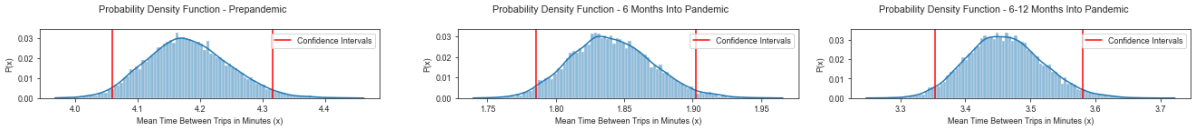


Figure 8: Bootstrap Sampling for Mean Time between Trips

5 Discussion and conclusions

Summary of findings In summary from the investigation, we found that overall usage has increased since the first lockdown in March 2020. We see this especially in the first 6 months of the pandemic, we

Table 2: Confidence Intervals for Mean Time Between Trips

Time Span	Lower Confidence Interval	Upper Confidence Interval
Prepandemic	4.060	4.216
6 Months Pandemic	1.786	1.902
6 - 12 Months Pandemic	3.353	3.580

see the mean number of trips per day increase from 350 to 800 roughly. This is a sizable increase but can be explained as these are warmer months and people are more inclined to go cycling in nicer weather. Then looking at months 6-12 of the pandemic we still see an increase to roughly 420 which is still a good indication that there is more usage of the hire-bikes.

When looking at the number of trips for each hour between each of the timespans we saw a change from peak times around the start and end of a business day to peak time from late morning to early evening. This indicates a change in the type of trip the bike are being used for. The peak times around the start and end of the business day indicate hire-bikes being used to commute, contrastingly with the peak usage from late morning to early afternoon we can infer that the hire-bikes are being used for recreational use predominantly. This change in behaviour can be seen when visualising station traffic throughout Edinburgh as prepandemic the highest concentration of traffic was in the city. This then changed after lockdown was introduced with higher concentrations of traffic in areas such as beaches and parks.

When modelling the data with distributions we found an indication of increased usage, when modelling with a Poisson distribution we found that months during the pandemic had a higher probability of higher trips. When modelling the time between each trip with the exponential distribution we found that the months during the pandemic had lower times between each trip which indicate higher frequency and therefore increased usage.

Evaluation of own work: strengths and limitations Some strengths of the investigation are the statistical methods used to provide key insights into the data, such as the use of bootstrap sampling and use of the Poisson Distribution. However some limitations of the study may be the heatmap visualisation, instead routes could be used to visualise the popularity of journeys which would give a better indication of trip behaviour.

Comparison with any other related work In comparison to other work done with this data set, previously mentioned there was a blog post by McArthur that had a rough look at hire-bike use in lockdown. The insights that he provided showed the beginning of an incline of usage as the first month of lockdown passed. This agreed with the investigation as usage did increase during the months of the pandemic.

Improvements and extensions An extension on this investigation is clustering of the GPS data to group trips together by type. This would visualise what routes were predominantly for recreational use or for commuting. Improvements that could be made would be surrounding the visualisations used such as more variety of graph and chart types.