

Cycle Hire Scheme Data Visualisation-led Investigation



By: Alba Haque Sultana

Table of Contents

Introduction to research topic.....	2
Data	2
Exploratory and explanatory data visualisation	3
Initial impact of Cycle Hire Scheme in 2010	4
Evaluation of the Cycle Hire Scheme a decade later in 2020	7
Conclusion and Evaluation	10
Bibliography:.....	11

Introduction to research topic

In 2007, London was suffering from major traffic congestion and its consequences, such as high levels of pollution and slow journey times. In 2010 it introduced a cycle hire scheme, which enabled cyclists to hire a bike from one of London's hundreds of docking stations (761 as of January 2016).

The cycle hire scheme and cycle superhighways were central to the Mayor's plan for a "cycling revolution" in London in 2010. The Transport Strategy set out a target to increase cycling trips by 400 per cent by 2026^[1]; the equivalent of 1.5 million cycling trips per day.

I intend to examine the initial impact of the cycle hire scheme including any issues arising from their early implementation and consider the solutions proposed. Currently, a decade later since the scheme launch, I want to research how far we've come to meet the Transport's Strategy target.

Research questions:

- 1: Initial impact of cycle hire scheme 2010 (30th July- 30th October)
 - a) Demographics
 - b) Out of regular cyclists, how many have registered to use the cycle hire scheme?
 - c) Cyclist's experiences whilst using this scheme
 - e) Time of hire per ride
- 2: A decade later, how far we've come the Transport's Strategy target to increase cycling trips by 400 per cent by 2026?
 - a) Numbers of bike hires per year
 - b) Difference of number of bike hires of 2020 compared to 2010
 - c) Forecast at least until 2026 to predict if the target of 1.5 million cycle trips per day will be met.
 - d) Average hire times in the last decade

Data

I will be using two datasets, one contains the necessary data to examine the initial impacts of the scheme, from <https://data.london.gov.uk/dataset/london-assembly-cycle-survey-responses>, available in csv format and undertaken over September to October 2010.. And the other dataset contains a timeseries of the number the bike hires, to forecast if the Transport's Strategy target will be met. The data was from <https://data.london.gov.uk/dataset/number-bicycle-hires>, it was available in xlsx format.

Both datasets are sourced from Data London, produced by the TfL. This source is known to be reliable and trustworthy, therefore I'm hoping that the findings from this research will be reliable.

The two images below show the raw data sets. I have used data pre-processing techniques to convert the raw data into usable clean data sets.

	A	B	C	D	E	F	G	H	I	J
1	RespondentID	Q1-Location	Q2-Age	Q3-Gender	Q4-BikeOwner	Q5-StartedCycling	Q6-WhenStarted	Q8-SuperhighwayUsed	Q9-SuperhighwayFrequency	Q10-Width
2	1203127707	Kent	30-39	Male	Yes	No, I already cycled	1 year	Superhighway 7, Merton to the City via A24-A3	Only tried it once	fair
3	1202170092	Camden	40-49	Male	No	Yes, because of the Cycle Hire Scheme				
4	1200582066	Richmond	18-29	Male	Yes	No, I already cycled	Longer			
5	1200133603		40-49	Male	No	Yes, because of the Cycle Hire Scheme	Longer	Superhighway 7, Merton to the City via A24-A3	Once a week	good
6	1199557604	Sutton	30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Occasionally	very good
7	1198461180	Lambeth	30-39	Male	Yes	No, I already cycled	Longer	Both	Several times a week	bad
8	1198250056		30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Once a week	good
9	1198050062	Merton	40-49	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Only tried it once	good
10	1198011867		30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Once a week	good
11	1197977212	Lambeth	30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Several times a week	fair
12	1197744842	lambeth	30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Several times a week	fair
13	1197561900	Islington	50	Male	Yes	Yes, because of the Cycle Hire Scheme				
14	1197323125	Richmond upon Thames	40-49	Male	Yes	No, I already cycled	Longer			
15	1196772000	Camden	40-49	Male	Yes	No, I already cycled	Longer			
16	1196771676	kingston	30-39	Male	Yes	No, I already cycled	Longer			
17	1196770485	Stockwell	18-29	Female	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Several times a week	good
18	1196767034	Westminster	30-39	Male	Yes	No, I already cycled	Longer			
19	1196759023	Harrow	50	Male	Yes	No, I already cycled	Longer			
20	1196757652	Merton	30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Only tried it once	bad
21	1196691925		30-39	Male	Yes	Yes, because of the Cycle Hire Scheme				
22	1196270317		40-49	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Only tried it once	fair
23	1196250068		50	Male	Yes	No, I already cycled	6 months ago			
24	1196188131		40-49	Male	Yes	No, I already cycled	Longer			
25	1196187421		40-49	Female	Yes	No, I already cycled	1 year	Both	Once a week	very good
26	1196184639	Merton	30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Several times a week	very good
27	1195882058		18-29	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Occasionally	very bad
28	1195861961		50	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Once a fortnight	very good
29	1195838523	Westminster	50	Male	Yes	Yes, because of the Cycle Hire Scheme				
30	1195818203		30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Several times a week	good
31	1195804669	Wandsworth	30-39	Male	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Only tried it once	good
32	1195777047	tower hamlets	18-29	Male	Yes	No, I already cycled	Longer			
33	1195770852	Hackney	30-39	Female	Yes	No, I already cycled	Longer	Superhighway 7, Merton to the City via A24-A3	Occasionally	good

Dataset no.1

	A	B	C	D	E	F	G	H	I	J	K
1	Day	Number of Bicycle Hires		Month	Number of Bicycle Hires		Year	Number of Bicycle Hires		Month	Average Hire Time (mins)
2	30/07/2010	6,897		Jul 10	12,461		2010	2,180,813		Jul 10	17
3	31/07/2010	5,564		Aug 10	341,203		2011	7,142,449		Aug 10	17
4	01/08/2010	4,303		Sep 10	540,859		2012	9,519,283		Sep 10	15
5	02/08/2010	6,642		Oct 10	544,412		2013	8,045,459		Oct 10	15
6	03/08/2010	7,966		Nov 10	456,304		2014	10,023,897		Nov 10	14
7	04/08/2010	7,893		Dec 10	285,574		2015	9,871,839		Dec 10	17
8	05/08/2010	8,724		Jan 11	403,178		2016	10,303,637		Jan 11	16
9	06/08/2010	9,797		Feb 11	398,292		2017	10,446,044		Feb 11	17
10	07/08/2010	6,631		Mar 11	556,155		2018	10,567,540		Mar 11	20
11	08/08/2010	7,854		Apr 11	672,629		2019	10,169,836		Apr 11	27

Dataset no.2

Some of the techniques to clean the data were: by assigning appropriate datatypes (dtypes) and suitable datetime formats when dealing with the time-series. Also, dealing with missing values appropriately, such as dropping columns with missing values or replacing data with 0's where appropriate. At times I've also cleaned the data by melting columns into single variable and creating a normalised cross tabulated table:

Exploratory and explanatory data visualisation

The variables of interest and their types in this research are:

- Age: (Numerical - Ratio)
- Gender: (Categorical - Nominal)
- Initial users of CHS: (Categorical - Nominal)
- Experience on various aspects ratings: (Categorical - Ordinal)
- Comments: (Categorical - Nominal)
- Time of hire in initial stages: (Numerical – Ratio)
- Number of bicycle hires: (Numerical - Ratio)
- Year: (Numerical - Ratio)
- Month: (Numerical - Ratio)
- Day: (Numerical - Ratio)
- Average time of hires: (Numerical - Ratio)

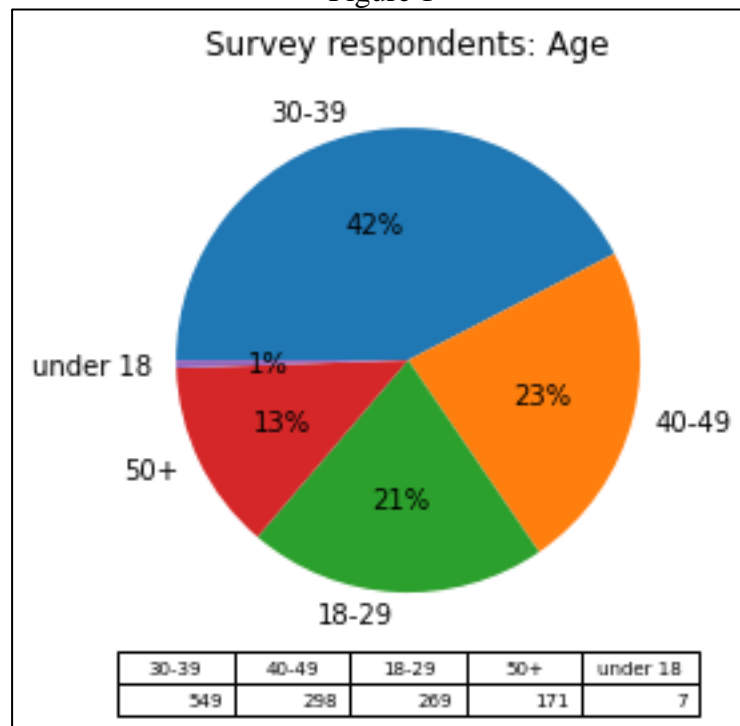
I have created basic graphs such as: a pie chart, a bar chart, a heat map, a box plot, a histogram and scatterplot. Some of the more advanced visualisations have also been included such as time-series analysis and qualitative data as a word cloud.

The key finding of this search are as follows:

Initial impact of Cycle Hire Scheme in 2010

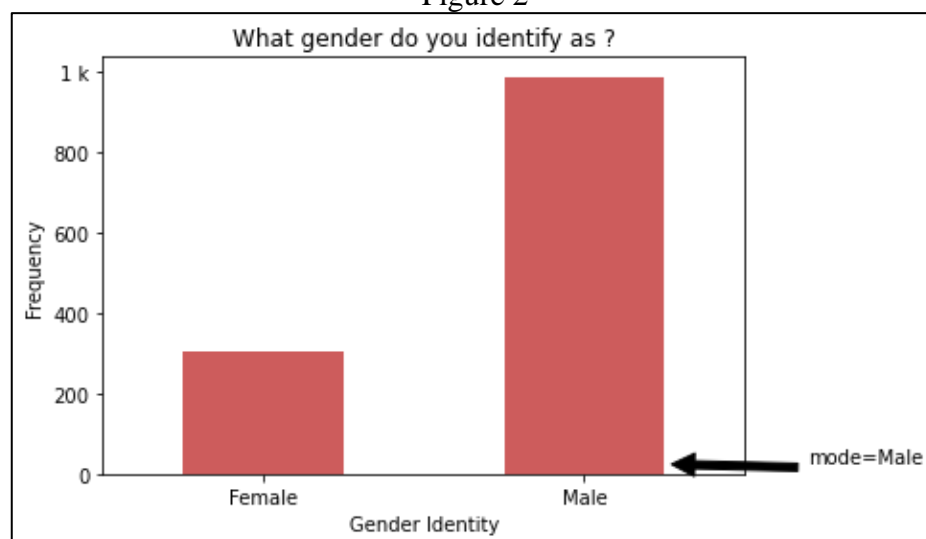
1 - Age: This survey was conducted on 1298 cyclist. The majority of the participants were between the ages 30-39.

Figure 1



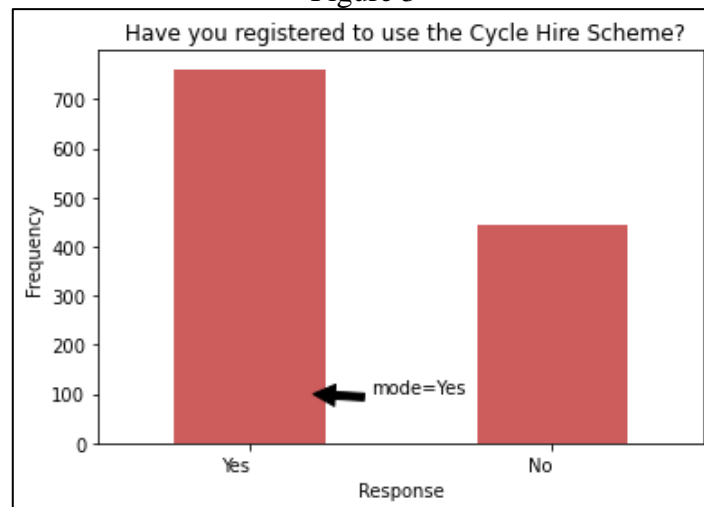
2- Gender: The majority of participants were male, with a 76.5%.

Figure 2



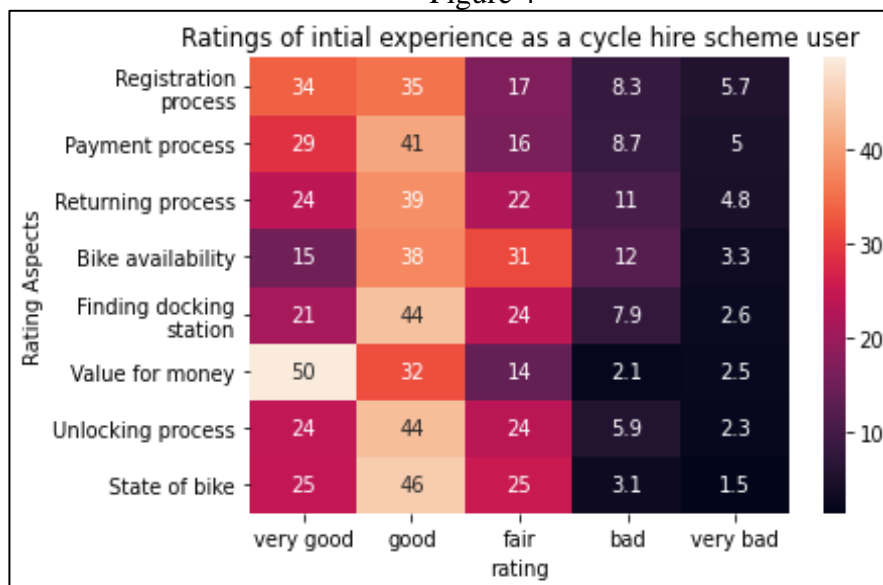
3 - Current cyclist who had registered to use cycle hire scheme during the initial launch. 63.2% of cyclists registered with the scheme out of the 1298 total survey participants.

Figure 3



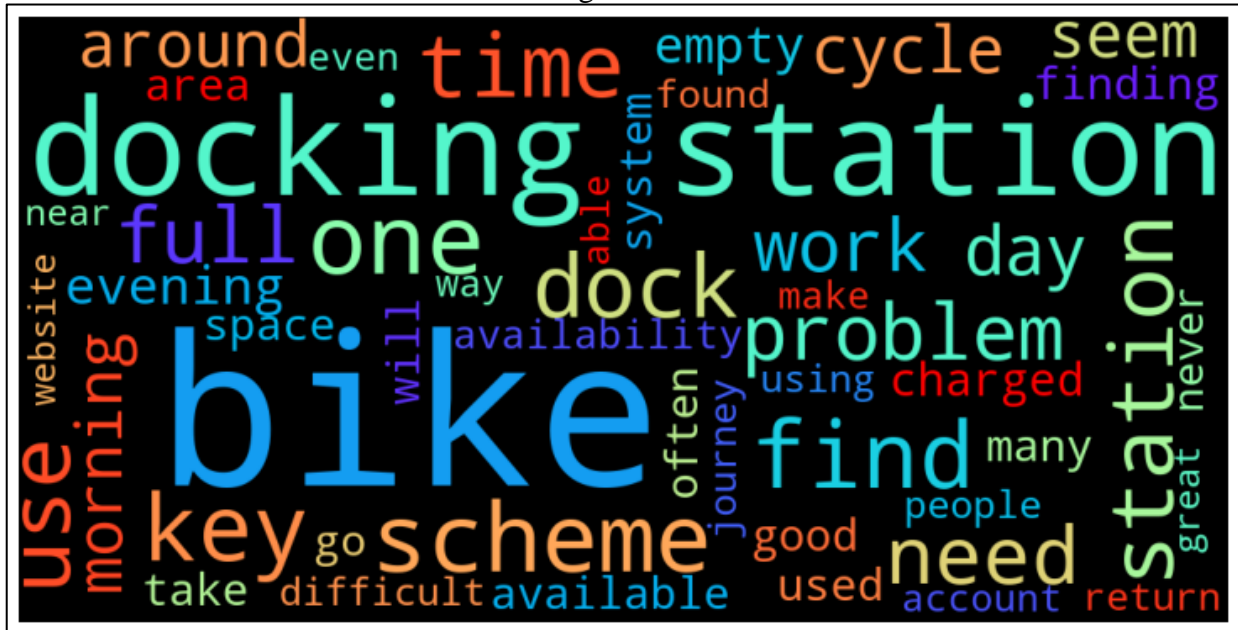
4 – The user experiences rating on different aspects of the cycle hire scheme were mostly positive. Such as, aspects as the value for money and the state of the bike got the most positive ratings in great percentages. But there were also issues which had to be dealt with, such as the registration process and the returning process which had some bad ratings.

Figure 4



Question 5 – Further Comments about Cycle hire scheme were also made by 523 participants. I have made a word cloud using the comments provided, where one of the biggest concerns seemed to be not finding enough docking stations. Hence, the most used words in the word cloud were: ‘docking station’, ‘find’, ‘dock’, ‘station’ and ‘full’.

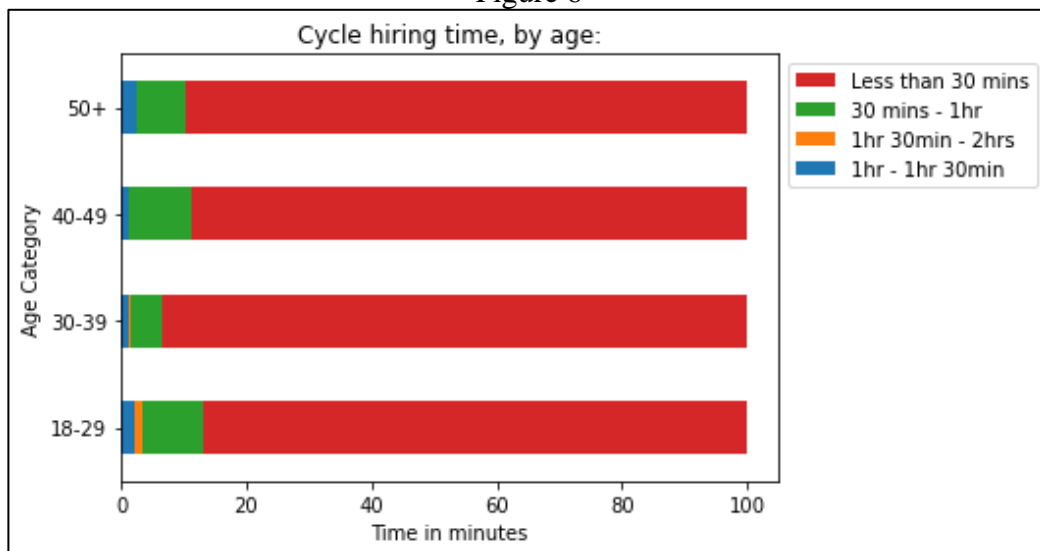
Figure 5



I did some research to find the number of docking stations available on the 10/10/10 [2] . And there were 344, they did intend to have 400 by 2011, a target which was met. Also, by 2016, there were 761 docking stations available in London. In the next research section below, it would be interesting to find if these improvements made upon user feedback had positively impacted on the growth of the scheme since its launch.

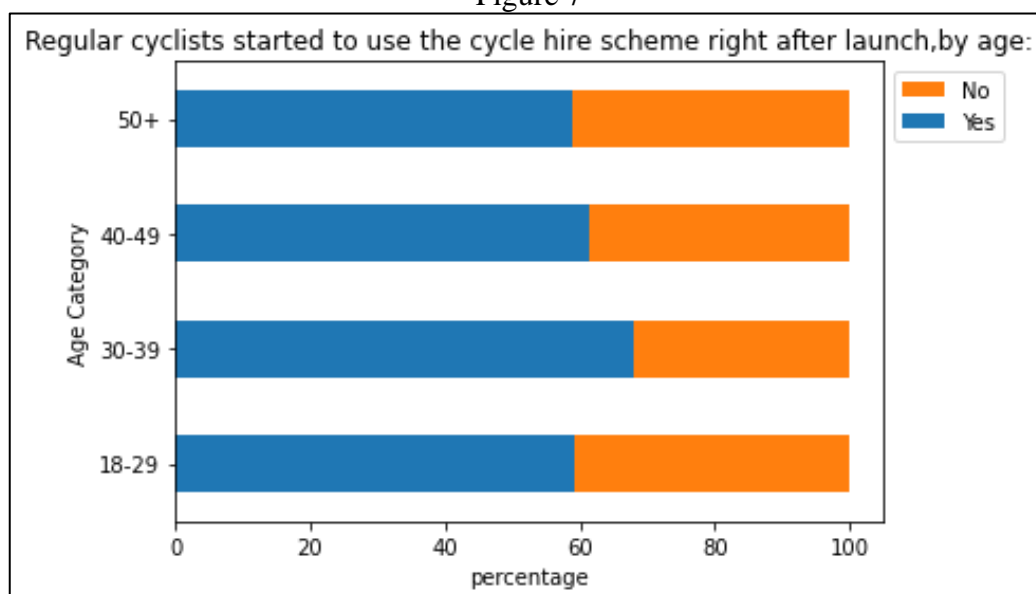
6 – This multivariate stacked graph shows hiring times per age during initial launch until Oct 2010. Over 93% of participants had said to hire a bike for less than 30 min. People within the ages 18-29 had said to hire the longest, less than 2 hours.

Figure 6



Question 7 – This multivariate graph shows that the majority of cyclists who had registered with the ‘Cycle Hire Scheme’ during it’s initial launch were between the ages 30-39.

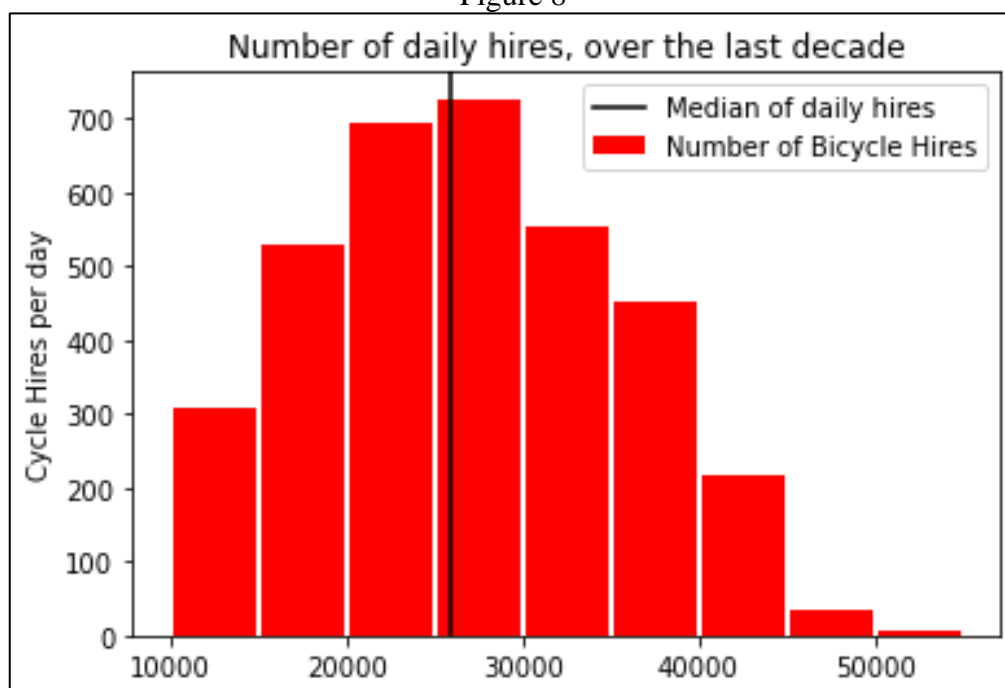
Figure 7



Evaluation of the Cycle Hire Scheme a decade later in 2020

8 – Figure 8 below displays the number of daily hires since the scheme launch as a histogram. The median of daily hires was at 26000.

Figure 8



9 – A box plot has been used to display the average hires times since 2010. The average median number of hire times is 18 min. But there are also some expectations hiring up to 36 minutes.

Figure 9



10 – In figure 10.0, I had used a seaborn scatterplot to view the relationship between the number of hire and the years. Overall, it does have a positive correlation. In figure 10.1, I plotted a line graph to see the rolling mean which is positive and the stationary deviation which seems overall stationary. The year with the highest number of bike hires was 2018, with over 1.2million hires.

Figure 10.0

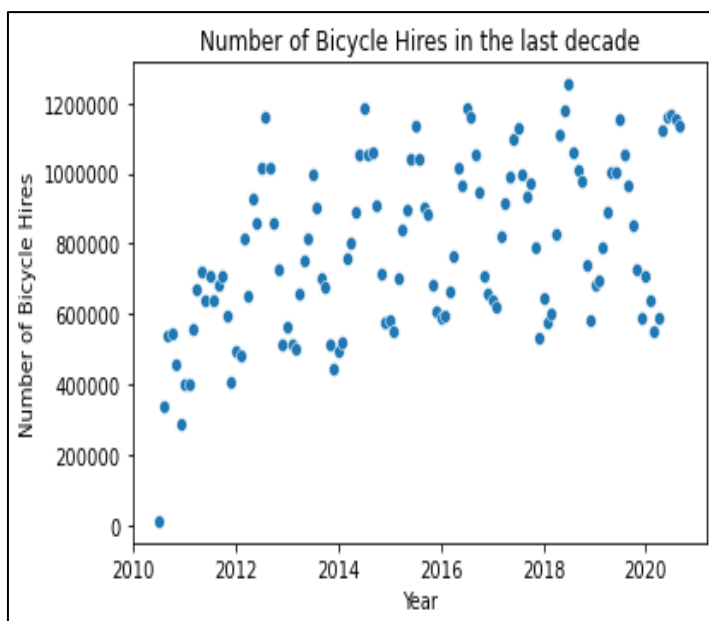
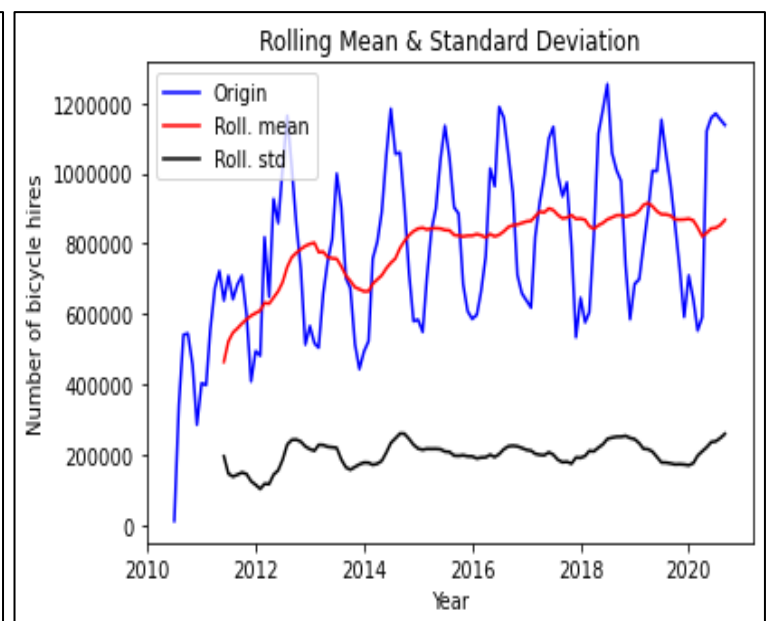


Figure 10.1



11 – I want to compare the difference in the scheme’s progress since 2010. Do to so, I have produced two graphs. Figure 11.0 displays the monthly hires of 2010 and figure 11.1 displays the monthly hires of 2020.

The improvements made by TfL after receiving feedback from initial users of the scheme, might have contributed to the growth positively. As, the maximum number of bicycle hires in 2020 was 1,169,418. And the maximum number of bicycle hires in 2010 was 544,412. The difference is of 625,006 of rides increase from 2010 to 2020.

Figure 11.0

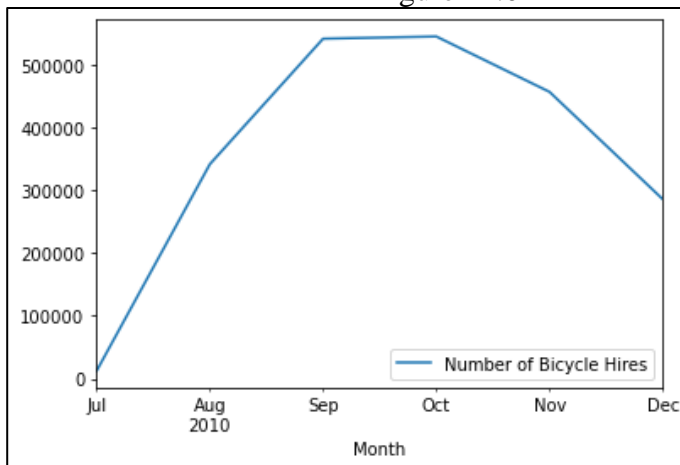
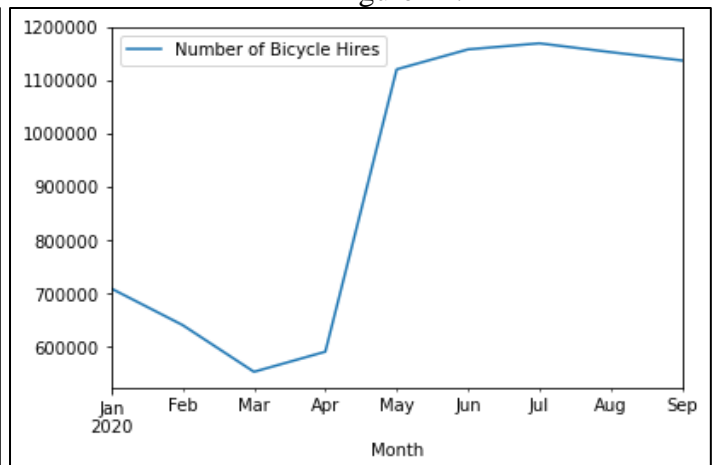


Figure 11.1



Other exterior contributing factors on the growth might have been present, which have not been examined in this research but there is research available on it. Such as people’s raising concerns about the climate change [\[3.0\]](#) and desire to cycle as a way to improve physical health [\[3.1\]](#) .

Also, in 2020 due to the COVID-19 pandemic, there has been drastic changes in transport usage, as people preferred using transport types which enabled social distancing, such as using the cycle hire scheme [\[4\]](#) .

I have used `.describe()` multiple times during this research to find statistical data, below is an example on how I used `.describe` to be able to find specific data about a particular year and used that to plot the graphs above, figures 11.0 and 11.1.

```
1 df.loc['2010'].describe().round(2)
```

Number of Bicycle Hires	
count	6.00
mean	363468.83
std	201293.79
min	12461.00
25%	299481.25
50%	398753.50
75%	519720.25
max	544412.00


```
1 df.loc['2010'].describe().max().round(2)
```

Number of Bicycle Hires	544412.0
dtype:	float64

12 – Future forecast: Can the Transport Strategy’s plan be achieved by 2026 of meeting 1.5 million cycling trips per day?

Figure 12.0

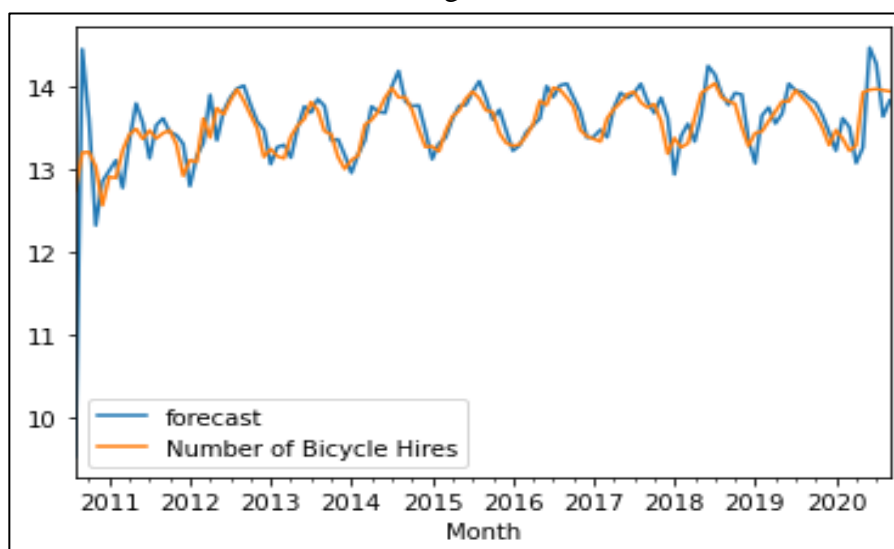
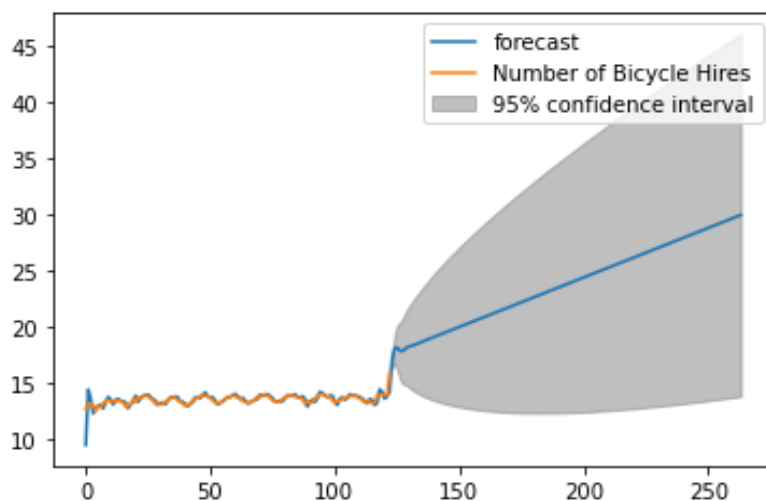


Figure 12.0 shows the forecast in detail until 2020. As there was data available the forecast seems to be pretty accurate for most years.

I don't have much knowledge on using the ARIMA model for forecasting, therefore I did research and adapted online examples. I was not able to rename the y-axis for the results to display the numbers appropriately. Number 13 represents 500,000 and 14 represents 1,000,000.

Figure 12.1



In figure 12.1, I have used the ARIMA model to forecast until 2030, which would be number 175 in the graph.

By the year 2026, which is at around number 160, 1.5 million rides will be met a month according to the forecast. But 1.5 million rides will not be met per day.

Conclusion and Evaluation

The key findings of this research have been that the initial impact of cycle hire scheme launched in 2010 (30th July- 30th October) was overall positive but required some improvements in areas such as easing the registration process and making more docking stations, so users have a wider range of choices to return the bike after a ride. From 2011 onwards work has been done to fulfil the improvements required.

Also, a decade later, the ‘Cycle Hire Scheme’ has increased its growth by more than double compared to the figures of the number of hires in 2010. And, according to forecasts, the Transport’s Strategy target to increase cycling trips by 1.5million a day by 2026, will not be met. Rather, the target might be met during a month in 2026. On the positive side, the scheme is gaining popularity and the numbers will keep increasing in the future, so even if it might take longer to meet the set target, it’s still achievable.

During this research, along with other things, I’ve learnt to create a word cloud and cleaning data containing time-series. For future improvements, I’d like to rename both the axis for the forecast graphs, as currently it does not as communicate clearly as I’d like it to.

Bibliography:

[1] 2010. [online] Available at:
<https://www.london.gov.uk/sites/default/files/gla_migrate_files_destination/FINAL%20REPORT.pdf, page no.9> [Accessed 11 December 2020].

[2] 2010. [online] Available at:
<https://www.london.gov.uk/sites/default/files/gla_migrate_files_destination/FINAL%20REPORT.pdf, page no.40> [Accessed 11 December 2020].

[3.0] Publications.parliament.uk. 2020. *Active Travel: Increasing Levels Of Walking And Cycling In England - Transport Committee - House Of Commons*. [online] Available at:
<<https://publications.parliament.uk/pa/cm201719/cmselect/cmtrans/1487/148704.htm>> [Accessed 11 December 2020].

[3.1] Wighton, K. and Campus, S., 2020. *Walking Or Cycling To Work Associated With Reduced Risk Of Early Death | Imperial News | Imperial College London*. [online] Imperial News. Available at: <<https://www.imperial.ac.uk/news/197618/walking-cycling-work-associated-with-reduced/>> [Accessed 11 December 2020].

[4] Wise, H., 2020. *Bike Sales Are Flourishing In Britain Due To Covid-19. Will It Last?*. [online] This is Money. Available at:
<<https://www.thisismoney.co.uk/money/markets/article-8315515/Bike-sales-flourishing-Britain-pandemic-last.html>> [Accessed 11 December 2020].

[5] Word Cloud source for code adaptation:
<https://www.datacamp.com/community/tutorials/wordcloud-python> [Accessed 07 December 2020].

[6] Forecasting source for code adaptation:
<https://www.youtube.com/watch?v=e8Yw4aIG16Q> [Accessed 07 December 2020].

[7] Time series dealing with indexes source:
<https://sergilehkyi.com/tips-on-working-with-datetime-index-in-pandas/> [Accessed 07 December 2020].