**An Exploratory Analysis on London Bike share**



Santander Cycles or (**Boris Bikes**) is London's self-service bike-sharing scheme for short journeys. The scheme is sponsored, with [Santander UK](https://en.wikipedia.org/wiki/Santander_UK) being the main sponsor from April 2015.

The scheme commenced operations as Barclays Cycle Hire on 30 July 2010 with 5,000 bicycles and 315 docking stations distributed across the [City of London](https://en.wikipedia.org/wiki/City_of_London) area and parts of eight [London boroughs](https://en.wikipedia.org/wiki/London_borough).

check more about Santander Cycles: <https://en.wikipedia.org/wiki/Santander_Cycles>

I performed an exploratory data analysis (EDA) on Historical data for bike sharing in London 2015 to find out if weather, season, holidays and other such variables have any effect on number of bike user.

The analysis in question was carried out with python & Excel. python library was used to gather data from Kaggle. I also explore, assess and manipulate the data using the pandas library in python.

The final step involves creating impactful visualization using excel to bringing all the insights together.

**Dataset Used:**

Historical data for bike sharing in London ‘Powered by TFL Open Data’ from Kaggle. The datasets contain data from year 2015 to 2017 with 17414 rows & 10 columns for every hour of the year. I’ve filtered the data only for year 2015 for my analysis.

**Metadata:**

**Timestamp** - timestamp field for grouping the data

**Cnt** - the count of a new bike shares

**T1** - real temperature in C

**T2** - temperature in C "feels like"

**Hum** - humidity in percentage

**Wind\_speed** - wind speed in km/h

**Weather\_code** - category of the weather

**Is\_holiday** - Boolean field - 1 holiday / 0 non-holiday

**Is\_weekend** - Boolean field - 1 if the day is weekend

**Season** - category field meteorological seasons: 0-spring; 1-summer; 2-fall; 3-winter.

Weather\_code category description:

1 = Clear; mostly clear but have some values with haze/fog/patches of fog/ fog in vicinity.

2 = scattered clouds / few clouds

3 = Broken clouds

4 = Cloudy

7 = Rain/ light Rain shower/ Light rain

10 = rain with thunderstorm

26 = snowfall

94 = Freezing Fog

**Data Cleaning**

Data cleaning & manipulation was done using python,python library was used to gather data from Kaggle. I also explore, assess and manipulate the data using the pandas library in python.

Finally, write the final data frame to an excel file that I used in my visualizations.

Check out my data cleaning & manipulation process in google Collab:

<https://colab.research.google.com/drive/1m454ewILZR7ijqAckQCbgNPlGf6AbIot?usp=sharing>

**Analysis**

**Part 1:**

This analysis shows the general statistics for London bike share dataset, by using excel to visualize the bike usage by seasons, months, weeks (weekdays & weekends), holidays & non-holidays and hours of the day.

**Total Number of Riders by Weekdays**

Monday, Tuesday, Wednesday & Thursday show the highest number of riders, while the number of riders is much lower in the weekend. people use bike mostly during weekdays i.e., to & from work.

**Number of Bikes Used on Weekdays & Weekends**

About 76% of the total bikes have been used on weekdays & 24% were used on weekends

**Bikes used on Holidays vs non-holidays.**

Almost 99% of the total bike was used on non-holidays, while about 1% were used on holidays.

**Number of Bikes on Holidays for the Year 2015**

The bars show the total number of bikes used for particular holiday. Spring bank holiday show the highest number of bikes used, it is a day off work for most people in London, businesses & organization.

**Riders Hours of the Day and Bike**

8 AM & 5 PM shows the highest number of bike riders, which correspond to the time that people go to & from work.

**Bike Riders in Weekdays and Weekends for Each Hour of The Day**

Most of the bike have been used during rush hour time between 8 AM & 5 PM in weekdays and 12 PM to 6PM in weekends.

**Bike used on weekends in various seasons.**

Summer had the highest number of bike users on weekends than any other seasons.

From our analysis People used the bikes mainly for work commutes but as can be seen from the graphs on weekdays, many people used them mostly from 8am and 7pm, people also used them for leisure as weather gets better in the weekend.

**Part 2:**

Investigating the effects of season, weather, temperature, humidity and wind speed over the bike usage.

Riding bike is an outdoor activity and naturally it is affected by weather and seasons equally. As a summer morning will draw more people out than a rainy or snowy day, it is important to consider the effect of seasons as well.

**Number of Riders in Four Seasons**

Around 3.2million people used London bike service in summer months alone, followed by 2.43million in Autumn and 2.4million in Spring, Winter shows the lowest number of riders in 2015 -1.69million

**Effect of humidity and windspeed on number of bike riders**

For most past of the year, London has a humidity between 50% - 80% which means there is high chance of precipitation. Number of bike riders gradually increase when the weather is clear as seen from the chart. Number of bikes change with humidity, windspeed, temperature and weather.

Graph showing relation between humidity, windspeed and number of bikes used on a **clear day.**

Humidity has negative correlation with bike usage. As humidity in the air increases, the number of bikes usage reduces gradually when the weather is rainy or broken cloud.

Graph showing relation between humidity, windspeed and number of bikes used on a **broken cloud day.**

Graph showing relation between humidity, windspeed and number of bikes used on a **rainy day.**

Graph showing relation between humidity, windspeed and number of bikes used on a **snowy day.**

**Relation between Temperature and Number of Bikes**

The bar chart shows the total number of riders for each month and the line chart shows the average temperature. As temperature increases so does the number of riders. It seems temperature has a positive correlation with number of riders.

**July was the warmest month in 2015**

**February is the coldest month in 2015**

**How does number of bikes change with humidity, temperature and windspeed**

As humidity decreases, the chances of precipitation decrease, naturally number of bike rides increases. The opposite is true for temperature. People are more likely to use bikes in summer and spring months when temperature is moderate, than in winter months when its freezing cold! There is no visible correlation between wind speed and number of bikes.

The vertical lines in Humidity, Temperature and Wind speed graphs show their average values.



**Conclusion:**

From the above analysis, the following conclusions can be drawn:

1. Weather and seasons have an influence over the number of bike riders in London. Temperate weather conditions as can be seen in summer and fall, draws more riders out than a rainy or snowy day.
2. Number of riders seems to increase as temperature increases, annually as well as daily. July and August had the highest number of riders, while these were also the warmest months in 2015.
3. Number of riders gradually decreases as humidity of the air increases. It coincides well with people heading out on clearer days than on cloudy or rainy days.
4. Most people tend to use bikes during the daytime rush-hours i.e., at 8 AM and 5 PM during weekdays, as they commute to and from workplaces. On weekends most people use bikes in early afternoon (1PM — 5PM).