**BUSINESS UNDERSTANDING**

It’s no secret that transportation is the [largest source](https://uspirg.org/news/usp/new-federal-data-show-transportation-sector-now-largest-source-carbon-pollution-united) of [carbon emissions](https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions) all over the world. It is therefore not surprising that, besides promoting the use of public transit, biking and walking, cities are also embracing [electric vehicles](https://www.nrdc.org/experts/madhur-boloor/electric-vehicles-101) (EVs) in their quest to cut carbon emissions. But some are thinking outside of the box: In addition to electrifying municipal fleets, many cities are also engaging with community and industry partners to support electrified options that don’t require owning your own EV, such as car sharing which has been demonstrated to [reduce car ownership](https://www.sciencedirect.com/science/article/abs/pii/S0967070X16307314) amongst users.

This reflects both a desire from cities to reduce carbon pollution and a desire more aggressively from residents for more varied, affordable, and sustainable transportation options. Additionally, several of these cities are striving to centre equity in the design and delivery of these programs, seeking to ensure that their benefits are available to communities traditionally most affected by transportation pollution.

**Desired outputs of the project**

The objective of this project for an electric car-sharing service company is to process stations data to understand electric car usage over time by solving for the following research question.

Research Question

* Identify the most popular hour of the day for picking up a shared electric car (Bluecar) in the city of Paris over the month of April 2018.

Bonus Questions (Optional)

* What is the most popular hour for returning cars?
* What station is the most popular?

1. Overall?
2. At the most popular picking hour?

· What postal code is the most popular for picking up Blue cars? Does the most popular station belong to that postal code?

1. Overall?
2. At the most popular picking hour?

· Do the results change if you consider Utilib and Utilib 1.4 instead of Blue cars?

The project plan included but not limited to

1. Initial assessment of tools and techniques where in this case SQL and python was used on google Collaboratory to carry out Data wrangling and analysis

2. Determining Data mining goals and business success criteria. The intended output of the project was to advise Electric car-sharing service company on the most popular hour of the day for picking up a shared electric car (Bluecar) in the city of Paris over the month of April 2018

3. Data understanding stage. Data used in this project was provided through the following URL ~[<http://bit.ly/autolib_dataset>]

The dataset comprised one CSV file. This stage also included loading, exploratory data analysis and verification of the quality of data provided in terms of completeness, accuracy and if there are missing values.

4. Data preparation. As mentioned in the previous stage, the data to be used had been provided which were four data sets. Here, among the activities carried out included data cleaning, integrating and merging datasets (if any).

5. Data analysis. This stage included analysing the data using SQL to generate insights and stating the assumptions made if any. An assessment of technique according to the data mining success criteria and desired test design was also done

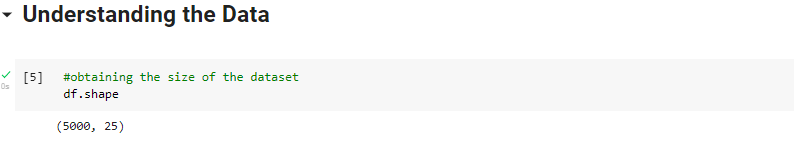
6. Lastly was the evaluation of results and giving recommendations. A thorough review of the data mining engagement was carried out in order to determine if there was any important factor or task that has somehow been overlooked.

**DATA UNDERSTANDING**

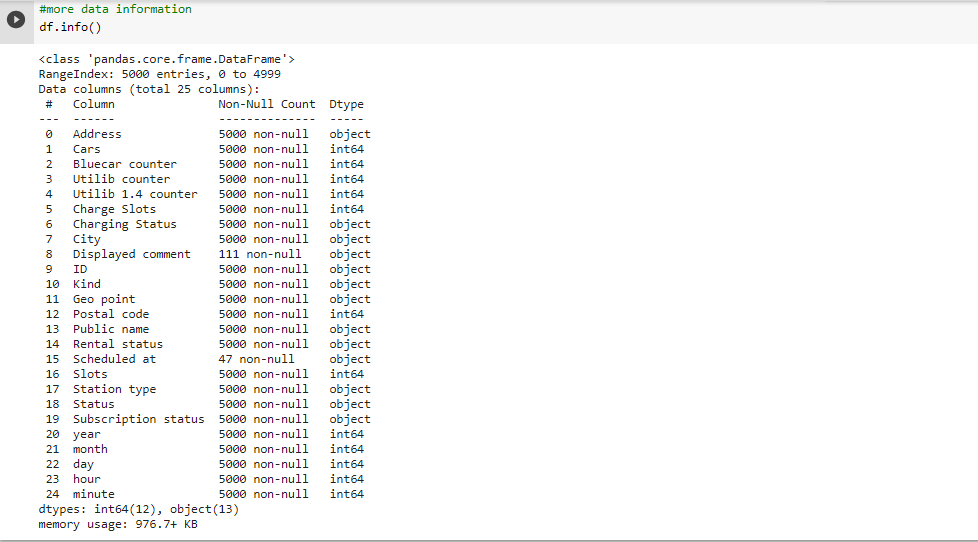
The dataset used in this project was provided using the following URL [<http://bit.ly/autolib_dataset>]. It was then downloaded and uploaded on the colab notebook by first importing the relevant libraries as below.



Using the .shape() function the data was found to be 5000 x 25 column as below

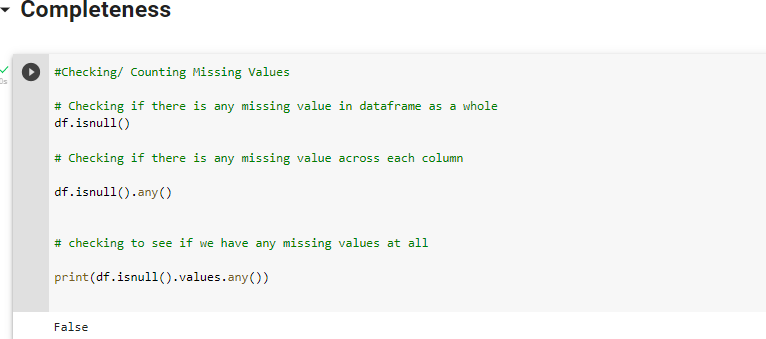


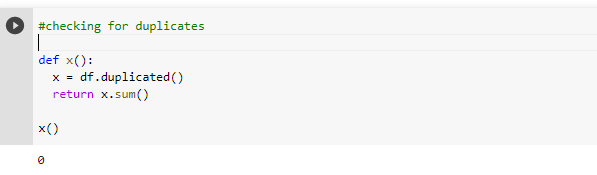
Further information using the .info() function in pandas was obtained as below showing that all columns in this dataset were of either object or integer type



The description of what each column represents was given in a separate documentation using the following link “<https://drive.google.com/file/d/13DXF2CFWQLeYxxHFekng8HJnH_jtbfpN/view>”.

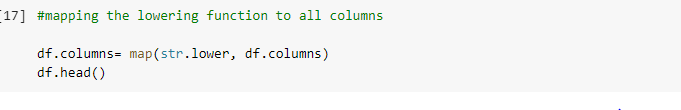
The data was then checked for missing values and duplicates as below and it was found to have no duplicates not missing values.



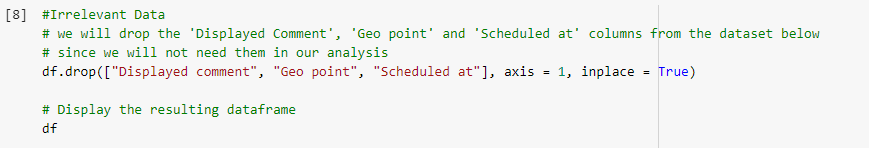


**DATA PREPARATION**

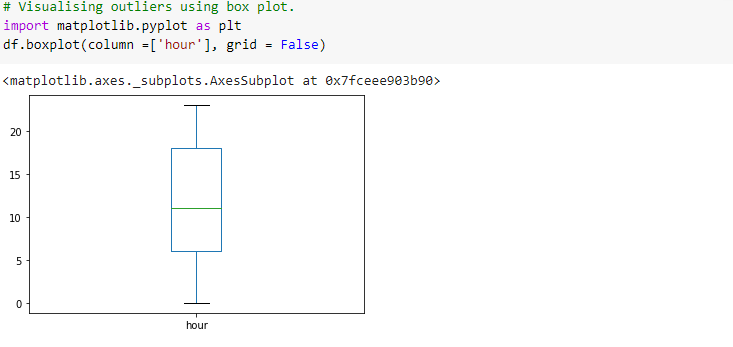
All the column names in the dataset were first converted to lowercase using the following code



There were irrelevant columns in this dataset and they were dropped as follows



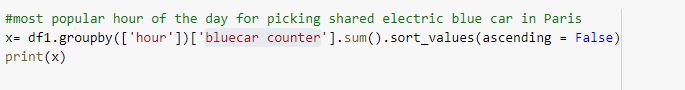
The dataset was consistent as it did not have repeated data points. While checking for outliers using a Box plot, the hour and cars column did not have any.



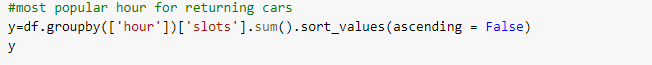
**DATA ANALYSIS**

To be able to identify the most popular hour of the day for picking up a shared electric car (Bluecar) in the city of Paris over the month of April 2018, a subset of the data was obtained and renamed df1. The subset contained data points only from the city of Paris over the month of April 2018.

Using the .groupby function by the hour column, the most popular hour was calculated.



To find the most popular hour for returning cars, the initial full dataset was used and the slots available for parking column was referenced.



The code below was used to find the most popular station.



To find the most popular postal code for picking up blue cars was calculated as follows



**RECOMMENDATION**

The results showed that the most popular hour of the day for picking up a shared electric car (Bluecar) in the city of Paris over the month of April 2018 was 21hrs.

The most popular hour for returning cars was 17hrs. The most popular station overall was Paris/Porte de Montrouge/8.

**EVALUATION**

The analysis should be re-evaluated to find if results change if you consider Utilib and Utilib 1.4 instead of Blue cars.