**1. Business Understanding**

**Business Objective**

The main objective of this report is to understand electric car usage over time by solving the following questions;

* 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.
* What is the most popular hour for returning cars?
* What station is the most popular?
  + Overall?
  + 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?
  + Overall?
  + At the most popular picking hour?
* Do the results change if you consider Utilib and Utilib 1.4 instead of Blue cars?

**Business Success Criteria**

Identify the most popular hour of the day for picking up a shared electric car that will increase the return on investment of the Electric company. We will conduct the following tasks;

**Assessing the Situation**

1. **Resource Inventory**
2. Dataset:

Autolib\_dataset : Link(“<http://bit.ly/autolib_dataset>”)

Autolib\_clean : Link(“<https://drive.google.com/open?id=10PX_tJt8F0TNphMtYK3inMd7hdN9pAq0>”)

1. Software( Github, Google colab)

2. **Assumptions**

1. The data provided is correct and up to date

3. **Constraints**

1. There are no constraints

**Data Mining Goals**

Our data mining goals for this project are as follows:

1. 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.
2. What is the most popular hour for returning cars?
3. What station is the most popular?
   1. Overall?
   2. At the most popular picking hour?
4. 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?
5. Do the results change if you consider Utilib and Utilib 1.4 instead of Blue cars?

**Data Mining Success Criteria**

Our success criteria will be measured by the following criteria;

* We Identify the most popular hour of the day for picking up a shared electric car (Bluecar) in the city of Paris.
* We target the most popular hour for returning cars?
* We target the station is the most popular

**2. Data Understanding**

**Data Understanding Overview**

For this project, we are using the availed dataset by the company.

* Autolib\_dataset

**Data Description**

We have one dataset available for this project. A detailed description of the dataset is

provided as follows:

Autolib\_dataset- This dataset, on the other hand, focuses on the data of an electric car-sharing service company. It contains eleven columns;

1. Address
2. Cars
3. Bluecar counter
4. Utilib counter
5. Utilib 1.4 counter
6. Charge slots
7. Charging status
8. City
9. Displayed comment
10. ID
11. Kind
12. Geo point
13. Postal code
14. Public name
15. Rental status
16. Scheduled at
17. Slots
18. Station type
19. Status
20. Subscription status
21. Year
22. Month
23. Hour
24. Minute

**Verifying Data Quality**

The **Autolib** dataset consists of columns “Displayed comment”, “Scheduled at” which contains Null values. The rest of the columns in the dataset do not have null values.

**3. Data Preparation**

These are the steps followed in preparing the data

1. Loading Data

Loaded the datasets from the CSV and then created an dataframes from

them.

1. Cleaning Data

Created our one data frame from the dataset provided.

Got rid of all the null value rows and worked with columns with values.

1. Dropping Unnecessary columns from our Dataset.

Dropping Unnecessary columns from our Dataset so that it would be easier to get good insight from the data.

1. Deriving New Attributes

Observed several key insights from the data;

* 1. 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.
  2. What is the most popular hour for returning cars?
  3. What station is the most popular?
     1. Overall?
     2. At the most popular picking hour?
  4. 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?
  5. Do the results change if you consider Utilib and Utilib 1.4 instead of Blue cars?

**4. Analysis**

During our analysis, we were able to single out the following cities;

1. 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 is - Hour 23.
2. What is the most popular hour for returning cars? - Hour 23.
3. What is the most popular car?
   1. Bluecar : 7
   2. Utilib : 2
   3. Utilib 1.4: 3
4. What station is the most popular?
   1. Overall? - 32 rue robert marchand : 50
   2. At the most popular picking hour? - 11 rue robert lavergne : 23
5. What postal code is the most popular for picking up Blue cars? Does the most popular station belong to that postal code?
   1. Overall? - 99 rue charles-de-gaulle
   2. At the most popular picking hour? - 65 boulevard de picpus
6. Do the results change if you consider Utilib and Utilib 1.4 instead of Blue cars?

Yes. The results change if I use Utilib and Utilib 1.4.

For example an instance where we want to get the most popular station (overall) from our dataset we get;

Blue Car - 32 rue robert marchand : 50

Utilib - 161 avenue de suffren : 2

Utilib\_1.4 - 123 grande rue : 3

The above analysis was done using Pandas. The full analysis can be found in the following notebook.

Github: [ [notebook Link](https://github.com/michaelteddy009/IP_week_4_moringa)]

**Recommendations**

From our analysis, we would recommend the most used station when using the blue car to be 32 rue robert marchand. Our main reason behind this recommendation would be that the stations listed above have the highest value compared to others.

Therefore, prioritizing these stations would not only increase the chances of increased revenue but also would increase the return on investment in the entire electronic company services plan and future investments.

1. **Evaluation**

From our business success criteria we have been able to determine the most frequently used station for electric cars. Our approach would be important for allowing the electric car sharing company to determine the most appropriate station to invest in development or upgrade of infrastructure. I have also provided a guide on how we performed data mining on our dataset in order to get such insights as from our recommendation.