**Electric car-sharing service company**

This is to process stations data to understand electric car usage over time by solving certain research questions. This is data which was produced in the month of May for 9 working days in the area

**Business understanding**

1. **Determine business objectives**

The company is trying to solve and understand the usage of electric cars overtime during the period of days given and

1. **Access the situation**

The company is trying to access the usage of electric cars of which we’ve been given ‘Bluecars’,’utilib’ and ‘utilib1.4’ of which all these are cars. We’ve been given the address of the cars, the charging\_stations, the postal code, The day, time, and hour. The rental status of which is either operational, broken, or non\_existent. There is also subscription status and GPS coordination of the station.

1. **Determine the data mining goals**

The main goal here is 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.

1. **Provide a project Plan**

To identify and plan ahead the hour of the day that people pick the shared electric car during the given period of time

**Data understanding**

1. **Collect**

We collect information from the data where the (Bluecar) column is and the hour and trying to figure out how the many people chose the car at that hour of the day

1. **Describe**

Next is to describe our collected data by knowing the object types and the values using the integer types and try to see if they are matched out together well

1. **Explore**

Here we explore our collected data information and figure out what we have to change, clean, or rename in the data so as to have a suitable workspace for us.

1. **Data quality**

Our collected data quality should be at top-notch where there are no added spaces, to remove Null values that are on the table so as to have an easy workflow

**Data preparation**

Here we make sure we clean our data where there are columns with null values we drop them, where there are duplicate values to remove them from the dataset so as to get accurate data.

Stripping white spaces and putting all of them in small capital letters so as to have an easy job with the data. To drop unnecessary columns which wouldn’t be applicable in this project

**Analysis**

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.**

The most popular hour of the day is 12 noon.

1. **What is the most popular hour for returning cars?**

The most popular hour of returning the cars is 1100 hrs

1. **What station is the most popular?**
   1. **Overall?**
   2. **At the most popular picking hour?**

The most popular station overall with the cars in it is ‘station’ and it also is the most popular picking hour

1. **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?**

The most popular postal code here is 75015 and it belongs to the station ‘station’.

1. **Do the results change if you consider Utilib and Utilib 1.4 instead of Blue cars?**

The results do not change whether it is Utilib or Utilib 1.4

**Recommendation**

According to our analysis, since ‘station’ is best used here in this data, our users go there to get the cars since they offer better quality service. I recommend the other stations to step up their game so as to provide good services to the people.

We should concentrate on the picking time and returning time and come up with ways to boost other stations so that people could use the cars more often.

**Evaluation**

**Evaluation of the results**

We see that *station* is the most picked station and the postal code 75015 was mostly used in this data.

**Next steps?**

Come up with more charging stations so that it can boost the number of cars picked and returned during the day