**TESLA**

ELECTRIC CAR USAGE

DATA REPORT

1. **Business Understanding.**

**Business Overview.**

Tesla Company is a company that deals with electric vehicles.These vehicles are shared amongst users. The vehicles are three in number: bluecar counter, utilib counter, utilib 1.4 counter. The electric vehicles have charging slots at different stations. After use, these electric The stations these vehicles are parked at are of different addresses and some are visited more frequently than others; the same applies to the vehicles.{some vehicles are used more than others}

**Business Objective**

The primary objective of this report 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.

Secondary objectives are:

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**

To identify the most popular hour of day for picking up a bluecar from any station.

**Assessing the situation**

1. Resource inventory

* The dataset (Autolib) is our resource inventory.
* Software(Github, Google Collab, SQL)

1. Assumptions

* The data provided is correct and up to date

1. Constraints

* There are no constraints.

**Data Mining Goals**

* Target most popular picking stations
* Identify most popular hour of day for returning bluecar
* Identify most popular postal code for picking up blue cars

**Data Mining Success Criteria**

Our success criteria will be measured by the following criteria:

* We target the most popular hour of the day for returning bluecar.

**2.Data Understanding**

Data Understanding Overview

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

* Autolib dataset

Data Description

We have one dataset available for this project. A detailed description of the dataset is provided as follows:

The dataset has 25 columns containing addresses of the pickup stations,types of electric cars,name of the stations,year,month and day etc.

Verifying Data Quality

* The dataset had missing values in two columns: Sceduled\_at and Displayed\_comment.
* No duplicate values existed in our dataset.
* The dataset had two similar columns: cars and bluecar counter.
* The dataset consisted of columns not named in an appropriate manner.

**3. Data Preparation**

These are the steps followed in preparing the data

**a. Loading Data**

Loaded the datasets from the CSV onto the google collaboratory

**b. Cleaning Data**

While doing data exploration, we noticed that in the columns cars and bluecar counter were similar therefore we decided to drop the column cars.

We also observed that two columns(displayed\_comment, scheduled at) consisted of mostly empty values and thus we dropped them as they were irrelevant.

The days of the week were named in numerical order; we changed them to represent the actual days of the week.

**4. Data Analysis**

From the analysis, we can see that the most popular hour for picking up bluecar counter was the 3rd hour of the day at a frequency of 7 times. The most popular hour of the day for returning a bluecar was the 0th hour at a frequency of 6 times.

It was observed that the most popular station was Apinay-sur Seine/Foch/53.

The above analysis was done using Python. The full analysis can be found in the following [Link](https://colab.research.google.com/drive/1DX8tBAKHS4RoYVqWxtLqQuhaFMHFJepH?usp=sharing)