**1)** **Business Understanding**

**Business Objectives**

**Business Overview**

Bluecar is an electric car sharing company based in the city of Paris. As a Data Scientist for this company, I am required to study and analyze the dataset aiming 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.

**Business Objectives**

The objective of this report is to draw insights from April 1 to April 9, 2018. First pilot: every 5 minutes from October 6, 2017, 11:13 AM to October 8, 2017, 10:21 AM, second Pilot: every minute from October 9, 2017, 15:53 PM to October 10, 2017, 15:31PM and production: every minute from October 30, 16:59 PM to July 31, 2018, 23:59 PM provided by opendataparis.com where the Autolib availability information was available in real time. The insight will assist in understanding electric car usage over time.

**Business Success Criteria**

To determine 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.

**Assessing the Situation**

Requirements, Assumptions and Constraints

a. Resources

i. Personnel (Technical support, Data mining experts)

ii. Project Datasets

1. url: [<http://bit.ly/autolib_dataset>]

2. Computing resources

**2)** **Data Understanding**

**Data Understanding overview**

The dataset file comprises data collected from opendataparis.com where the Autolib availability information was available in real-time. No additional dataset required to meet the needs of the study. The data file needed for this project will be:

1. url: [<http://bit.ly/autolib_dataset>

**Collecting Initial Data**

The data collected was sourced from opendataparis.com. The Autolib availability information was available in real time.

**Describing and Exporting Data**

Only one dataset is available descriptive of the columns and values of the rest of the dataset.

A further description of the provided dataset is as follows:

url: [https://drive.google.com/file/d/13DXF2CFWQLeYxxHFekng8HJnH\_jtbfpN/view]

The dataset provides outlines to assist in understanding the data provided. The data set is presented in a pdf format with statistical presentations such as graphs.

url: [<http://bit.ly/autolib_dataset>]

The dataset provides information about (Bleucar) an electric car-sharing service company as from from April 1 to April 9, 2018. Data types are objects and int64. The number of columns is 25 and the number of rows is 5000. There are nan values of int64 datatype.

**Verifying Data Quality**

Most of the datasets contains null values with the datatypes being integers and objects with 125,000 entries.

**3)** **Data Preparation**

Steps were taken during data exploration are as follows:

**Selecting Data**

The following dataset was used for analysis on this projects based on the relevance of our goal and data quality.

· url: [<http://bit.ly/autolib_dataset>

Used data Frame to load a file, examine basics statistics of the data and finally output the results.

**Cleaning Data**

Data cleaning procedure used during the analysis is:

· Dropping the Displayed \*comment\* and \*Scheduled at\* columns due to lots of NaN values.

**Constructing New data**

New data was created after data cleaning.

**Integrating and formatting Data**

No merging done since the dataset was only one.

Dropping the Displayed \*comment\* and \*Scheduled at\* columns due to lots of NaN values.

A new data frame called \*subsetDataFrame\* to help with the major project objective.

**4)** **Analysis**

During our analysis, the following questions were answered.

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 for picking up a shared electric car (Bluecar) in the city of Paris over the month of April 2018. IS <------------- HOUR 21 with a frequency count of 147------------->

Bonus Questions (Optional)

2. What is the most popular hour for returning cars?

3. What station is the most popular?

· Overall?

· 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?

· Overall?

· At the most popular picking hour?

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

The above analysis was done using the python programming language. This analysis can be found in the following python notebook.

link:<https://colab.research.google.com/drive/1YhabsZwtGj5CfTXAgnc0lhQOhdqKb9Bf?usp=sharing>

**5)** **Recommendation**

Following our analysis, the following recommendations were provided:

1. Avail more cars for picking up people at Paris on the \*21 hour\*

2. Avail more cars for picking up people at Paris on the \*4th day\* of the month.

**6)** E**valuation**

From our business success criteria, we have been able to determine 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 the \*hour 21\*. For purposes of further investigation, we have also provided further questions that can provide more insights while performing data mining.

The following is a link to my GitHub project repository

Link: <https://drive.google.com/drive/folders/16s4MpUJQPV2XefOcqOB4RLWpl4iyDCH2?usp=sharing>

The following is a link to JIRA

Link: <https://team-1621603387531.atlassian.net/browse/FIPW4>