Data Report

**Business Understanding**

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.

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, 15:31 PM 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.

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

Requirements, Assumptions and Constraints:

Resources

· Personnel (Technical support, Data mining experts)

· Project Datasets-<http://bit.ly/autolib_dataset>

- Computing resources

**Data Understanding**

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:<http://bit.ly/autolib_dataset>

The data collected was sourced from opendataparis.com. The Autolib availability information was available in real time. 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:<https://drive.google.com/file/d/13DXF2CFWQLeYxxHFekng8HJnH_jtbfpN/view>

The dataset provides information about (Bleucar) an electric car-sharing service company as 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 data type. Most of the datasets contains null values with the data types being integers and objects with 125,000 entries.

**Data Preparation**

Steps taken during data exploration are as follows:

Selecting Data- The following dataset was used for analysis on this project based on the relevance of our goal and data quality.<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.

**Analysis**

During the analysis, the following questions were answered.

· 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.

Answer: Hour 21 with a frequency count of 147

Bonus Questions (Optional)

· What is the most popular hour for returning cars?

· What station is the most popular?

v 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?

v Overall?

v At the most popular picking hour?

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

**Recommendation**

Following the analysis, the following recommendations are provided:

· Avail more cars for picking up people in Paris on the 21st hour

· Avail more cars for picking up people in Paris on the 4th day of the month.

**Evaluation**

From the analysis, I have determined 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 the 21st hour.

Github Repository Link: <https://github.com/PETER-BIKO-OTIENO/Friday-Independent-Project---Week-4>

Jira KANBAN Link: <https://yves-moringa.atlassian.net/jira/software/projects/FIPW4/boards/2>