# 1.Business Understanding

## Business Overview

Car-sharing is a model of car rental where people rent cars for short periods, often by the hour. It differs from traditional car rental in that the owners of the cars are often private individuals themselves, and the car-sharing facilitator is generally distinct from the car owner. Car-sharing is part of the larger trend of shared mobility.

As car-sharing continues to influence how we get from point A to B in our communities, electric car-sharing could help accelerate our transition to cleaner vehicles as they work to meet their climate goals and reduce their carbon footprint.

## Business Objectives

The main objective of this research 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 9 days in April 2018.

## Business Success Criteria

The research will be considered a success if we can identify which hours the Electric cars are picked up most in Paris in April 2018. This helps to ensure availability when needed.

## Assessing the situation

### Resources

**1.Datasets:**

Autolib: It contains information on the stations and types of cars available as well as their location.[<http://bit.ly/autolib_dataset>]

Autolib Description: It contains information on the type of values in respective columns of the dataset.[[Link](https://drive.google.com/file/d/13DXF2CFWQLeYxxHFekng8HJnH_jtbfpN/view)]

Software: Python libraries, Github, Jira, Google Collaboratory.

**2.Assumptions**

We assume that the information provided is correct and no errors were made during the collection process

**3.Constraints**

The data provided is only for 9 days in April hence may not be a true reflection of the situation at the stations.

## Data Mining Goals

The data mining approach for this project is as follows:

* Identify the most popular hour of the day for picking up a shared electric car (Bluecar) in the city of Paris over April 2018.
* What is the most popular hour for returning cars?
* What station is the most popular overall?
* What station is the most popular at the most popular pickup hour?
* What postal code is the most popular for picking up Blue cars?
* Does the most popular station overall belong to that postal code?
* Does the most popular station at the most popular pickup hour belong to that postal code?
* Do the results change if you consider Utilib and Utilib 1.4 instead of Blue cars?

### Data Mining Success Criteria

The research will be considered a success if all the above questions are answered successfully.

Project Plan

Phase

Time

Resources

Risk

Business Understanding

# 2.Data Understanding

## Data Understanding Overview

For this research, we are using the availed dataset by the company. These datasets are:

Autolib: This dataset gives the electric cars, charging slots, City, Geo points, Postal codes, Station types, and Time(month, day, hour, and minute).

Autolib description: This dataset describes the collected Autolib dataset. Provides more detailed source and license, collected data, and fields.

## Data Description

There is one dataset available and its description contains Delberg licenses for data collection:

Autolib Dataset: This dataset contains different types of cars across different cities, there are 5000 entries. It also contains the time and active location of said vehicles.

## Verifying Data Quality

The dataset has some missing values but those values are in redundant columns that are not relevant to our research hence were dropped. The remaining data has not been compromised by this action.

# 3.Data Preparation

Loading the Dataset : We load the Autolib dataset to a Python notebook in order to begin our analysis.

Cleaning the Data: We use Python libraries to clean the data.

During data cleaning, several steps were undertaken to ensure reliable quality data for our analysis. These were:

Dropped all columns irrelevant for our analysis ensuring we remain with the most relevant dataset for our analysis.

Dropped all columns with missing data since most of the data was missing and the columns also were relevant for our analysis.

We created a new column that combined the year, month, day, hour, and minutes columns to form a Date-Time column.

# 4.Analysis

Identify the most popular hour of the day for picking up a shared electric car (Bluecar) in the city of Paris over April 2018.

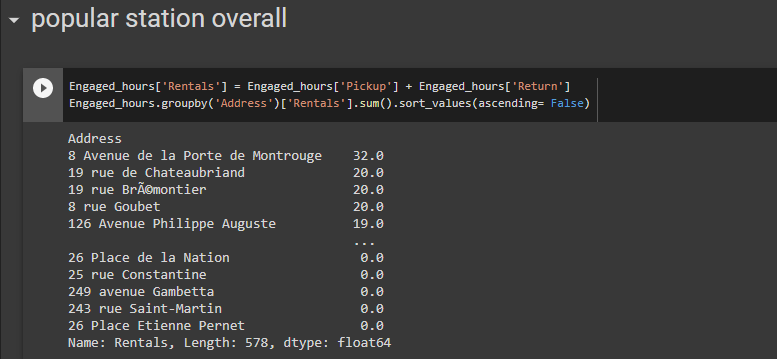
3:00Hrs

What is the most popular hour for returning cars?

22:00Hrs

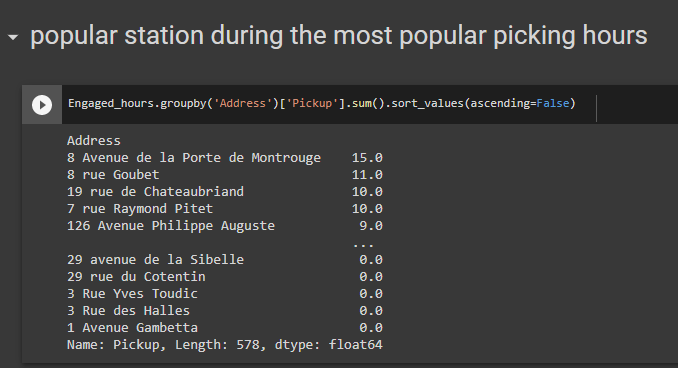
What station is the most popular overall?

8 Avenue De la Porte de Montrouge



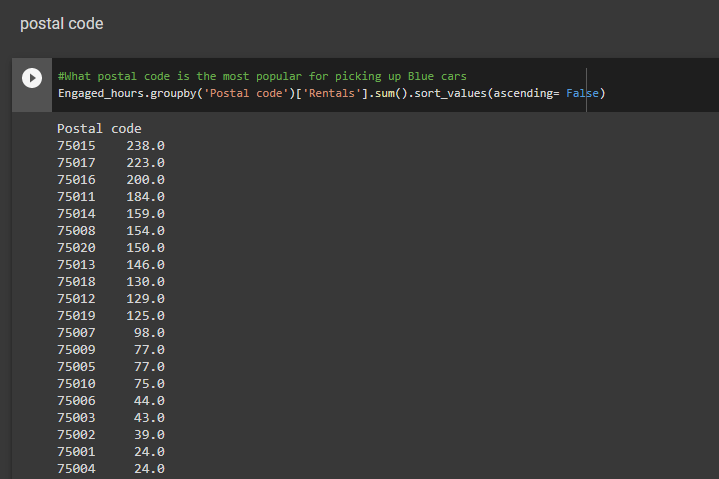
What station is the most popular at the most popular pickup hour?

8 Avenue De la Porte de Montrouge



What postal code is the most popular for picking up Blue cars?

75015



Does the most popular station overall belong to that postal code?

It doesn’t belong

Does the most popular station at the most popular pickup hour belong to that postal code?

It doesn’t belong

# 5.Recommendations

The most popular hour for pickup is 03:00 am and the most popular hour of return is 10:00 pm. It is also seen that most pickups occur during the weekdays. More cars should be made available during working hours. The most frequented station is 8 Avenue De la Porte de Montrouge. There shouldn’t lack cars at this station at any given time.

The analysis was done in this python Notebook:

Git Repository

https://github.com/njonge-nathan/Moringa\_datascience-\_week4-\_python-notebook

Jira

https://nathannjonge.atlassian.net