## **Business Understanding**

### **Business Overview**

Autolib was an electric car sharing service in Paris, France. It was operated by the Bollore industry and complemented by the city bike sharing service.The autolib maintained a fleet of all electric Bollore Bluecars for public transport on a paid subscription .Bollore began leasing the Bluecar to individuals and corporate customers in october 2012 .Rolled out utillib and utilib 1.4 which are the delivery versions with two seats instead of four.

### **Business Objective**

* 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 Success Criteria**

This forecast will be used to figure out customer behavior: what time and places our cars are most used and allocate resources effectively.

### **Assessing the Situation**

1. **Resource Inventory**
   1. Data - Autolib\_dataset
   2. Software( Github, Google Collaboratory)
2. **Assumptions**
   1. The data provided is correct and up to date
3. **Constraints**
   1. There are no constraints

### **Data Mining Goals**

Our data mining goals for this project are as follows:

* 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?

**Data Mining Success Criteria**

Our success criteria will be measured by the following criteria;

* We target the locations with the most usage of our cars

## **Data Understanding**

### **Data Understanding Overview**

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

* Autolib dataset - This dataset contains the variety of cars their destination information and time information activity over a period of 6 days.

### **Data Description**

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

* **Autolib dataset -** This dataset contains the number of cars either a blue car, utilib or utilib 1.4 with its destination information through addresses, cities, longitude and latitude information , postal code and charging station information

### **Verifying Data Quality**

None of the two datasets had any missing values. There were also no known data errors in the datasets.

## **Data Preparation**

These are the steps followed in preparing the data

#### **Loading Data**

Loaded the datasets into google colab and using pandas created data frames.

#### **Cleaning Data**

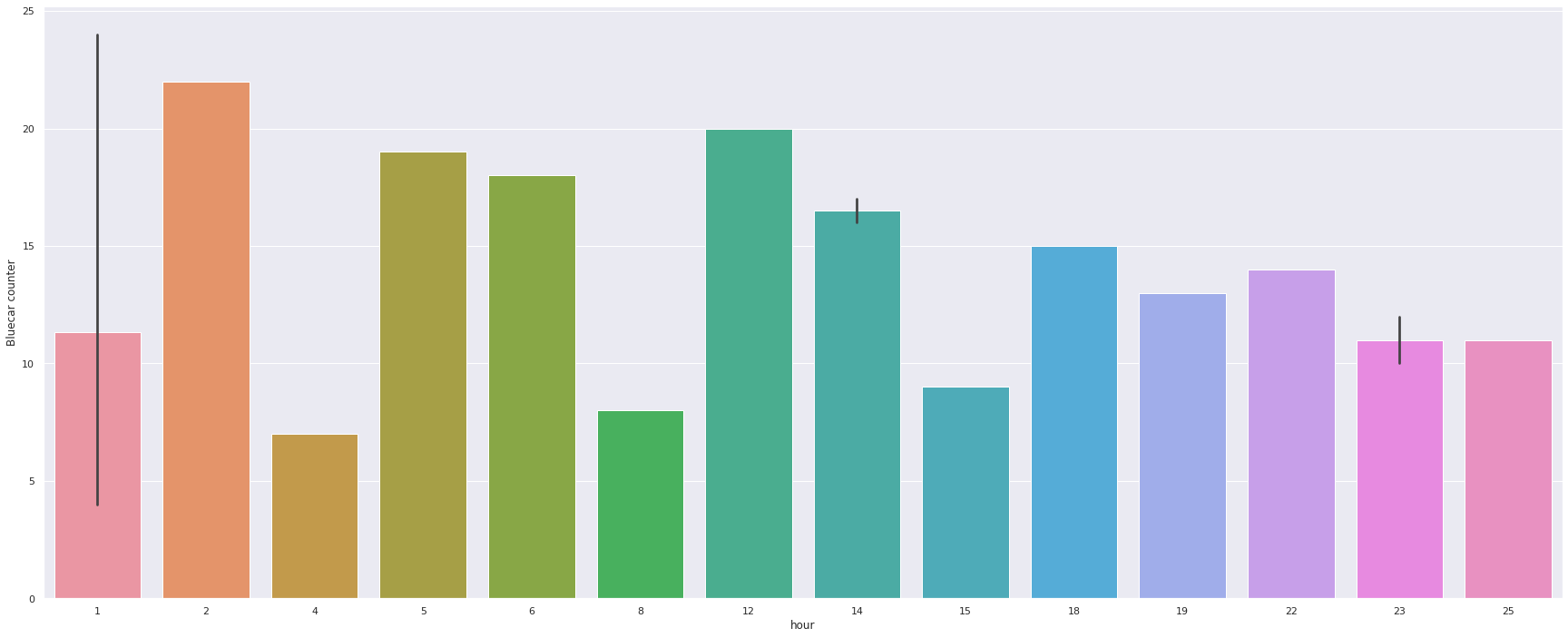
1. The dataset had no column name errors
2. We had close to none duplicate values
3. Dropped columns scheduled at and displayed comment due to 99 percent of the data was missing

## **Analysis**

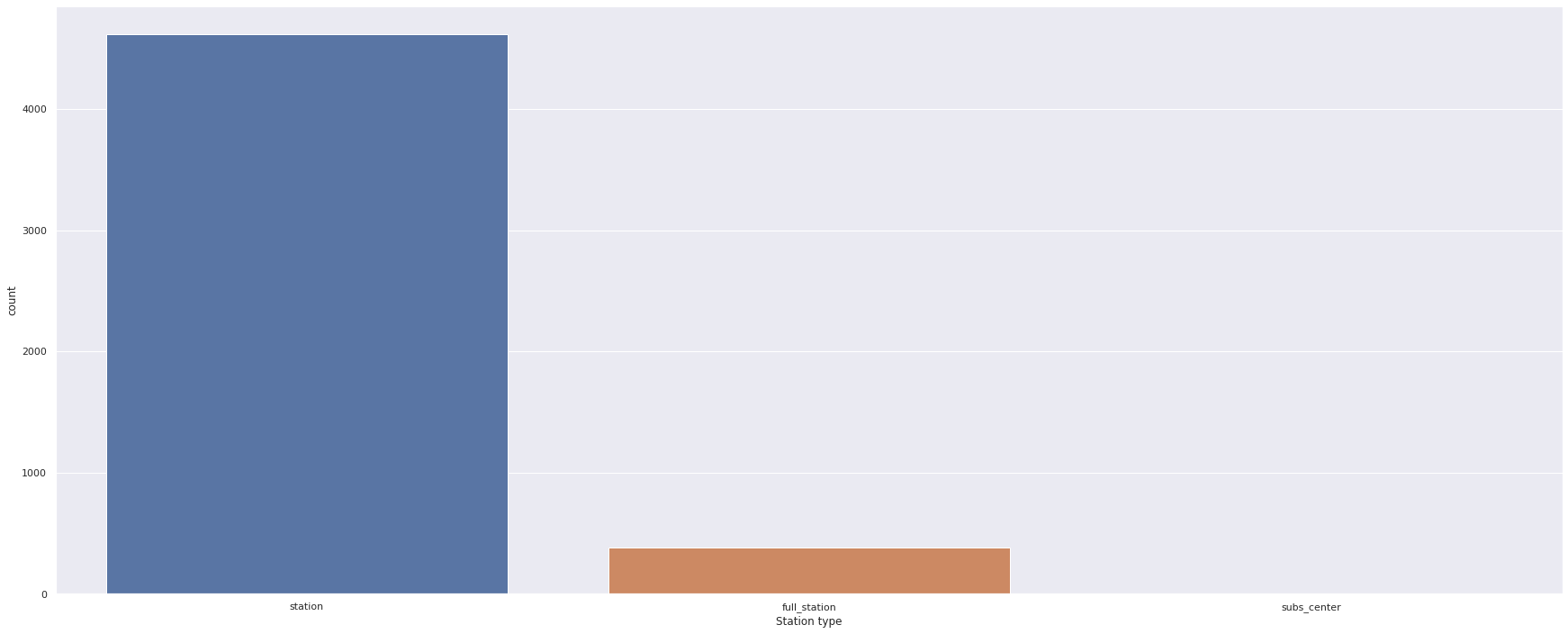
1. The most popular hour overall was 6am in the morning overally

## 

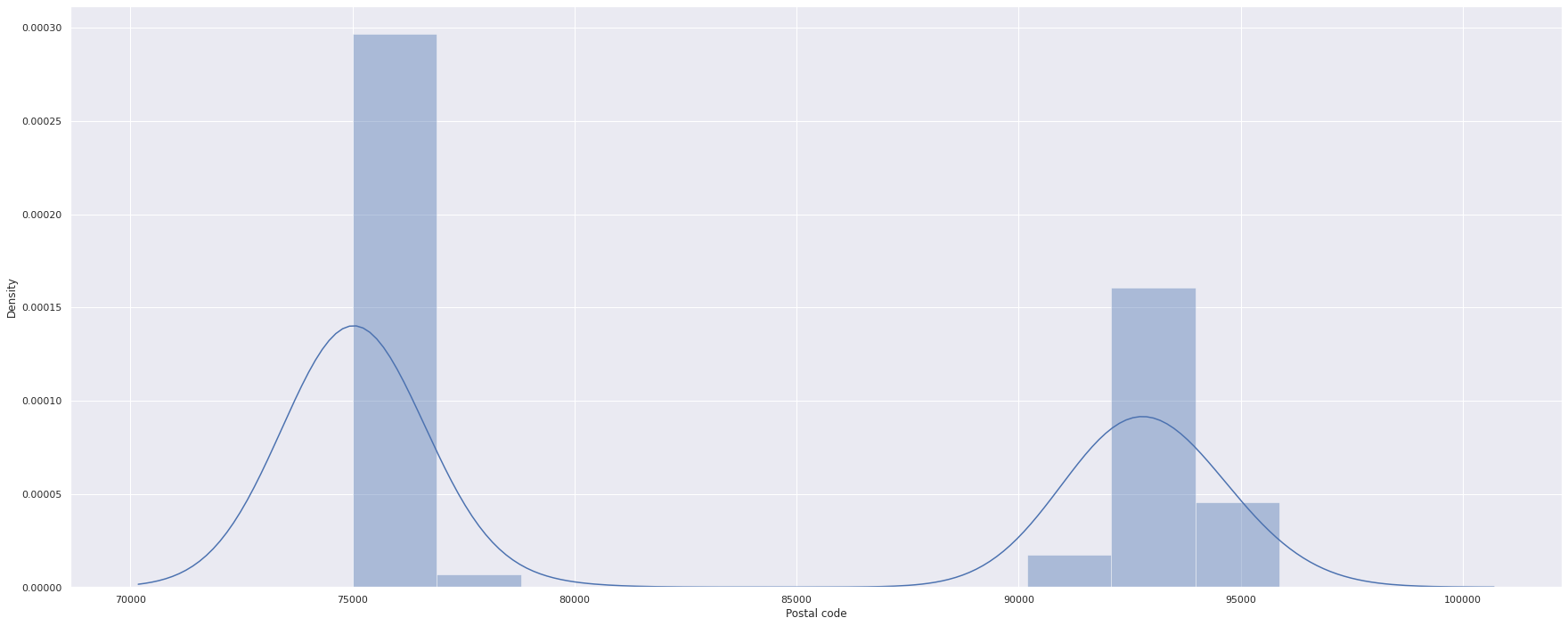
1. The hour that had the most pickups by bluecars was 2am and preceded by around noon



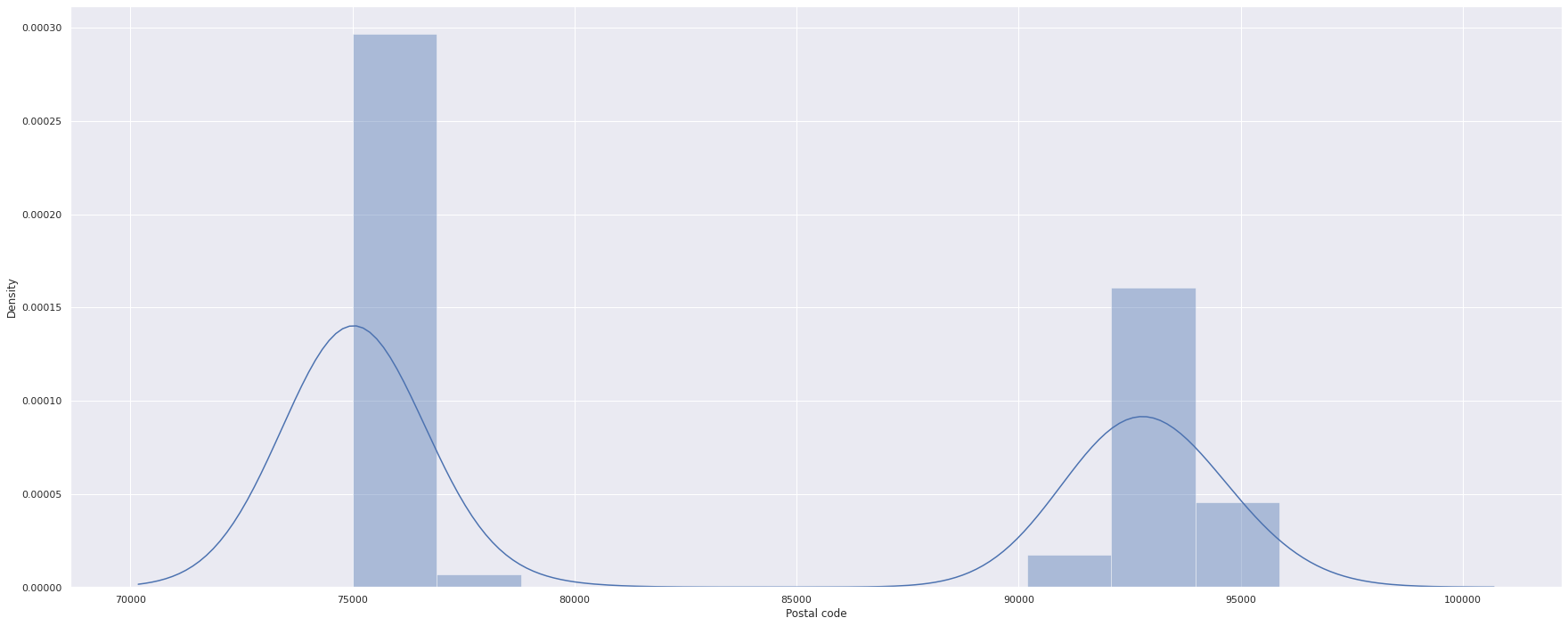
1. 9pm was the most popular time to return cars
2. Most popular station station 1



1. The most visited postal code is 75015



1. The most visited postal code is 75015 by bluecars



## Github link