**DATA REPORT FOR AUTOLIB ELECTRIC CAR-HIRE SERVICE COMPANY**

## **Problem Statement**

### **1.1 Business Overview**

Autolib' was an electric car sharing service which was inaugurated in Paris, France, in December 2011. The Autolib' service maintained a fleet of all-electric Bollore Bluecars for public use on a paid subscription basis, employing a citywide network of parking and charging stations. As of 3 July 2016, 3,980 Bluecars had been registered for the service and had more than 126,900 registered subscribers Autolib' furthermore offered 1,084 electric car stations in Paris agglomeration with 5,935 charging points.

## **Data Description**

For this project, a dataset was provided and also the description of the dataset. THe dataset size was 16085 rows and 13 columns.

The dataset columns consist of 3 types of cars (bluecar, utilib and utilib 14). For each of the car types, summation of the number of cars taken, summation of the cars returned was provided, summation of the freed slots and that of taken slots.

Also, provided was the date, day of the week and the postal code from where the car was picked.

## **Hypothesis Testing Procedure**

#### **3.1 Specifying the null hypothesis**

Null hypothesis is a statement of no effect or difference between two or more groups. In research studies, a researcher is usually interested in disapproving the null hypothesis.

For this research, the null hypothesis is that there is no difference in the sum of bluecars taken from postal code 75015 and postal code 75017 during weekdays.

#### **3.2 Specifying the alternative hypothesis**

Alternative hypothesis is the statement that there is an effect or difference. It is usually the hypothesis that the researcher is proving.

For this research, the alternative hypothesis is that, the sum of bluecars taken from postal code 75015 is not equal to that from postal code 75017 during weekdays.

#### **3.3 Set the significance level**

For this research, the significance level was set to 0.05.

This means there is a 5% chance that the alternative hypothesis will be accepted, when the null hypothesis is true.

#### **3.4 Calculate the test statistic and the corresponding p-value**

For this research, z score test statistic was used. This is because the population standard deviation was known (calculated from the dataset given) .

P-value describes the probability of obtaining a sample statistic as or more extreme by chance alone if the null hypothesis was true.

## **Sampling Technique and sample size**

**4.1 Sampling technique**

Simple random sampling technique was used. In this technique, every item in the population has an even chance and likelihood of being selected in the sample.Therefore it’s free from biasness. The postal codes (75015 and 75017) used in hypothesis testing will be picked randomly. These postal codes should be those on service during the weekdays.

**4.2 Sample size**

A good maximum sample size is usually around 10% of the population. Most statisticians agree that the minimum sample size to get any kind of meaningful result is 100. If the population size is less than 100, then it is better to survey all of them.

For this research, the population size of the postal codes(75015 and 75017) operating on weekdays is 224. 10% of this population gives a sample size of 22.

## **Hypothesis Testing Results.**

Statistical summaries of the population were:

* population mean = 305.14285714285717
* Population standard deviation = 24.89510987796604

Statistical summaries of the sample were:

* sample mean = 295.75
* sample standard deviation= 55.1123851779253

The results of the hypothesis test was as follows;

* Z-score value = -0.3772972760072258
* P-value = 0.7059526962957161

## **Summary and Conclusions**

If the p value found from the hypothesis test is lesser than the set significance level, then the results are statistically significant to reject the null hypothesis.

From the results of the hypothesis test, p value was 0.70, which is greater than the significant value 0.05.

Therefore, we fail to reject the null hypothesis.

Just because the sample wasn't extreme enough to reject the null hypothesis doesn't mean that there wasn't another sample that exists that is extreme enough to reject the null.

To avoid making a Type II error, we would usually say that we "fail to reject the H0".