**Report on Hypothesis Testing**

**Problem Statement**

We will be using the Autolib electric car sharing service company dataset. This dataset has very useful information on cars being picked up from and returned to their stations. It has particular dates as well as day times i.e weekend or weekday.

We will be looking at the mean of the sum of blue cars taken in two postal codes, 75015 and 75017.

The null Hypothesis we will be testing is that the mean of the blue cars taken in postal area 75015 is greater than or equal to the mean of the blue cars taken from postal code 75017. We will get the mean of the 75017 blue cars taken and we will use it as the reference. Ho : u >= 734.5 .

The alternative hypothesis will try to oppose the null hypothesis and it will therefore read, the mean of the blue cars taken from 75015 area is less than the mean of blue cars taken from 75017 which is 734.5. Ha : u < 734.5.

We will run this hypothesis test so that we can gauge between the two postal codes, just so as to read into which area has utilized the electric car sharing culture.

**Data Description**

The data we have has thirteen variables, two of which are categorical while the rest are numerical. From the categorical variables, we have the date and the day type. These help us to know the dates of the electric car transactions and whether the transactions are during the weekdays or weekend.

We will be looking at the blue cars that are taken, which highly correlates with the blue cars returned. We shall get the mean of the sum of blue cars taken from postal code 75017 and use it as our reference for us to test if the mean of the sum of blue cars taken is greater than or equal to our found mean.

Autolib' was an electric car sharing service which was inaugurated in Paris, France, in December 2011 and closed on July 31st 2018. Our dataset was collected in different postal areas in Paris and the postal areas recorded in postal codes. The data was collected in the period of January to June 2018.

The mean of the daily data collected is 1431.330619 and the standard deviation is 33.212050.

**Hypothesis Testing Procedure**

We will group our sample into postal codes 75015 and 75017 as these are our areas of interest. We will then get the mean of the sum of blue cars taken in these two codes. This mean will be our sample mean when getting our test statistic s. We will also get the mean of our blue cars taken from postal code 75017.

Our null hypothesis will be that the mean of blue cars taken from 75015 is greater than or equal to the mean of blue cars taken from 75017. The alternative will be that the mean of blue cars taken from 75015 is less than the mean of cars taken from 75017. We would like to look into the awareness of electric car sharing in these two areas by comparing the means of blue car transactions.

Since we are testing the hypothesis about the mean, our population standard deviation is unknown, but we have our sample standard deviation, we will use the T test to find our p value. Our distribution is also normal given the mean. We have satisfied most conditions for our T test.

We will use a confidence level of 95% which leaves us with a significance level of 5%.

Therefore our alpha will be 0.05.

**Hypothesis Testing Results**

Given our sample mean as 815.7, our claim mean as 734.5, sample standard deviation as 180.17 and our sample size is 312, we can comfortably calculate our T test.

The T test is 7.96 . From this we can work out our p value which we get as 0.999.

The p value is greater than our significance level of 0.05 which means we fail to reject our null hypothesis.

**Summary and Conclusion**

We can conclude that our result is not statistically significant as we failed to reject our null hypothesis. We can therefore say that we have sufficient evidence to prove that the mean of blue cars taken in 75015 is greater than the mean of blue cars taken in 75017.