Hypothesis Testing Report

1. **Problem Statement**

The dataset is a daily aggregation, by date and postal code, of the number of events on the Autolib network (car-sharing and recharging). The cars in the dataset include blue cars, utilib, and utilib\_14.

We will investigate the blue cars taken and returned during the weekend.

The null hypothesis is :There is no difference between the blue cars taken and blue cars returned on a weekend for a specific postal code

The alternate hypothesis is:There is a difference between the blue cars taken and blue cars returned on a weekend for a specific postal code

The hypothesis is interesting because it will be good to know whether there is any difference between the bluecars taken and returned on the weekends.

1. **Data Description**

Data cleaning was done by removing the outliers and the null values.

I created a new dataset containing only weekend data with 4541 records down from the original 16085 records.

We will use stratified sampling to take a sample of approximately 20% of the records.

1. **Hypothesis Testing Procedure**

We used an alpha value of 5%.

We are also using a two tailed test.

We use a z statistic since the data meets the following conditions:

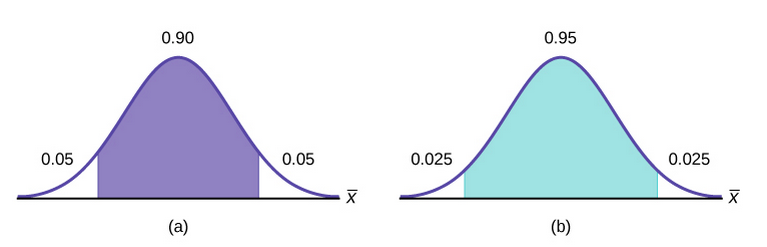
* Your sample size is greater than 30. ...
* Data points should be independent from each other. ...
* Your data should be normally distributed. ...
* Your data should be randomly selected from a population, where each item has an equal chance of being selected.
* Sample sizes should be equal if at all possible.

The z statistic is 2.105 which means that it is 2 standard deviations above the mean.

If we get a p-value less than 0.05, we fail to reject the null hypothesis. Since the P value is 0.035, we fail to reject the null hypothesis

1. .**Hypothesis Testing Results**

The confidence level is as follows:



We got the following results for ztest and p-value:(-0.64 and 0.50)Since the P-Value is greater than 0.05(the alpha level), we fail to reject the null hypothesis and conclude that there is no difference between the blue cars taken and blue cars returned on a weekend for a specific postal code

1. **Discussion of Test Sensitivity**

There were some z-scores which were above 0.05 and others below 0.05. However, the sample test shows that we should fail to reject the null hypothesis.

6. **Summary and Conclusions**

This project investigated whether there is a difference between the blue cars taken and blue cars returned on the weekend for a specific postal code. We concluded that there is no difference between the blue cars taken and the blue cars returned on the weekend for a specific postal code. The company should strategize to get the uptake of the blue cars taken during the weekend to increase.