HYPOTHESIS TESTING REPORT.

Problem Statement

An electric car company in France has provided data on the different electrical cars, in various postal codes but we do not fully understand the usage of the bluecar. I would like to understand the usage of the bluecar, especially over the weekend focusing on comparing the number of bluecars taken and returned over the weekend in a given postal code.

Null Hypothesis

Ho: The probability of the bluecars being taken during the weekend in greater than or equal to the odds of them being returned during the weekend.

Ho: U1=>U2

Alternate Hypothesis

The probability of bluecars being taken during the weekend is less than the odds of them being returned.

Ho: U1<U2

Data description.

The Autolib car sharing provided the dataset which is secondary in nature. It provides details on the type of car taken, the date, day of the week, whether it was taken over the weekend or weekday, postal code, sum of cars returned or taken and information or freed and taken slots.

The variables that were analysed were bluecars_taken_sum and bluecars_taken_sum. They provide data on the sum of bluecars taken and bluecars returned on a given date in a particular postal code.

Hypothesis Testing Procedure

The hypothesis testing was done to help the Autolib car sharing company determine bluecar usage over the weekend to enable them to put measures in place that will increase the number of bluecars being taken in a given postal code.

Poisson distribution was used to test the hypotheses because we were comparing the probability of the bluecar being taken and the bluecar being returned within a given time period.

lambda_ = ((events/time)*time_period)

poisson.pmf(lambda_,mean)

The probability of the two occurrences will be compared then forma basis for the rejection or acceptance of the null hypothesis.

Hypothesis Testing Results

These are the test results:

Probability of bluecar being taken: 2.04243515e-38

Probability of the bluecar being returned: 5.16128439e-39

Since the probability of the bluecar being returned is higher than that of the bluecar being taken, we reject the null hypothesis.

Discussion of Test Sensitivity.

An increase in the sample size would mean more accuracy because the statistical power of the hypothesis test will also increase; however, that increases the chances of a type 2 error.

Summary and Recommendations

The hypothesis test shows that more bluecars are returned than taken from the station. This means that the postal code in our test is more popular as a destination that starting point of the journey.

- We recommend that more bluecars be availed for use by those taking a round trip via this postal code. This way the number of bluecars being taken will increase.
- The lower number of taken bluecars could be because of unavailability, therefore, the electric car sharing company can increase the number of bluecars.