

Autolib electric car-sharing service company

Business overview

Autolib is an electric car sharing service and it was launched for ‘beta testing’ in 2011 in Paris. Bolloré industrial group was the one who supplied the electric car system. At that time there were 66 cars and 33 car rental stations.

Business objective

In this hypothesis report we used the data from the Autolib electric car-sharing service company. The variable that we mostly dealt with is the number of blue cars that were returned on Monday and Friday.

The objective of this hypothesis was to find whether there is no statistical difference between the mean of blue cars taken on Monday and Friday and this was our null hypothesis. Our alternate hypothesis was a statistical difference between the mean of blue cars taken on Monday and Friday.

This hypothesis was chosen because we wanted to know the day which we had most customers so as to increase the availability of resources on that day and also to find which day was critical to the company, is it at the beginning of the weekday or at the end of the weekday.

Data Description.

In this project we are provided with two datasets: one for description of the columns in our data

Column name	Explanation
Postal Code	postal code of the area (in Paris)
date	date of the row aggregation
n_daily_data_points	number of daily data points that were availabl...
dayOfweek	identifier of weekday (0: Monday -> 6: Sunday)
day_type	weekday or weekend
BlueCars_taken_sum	Number of bluecars taken that date in that area
BlueCars_returned_sum	Number of bluecars returned that date in that ...
Utilib_taken_sum	Number of Utilib taken that date in that area
Utilib_returned_sum	Number of Utilib returned that date in that area

Utilib_14_taken_sum	Number of Utilib 1.4 taken that date in that area
Utilib_14_returned_sum	Number of Utilib 1.4 returned that date in tha...
Slots_freed_sum	Number of recharging slots released that date ...
Slots_taken_sum	Number of recharging slots taken that date in ...

The other dataset had 13 columns which are Postal code, date, n_daily_data_points, dayOfWeek, day_type, BlueCars_taken_sum, BlueCars_returned_sum, Utilib_taken_sum, Utilib_returned_sum, Utilib_14_taken_sum, Utilib_14_returned_sum, Slots_freed_sum, Slots_taken_sum,

The source of our data was from the following link: for autolib dataset is [\[http://bit.ly/DSCoreAutolibDataset\]](http://bit.ly/DSCoreAutolibDataset) and

for the description of the data is ("<http://bit.ly/DSCoreAutolibDatasetGlossary>")

Hypothesis Testing Procedure

This hypothesis followed the following procedure where we first get the variable for both monday and friday, after that we then do the random sampling for both. We take a sample size of 100 for blue cars taken on both Monday and Friday.

The logic of having the hypothesis was to find which was the critical day for the company between the end of the weekday and the beginning of the weekday. This was to enable the company to make the right decision on when to put more efforts.

Since the sample size was greater than 30, we can calculate the standard deviation also both populations followed the normal distribution principle, in which both fall under light tailed normal distribution this led us to using the z-test method.

Alpha level is the probability of making a Type I error, the smaller the alpha level, the smaller the area this will lead to rejecting the null hypothesis. The more we have a tiny area, this increases the chances that we will not reject the null, when in fact you should. This is a Type II error. If we try to avoid a Type I error, the more likely to have Type II error. We decided to use an alpha level of 0.05 so as to balance between these two issues.

Hypothesis Testing Results

The results of our hypothesis is that we fail to reject a null hypothesis, because we found that the p-value was greater than significant level. The value for the test statistics was -

0.5738722676873629 and the p-value was 0.5660542407153357. The point of estimate is 12.71, this means that more people take blue cars on friday than of monday

Discussion of Test Sensitivity

We can falsely conclude that there is a significant difference between monday and friday (type I error). The probability of a type I error is α . We can falsely conclude that there is no difference between monday and friday (type II error). The probability of a type II error is β . Statistical power indicates that the ability of a project to detect a significant difference between populations, when a significant difference truly exists. Power equals $1 - \beta$ was used.

Summary and Conclusions

The project was conducting an hypothesis testing to find whether there is no statistical difference between the mean of blue cars taken on monday and friday. This was successfully done because we were able to determine the hypothesis, which we failed to reject the null hypothesis because the p-value was greater than a significant level of 0.05.