TEST OF HYPOTHESIS PROCESS

preface

The problem question involves an Electric Carsharing Company in the city of Paris France. The task revolves around lending cars and returning them. Using Data Science techniques, we will be scheming and planning towards a strategy to achieve reliable results.

i).Problem statement

Lending vehicles at a is a risky job especially if the vehicles are not being returned to the same station as Autolib does. Clients in this business may surpass the agreed time with the vehicles or even steal the vehicles. For this business to succeed, clients are expected to pick vehicles and return them in time. Hence our study will look into each of the groups of Vehicles given to clients and the respective vehicles returned.

The research will investigate if the means of vehicles lended and returned is the same or if the latter is lower. (Of Course the mean of vehicles returned cannot be higher than the vehicles taken for no one can return a vehicle they did not take). Hence we will be dealing with a two sample t - test in the experiment.

We will use the means obtained from univariate to perform the hypothesis testing. These means will help us compare how the Cars taken and Cars returned operate . We will check if there is a significant difference between the two.

2.Data Description

These are the columns of the datasets Autolib.

1.'Postal_code', -This describes the position the Codes of where the Vehicles are taken and returned

2.'Date' - This records the date of Picking up the vehicle

3. 'n_daily_data_points', - These are the points accumulated daily by the clients for vehicles picked.

4.'dayOfWeek',- This is the column for the Day of the week for vehicle taken

5.'Day_type',- This is the day when vehicles are taken and vehicles are returned.

6. 'BlueCars_taken_sum', This is the summation are the vehicles of blue Cars taken to the slots
7.'BlueCars_returned_sum',- This is is the summation are the vehicles of blue_Cars returned to the slots

8. 'Utilib_taken_sum', - This is is the summation are the vehicles of utilib taken to the slots

9.'Utilib_returned_sum', - This is is the summation are the vehicles of utilib returned to the slots

10.'Utilib_14_taken_sum', - This is is the summation are the vehicles of utilib -14 taken to the slots

11.'Utilib_14_returned_sum',- This is is the summation are the vehicles of utilib -14 returned to the slots

12. 'Slots_freed_sum', - These are the slots that are awaiting slots from vehicles taken

13. 'Slots_taken_sum'- These are the parking places that have been filled by the vehicles that are awaiting to be picked.

Starting with the Blue Cars, we will check the difference between the cars taken and Cars returned.We will also perform a t-test for Utilib and Utilib 14 to check the differences in the means too.

3.Hypothesis testing procedure

We will start by defining the null and Alternative hypothesis for the three tests

Key : Null hypothesis = Ho(The Status Quo)

Alternative hypothesis = H1(What we are trying to prove right or wrong)

We will be using the 95 % Confidence Level hence our level of significance of 5%

For our 3 cases below:

I.Blue Cars

Ho : There is no difference in means of Blue Cars Taken and Blue Cars Returned

H1 : The means of Blue Cars Cars Returned is either equal or less than Blue Cars Cars returned

From the Experiment of Differentiating means of Blue Cars taken and returned the P-Value(The value to compare with 0.05) is 0.994507330804926

ii. Utilib Cars

Ho : There is no difference in means of Utilib Cars Taken and Utilib Cars Returned.

H1 : The means of Utilib Cars Returned is either equal or less than Utilib Cars returned

From the Experiment of Differentiating means of Utilib Cars taken and returned the P-Value(The value to compare with 0.05) is 0.0

iii. Utilib -14 Cars

Ho : There is no difference in means of Utilib 14 Cars Taken and Utilib 14 Cars Returned.

H1 : The means of Utilib 14 Cars Returned is either equal or less than Utilib 14 Cars returned

From the Experiment of Differentiating means ofUtilib 14 Cars taken and returned the P-Value(The value to compare with 0.05) is 0.9923956158598111

Hypothesis Testing results

1.Conclusion for Blue Cars

From the Experiment of Differentiating means of Blue Cars taken and returned the P-Value(The value to compare with 0.05) is 0.994507330804926

We therefore fail to reject the null hypothesis and conclude that the BLue Cars returned have a substantial difference from Blue cars taken at 5% level of significance

This means that most Blue Cars taken from the Slots were not returned hence this must have led to losses in the Car lending company.

ii).Conclusion for Utilib Vehicles

From the Experiment of Differentiating means of Utilib Cars taken and returned the P-Value(The value to compare with 0.05) is 0.0

We therefore reject the null hypothesis and conclude that Utilib taken is not any different from Utilib returned at 5% level of significance

This means that the Utilib Car Vehicles taken were returned to the freed Slots hence the company was able to gain profits from these cars

iii).Conclusion for Utilib 14 Vehicles

From the Experiment of Differentiating means ofUtilib 14 Cars taken and returned the P-Value(The value to compare with 0.05) is 0.9923956158598111

We therefore fail to reject the null hypothesis and conclude that the Utilib 14 returned have a substantial difference from utilib 14 taken at 5% level of significance.

This means that most of Utilib 14 vehicles were failed to be returned hence the difference in means of the vehicles taken and returned.

Hypothesis Sensitivity

Power of a Statistical test gives the likelihood of rejecting the Null hypothesis when the Null hypothesis is false.

- A high Statistical power means that the test results are likely valid. As the power increases , the probability of making a type II error decreases.

  To increase the power of a test, you may do the following

  1. Increase the effect size(the difference between Null and alternative Hypothesis.
  2. Increase the sample sizes.
  3. Decrease the variability in the samples.
  4. Increase the significance Level of the test.

Summary and Conclusions

For Blue Cars and Utilib 14 Cars , the vehicles had a fairly average return rate of vehicles .

The Utilib Cars had a good return rate and hence were better profitable. To improve this results for the company , the Company should have better measures and strict regulations to ensure returning vehicles.