**HYPOTHESIS TESTING REPORT**

**INTRODUCTION**

This is a report for hypothesis test to find out whether there is a difference between the averages of blue cars taken on a weekday and those taken on a weekend. From the given dataset, after all necessary data cleaning procedures, samples were drawn and the Two sample / Independent t test was used for the test, which I will cover more of in the steps on this report. The report is divided into several sections as follows; first is the problem statement and hypothesis statement, data description and analysis, hypothesis testing and result, then lastly conclusions and summaries follow.

**Problem statement**: To check if there is a difference between the averages of blue cars taken on a weekday and those taken on a weekend, taking postal code 95870 as our case study.

**Stating the Hypothesis**:

My claim is that there is a difference between blue cars are taken on a weekend and those taken on a weekday.

Ho: µ1  = µ2

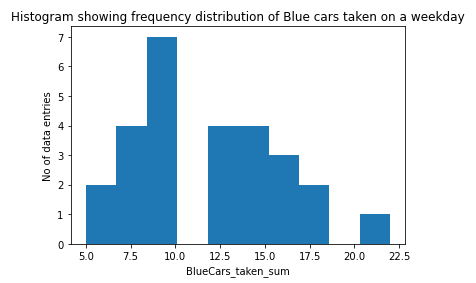
H1: µ1  ≠ µ2 (Claim)

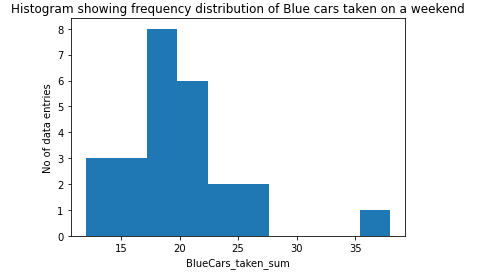
(µ1 = weekday average/mean, µ2  =  weekend average/mean)

**DATA DESCRIPTION**

The data is dated and across different postal codes/areas. For the purpose of the testing the claim I will be focusing on code/area 95870, which was further split into weekend and weekday where samples were drawn from. This was done after all necessary the data cleaning procedures where I removed outliers and missing data before doing some exploratory data analysis.

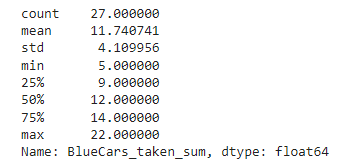
I was interested in how the number of blue cars taken on weekdays and weekends differ. But first sample analysis of the distribution of blue cars taken depending on day type can be seen in the below histograms:





The distributions for both samples seem to have a normal distribution at first glance. Further normality test will be done to confirm this.

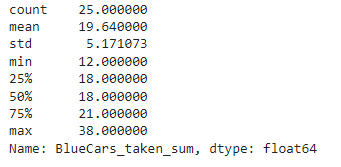
The following is descriptive statistics for my weekday sample:



Sample size = 27, sample mean = 11.74, sample standard deviation = 4.11

Minimum value is 5 while Maximum value is 22

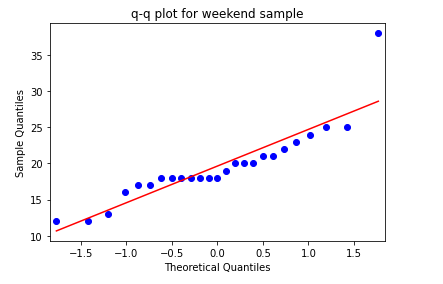
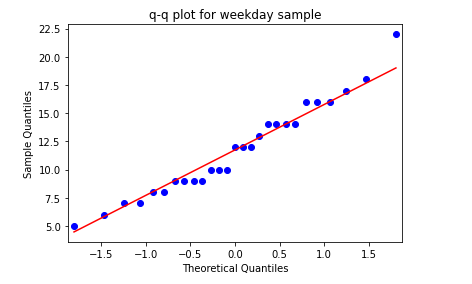
The following is descriptive statistics for my weekend sample:



Sample size = 25, sample mean = 19.64, sample standard deviation = 5.17.

Minimum value is 12 while Maximum value is 38.

Q-Q plot for both samples testing for normality before hypothesis testing:



This shows both samples have a normal distribution although not perfect.

Checking for unequal variances:

S1  4.11 = vs S2 = 5.17

4.11/5.17 = 0.79 which approx. 1. We can do a f test to further confirm this. F test results from python are as shown below (f-statistic, p-value):



p-value > 0.05 therefore we can confirm equal variances.

Our independent samples have a normal distribution and equal variances, we can now proceed to the hypothesis testing.

**HYPOTHESIS TESTING**

Since there are no assumption violations we can go ahead and use the Two sample/Independent t test. This is a method used to test whether the unknown population means of two groups are equal or not hence appropriate.

**Step 1**: State the hypothesis and identify the claim.

I think the average number of blue cars taken on weekends is different than the average of blue cars taken on weekdays.

OR

OR (Claim)

(µ1 = weekday average/mean, µ2  = weekend average/mean)

This is a two tailed test.

**Step 2**: Find the critical value

Taking significance level as 5% or 0.05 (α = 0.05). Since it’s a two tailed test we divide alpha by two; 0.025.

The degrees of freedom given by the formula below,

hence = 50.

t-value for α = 0.05 i.e.

p-value = 0.049787

form the boundaries of the critical region

**Step 3:** Compute test statistic

where *sp* is the pooled standard deviation

Pooled standard deviation for samples with different sample sizes can be calculated by the following formula

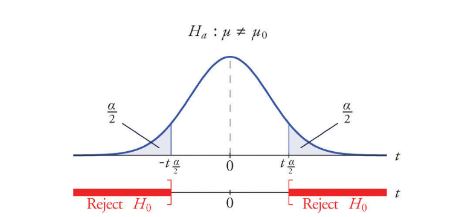
*Sp* = 4.633, therefore:

t statistic = -6.1435

similar results from python output:



**Step 4:** Compare values and decide to reject or fail to reject null hypothesis



Comparing the test statistic and t value: -2.011 > -6.1435 which places it in the rejection region of the left side of the tail.

Since our alpha is greater than the p-value, we can reject the null hypothesis as we have sufficient evidence to say that the means of the two populations are different.

**Step 5**: Conclusion

In our case there is sufficient evidence to say that there’s a difference in the means of blue cars taken on a weekday and blue cars taken on a weekend.

Calculating standard Error:

Confidence Interval /CI:

Difference of two sample means:

Hence CI at 95% confidence level

CI = 5.346 , 10.454

In this case there’s a 95% confidence that the difference between the two means (μ1 - μ2) lies between 5.346 and 10.454, that is 95% of all sample mean differences can be found in within the interval.

**DISCUSSION OF TEST SENSITIVITY**

The probability of committing a type I error is the same as our level of significance (in this case 0.05) and it represents our willingness to reject a true null hypothesis, while the probability of committing a type II error or beta (ß) represents not rejecting a false null hypothesis. It is ideal to minimize both.

Power of test is the probability of correctly rejecting the null hypothesis. The power of any test is 1 - ß, since rejecting the false null hypothesis is our goal. It stands that the smaller the type 2 error value the higher the power of a test.

However, if we decrease our significance level (e.g., from 0.05 to 0.01) the power of the test also decreases. This is because the region of acceptance gets bigger and we are less likely to reject the null hypothesis when it is false, hence more likely to make a Type II error

Increasing our sample size would also increase the power of the test.

**SUMMARY AND CONCLUSION**

My hypothesis test was about whether there was a difference in means of blue cars being taken on a weekday and blue cars taken on a weekend, my claim is that there is a difference between the two. My two samples (taken from one area code) resulted in different means; blue cars taken on a weekday mean = 11.74 while mean of blue cars taken on a weekend = 19.64. From the test result we found that there is enough evidence to reject the null of equal means and conclude that there is a difference. Further testing can be done to ascertain the nature of this difference, that is whether one is greater than the other by doing a one tail test.

Sample size for each was less than 30; there is a possibility the if I had larger sample sizes I could have gotten different results, and as discussed above increased sample size increases power of the test. This also goes for the level of significance chosen.

The test samples were only drawn from one area code; 95870. Other area codes likely vary in values and would have given different results if considered so caution should be taken while using the results of this test to make inferences about blue cars taken depending on day type in general.

(Google colab notebook: [link](https://colab.research.google.com/drive/1LQz-GbfNjRXQpVcORJPe5fYBahfg_PY9?usp=sharing))