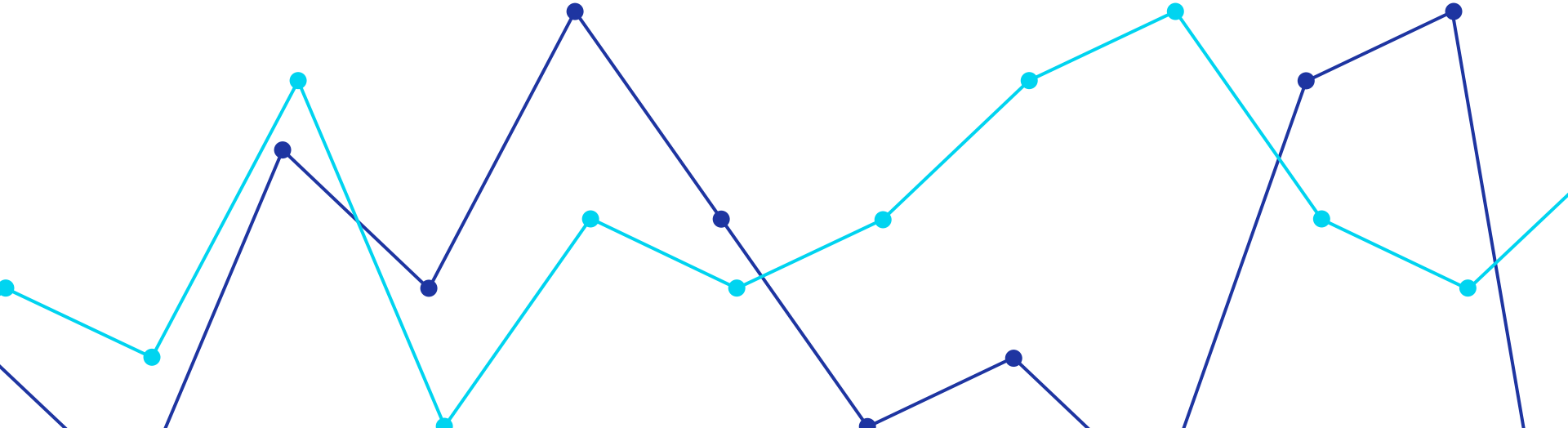
TESTING HYPOTHESIS

~DHRUVAL PATEL



HYPOTHESIS

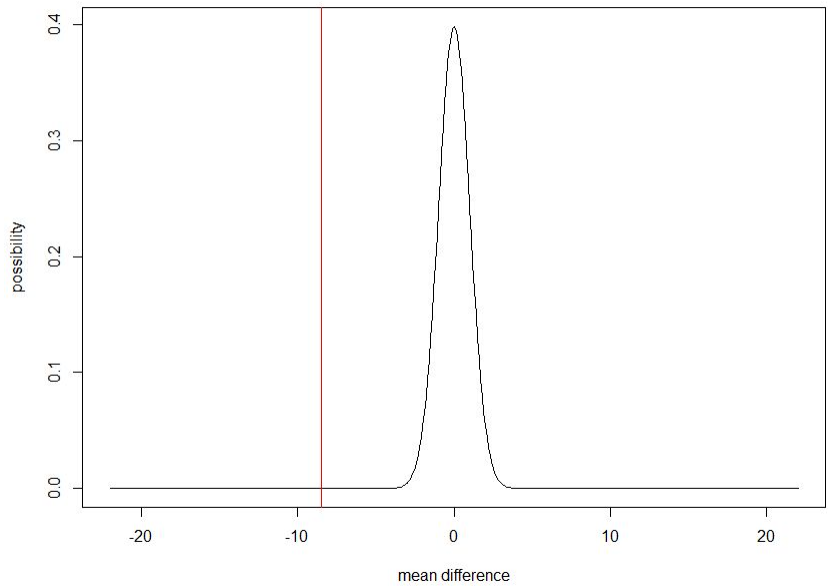
NULL HYPOTHESIS :
Subscribers of our company have EQUAL ride duration on our cycle than our random customers.

ALTERNATIVE HYPOTHESIS: Subscribers of our company have more or equal ride duration on our cycle than our customer ride duration

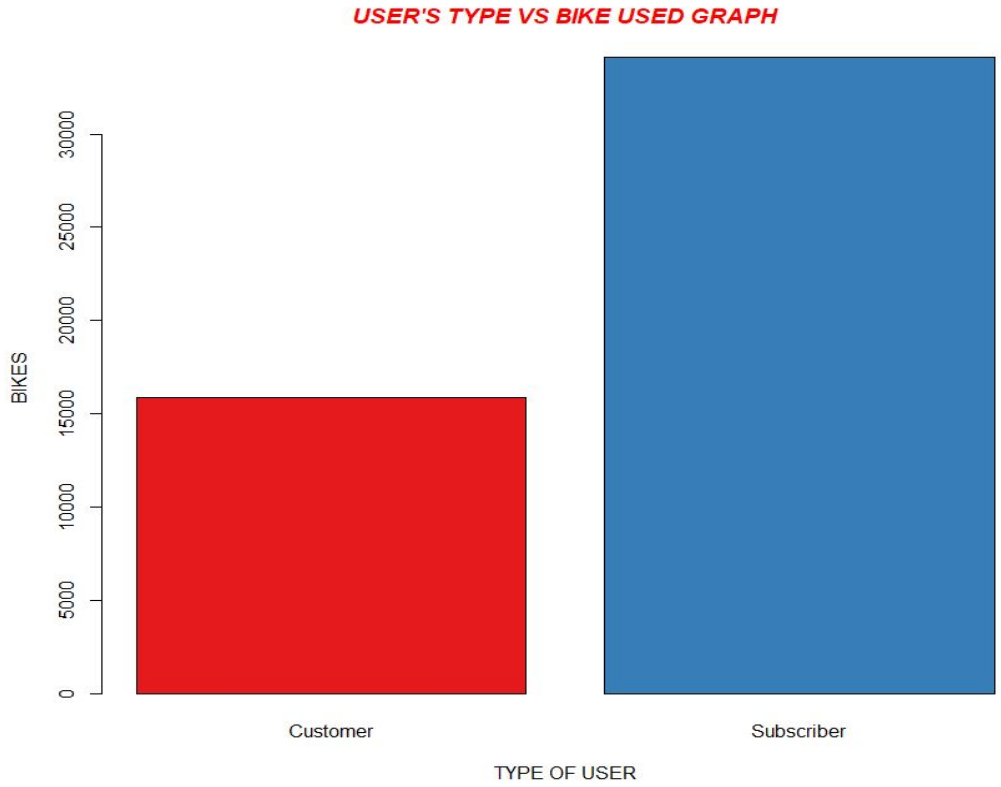
Values

#Z.score=(mean.Customer-mean.Subscriber)/sd.Sb.Cu
Z.score:
= -8.464475
pnorm(Z.score)
= 1.286544e-17
p=-pnorm(Z.score)(Because Z score is negative, and left tailed)
P = -1.287374e-17

Data Charts Infographics



The p value is significantly lower and with high negative z score which was recorded as -8.5. Thus, there is sufficient evidence to support the claim and we reject the null hypothesis and we can say that our Subscribers had more ride time on our bike than The random Customers coming in and using or bike.



- We can use this chart as a reference to accept the claim that the subscribers has more ridetime.

This comparison of bar plots between the customer and number of bikes tell us that Subscribers used out Cycle more and thus we had more ridetime associated with Subscribers than Customers

*Code given in R file

Second Hypothesis testing

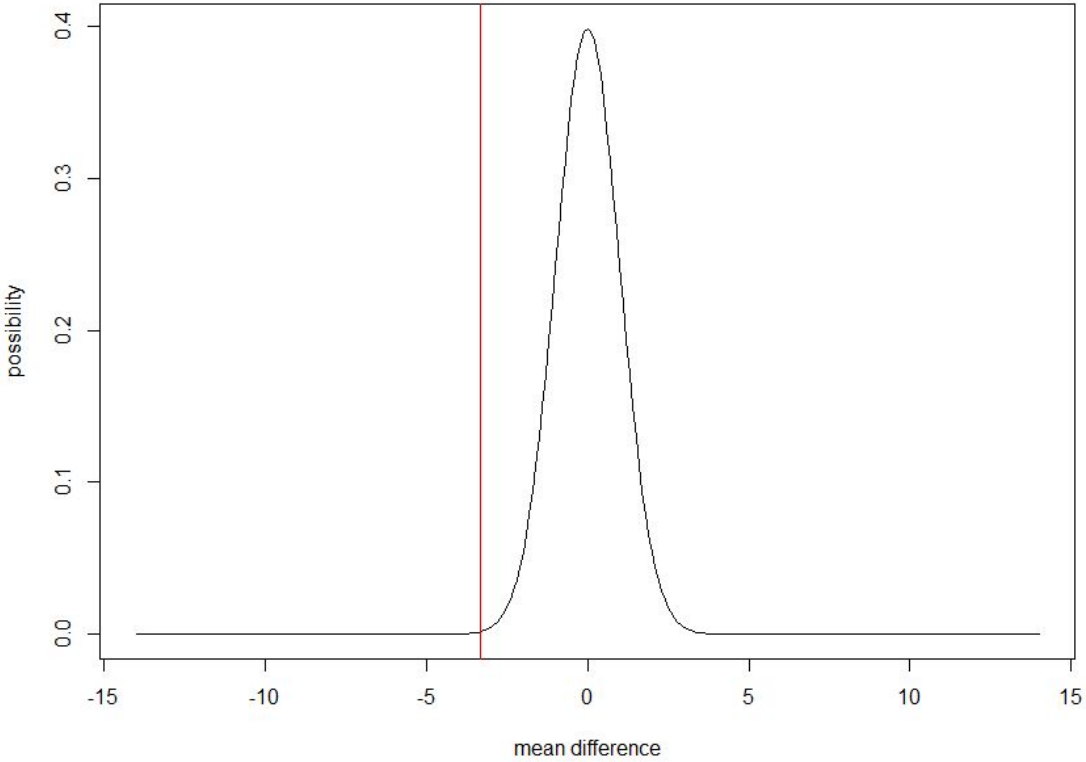
NULL HYPOTHESIS : The ride duration made by Younger age group(<30age) of people who used our bike is equal to the ride duration of city bike by older age group (>30 and <80age)

ALTERNATIVE HYPOTHESIS: The ride duration made by Younger age group(<30age) of people who used our bike is greater than the ride duration of city bike by older age group (>30 and <80age)

Values

Z.score
= -3.306567
pnorm(Z.score)
= 0.0004722342
p= -pnorm(Z.score)
P =-0.0004722338

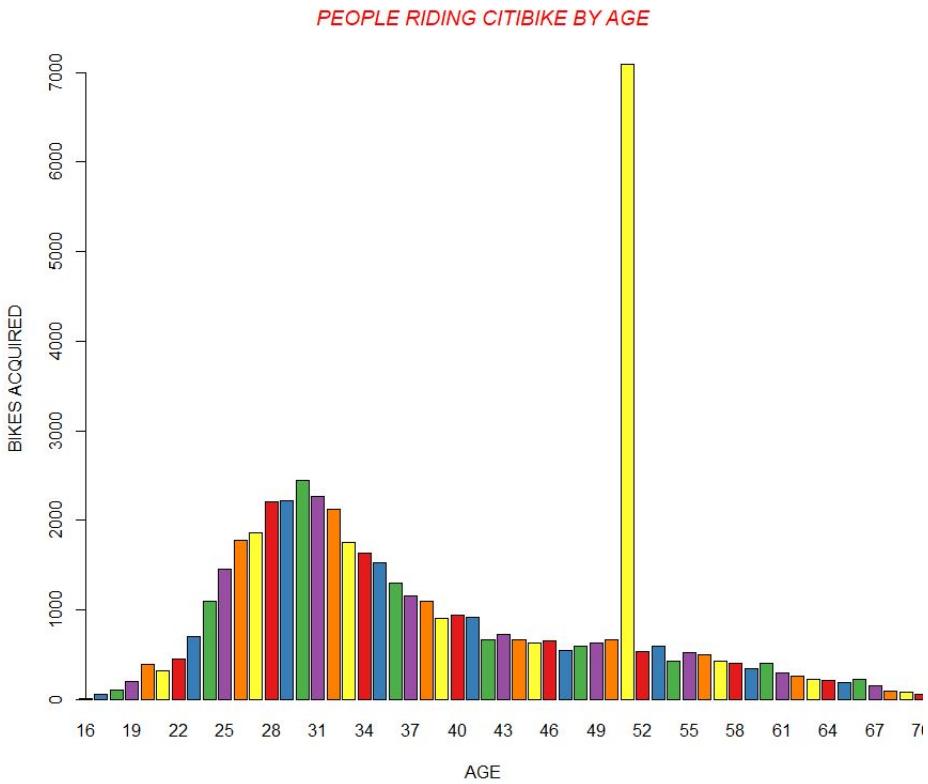
Data Charts Infographics



We have a value of P which is Significantly less and thus with a Z value of -3.3.

Thus we reject the null hypothesis and we can evidence to prove that Older age grouped people used most of our cycle and as a result have more ride duration than than the Younger age group

Other reference



- We can also use this plot as a reference to prove our claim, because we can see that the graph favours elder group with 51 being highest with bike acquired of 7000 which falls in Older age group.

From this we can get a reference that there were more Older age group bike uses and thus makes a sense that they have more ride duration that the younger group