**Preface –**

First of all Titan insurance company has introduced a new scheme and wants to measure whether the scheme has made a significant difference. Since the company is paying high incentives related to the salesperson output in terms of sum assured sold, it wants to compensate the expenditure by sales increase. So, in order to be successful the scheme sales should show an increase in sales for our measurement.

**Action Area –**

We would ideally like to check if the sales data shows a definite indication of increase after 4 months of launch of the scheme.

**Decision process –**

As we examine the data, we can see that - for the same salesman the

output data of old scheme sales is put up with the same salesman output data with the new scheme after 4 months or so. Hence, if we check how many measurements are taken of each subject, the value comes out to be two in each case. On this evidence, we would like to go with the paired t test.

Let's check for normality assumption to go ahead with the paired t test. Since the sample size is 30, we do have 30 pairs here and hence 30 differences on which we can do the test. So, we can go ahead with the t test based on our

Normality assumption.

Still, we additionally do a round of checks for the normality assumption.

> scheme <- read.csv("Insurance\_GA.csv")

> attach(scheme)

> summary(Difference)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-34.00 -6.75 7.50 4.00 16.00 25.00

> sd(Difference)

[1] 14.08105

Thus, we can see that – 100% values are falling with in the Mean +- StdDev\*3 range here confirming the normality assumption of the paired differences.

> shapiro.test(Difference)

Shapiro-Wilk normality test

data: Difference

W = 0.93964, p-value = 0.08895

Thus even the Shapiro test also proves the same(null hypothesis of normality can not be rejected as the p-value is greater than 0.05).

**Form the test –**

So, let's now form our null hypothesis.

If, d = difference between the sales output across 2 different schemes for

each sales person.

Mean of the difference of 30 samples comes out to be μd.

So, our hypothesis form will be like –

H0 : μd = 0

Ha : μd > 0

We will do our test based on 95% confidence level, i.e., a level of significance alpha = 0.05

Let's run the t test in R and try to understand the results.

**Run the test –**

> attach(scheme)

> summary(Old.Scheme)

Min. 1st Qu. Median Mean 3rd Qu. Max.

28.00 54.00 67.00 68.03 81.50 110.00

> summary(New.Scheme)

Min. 1st Qu. Median Mean 3rd Qu. Max.

32.00 55.00 74.00 72.03 85.75 122.00

**> t.test(New.Scheme,Old.Scheme,paired=TRUE,alternative = "greater")**

Paired t-test

data: New.Scheme and Old.Scheme

t = 1.5559, df = 29, p-value = 0.06529

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

-0.3681762 Inf

sample estimates:

mean of the differences

4

**Interpret our test result**

> ## As we can see here that the p-value is bigger than level of significance 0.05 at DF 29(30-1). Thus we won't be able to reject the null hypothesis that the difference in means is 0.

**Conclusion –**

> ## Thus the conclusion is : There is no statistical evidence to say that

the true difference in means of sales figure is greater than 0. Hence, We can conclude that the new scheme is not able to increase the average sales as expected as of now.