***ASSIGNMENT DAY 12***

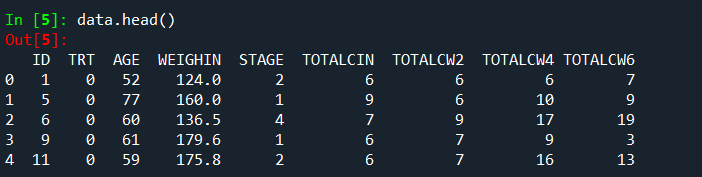
***NON PARAMETRIC TESTS***

*1>Wilcoxon sign test*

*import pandas as pd*

*data=pd.read\_excel("1 Wilcoxon.xlsx")*

*data.head()*

**

*H0 = There is no significant difference between the calcium levels of patients from initial to first 2 weeks*

*HA = There is no significant difference between the calcium levels of patients from initial to first 2 weeks*

*from scipy.stats import Wilcoxon*

*stats,p=wilcoxon(data.TOTALCIN,data.TOTALCW2)*

**

*As the p value is less than .05 the null hypothesis is rejected .*

*So there is significant difference in calcium levels of patients in 2 weeks.*

*2>FRIDMAN TEST*

*H0 = There is no significant difference between the calcium levels of patients from initial to first 2 weeks and to 4 weeks*

*HA = There is no significant difference between the calcium levels of patients from initial to first 2 weeks and to 4 weeks*

*from scipy.stats import friedmanchisquare*

*d1=data.TOTALCIN*

*d2=data.TOTALCW2*

*d3=dataset.TOTALCW4*

*stat,p=friedmanchisquare(d1,d2,d3)*

**

*As the p value is less than .05 the null hypothesis is rejected .*

*So there is significant difference in calcium levels of patients in 2 weeks and 4 weeks.*

*3>MANN WHITNEY TEST*

*H0 = There is no significant difference in sales due to design 1 and design 2*

*HA = There is significant difference in sales due to design1 and design2.*

*from scipy.stats import mannwhitneyu*

*datat1=pd.read\_excel("3 Mann Whitney.xlsx",shee\_tname=1*

*a1=dataset1.Design1*

*a2=dataset1.Design2*

*stat,p=mannwhitneyu(a1,a2)*

*print(stat,p)*

**

*As the p value is more than .05 the null hypothesis is accepted .*

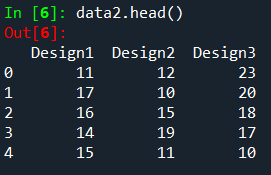
*So there is no difference in design1 and design 2 on sales.*

*4>KRUSHAL WALLIS TEST*

*from scipy.stats import Kruskal*

*data2=pd.read\_excel("4 Kruskal Wallis.xlsx",sheet\_name=0)*

*data2.head()*

**

*H0 = There is no significant difference in sales due to design 1 and design 2 and design 3.*

*HA = There is significant difference in sales due to design1 and design2 and design 3.*

*b1=data2.Design1*

*b2=data2.Design2*

*b3=data2.Design3*

*stat,p=kruskal(b1,b2,b3)*

**

*As the p value is more than .05 the null hypothesis is accepted .*

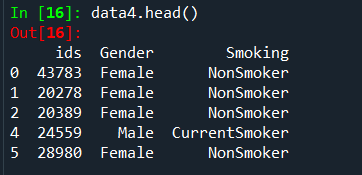
*So there is no difference in design1 and design 2 and design 3 on sales.*

*5>CHI SQUARE TEST*

*data3=pd.read\_excel("5 Chi square Test.xlsx")*

*data4=data3.dropna()*

*data4.head()*

**

*H0 = There is no significant dependency between in gender and smoking.*

*HA = There is significant dependency between in gender and smoking.*

*from scipy.stats import chi2\_contingency*

*chitable=pd.crosstab(data4.Gender,data4.Smoking)*

*stats,p,dof,expected=chi2\_contingency(chitable)*

**

*As the p value is more than .05 the null hypothesis is accepted .*

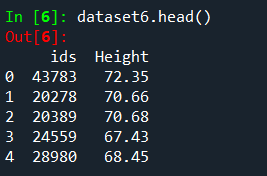
*So there is no dependency in gender for smoking.*

***PARAMETRIC TEST***

*1>ONE SAMPLE T-TEST*

*from scipy.stats import ttest\_1samp*

*dataset6=pd.read\_excel("1. One Sample.xlsx",sheet­\_name=0)*

**

*H0 = There is no significant difference between sample mean and population mean.*

*HA = There is significant difference between sample mean and population mean.*

*h1=dataset6.Height*

*stat,p=ttest\_1samp(h1,65)*

**

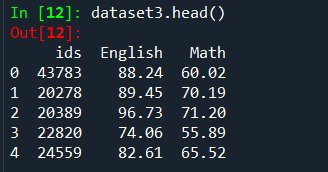
*As the p value is less than .05 the null hypothesis is rejected .*

*So there is difference in sample and population mean .*

*2> TWO SAMPLE PAIRED T-TEST*

*from scipy.stats import ttest\_rel*

*dataset3=pd.read\_excel("2. Paired Sample.xlsx",sheet\_name=0)*

**

*H0 = There is no significant difference between mean of Maths and English marks.*

*HA = There is significant difference between mean of Maths and English marks.*

*p1=dataset3.English*

*p2=dataset3.Math*

*stat,p=ttest\_rel(p1,p2)*

**

*As the p value is less than .05 the null hypothesis is rejected .*

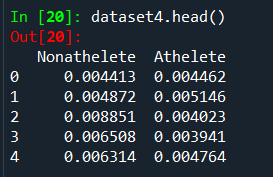
*So there is difference in mean of English and maths marks.*

*3>TWO SEPARATE/INDEPENDENT T-TEST*

*from scipy.stats import ttest\_ind*

*dataset4=pd.read\_excel("3. Independent Sample.xlsx",sheetname=3)*

*dataset4.head()*

**

*H0 = There is no significant difference between mean of duration by atheletes and non atheletes.*

*HA = There is significant difference between mean of duration by atheletes and non atheletes.*

*z1=dataset4.Nonathelete*

*z2=dataset4.Athelete*

*stat,p=ttest\_ind(z1,z2)*

*print(stat,p)*

**

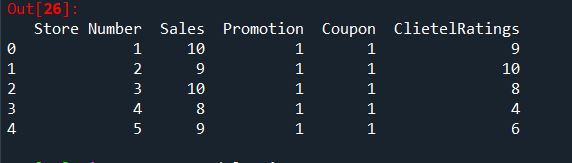
*As the p value is less than .05 the null hypothesis is rejected .*

*So there is difference in mean of duration of atheletes and non atheletes.*

*4> ONE SAMPLE F-TEST(ONE WAY ANOVA)*

*data6=pd.read\_excel("ANCOVA1.xlsx")*

*data6.head()*

**

*H0 = There is no significant difference in promotion of low, medium, high on sales*

*HA = There is significant difference in promotion of low, medium, high on sales*

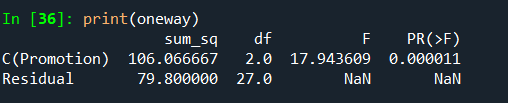
*import statsmodels.api as sm*

*from statsmodels.formula.api import ols*

*mode=ols('Sales~C(Promotion)',data6).fit()*

*oneway=sm.stats.anova\_lm(mode,typ=2)*

*print(oneway)*

**

*As the p value is less than .05 the null hypothesis is rejected .*

*So there is difference in promotion of low, medium, high on sales.*

*5> TWO SAMPLE F-TEST(TWO WAY ANOVA)*

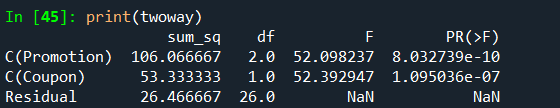
*H0 = There is no significant difference in coupon and promotion of low, medium, high on sales*

*HA = There is significant difference in coupon and promotion of low, medium, high on sales*

*mode=ols('Sales~C(Promotion)+C(Coupon)',data6).fit()*

*twoway=sm.stats.anova\_lm(mode,typ=2)*

*print(twoway)*

**

*As the p value is less than .05 the null hypothesis is rejected .*

*So there is difference in coupon and promotion on sales.*