**Applied Statistics - Lab 7**

**SAP Id:** 500083382

**Name:** Anurag Singh

**Batch:** AI&ML B2

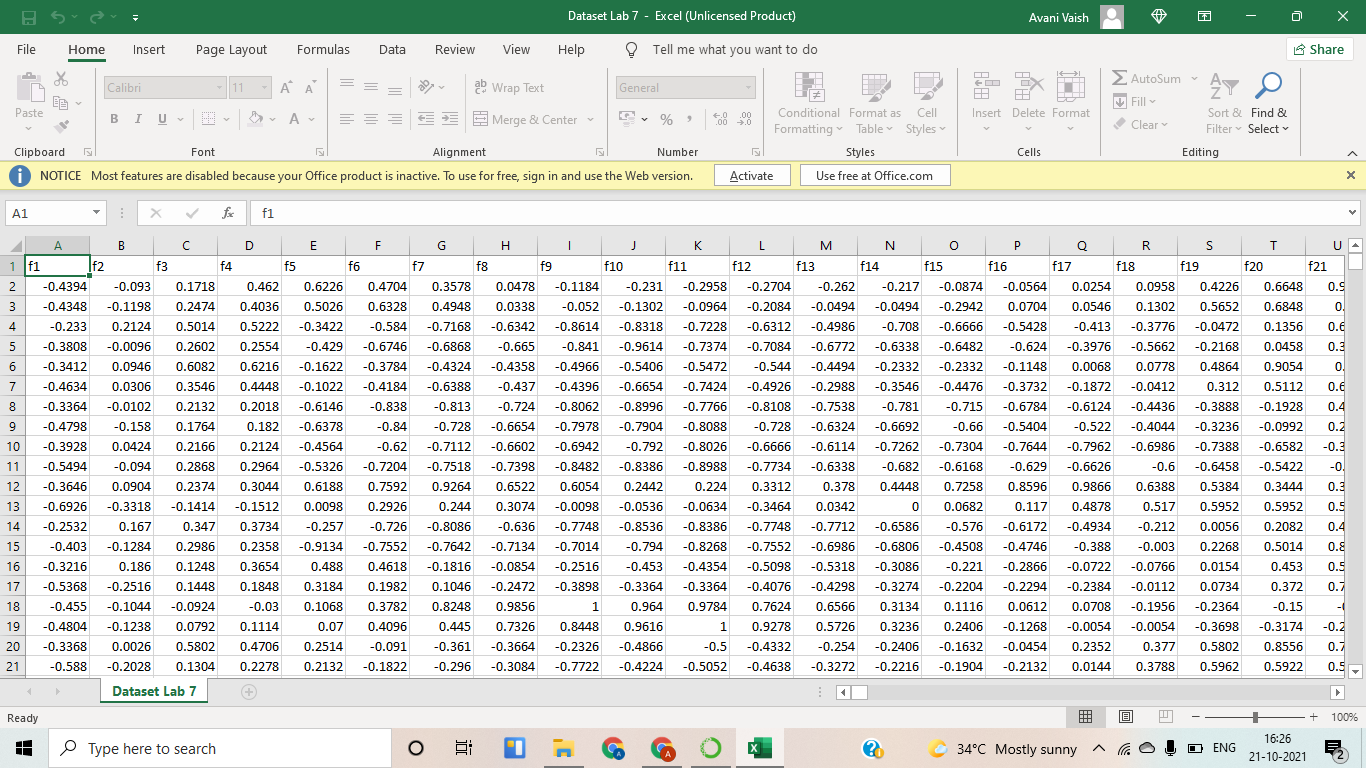
**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Topic: Statistic Measures**

Observe the dataset: https://www.kaggle.com/subhashinimariappan/numerical-dataset

Compute all the statistical measures learnt till the last class: central tendency, dispersion, covariance, correlation, eq of a straight line for all comb. where r>0.6

**Dataset:**



**Code:**

import pandas as pd

import statistics

import pandas as pd

import random

data=pd.read\_csv("Dataset Lab 7.csv")

data1=data.head(100)

for i in ['f1','f2','f3','f4','f5','f6']:

print("Mean of ",i,"is: ",statistics.mean(data1[i]))

print("Median of ",i,"is: ",statistics.median(data1[i]))

print("Mode of ",i,"is: ",statistics.mode(data1[i]))

print("Standerd Deviation of ",i,"is: ",statistics.stdev(data1[i]))

print()

ran1=random.choice(['f1','f2','f3','f4','f5','f6'])

ran2=random.choice(['f1','f2','f3','f4','f5','f6'])

sum=0

for j in range(len(data1)):

sum = sum + (data1[ran1][j]-statistics.mean(data[ran1]) \* data1[ran2][j]-statistics.mean(data[ran2]))

covar=sum/len(data1)

corr=covar/(statistics.stdev(data1[ran1])\*statistics.stdev(data1[ran2]))

print("Covariance of ",ran1,"and ",ran2,"is ",covar)

print("Correlation of ",ran1,"and ",ran2,"is ",corr)

print()

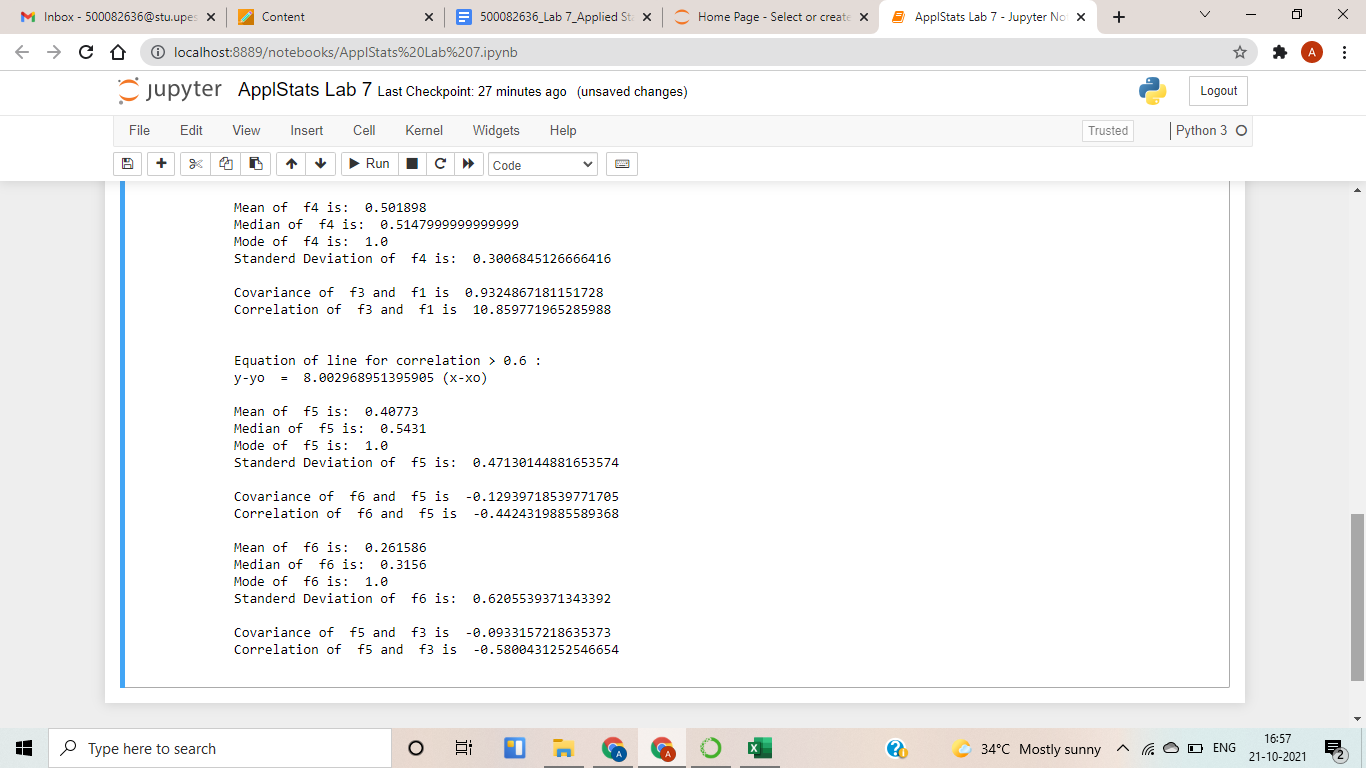
if (corr>0.6):

print()

print("Equation of line for correlation > 0.6 : ")

print("y-yo = ", covar/(statistics.stdev(data1[ran1])\*\*2), "(x-xo)")

print()

**Screenshot:  
 **