LAB QUESTIONS AND ALGORITHMS

1) Design a data frame for manipulating the student record it consist of roll no, name, mark of five subject .List the details of student those who have score maximum mark and score the toppest mark in each and every subject.

Algorithm:

Step 1: Start

Step2:Import the package we needed in this program(import pandas as pd)

Step3:Design a data frame name data containing record of student consist of name, roll no, mark of five subjects.

Step4:Store the data in df that means df=pd.DataFrame(Data).

Step5:Then print our dataframe.

Step6:print who have to score maximum mark in each and every subject so we do df[“name”].loc[df[“subject name”].idxmax()],df[“subject name”].loc[df[“subject name”].idxmax()]/*this will always print the name of student who score maximum mark in subject/*

Step7:Stop

Program:

```
import pandas as pd
Data={
    "rollno":[1,2,3,4,5],
    "name":['priya','manu','gopika','pavi','reena'],
    "cs":[89,56,78,90,45],
    "maths":[83,46,89,78,46],
    "ai":[34,56,78,90,20],
    "graphics":[43,56,78,23,90],
    "english":[34,89,78,65,90],
}
df=pd.DataFrame(Data)
print(df)
print("maximum value score student in cs
is:",df["name"].loc[df["cs"].idxmax()],df["cs"].loc[df["cs"].idxmax()])
print("maximum value score student in cs
is:",df["name"].loc[df["maths"].idxmax()],df["maths"].loc[df["maths"].idxma
x()])
print("maximum value score student in cs
is:",df["name"].loc[df["ai"].idxmax()],df["ai"].loc[df["ai"].idxmax()])
print("maximum value score student in cs
is:",df["name"].loc[df["graphics"].idxmax()],df["graphics"].loc[df["graphic
s"].idxmax()])
print("maximum value score student in cs
```

```
is:",df["name"].loc[df["english"].idxmax()],df["english"].loc[df["english"]
.idxmax()])
```

2)Download and read any numeric dataset available in a public space .Find the standard deviation of each and every column

Algorithm

Step1:start

Step2:download a csv file

Step3:import pandas as pd

Step4:from math import sqrt(in this program some mathematical functions are performed so we want to import this)

Step4:read the csv file and store in a variable

Step5:print csv file

Step6:column=file.columns

Step7:for I in colum,set sum=0,if i==”unnamed==0”,continue

Step8:for j in file[i],sum+=j,

Step9:repeat step7 and 8 until the whole column is traversed:Then go to step10

Step 10:mean=sum/len(file[i])

Step11:set var=0

Step12:for j in file[i]:

Step13:sum+=j

Step14:mean=sum/length(file[i]-1),std=sqrt(variance)

Step15:print standared deviation

Step 6:stop

Program:

```
import pandas as pd
from math import sqrt
file = pd.read_csv("ArgentinaCPI.csv")
print(file)
column = file.columns
for i in column:
```

```
sum = 0
if i == "unnamed:0":
    continue
for j in file[i]:
        sum += j
mean = sum/len(file[i])
var=0
for j in file[i]:
    var+=(j-mean)**2
    variance=var/len(file[i]-1)
std=sqrt(variance)
print("standard deviation",std)
```

3)Download and read any numeric data anf find min,maximum,meadian in each and every column

Step1:start

Step 2:import pandas.pd

Step3:read csv file

Step4:calculate mean,min, max

Step5:print min max mean

Step6:stop

4)Download any numeric dataset.Normalized the dataset with any function for normalization

Step1:start

Step2:from sklearn import preprocessing

Step3:import pandas as pd

Step4:read csv file and store in variable

Step5:preprocessing.normalize(csvfile)

Step6:normalize the columns in csv file

Step 6:print the normalize dataset.

Step6:stop

5)download any chategorical dataset and convert it in to neumerical data by applying some mechanism.

Step1:start

Step2:import pandas as pd

Step3:read the download csv file

Step4:pd.get_dummies(df[“purchased”])

Step5:pd.conct([df,df1],axis=1).reindex(df.index)

Step6:df.drop(“purchased”,axis=1,inplace=”True”)

Step6:print the new dataset

Step7:stop