Practical-4

Download “heart_2020_cleaned.csv” dataset from “https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease “and perform all the descriptive statistics on above dataset using statistics module of python and scipy.stats package (Measures of central tendency, measure of dispersion/variation, measure of location, measure of shape and symmetry).

```
import csv
import pandas as pd
from scipy.stats import skew
from termcolor import colored

class color:
   BOLD = '\033[1m'
   END = '\033[0m'

print("12002040701067")
def Seperator():
  print("----------------------------------------------------------------------------")

df=pd.read_csv("/content/heart_2020_cleaned.csv")
Seperator()
print(color.BOLD+"Head Of DataFrame : \n\n"+color.END,df.head(5))
Seperator()

print(color.BOLD+"Tail Of DataFrame : \n\n"+color.END,df.tail(5))
Seperator()

print(color.BOLD+"Mean Of BMI Column Is : "+color.END,df["BMI"].mean())
Seperator()

print(color.BOLD+"Median Of BMI Column Is : "+color.END,df["BMI"].median())
Seperator()

print(color.BOLD+"Mode : \n\n"+color.END,df.mode())
Seperator()

print(color.BOLD+"Measure Of Standard Deviation : \n\n"+color.END,df.std())
Seperator()

print(color.BOLD+"Measure Of Variance : \n\n"+color.END,df.var())
Seperator()

# Skewness = 0: Then normally distributed.
# Skewness > 0: Then more weight in the left tail of the distribution.
# Skewness < 0: Then more weight in the right tail of the distribution.

print(color.BOLD+"Shape Of Dataset : "+color.END,df.shape)
```

```
Seperator()

a=skew(df["BMI"], axis=0, bias=True)
print(color.BOLD+"Value Of Skewness Is : "+color.END,a)
print(color.BOLD+"According To Value Of Skewness, The Graph"+color.END)
if(a==0):
  print(color.BOLD+"Is Normally Distributed."+color.END)
elif(a>0):
  print(color.BOLD+"Has More Weight In The Left Tail Of The Distribution."+color.END)
else:
  print(color.BOLD+"Has More Weight In The Right Tail Of The Distribution."+color.END)
Seperator()
```

```
12002040701067
--------------------------------------------------------------------------------
Head Of DataFrame :

   HeartDisease    BMI Smoking AlcoholDrinking Stroke  PhysicalHealth  \
0            No  16.60     Yes              No     No             3.0
1            No  20.34      No              No    Yes             0.0
2            No  26.58     Yes              No     No            20.0
3            No  24.21      No              No     No             0.0
4            No  23.71      No              No     No            28.0

   MentalHealth DiffWalking     Sex  AgeCategory   Race Diabetic  \
0          30.0          No  Female        55-59  White      Yes
1           0.0          No  Female  80 or older  White       No
2          30.0          No    Male        65-69  White      Yes
3           0.0          No  Female        75-79  White       No
4           0.0         Yes  Female        40-44  White       No

  PhysicalActivity  GenHealth  SleepTime Asthma KidneyDisease SkinCancer
0              Yes  Very good        5.0    Yes            No        Yes
1              Yes  Very good        7.0     No            No         No
2              Yes       Fair        8.0    Yes            No         No
3               No       Good        6.0     No            No        Yes
4              Yes  Very good        8.0     No            No         No
--------------------------------------------------------------------------------
Tail Of DataFrame :

        HeartDisease    BMI Smoking AlcoholDrinking Stroke  PhysicalHealth  \
319790          Yes  27.41     Yes              No     No             7.0
319791           No  29.84     Yes              No     No             0.0
319792           No  24.24      No              No     No             0.0
319793           No  32.81      No              No     No             0.0
319794           No  46.56      No              No     No             0.0

        MentalHealth DiffWalking     Sex  AgeCategory      Race Diabetic  \
319790           0.0         Yes    Male        60-64  Hispanic      Yes
319791           0.0          No    Male        35-39  Hispanic       No
319792           0.0          No  Female        45-49  Hispanic       No
319793           0.0          No  Female        25-29  Hispanic       No
319794           0.0          No  Female  80 or older  Hispanic       No

       PhysicalActivity  GenHealth  SleepTime Asthma KidneyDisease SkinCancer
319790               No       Fair        6.0    Yes            No         No
319791              Yes  Very good        5.0    Yes            No         No
319792              Yes       Good        6.0     No            No         No
319793               No       Good       12.0     No            No         No
```

```
319794              Yes       Good        8.0      No             No          No
--------------------------------------------------------------------------------
Mean Of BMI Column Is :  28.325398520927465
--------------------------------------------------------------------------------
Median Of BMI Column Is :  27.34
--------------------------------------------------------------------------------
Mode :

   HeartDisease    BMI Smoking AlcoholDrinking Stroke  PhysicalHealth  \
0            No  26.63      No              No     No             0.0
```

✓ 2s completed at 6:53 PM