

Practical No. 01

Code:

```
import pandas as pd

# Read data into Python

education = pd.read_csv("D:\Desktop files\DESKTOP DEC 2021\Basi
Statistics_v3\education.csv")

type(education)
education.info\(\)

# C:\Users\education.csv - this is windows default file path with a '\'
# C:\\Users\\education.csv - change it to '\\' to make it work in Python

# Exploratory Data Analysis
# Measures of Central Tendency / First moment business decision
education.workex.mean()

education.gmat.mean() # '.' is used to refer to the variables within object
education.workex.median()
education.workex.mode()
max(education.workex)
min(education.workex)
max(education.gmat)
min(education.gmat)
r = max(education .workex)-min (education .workex)
print(r)

# pip install numpy
from scipy import stats
stats.mode(education.workex)

# Measures of Dispersion / Second moment business decision
education.workex.var() # variance
education.workex.std() # standard deviation
education.gmat.var()
education.gmat.std()
range = max(education.workex) - min(education.workex) # range
```

```
range
y=min(education.workex)
print(y)
x=max(education.workex)
print(x)
# Third moment business decision
education.workex.skew()
education.gmat.skew()

# Fourth moment business decision
education.workex.kurt()

# Data Visualization
import matplotlib.pyplot as plt
import numpy as np

education.shape

plt.bar(height = education.gmat, x = np.arange(1, 774, 1)) # initializing the parameter

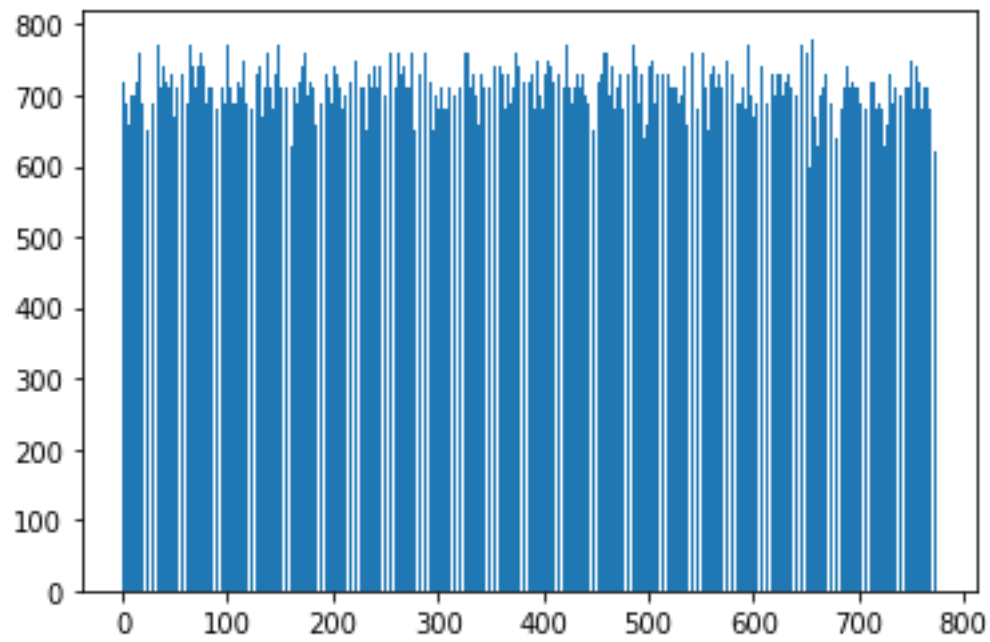
plt.hist(education.gmat) #histogram
plt.hist(education.workex)
plt.hist(education.workex, color='red')

help(plt.hist)

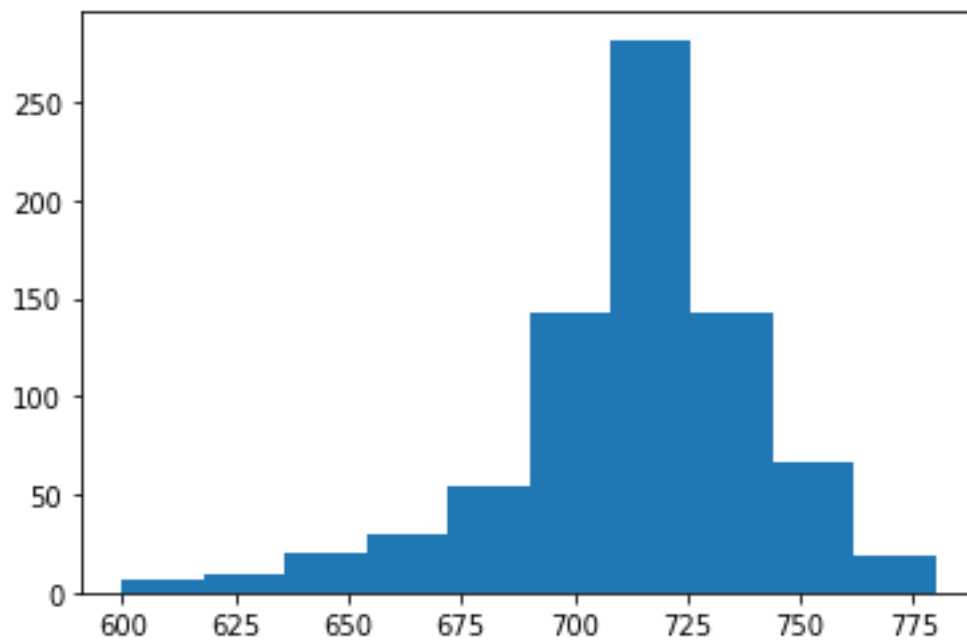
plt.boxplot(education.gmat) #boxplot

help(plt.boxplot)
```

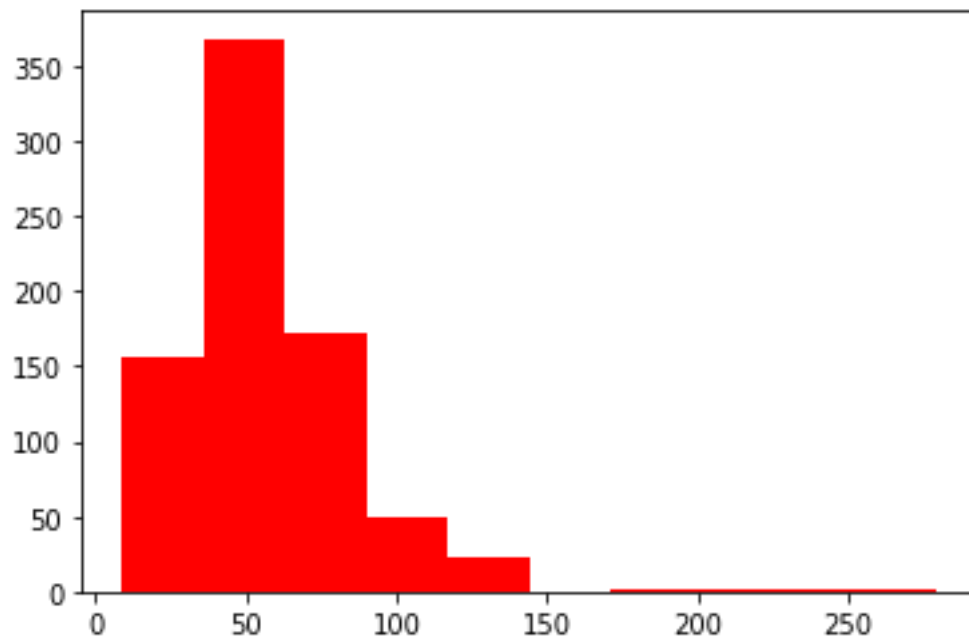
Outputs:



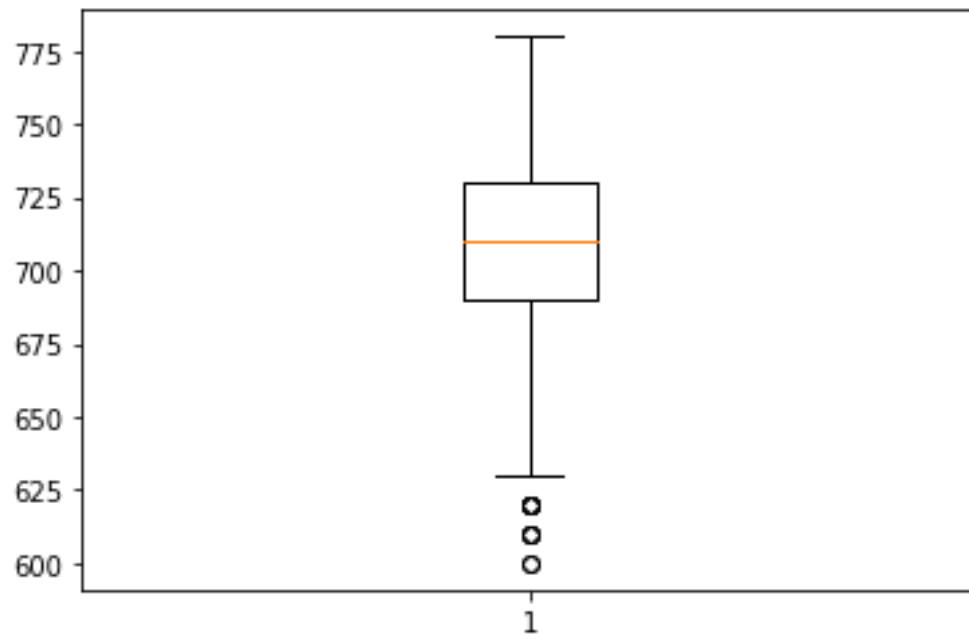
Bar graph for Gmat Scores



Histogram for Gmat Scores



Histogram for Workex



Box-Plot for Gmat Scores