Task 8

Created	@May 15, 2020 2:14 PM
Last Edited Time	@May 15, 2020 2:17 PM
Reviewed	☐

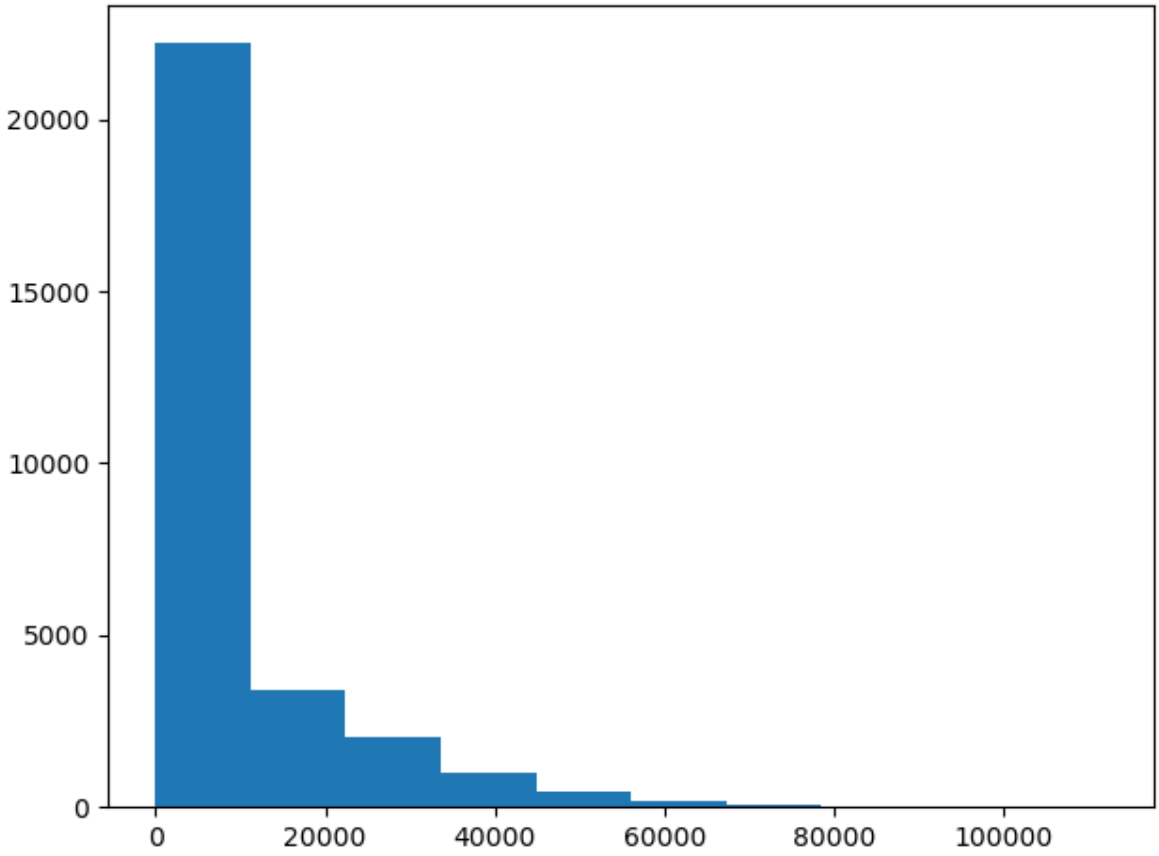
2017038106 권용균

```
import matplotlib.pyplot as plt
import pymongo
import numpy
from scipy import stats

myclient= pymongo.MongoClient("mongodb://localhost:27017/")
mydb= myclient ["test"]
mycol=mydb["zipcodes"]

#배열선언
x_values= []
y_values=[]

for item in mycol.find():
   x_values.append(item["state"])
   y_values.append(item["pop"])


#No1 Measure mean, median, and mode of any attribute in Python

#mean
x=numpy.mean(y_values)
#median
y=numpy.median(y_values)
#mode
z=stats.mode(x_values)

print("Zipcodes Population Mean:",x)
print("Zipcodes Population Median:",y)
print("Zipcodes State Mode:", z)

#No2 Measure spread using standard deviation of any attribute in Python

#Spread
x1=numpy.std(y_values)

print("Zipcodes Spread:", x1)
```

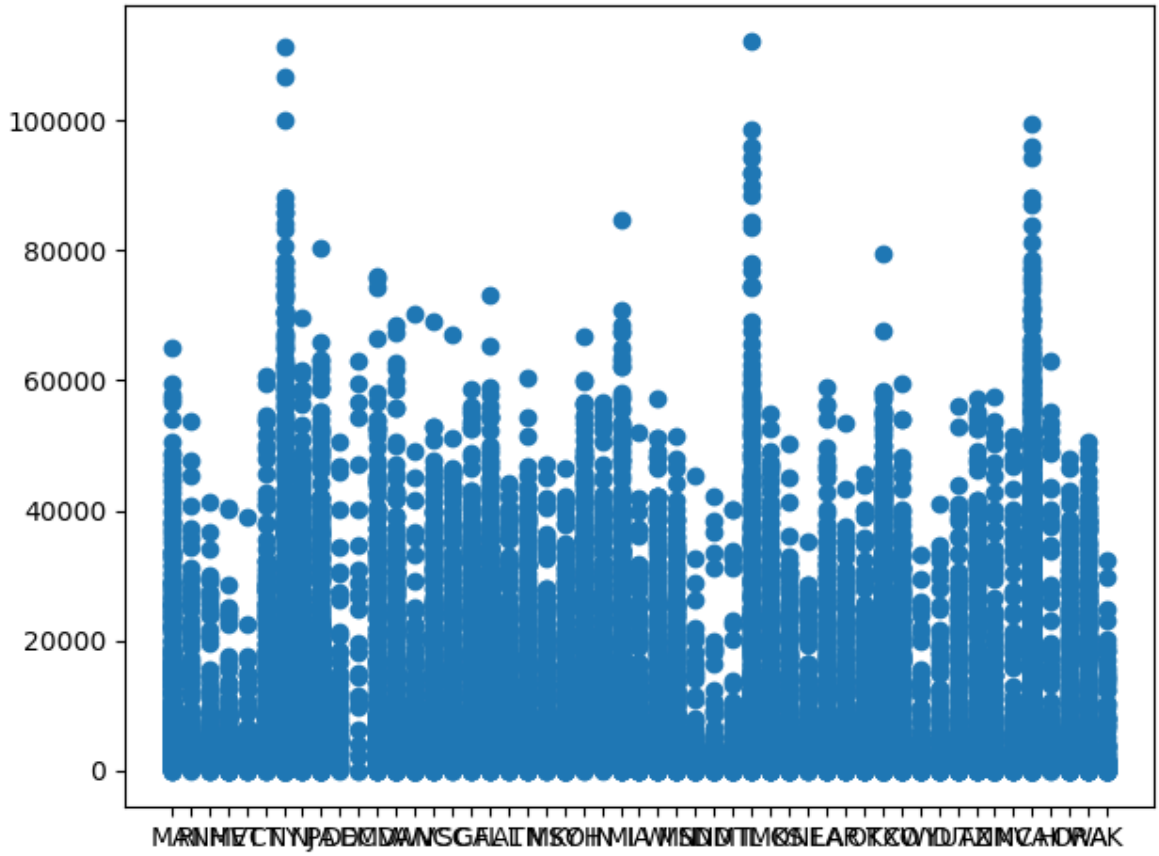
```
#No3 Visualize any numeric attributes using Histogram and Scatter Plot in Python
#Histogram
plt.hist(y_values)
plt.show()
#Scatter Plot
plt.scatter(x_values,y_values)
plt.show()

#No4 Find outliers of any attributes using Boxplot in Python

plt.boxplot(y_values)
plt.show()
```

```
Zipcodes Population Mean: 8462.794262937348
Zipcodes Population Median: 2797.0
Zipcodes State Mode: ModeResult(mode=array(['TX'], dtype='<U2'), count=array([1671]))
Zipcodes Spread: 12329.470279871251

Process finished with exit code 0
```

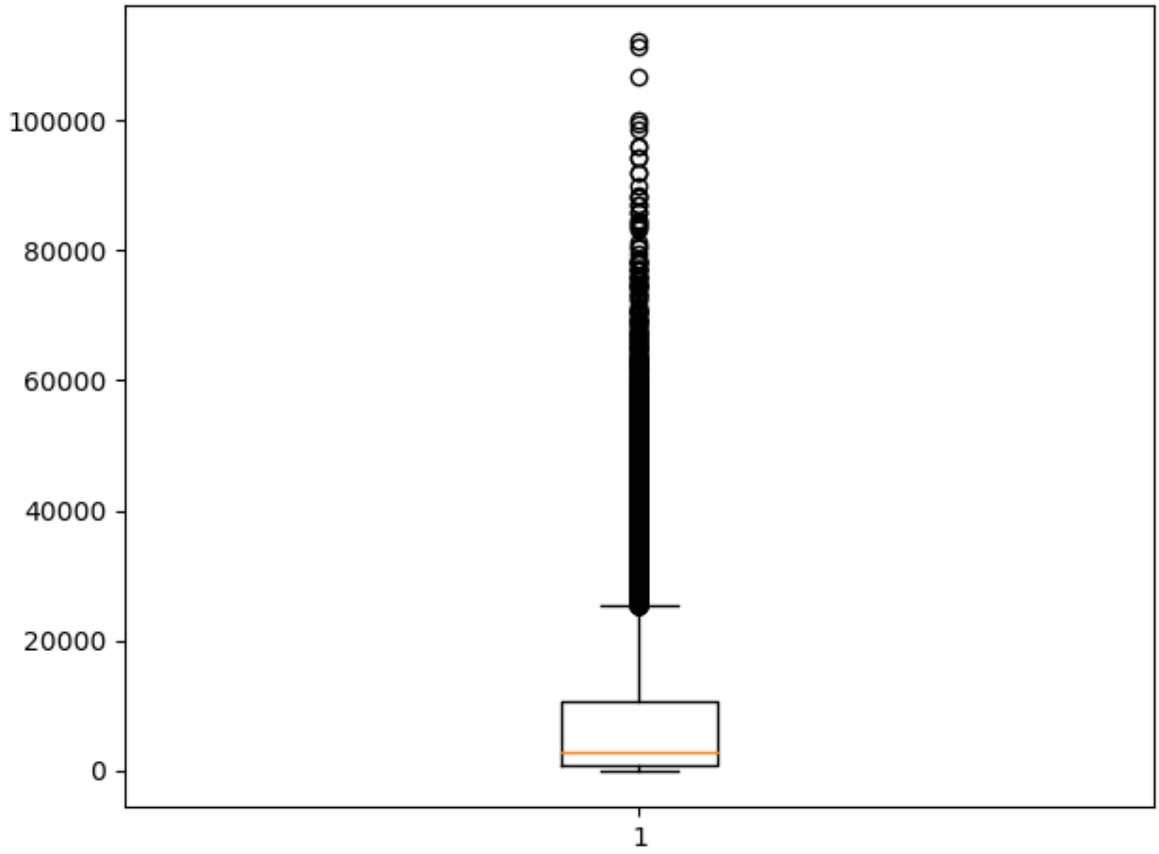
Histogram



Scatter Plot



Boxplot


100000
80000
60000
40000
20000
0
1