REG.NO:1055
NAME: Venkata Narayan Reddy
DATE: 14-10-2022

```python
from pandas import *
from numpy import *
import statistics as stat
import plotly.express as exp
import matplotlib.pyplot as plt

a=read_csv('/content/Enrollments_28092022.csv')
a
```

```
     StudentNo  DEGREE  INTERMEDIATE  SSC
INTERNSHIP
0        1001    8.10          76.0  92.0                       Data
Science
1        1002    8.10          76.0  92.0      MEAN Stack Web
Development
2        1003    7.80          94.6  92.0      MEAN Stack Web
Development
3        1004    9.03          89.5  89.0                       Data
Science
4        1005    8.38          87.0  90.0      MEAN Stack Web
Development
..        ...     ...           ...   ...
...
292      2188    8.70          94.1  93.0                       Data
Science
293      2189    8.45          90.0  93.0                       Data
Science
294      2190    8.40          94.9  98.0                       Data
Science
295      2191    7.06          90.6  88.0  Cloud Computing Services
(AWS)
296      2192    7.50          95.5  95.0  Cloud Computing Services
(AWS)

[297 rows x 5 columns]
```

```python
a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 297 entries, 0 to 296
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   StudentNo     297 non-null    int64
 1   DEGREE        297 non-null    float64
 2   INTERMEDIATE  297 non-null    float64
 3   SSC           297 non-null    float64
 4   INTERNSHIP    297 non-null    object
dtypes: float64(3), int64(1), object(1)
memory usage: 11.7+ KB
```

2 Size of data

```
print('No.of rows',len(a))
print('No.of columns',len(a.axes[1]))
    # (OR)
print('No. of rows and columns are',a.shape)
```
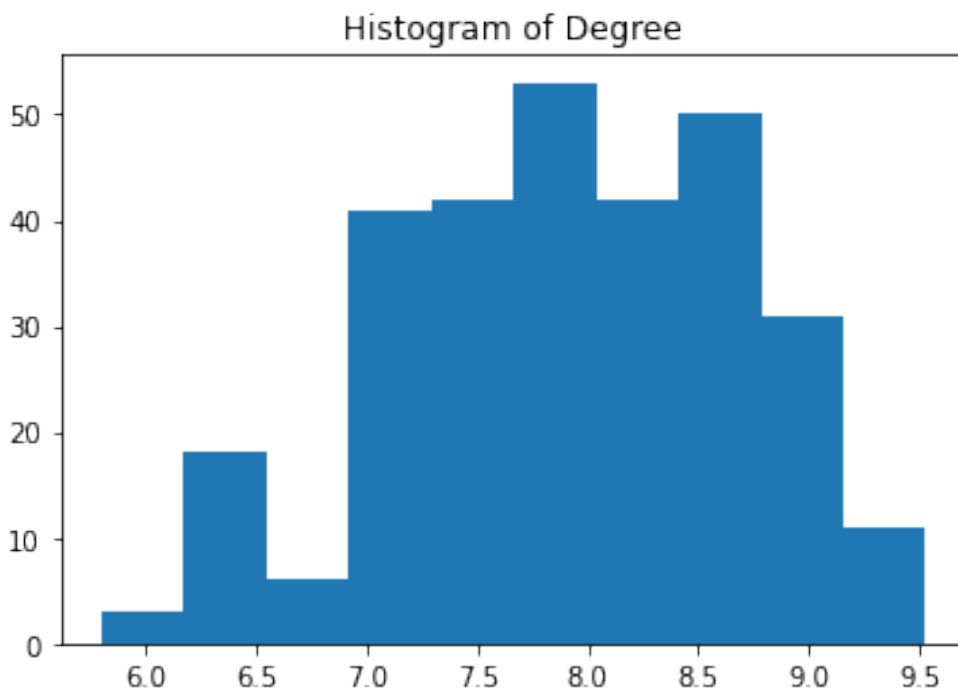
```
No.of rows 297
No.of columns 5
No. of rows and columns are (297, 5)
```
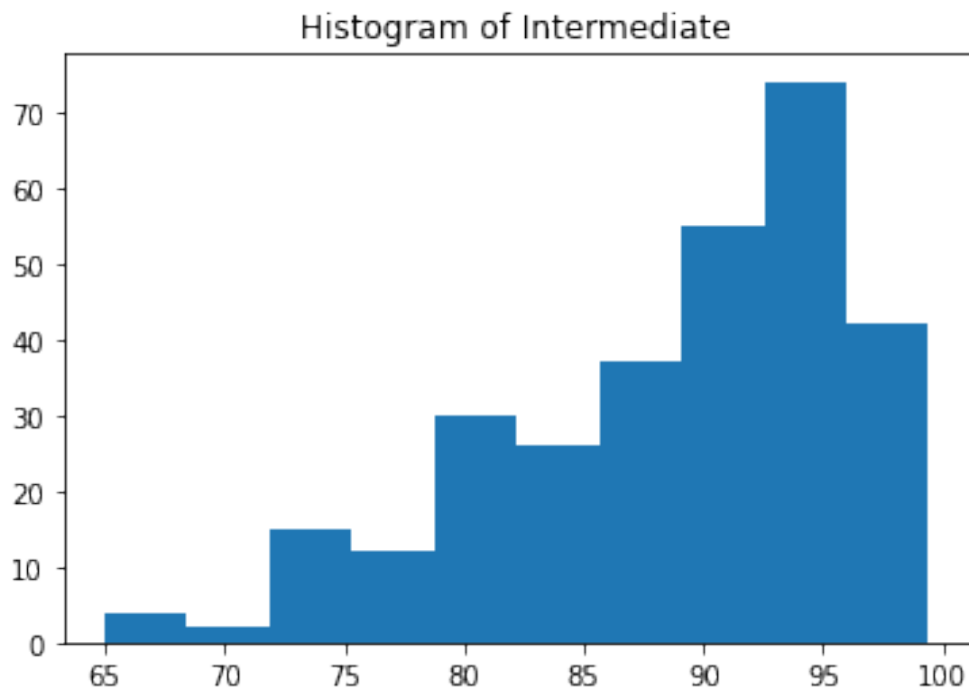
3 Histogram of "Degree"

```
plt.hist(a['DEGREE'])
plt.title('Histogram of Degree')
```

```
Text(0.5, 1.0, 'Histogram of Degree')
```
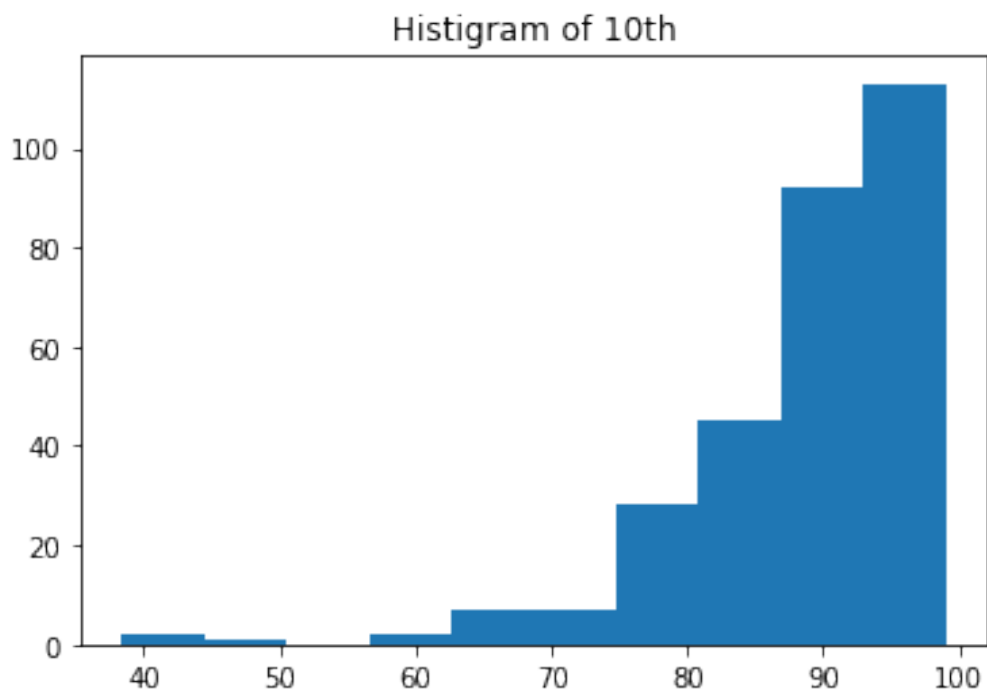

3 Histogram of "Intermediate"

```
plt.hist(a['INTERMEDIATE'])
plt.title('Histogram of Intermediate')
```

```
Text(0.5, 1.0, 'Histogram of Intermediate')
```

Histogram of Intermediate

3 Histogram of "SSC"

```
plt.hist(a['SSC'])
plt.title('Histigram of 10th')
```

```
Text(0.5, 1.0, 'Histigram of 10th')
```


Histigram of 10th

4 Pie Chart representing number of enrollments in differnt internship programs

```python
internshipprograms=['Data Science','Cloud Computing Services
(AWS)','MEAN Stack Web Development']
NoOfEnrollments=[156,90,51]
exp.pie(values=NoOfEnrollments,names=internshipprograms,title='Number
of Enrollments in differnt InternshipPrograms')
```

5 No.Of Enrollements in different Internship Programs

```python
DataFrame(a['INTERNSHIP'].value_counts())
```

```
                              INTERNSHIP
Data Science                         156
Cloud Computing Services (AWS)        90
MEAN Stack Web Development             51
```

6 MEAN,MODE,MEDIAN OF DEGREE , INTERMEDIATE , 10th

```python
b=DataFrame({'MEAN':
[mean(a['DEGREE']),mean(a['INTERMEDIATE']),mean(a['SSC'])],'MEDIAN':
[median(a['DEGREE']),median(a['INTERMEDIATE']),median(a['SSC'])],'MODE
':
[stat.mode(a['DEGREE']),stat.mode(a['INTERMEDIATE']),stat.mode(a['SSC'
])]},index=['DEGREE','INTERMEDIATE','SSC'])
b
```

```
                   MEAN   MEDIAN   MODE
DEGREE         7.928081      8.0    7.0
INTERMEDIATE  88.662626     90.8   95.0
SSC           88.106734     90.0   95.0
```

7 Measures of Varience Of "Degree"

```python
cv= lambda x: std(x, ddof=1)/ mean(x)*100
print("DEGREE")
print("Range=",max(a['DEGREE'])-min(a['DEGREE']))
print("Co-efficient of variation=",cv(a['DEGREE']))
a['DEGREE'].describe()
```

```
DEGREE
Range= 3.7299999999999995
Co-efficient of variation= 9.90881225818308

count    297.000000
mean       7.928081
std        0.785579
min        5.800000
25%        7.400000
50%        8.000000
75%        8.560000
```

```
max          9.530000
Name: DEGREE, dtype: float64
```

7 Measureds of Varience of "Intermediate"

```
print("INTERMEDIATE")
print("Range=",max(a['DEGREE'])-min(a['INTERMEDIATE']))
print("Co-efficient of variation=",cv(a['INTERMEDIATE']))
a['INTERMEDIATE'].describe()
```

```
INTERMEDIATE
Range= -55.47
Co-efficient of variation= 8.29631726338337

count    297.000000
mean      88.662626
std        7.355733
min       65.000000
25%       83.000000
50%       90.800000
75%       94.600000
max       99.400000
Name: INTERMEDIATE, dtype: float64
```

7 Measures of Varience of "10th"

```
print("SSC")
print("Range=",max(a['DEGREE'])-min(a['SSC']))
print("Co-efficient of variation=",cv(a['SSC']))
a['SSC'].describe()
```

```
SSC
Range= -28.869999999999997
Co-efficient of variation= 10.24664491920062

count    297.000000
mean      88.106734
std        9.027984
min       38.400000
25%       85.000000
50%       90.000000
75%       95.000000
max       99.000000
Name: SSC, dtype: float64
```

8 Standard Scores of Degree , Intermediate , SSC

```
import scipy.stats as stats
c=DataFrame({'DEGREE':stats.zscore(a['DEGREE']),'INTERMEDIATE':stats.z
score(a['INTERMEDIATE']),'10th':stats.zscore(a['SSC'])})
c
```

```
       DEGREE   INTERMEDIATE       10th
0     0.219213      -1.724369  0.431972
1     0.219213      -1.724369  0.431972
2    -0.163315       0.808539  0.431972
3     1.405052       0.114032  0.099111
4     0.576240      -0.226413  0.210065
..         ...            ...       ...
292   0.984271       0.740450  0.542926
293   0.665497       0.182121  0.542926
294   0.601742       0.849392  1.097694
295  -1.106886       0.263827 -0.011843
296  -0.545844       0.931099  0.764833

[297 rows x 3 columns]
```

8 Inter Quartile Range of Degree, Intermediate , 10th

```
iqr=DataFrame({'Degree':quantile(a['DEGREE'],0.75)-
quantile(a['DEGREE'],0.25),'Intermediate':quantile(a['INTERMEDIATE'],0
.75)-quantile(a['INTERMEDIATE'],0.25),'10th':quantile(a['SSC'],0.75)-
quantile(a['SSC'],0.25)},index=['Inter Quartile Range'])
iqr
```

```
                      Degree  Intermediate  10th
Inter Quartile Range    1.16          11.6  10.0
```
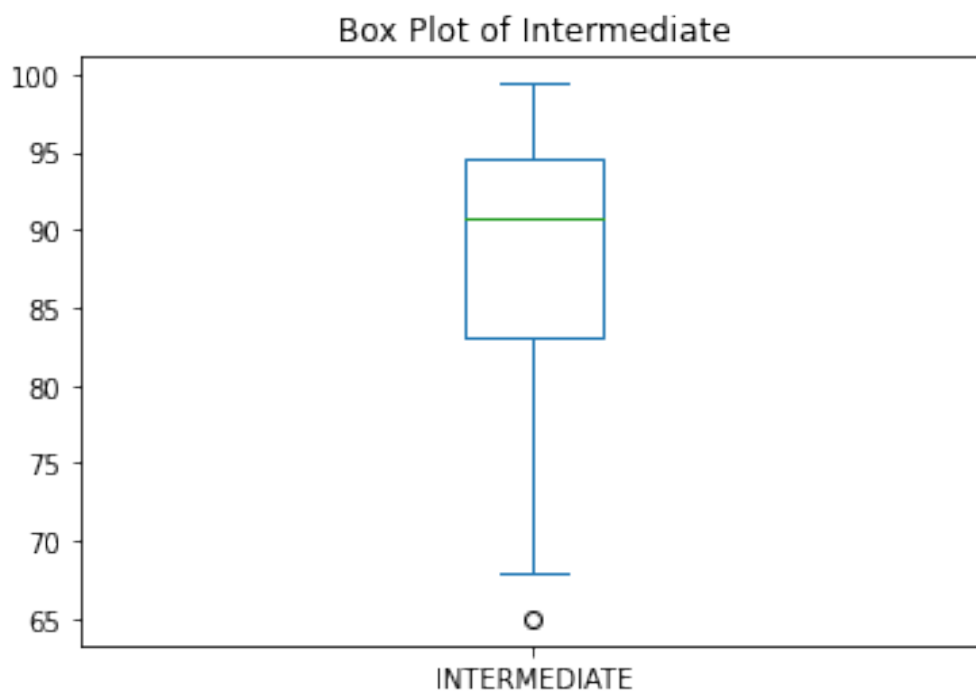
9 Box plot of Degree

```
a['DEGREE'].plot(kind='box',title='Box Plot of Degree')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f23b276c290>
```
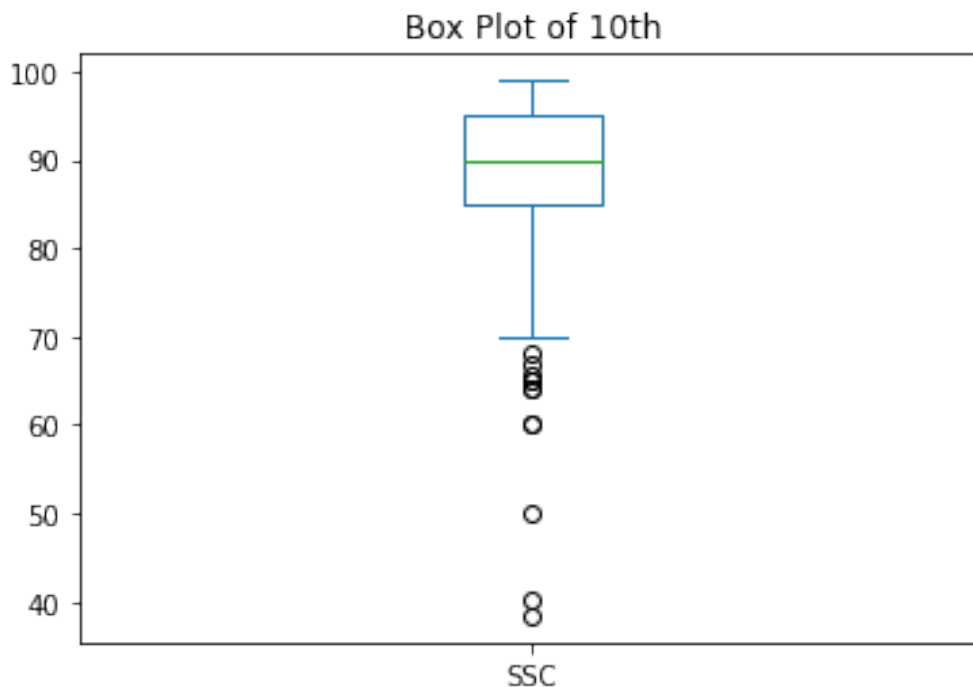
Box Plot of Degree
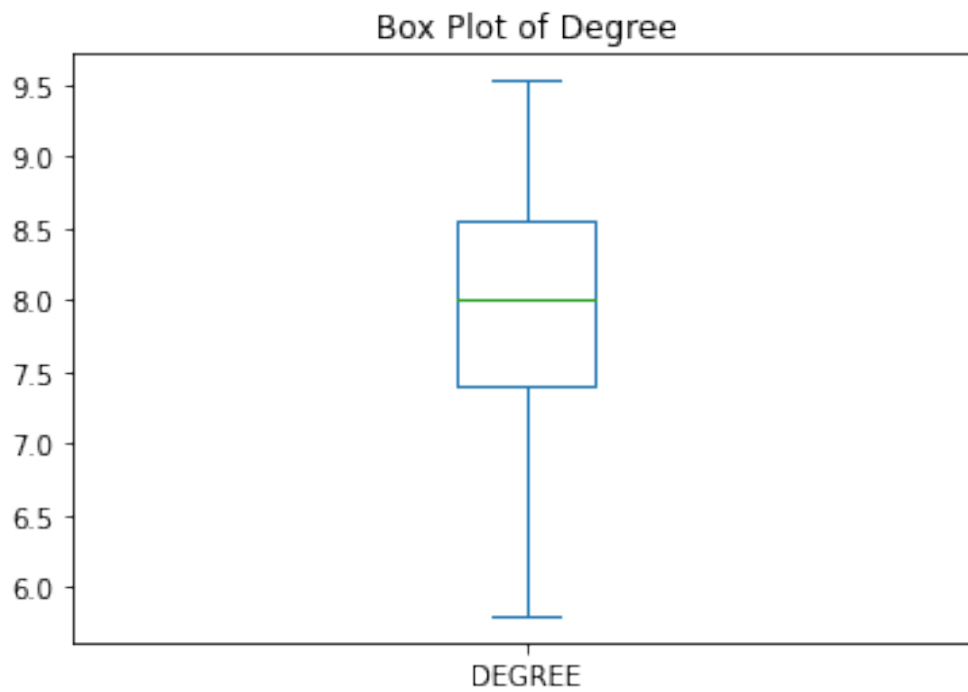
9 Box plot of Intermediate

```
a['INTERMEDIATE'].plot(kind='box',title='Box Plot of Intermediate')
<matplotlib.axes._subplots.AxesSubplot at 0x7f23b2630f50>
```


Box Plot of Intermediate

9 Box plot of 10th

```
a['SSC'].plot(kind='box',title='Box Plot of 10th')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f23b259e0d0>
```


Box Plot of 10th

9 Detecting Outliers

```python
def Outlier(b):
 q2=quantile(b,0.75)
 q1=quantile(b,0.25)
 upper_bound=q1+(1.5*(q2-q1))
 lower_bound=q2-(1.5*(q2-q1))
 outlier=b[(b<=lower_bound)|(b>=upper_bound)]
 print('Outliers in Degree box plot: ',outlier)
```

9 Outliers in "Degree

```
Outlier(a['DEGREE'])
```

```
Outliers in Degree box plot:  16      9.53
19       9.16
40       9.34
96       9.20
101      6.30
115      6.80
119      6.50
121      6.80
123      6.80
127      9.20
138      9.38
144      9.16
```

```
154     9.20
174     6.33
182     6.50
196     6.50
198     6.50
209     6.60
211     6.51
216     6.00
217     6.20
221     9.28
223     6.50
231     6.50
232     6.60
250     6.50
252     9.45
258     9.34
261     6.50
262     6.20
263     6.50
264     5.80
265     6.50
271     6.50
272     6.30
278     6.00
281     6.80
289     6.50
Name: DEGREE, dtype: float64
```

9 Outliers in "Intermediate"

```
Outlier(a['INTERMEDIATE'])
```

```
Outliers in Degree box plot:  0      76.0
1       76.0
27      75.0
47      73.6
51      68.0
58      70.0
72      76.0
74      74.0
81      75.2
82      74.2
97      75.0
105     77.0
111     75.4
123     68.0
126     72.0
128     76.0
136     75.1
170     75.0
175     75.0
```

```
176      73.1
237      72.0
245      74.7
250      75.0
261      75.0
266      68.0
267      77.0
271      65.0
272      69.0
281      76.0
288      77.0
Name: INTERMEDIATE, dtype: float64
```

9 Outliers in "10th"

```
Outlier(a['SSC'])

Outliers in Degree box plot:  5       64.00
7        70.00
18       78.85
31       60.00
36       75.00
37       78.00
47       75.00
49       75.00
51       68.00
55       80.00
58       75.00
69       60.00
72       77.00
75       73.00
79       75.00
82       65.60
86       50.00
97       73.00
98       75.00
101      78.00
104      79.00
107      64.00
113      80.00
135      75.00
157      75.00
165      78.00
170      79.40
175      79.00
176      80.00
178      71.00
195      80.00
200      73.00
202      80.00
209      75.00
```

```
218     80.00
228     80.00
235     71.00
236     38.40
237     67.00
240     72.00
243     40.20
249     78.00
270     65.00
271     80.00
276     75.00
287     77.00
288     65.00
Name: SSC, dtype: float64
```

10 Identifying No.of Students with 90 percentile

```python
def percentile(c):
  q=quantile(c,0.9)
  r=c[c==q]
  print('No.of Students with 90 percentile',r.count())
```

10 No.Of students with 90 percentile in Degree

```python
percentile(a["DEGREE"])
```

```
No.of Students with 90 percentile 3
```

10 No.Of Students with 90 percentile in Intermediate

```python
percentile(a["INTERMEDIATE"])
```

```
No.of Students with 90 percentile 3
```

10 No.Of Students with 90 percentile in 10th

```python
percentile(a["SSC"])
```

```
No.of Students with 90 percentile 19
```