```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
df= pd.read_csv('studentData.csv')
```

```
df.head()
```

Email Address	Name	Email	Roll no	
nbare19@gmail.com	Sejal Zambare	sejal.zambare19@gmail.com	TECOC359	72
thorat19@gmail.com	Rushikesh Vilas Thorat	rushikesh.thorat19@gmail.com	TECOC347	72
takke19@gmail.com	Atharv Sontakke	atharv123sontakke@gmail.com	TECOC340	72(
rekar19@gmail.com	Amisha Sunil Sherekar	amisha.sherekar19@gmail.com	TECOC328	72(
dekar19@gmail.com	Saurabh Raju Sawardekar	saurabh.sawardekar19@gmail.com	TECOC326	72(

```
df.isnull()
```

	Timestamp	Email Address	Name	Email	Roll no	PRN No.	Mobile No.	First year: Sem 1	First year: Sem 2	S
0	False	False	False	False	False	False	False	False	False	
1	False	False	False	False	False	False	False	False	False	

```
df.isnull().sum()
```

```
Timestamp               0
Email Address           0
Name                    0
Email                   0
Roll no                 0
PRN No.                 0
Mobile No.              0
First year:   Sem 1     0
First year:   Sem 2     0
Second year:   Sem 1    0
Second year:   Sem 2    0
dtype: int64
```

```
df.describe()
```

	Mobile No.	First year: Sem 1	First year: Sem 2	Second year: Sem 1	Second year: Sem 2
count	6.400000e+01	64.000000	64.000000	64.000000	64.000000
mean	8.623097e+09	8.834219	9.095469	9.292031	9.377187
std	9.132070e+08	11.187839	11.171986	0.528523	0.495185
min	7.028870e+09	0.000000	0.000000	6.900000	7.200000
25%	7.766559e+09	7.237500	7.655000	9.050000	9.140000
50%	8.805720e+09	8.260000	8.400000	9.445000	9.450000
75%	9.335094e+09	8.802500	9.115000	9.645000	9.725000

```
plt.hist(df['Second year:   Sem 1'])
plt.show()
```


```
lower_bound=.69
upper_bound=.991
res=df['Second year:   Sem 1'].quantile([lower_bound,upper_bound])
res
```

```
0.690    9.59940
0.991    9.88165
Name: Second year:   Sem 1, dtype: float64
```


```
trueindex=(res.loc[lower_bound]<df['Second year:   Sem 1'].values) & (df['Second year:   S
trueindex
```

```
array([ True, False, False, False, False, False,  True, False, False,
       False, False,  True,  True, False, False, False,  True,  True,
       False, False,  True, False,  True,  True, False, False, False,
       False, False, False, False,  True, False, False, False,  True,
        True,  True, False,  True, False, False, False, False,  True,
       False,  True, False, False, False,  True, False, False, False,
       False,  True, False, False,  True, False, False, False, False,
       False])
```

```
falseindex=~trueindex
falseindex
```

```
array([False,  True,  True,  True,  True,  True, False,  True,  True,
        True,  True, False, False,  True,  True,  True, False, False,
        True,  True, False,  True, False, False,  True,  True,  True,
        True,  True,  True,  True, False,  True,  True,  True, False,
       False, False,  True, False,  True,  True,  True,  True, False,
        True, False,  True,  True,  True, False,  True,  True,  True,
        True, False,  True,  True, False,  True,  True,  True,  True,
        True])
```

```
df['Second year:   Sem 1'][trueindex]
```

```
0     9.80
6     9.68
11    9.77
12    9.77
16    9.68
17    9.86
20    9.64
22    9.82
23    9.86
31    9.86
35    9.64
36    9.75
37    9.66
39    9.86
44    9.70
46    9.67
50    9.61
```

```
55    9.73
58    9.63
Name: Second year:   Sem 1, dtype: float64
```

```
mid=np.median(df['Second year:   Sem 1'][trueindex])
mid
```

```
9.73
```

✓ 0s completed at 10:10 PM