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

```
data=pd.read_excel("/content/add.xlsx")
data.head()
```

	Serial No.	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
0	1	337	118	4	4.5	4.5	9.65	1	0.92
1	2	324	107	4	4.0	4.5	8.87	1	0.76
2	3	316	104	3	3.0	3.5	8.00	1	0.72
3	4	322	110	3	3.5	2.5	8.67	1	0.80
4	5	314	103	2	2.0	3.0	8.21	0	0.65

display summary statistics

```
data.describe()
```

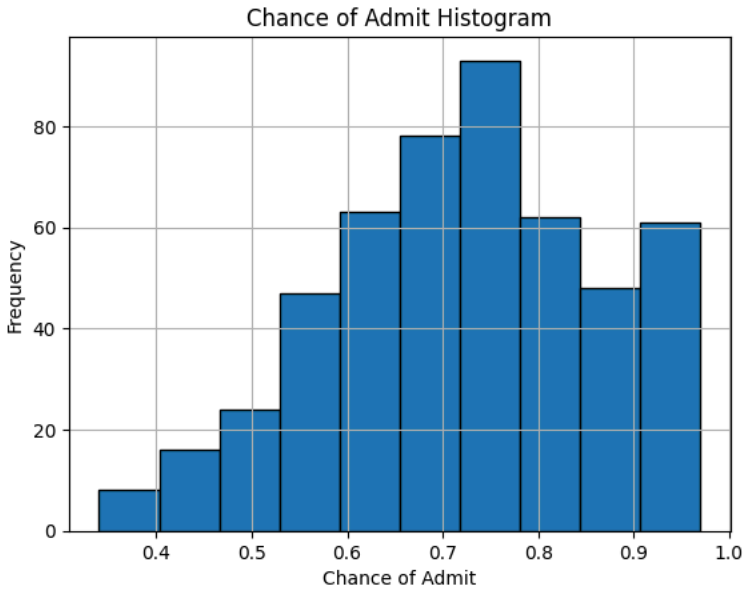
	Serial No.	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
count	500.000000	500.000000	500.000000	500.000000	500.000000	500.00000	500.000000	500.000000	500.00000
mean	250.500000	316.472000	107.192000	3.114000	3.374000	3.48400	8.576440	0.560000	0.72174
std	144.481833	11.295148	6.081868	1.143512	0.991004	0.92545	0.604813	0.496884	0.14114
min	1.000000	290.000000	92.000000	1.000000	1.000000	1.00000	6.800000	0.000000	0.34000
25%	125.750000	308.000000	103.000000	2.000000	2.500000	3.00000	8.127500	0.000000	0.63000
50%	250.500000	317.000000	107.000000	3.000000	3.500000	3.50000	8.560000	1.000000	0.72000
75%	375.250000	325.000000	112.000000	4.000000	4.000000	4.00000	9.040000	1.000000	0.82000
max	500.000000	340.000000	120.000000	5.000000	5.000000	5.00000	9.920000	1.000000	0.97000

```
data.var()
```

```
Serial No.           20875.000000
GRE Score              127.580377
TOEFL Score             36.989114
University Rating        1.307619
SOP                      0.982088
LOR                      0.856457
CGPA                     0.365799
Research                 0.246894
Chance of Admit          0.019921
dtype: float64
```

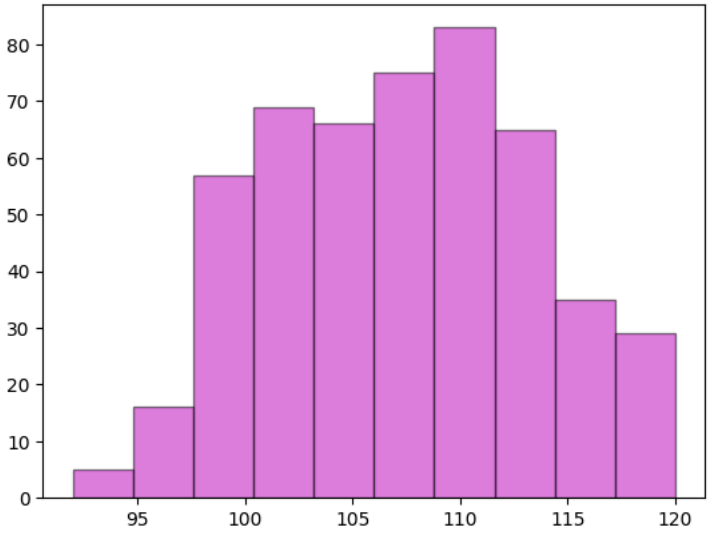
histograms of a each feature

```
chance_of_admit = data['Chance of Admit']
plt.hist(chance_of_admit, edgecolor='black')
plt.title('Chance of Admit Histogram')
plt.xlabel('Chance of Admit')
plt.ylabel('Frequency')
plt.grid(True)
plt.show()
```

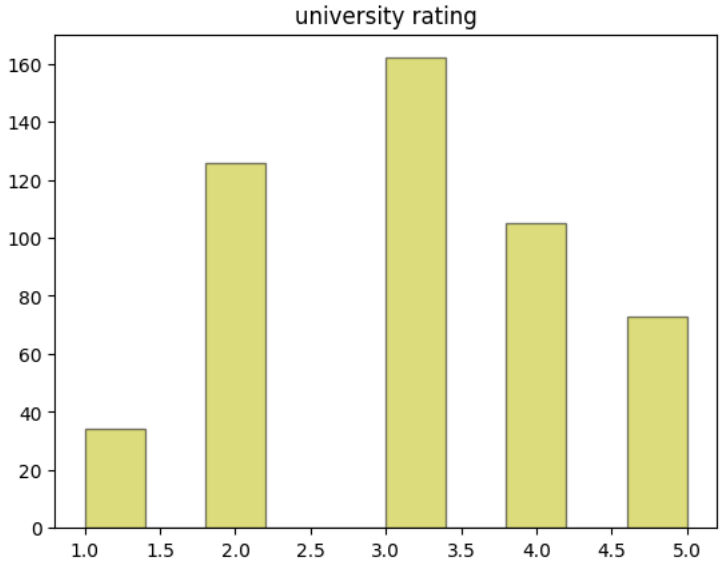



```
gre_scores = data['GRE Score']
plt.hist(gre_scores,edgecolor="black",color="red",alpha=0.7)
plt.title("GRE score")
plt.xlabel("GRE score")
plt.show()
```

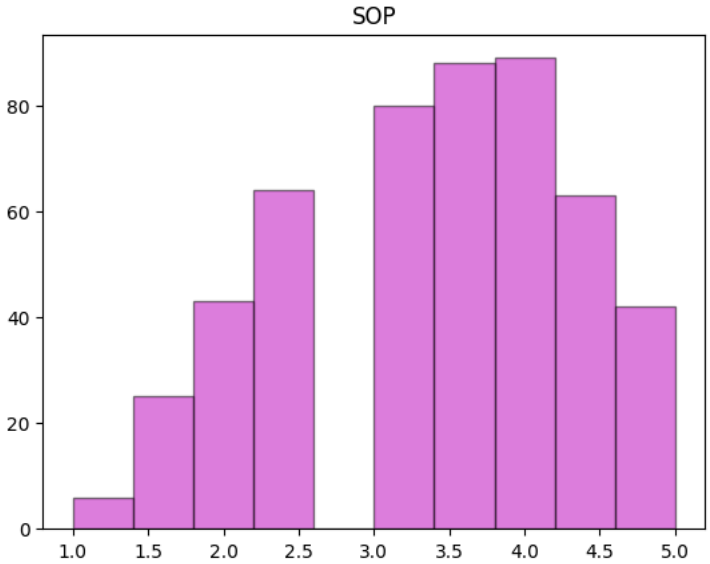

```
ts = data['TOEFL Score']
plt.hist(ts,edgecolor="black",color="m",alpha=0.5)
plt.show()
```




```
ur = data['University Rating']
plt.hist(ur,edgecolor="black",color="y",alpha=0.5)
plt.title("university rating")
plt.show()
```




```
s = data['SOP']
plt.hist(s,edgecolor="black",color="m",alpha=0.5)
plt.title("SOP")
plt.show()
```



```
c=data["CGPA"]
plt.hist(c,edgecolor="r",color='b',linewidth=0.5)
plt.title("cgpa")
plt.show()
```