In [1]:
```
import pandas as pd
import numpy as np
```

In [27]:
```
df = pd.read_csv("D:Courses/Projects/Collage.txt")
df
```

Out[27]:

	Unnamed: 0	Private	Apps	Accept	Enroll	Top10perc	Top25perc	F.Undergrad	P.Undergrad	Outstate	Room.Board	Books	Personal	PhD	Terminal	S.
0	Abilene Christian University	Yes	1660	1232	721	23	52	2885	537	7440	3300	450	2200	70	78	
1	Adelphi University	Yes	2186	1924	512	16	29	2683	1227	12280	6450	750	1500	29	30	
2	Adrian College	Yes	1428	1097	336	22	50	1036	99	11250	3750	400	1165	53	66	
3	Agnes Scott College	Yes	417	349	137	60	89	510	63	12960	5450	450	875	92	97	
4	Alaska Pacific University	Yes	193	146	55	16	44	249	869	7560	4120	800	1500	76	72	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
772	Worcester State College	No	2197	1515	543	4	26	3089	2029	6797	3900	500	1200	60	60	
773	Xavier University	Yes	1959	1805	695	24	47	2849	1107	11520	4960	600	1250	73	75	
774	Xavier University of Louisiana	Yes	2097	1915	695	34	61	2793	166	6900	4200	617	781	67	75	
775	Yale University	Yes	10705	2453	1317	95	99	5217	83	19840	6510	630	2115	96	96	
776	York College of Pennsylvania	Yes	2989	1855	691	28	63	2988	1726	4990	3560	500	1250	75	75	

777 rows × 19 columns

In [6]:
```
df.describe()
```

Out[6]:

	Apps	Accept	Enroll	Top10perc	Top25perc	F.Undergrad	P.Undergrad	Outstate	Room.Board	Books	Personal	
count	777.000000	777.000000	777.000000	777.000000	777.000000	777.000000	777.000000	777.000000	777.000000	777.000000	777.000000	77
mean	3001.638353	2018.804376	779.972973	27.558559	55.796654	3699.907336	855.298584	10440.669241	4357.526384	549.380952	1340.642214	7
std	3870.201484	2451.113971	929.176190	17.640364	19.804778	4850.420531	1522.431887	4023.016484	1096.696416	165.105360	677.071454	1
min	81.000000	72.000000	35.000000	1.000000	9.000000	139.000000	1.000000	2340.000000	1780.000000	96.000000	250.000000	
25%	776.000000	604.000000	242.000000	15.000000	41.000000	992.000000	95.000000	7320.000000	3597.000000	470.000000	850.000000	6
50%	1558.000000	1110.000000	434.000000	23.000000	54.000000	1707.000000	353.000000	9990.000000	4200.000000	500.000000	1200.000000	7
75%	3624.000000	2424.000000	902.000000	35.000000	69.000000	4005.000000	967.000000	12925.000000	5050.000000	600.000000	1700.000000	8
max	48094.000000	26330.000000	6392.000000	96.000000	100.000000	31643.000000	21836.000000	21700.000000	8124.000000	2340.000000	6800.000000	10

In [12]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 777 entries, 0 to 776
Data columns (total 19 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   Unnamed: 0   777 non-null    object
 1   Private      777 non-null    object
 2   Apps         777 non-null    int64
 3   Accept       777 non-null    int64
 4   Enroll       777 non-null    int64
 5   Top10perc    777 non-null    int64
 6   Top25perc    777 non-null    int64
 7   F.Undergrad  777 non-null    int64
 8   P.Undergrad  777 non-null    int64
 9   Outstate     777 non-null    int64
 10  Room.Board   777 non-null    int64
 11  Books        777 non-null    int64
 12  Personal     777 non-null    int64
 13  PhD          777 non-null    int64
 14  Terminal     777 non-null    int64
 15  S.F.Ratio    777 non-null    float64
 16  perc.alumni  777 non-null    int64
 17  Expend       777 non-null    int64
 18  Grad.Rate    777 non-null    int64
dtypes: float64(1), int64(16), object(2)
memory usage: 115.5+ KB
```

In [19]:
```
df.dtypes
```

Out[19]:
```
Unnamed: 0     object
Private        object
Apps            int64
Accept          int64
Enroll          int64
Top10perc       int64
Top25perc       int64
F.Undergrad     int64
P.Undergrad     int64
Room.Board      int64
Books           int64
Personal        int64
PhD             int64
Terminal        int64
S.F.Ratio     float64
perc.alumni     int64
Expend          int64
Grad.Rate       int64
dtype: object
```

In [22]:
```
df.groupby("Accept").Apps.mean()
```

Out[22]:
```
Accept
72          81.0
90         100.0
118        141.0
128        152.0
130        158.5
           ...
13007    20192.0
13243    16587.0
15096    18114.0
18744    21804.0
26330    48094.0
Name: Apps, Length: 693, dtype: float64
```

In [24]:
```
df.groupby("Enroll").Accept.mean()
```

Out[24]:
```
Enroll
35          90.0
46         130.0
51          72.0
55         132.0
63         165.0
           ...
5705     11652.0
5873     13243.0
5874     18744.0
6180     15096.0
6392     10519.0
Name: Accept, Length: 581, dtype: float64
```

In [25]:
```
df.drop(columns='Enroll', inplace=True)
df
```

	Unnamed: 0	Private	Apps	Accept	Top10perc	Top25perc	F.Undergrad	P.Undergrad	Room.Board	Books	Personal	PhD	Terminal	S.F.Ratio	perc.alumni	Expend	Grad.Rate
	Abilene Christian University	Yes	1660	1232	23	52	2885	537	3300	450	2200	70	78	18.1	12	7041	60
	Adelphi University	Yes	2186	1924	16	29	2683	1227	6450	750	1500	29	30	12.2	16	10527	56
	Adrian College	Yes	1428	1097	22	50	1036	99	3750	400	1165	53	66	12.9	30	8735	54
	Agnes Scott College	Yes	417	349	60	89	510	63	5450	450	875	92	97	7.7	37	19016	59
	Alaska Pacific University	Yes	193	146	16	44	249	869	4120	800	1500	76	72	11.9	2	10922	15
	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
	Worcester																

In [29]:
```
print(df.groupby('Enroll').Accept.max())
```

```
Enroll
35        90
46       130
51        72
55       146
63       165
        ...
5705   11652
5873   13243
5874   18744
6180   15096
6392   10519
Name: Accept, Length: 581, dtype: int64
```

In [33]:
```
df.groupby('Unnamed: 0').Accept.agg(['min', 'max'])
```

Out[33]:

	min	max
Unnamed: 0		
Abilene Christian University	1232	1232
Adelphi University	1924	1924
Adrian College	1097	1097
Agnes Scott College	349	349
Alaska Pacific University	146	146
...	...	...
Worcester State College	1515	1515
Xavier University	1805	1805
Xavier University of Louisiana	1915	1915
Yale University	2453	2453
York College of Pennsylvania	1855	1855

777 rows × 2 columns

In [41]:
```
df[["Unnamed: 0","Apps","Accept","Enroll"]].head()
```

Out[41]:

	Unnamed: 0	Apps	Accept	Enroll
0	Abilene Christian University	1660	1232	721
1	Adelphi University	2186	1924	512
2	Adrian College	1428	1097	336
3	Agnes Scott College	417	349	137
4	Alaska Pacific University	193	146	55

In [50]:
```
df[["Unnamed: 0","Apps","Accept","Enroll"]].tail()
```

Out[50]:

	Unnamed: 0	Apps	Accept	Enroll
772	Worcester State College	2197	1515	543
773	Xavier University	1959	1805	695
774	Xavier University of Louisiana	2097	1915	695
775	Yale University	10705	2453	1317
776	York College of Pennsylvania	2989	1855	691

In [36]:
```
print(len(df.columns))
```

```
19
```

In [38]:
```
print(len(df.Apps))
```

```
777
```

In [39]:
```
df["Enroll"]>500
```

Out[39]:
```
0       True
1       True
2      False
3      False
4      False
       ...
772     True
773     True
774     True
775     True
776     True
Name: Enroll, Length: 777, dtype: bool
```

In [49]:
```
df.loc[: , "Apps" : "PhD"]
```

Out[49]:

	Apps	Accept	Enroll	Top10perc	Top25perc	F.Undergrad	P.Undergrad	Outstate	Room.Board	Books	Personal	PhD
0	1660	1232	721	23	52	2885	537	7440	3300	450	2200	70
1	2186	1924	512	16	29	2683	1227	12280	6450	750	1500	29
2	1428	1097	336	22	50	1036	99	11250	3750	400	1165	53
3	417	349	137	60	89	510	63	12960	5450	450	875	92
4	193	146	55	16	44	249	869	7560	4120	800	1500	76
...	...	...	...	...	...	...	...	...	...	...	...	...
772	2197	1515	543	4	26	3089	2029	6797	3900	500	1200	60
773	1959	1805	695	24	47	2849	1107	11520	4960	600	1250	73
774	2097	1915	695	34	61	2793	166	6900	4200	617	781	67
775	10705	2453	1317	95	99	5217	83	19840	6510	630	2115	96
776	2989	1855	691	28	63	2988	1726	4990	3560	500	1250	75

777 rows × 12 columns

In [51]:
```
df.select_dtypes("number")*10
```

Out[51]:

	Apps	Accept	Enroll	Top10perc	Top25perc	F.Undergrad	P.Undergrad	Outstate	Room.Board	Books	Personal	PhD	Terminal	S.F.Ratio	perc.alumni
0	16600	12320	7210	230	520	28850	5370	74400	33000	4500	22000	700	780	181.0	120
1	21860	19240	5120	160	290	26830	12270	122800	64500	7500	15000	290	300	122.0	160
2	14280	10970	3360	220	500	10360	990	112500	37500	4000	11650	530	660	129.0	300
3	4170	3490	1370	600	890	5100	630	129600	54500	4500	8750	920	970	77.0	370
4	1930	1460	550	160	440	2490	8690	75600	41200	8000	15000	760	720	119.0	20
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
772	21970	15150	5430	40	260	30890	20290	67970	39000	5000	12000	600	600	210.0	140
773	19590	18050	6950	240	470	28490	11070	115200	49600	6000	12500	730	750	133.0	310
774	20970	19150	6950	340	610	27930	1660	69000	42000	6170	7810	670	750	144.0	200
775	107050	24530	13170	950	990	52170	830	198400	65100	6300	21150	960	960	58.0	490
776	29890	18550	6910	280	630	29880	17260	49900	35600	5000	12500	750	750	181.0	280

777 rows × 17 columns