

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import scipy.cluster.hierarchy as sch
from sklearn.cluster import AgglomerativeClustering

```

```

data = pd.read_csv('Downloads/Universities.csv')
data

```

	Univ	SAT	Top10	Accept	SFRatio	Expenses	GradRate
0	Brown	1310	89	22	13	22704	94
1	CalTech	1415	100	25	6	63575	81
2	CMU	1260	62	59	9	25026	72
3	Columbia	1310	76	24	12	31510	88
4	Cornell	1280	83	33	13	21864	90
5	Dartmouth	1340	89	23	10	32162	95
6	Duke	1315	90	30	12	31585	95
7	Georgetown	1255	74	24	12	20126	92
8	Harvard	1400	91	14	11	39525	97
9	JohnsHopkins	1305	75	44	7	58691	87
10	MIT	1380	94	30	10	34870	91
11	Northwestern	1260	85	39	11	28052	89
12	NotreDame	1255	81	42	13	15122	94
13	PennState	1081	38	54	18	10185	80
14	Princeton	1375	91	14	8	30220	95
15	Purdue	1005	28	90	19	9066	69
16	Stanford	1360	90	20	12	36450	93

```

def norm_func(i):
    x=(i-i.min())/(i.max()-i.min())
    return(x)

```

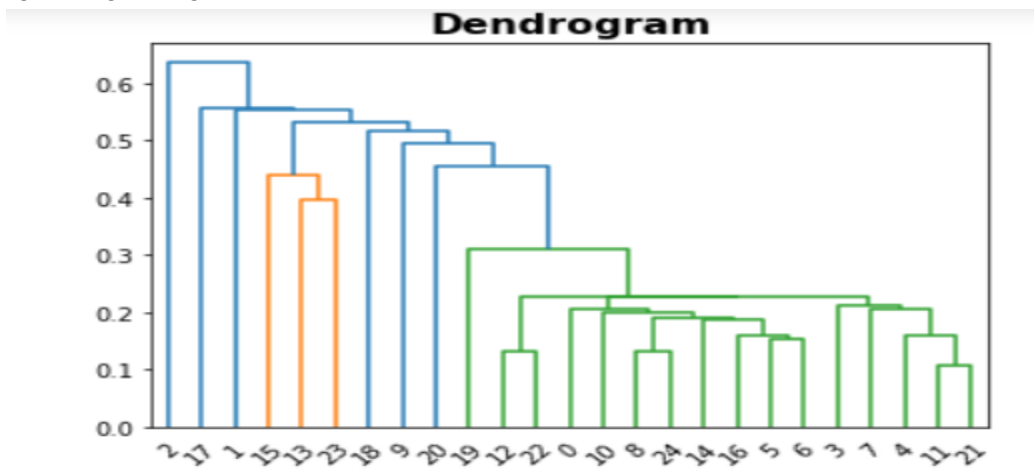
```

divide = norm_func(data.iloc[:,1:])
divide

```

	SAT	Top10	Accept	SFRatio	Expenses	GradRate
0	0.743902	0.847222	0.105263	0.368421	0.255144	0.900000
1	1.000000	1.000000	0.144737	0.000000	1.000000	0.466667
2	0.621951	0.472222	0.592105	0.157895	0.297461	0.166667
3	0.743902	0.666667	0.131579	0.315789	0.415629	0.700000
4	0.670732	0.763889	0.250000	0.368421	0.239835	0.766667
5	0.817073	0.847222	0.118421	0.210526	0.427512	0.933333
6	0.756098	0.861111	0.210526	0.315789	0.416996	0.933333
7	0.609756	0.638889	0.131579	0.315789	0.208161	0.833333
8	0.963415	0.875000	0.000000	0.263158	0.561699	1.000000
9	0.731707	0.652778	0.394737	0.052632	0.910991	0.666667
10	0.914634	0.916667	0.210526	0.210526	0.476864	0.800000
11	0.621951	0.791667	0.328947	0.263158	0.352609	0.733333
12	0.609756	0.736111	0.368421	0.368421	0.116965	0.900000
13	0.185366	0.138889	0.526316	0.631579	0.026991	0.433333
14	0.902439	0.875000	0.000000	0.105263	0.392120	0.933333
15	0.000000	0.000000	1.000000	0.684211	0.006597	0.066667
16	0.865854	0.861111	0.078947	0.315789	0.505659	0.866667

```
dendrogram = sch.dendrogram(sch.linkage(divide,method='single'))
plt.title('Dendrogram',fontsize=16,fontweight='bold')
fig=plt.figure(figsize=(16,8))
```



<Figure size 1152x576 with 0 Axes>

```
a = AgglomerativeClustering(n_clusters=4,affinity='euclidean',linkage='single')
a
AgglomerativeClustering(linkage='single', n_clusters=4)

predict = a.fit_predict(divide)
clusters = pd.DataFrame(predict,columns=['clusters'])
clusters
```

clusters	
0	0
1	3
2	1
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0