```
import pandas as pd
import pandas as pd
import seaborn as sns
```

```
dataset = pd.read_excel (r'/content/Admission_St.xlsx')
```

```
dataset
```

	Admit	GRE	GPA	RANK
0	0	380	3.61	3
1	1	660	3.67	3
2	1	800	4.00	1
3	1	640	3.19	4
4	0	520	2.93	4
...	...	...	...	...
395	0	620	4.00	2
396	0	560	3.04	3
397	0	460	2.63	2
398	0	700	3.65	2
399	0	600	3.89	3

400 rows × 4 columns

Next steps: View recommended plots

```
X = dataset.iloc[:,1:4]
```

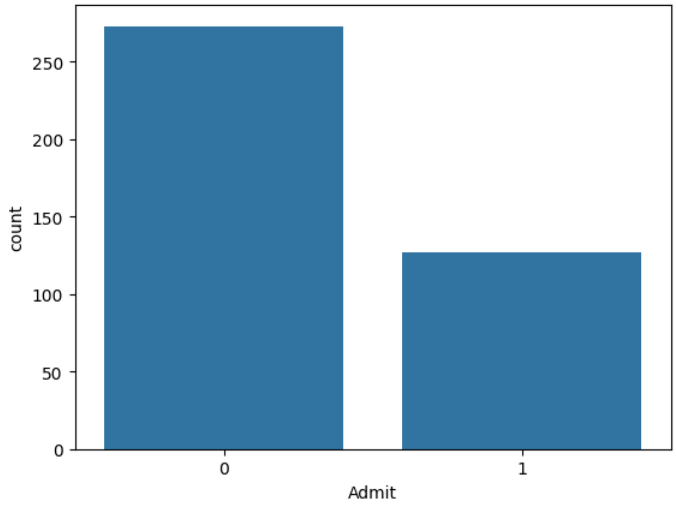
```
Y = dataset.iloc[:,0:1]
```

```
Y.value_counts()
```

```
Admit
0        273
1        127
dtype: int64
```

```
sns.countplot(x="Admit", data=dataset)
```

```
<Axes: xlabel='Admit', ylabel='count'>
```




```
from sklearn.model_selection import train_test_split
X_train,X_test, Y_train, Y_test = train_test_split(X,Y,test_size=0.3,random_state=0)
```

```
len(X_train)
len(Y_train)
```

```
280
```

```
len(X_test)
len(Y_test)
```

```
120
```

```
from imblearn.over_sampling import RandomOverSampler
ros = RandomOverSampler()
X_ros,Y_ros = ros.fit_resample(X_train,Y_train)
```

```
len(Y_ros)
```

```
382
```

```
Y_ros.value_counts()
```

```
Admit
0        191
1        191
dtype: int64
```

```
from imblearn.over_sampling import SMOTE
X_smote,Y_smote=SMOTE(k_neighbors=3).fit_resample(X_train, Y_train)
Y_smote.value_counts()
```

```
Admit
0        191
1        191
dtype: int64
```

Conclusion:

SMOTE algorithm is used as a solution for imbalanced data in the experiment. SMOTE is an intelligent alternative to oversampling: rather than creating duplicates of the minority class, it creates synthetic data points that are relatively similar to the original ones.