ASSIGNMENTDATE	22SEPTEMBER2022
STUDENTNAME	KOMMI CHENCHU KISHORE
STUDENT ROLLNO	110719104024
MAXIMUMMARKS	2 MARKS

ASSIGNMENT:2

DataVisualizationandPre-processing

liver_data_train['HasDisease'].value_counts()

yes   416
no   167
Name:HasDisease,dtype:int64

Plottingsomecolumnstogaininsightsfromthedata

!condainstall-canacondaseaborn-y

First,therelationshipbetweengenderandage

importseaborn**as**sns
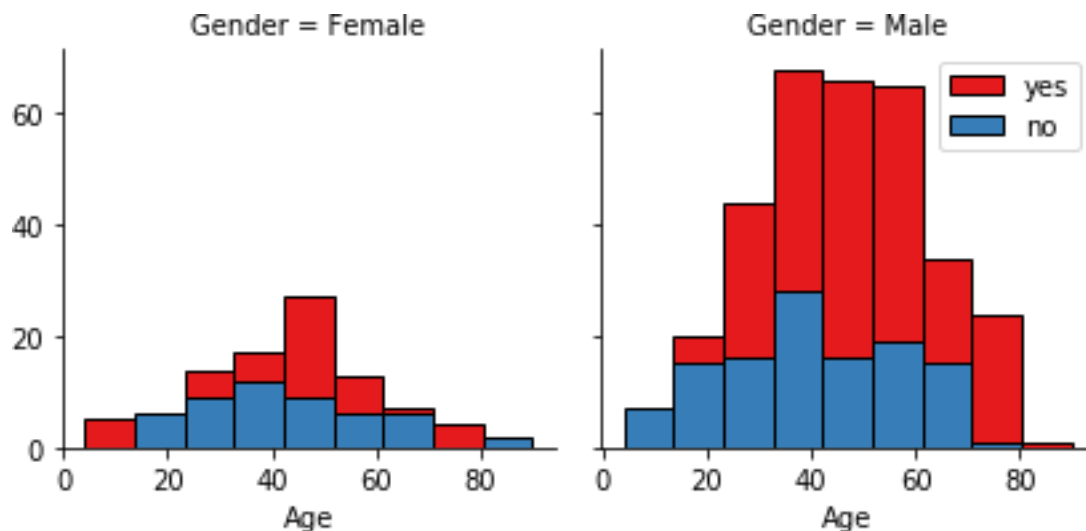bins = np**.**linspace(liver_data_train**.**Age**.**min(),
liver_data_train**.**Age**.**max(),10)grids = sns**.**FacetGrid(liver_data_train,
col="Gender", hue="HasDisease",palette="Set1",col_wrap=2)
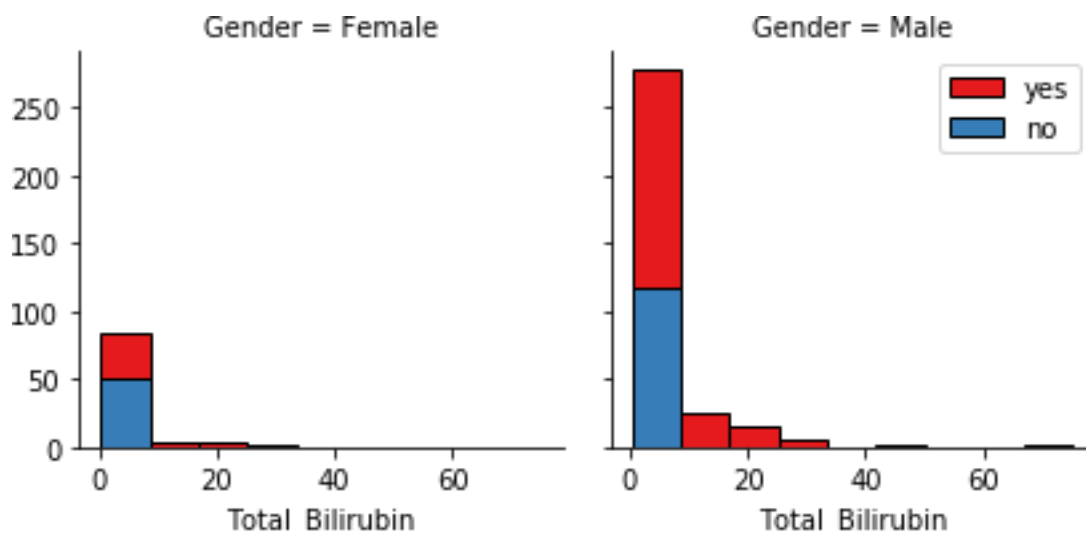grids**.**map(plt**.**hist,'Age',bins=bins,ec="k")

grids**.**axes[**-**
1]**.**legend()plt**.**show()

Amountofbilirubinbasedongender

```
bins =
np.linspace(liver_data_train.Total_Bilirubin.min(),liver_
data_train.Total_Bilirubin.max(),10)
grids = sns.FacetGrid(liver_data_train, col="Gender",
hue="HasDisease",palette="Set1",col_wrap=2)
grids.map(plt.hist,'Total_Bilirubin',bins=bins,ec="k")

grids.axes[-
1].legend()plt.show()
```
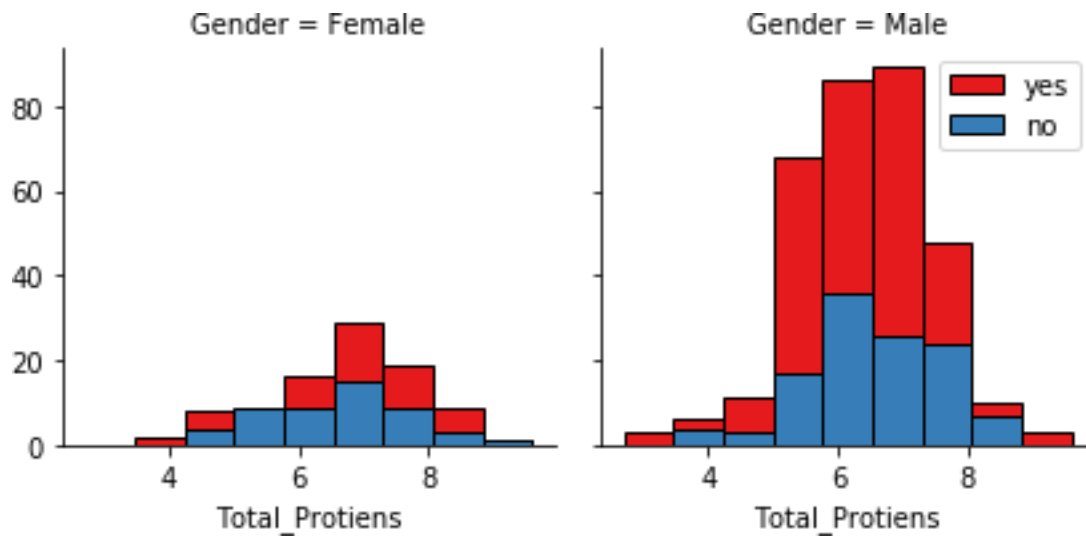

Totalprotiensplayamajorroleinliverdisease.So plottingitgenderwise

```
bins =
np.linspace(liver_data_train.Total_Protiens.min(),liver_
data_train.Total_Protiens.max(),10)
grids = sns.FacetGrid(liver_data_train, col="Gender",
hue="HasDisease",palette="Set1",col_wrap=2)
grids.map(plt.hist,'Total_Protiens',bins=bins,ec="k")
```

grids.axes[-1].legend()plt.show()

convertmales andfemales to0and1

liver_data_train['Gender'].replace(to_replace=['Male','Female'],value=[0,1],inplace=**True**)

Checkifanycolumnhas null valuesandreplacewithmean

liver_data_train.isna().sum()

```
Age                      0
Gender                   0
Total_Bilirubin          0
Direct_Bilirubin         0
Alkaline_Phosphotase     0
Alamine_Aminotransferase     0
Aspartate_Aminotransferase   0
Total_Protiens           0
Albumin                  0
Albumin_and_Globulin_Ratio   4
Dataset                  0
HasDisease               0
dtype:
```
int64liver_data_train['Albumin_and_Globulin_Ratio'].fillna((liver_data_train['Albumin_and_Globulin_Ratio'].mean()),inplace=**True**)