```
pip install pandas
In [ ]:
         import pandas as pd
In [ ]:
         pd.read_csv("https://raw.githubusercontent.com/sitmbadept/sitmbadept.github.io/main/BDTM/R/IPL.csv")
In [ ]:
         pd.read_excel("https://github.com/sitmbadept/sitmbadept.github.io/blob/main/BDTM/R/orders_data.xlsx?raw=true")
In [ ]:
         ipl = pd.read_csv("https://raw.githubusercontent.com/sitmbadept/sitmbadept.github.io/main/BDTM/R/IPL.csv")
In [ ]:
         # View DataFrame Structure
         ipl.info()
In [ ]:
         # View No of Rows & Columns
         ipl.shape
         # 132 rows & 26 columns
In [ ]:
         # Display First & Last few records
         ipl.head() #Default is 5
In [ ]:
         ipl.head(2) #Custom number of records
In [ ]:
         ipl.tail() #Default is 5
In [ ]:
         # Frequency count of unique value present in Columns
         ipl['COUNTRY'].value_counts()
In [ ]:
         # Cross tab for multiple columns
         pd.crosstab(ipl['COUNTRY'], ipl['PLAYING ROLE'])
         ipl=ipl[ipl['AGE'].isnull()==False]
```

```
In [ ]:
         # Converting Data From one data type to another
         ipl['AUCTION YEAR'] = ipl['AUCTION YEAR'].astype('int')
In [ ]:
         ipl[['COUNTRY', 'TEAM', 'PLAYING ROLE', 'AUCTION YEAR', 'BASE PRICE', 'SOLD PRICE']].sort_values(['SOLD PRICE'])
In [ ]:
         ipl[['COUNTRY', 'TEAM', 'PLAYING ROLE', 'AUCTION YEAR', 'BASE PRICE', 'SOLD PRICE']]\
             .sort_values(['SOLD PRICE'], ascending=False)
In [ ]:
         # Creating new columns
         ipl['AUCTION_BY'] = "BCCI" #"IPL board"
In [ ]:
         ## Grouping & Aggregation(Summary)
         ipl['BASE PRICE'].sum() #2.49 Crore
In [ ]:
         # Base Price summary by Team
         ipl.groupby(['TEAM']).agg({"BASE PRICE":"sum"})
In [ ]:
         # Base Price summary by Team
         ipl.groupby(['TEAM']).agg({"BASE PRICE":["sum","mean", "max","min"]})
In [ ]:
         # Base Price summary by Team
         ipl.groupby(['TEAM']).agg({"BASE PRICE":["sum","mean", "max","min"],
                                     "SOLD PRICE":["sum", "mean", "max", "min"]})
In [ ]:
         # Renaming columns
         ipl.rename(columns={'PLAYER NAME':'PLAYER',
                              'AGE': 'AgeGroup',
                              'PLAYING ROLE': 'ROLE'},
                   inplace=True
In [ ]:
         ipl.drop(["ROLE", "PLAYER", "AgeGroup"],
                  inplace=True,
                  axis=1)
```

```
In [ ]:
         # Select Row based on Index
         ipl[0:10]
In [ ]:
         # iloc provide two index Row Index & Column index so we can slice dataframe based on Row & Col index
         ipl.iloc[1,2] #2nd row & 3rd column value...index starts from 0 so 0 is first column & row in DF
In [ ]:
         ipl.iloc[0:10] #0-10 rows & all the columns
         ipl.iloc[0:10, 1:5] #0-10 rows & 1-5 columns
In [ ]:
         ipl.iloc[[0,10,20],[1,10,2,10]]
In [ ]:
         ipl[ipl['COUNTRY']=='SL']
In [ ]:
         ipl[
             (ipl['COUNTRY']=='SL') &
             (ipl['TEAM']=='MI') &
             (ipl['AGE']==2.0)
         ]
In [ ]:
```