

LAB ASSIGNMENT -02

Using pandas and Matplotlib demonstrate the following operations for the sample dataset given, i) Bar chart and Histogram ii) Comparing Distribution iii) Box plot and mention quartiles iv) Correlation using pairplot and heatmap

"""

Commented out IPython magic to ensure Python compatibility.

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline

ipldf = pd.read_csv('IPL IMB381IPL2013.csv')

ipldf.head()

"""BAR CHART"""

sns.barplot(x='AGE',y='SOLD PRICE',data = ipldf)

sns.barplot(x='AGE',y='SOLD PRICE',data = ipldf,hue='PLAYING ROLE')

"""HISTOGRAM"""

plt.hist(ipldf['SOLD PRICE'])

plt.hist(ipldf['SOLD PRICE'],bins = 20)

"""Distribution or Density Chart"""

sns.distplot(ipldf['SOLD PRICE'])

"""Box Plot"""

box = plt.boxplot(ipldf['SOLD PRICE'])

[item.get_ydata()[0] for item in box['caps']]

[item.get_ydata()[0] for item in box['whiskers']]

[item.get_ydata()[0] for item in box['medians']]

ipldf[ipldf['SOLD PRICE']>1350000.0][['PLAYER NAME','PLAYING ROLE','SOLD PRICE']]

"""Comparing Distributions"""

sns.distplot(ipldf[ipldf['CAPTAINCY EXP']==1]['SOLD PRICE'],color='y',label = 'Captaincy Experience')

sns.distplot(ipldf[ipldf['CAPTAINCY EXP']==0]['SOLD PRICE'],color='r',label = 'No Captaincy Experience')

plt.legend()

sns.boxplot(x='PLAYING ROLE',y='SOLD PRICE',data = ipldf)

```
"""Scatter Plot"""
```

```
ipldf_batsman = ipldf[ipldf['PLAYING ROLE']=='Batsman']  
plt.scatter(x=ipldf_batsman.SIXERS,y=ipldf_batsman['SOLD PRICE'])
```

```
sns.regplot(x='SIXERS',y='SOLD PRICE',data =ipldf_batsman)
```

```
"""Pair Plot"""
```

```
infl_features = ['SR-B','AVE','SIXERS','SOLD PRICE']  
sns.pairplot(ipldf[infl_features],size = 2)
```

```
"""Correlation and HeatMap"""
```

```
ipldf[infl_features].corr()
```

```
sns.heatmap(ipldf[infl_features].corr(),annot = True)
```