

# Sports Analytics using Pandas

## Import Libraries

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.graph_objs as go
import plotly.express as px
import warnings
warnings.simplefilter("ignore")
```

## Loading up the data

```
In [2]: data = pd.read_csv("players_stats_by_season_full_details.csv")
data.head()
```

```
Out[2]:
```

	League	Season	Stage	Player	Team	GP	MIN	FGM	FGA	3PM	...	birth_date	he
0	NBA	1999 - 2000	Regular_Season	Shaquille O'Neal	LAL	79	3163.0	956	1665	0	...	Mar 6, 1972	
1	NBA	1999 - 2000	Regular_Season	Vince Carter	TOR	82	3126.0	788	1696	95	...	Jan 26, 1977	
2	NBA	1999 - 2000	Regular_Season	Karl Malone	UTA	82	2947.0	752	1476	2	...	Jul 24, 1963	
3	NBA	1999 - 2000	Regular_Season	Allen Iverson	PHI	70	2853.0	729	1733	89	...	Jun 7, 1975	
4	NBA	1999 - 2000	Regular_Season	Gary Payton	SEA	82	3425.0	747	1666	177	...	Jul 23, 1968	

5 rows × 34 columns



```
In [3]: data.shape
```

```
Out[3]: (53949, 34)
```

## Leagues

```
In [4]: values = data['League'].value_counts().tolist()
names = list(dict(data['League'].value_counts()).keys())

fig = go.Bar(x = names,
             y = values,
             marker = dict(color = 'rgba(238, 103, 48, 0.85)',
                           line=dict(color='rgb(25, 20, 20)',width=1.0)))

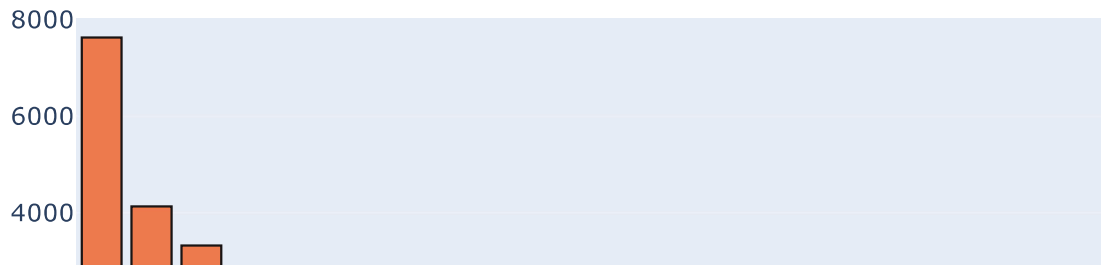
layout = go.Layout()
```

```
fig = go.Figure(data = fig, layout = layout)

fig.update_layout(title_text='Basketball Leagues')

fig.show()
```

## Basketball Leagues



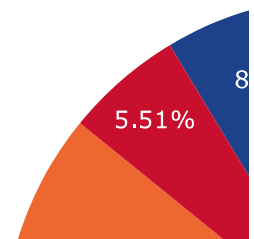
## Match Stages

In [5]:

```
values = data['Stage'].value_counts().tolist()
names = list(dict(data['Stage'].value_counts()).keys())

px.pie(data, values=values, names=names, title="Basketball Match Stages",
        color_discrete_sequence=["#ee6730", "#1d428a", "#c8102e"])
```

## Basketball Match Stages



## Seasons

```
In [6]: values = data['Season'].value_counts().tolist()
names = list(dict(data['Season'].value_counts()).keys())

fig = go.Bar(x = names,
             y = values,
             marker = dict(color = 'rgba(29, 66, 138, 0.75)',
                           line=dict(color='rgb(25, 20, 20)',width=1.25)))

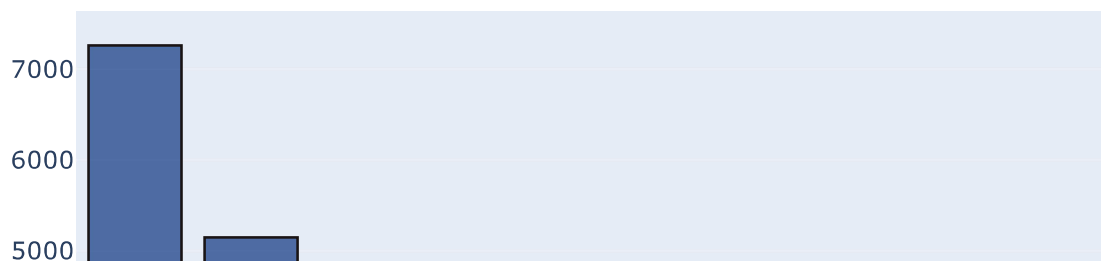
layout = go.Layout()

fig = go.Figure(data = fig, layout = layout)

fig.update_layout(title_text='Basketball Seasons')

fig.show()
```

### Basketball Seasons



## Top Teams

In [7]:

```
values = data['Team'].value_counts().tolist()[:10]
names = list(dict(data['Team'].value_counts()).keys())[:10]

fig = go.Bar(x = names,
             y = values,
             marker = dict(color = 'rgba(85, 37, 130, 0.85)',
                           line=dict(color='rgb(253, 185, 39)',width=1.25)))

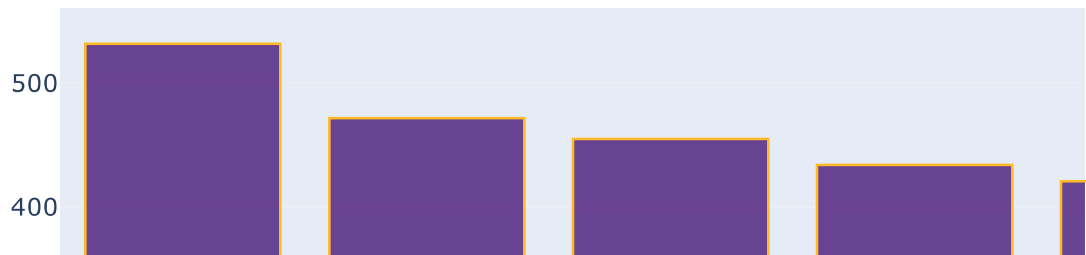
layout = go.Layout()

fig = go.Figure(data = fig, layout = layout)

fig.update_layout(title_text='Top-10 Teams')

fig.show()
```

### Top-10 Teams



## Top Players

In [8]:

```
values = data['Player'].value_counts().tolist()[:10]
```

```

names = list(dict(data['Player'].value_counts()).keys())[:10]

fig = go.Bar(x = names,
             y = values,
             marker = dict(color = 'rgba(206, 17, 65, 0.85)',
                           line=dict(color='rgb(6, 25, 34)',width=1.25)))

layout = go.Layout()

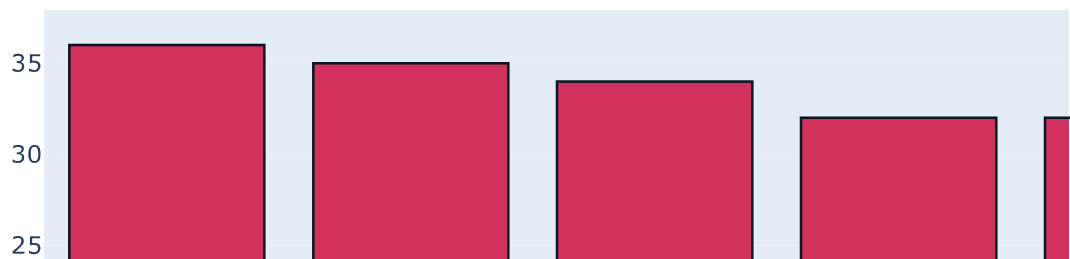
fig = go.Figure(data = fig, layout = layout)

fig.update_layout(title_text='Top-10 Players')

fig.show()

```

## Top-10 Players



In [9]:

```

plt.style.use("seaborn")
fig, ax = plt.subplots(2,1, figsize=(20,15))

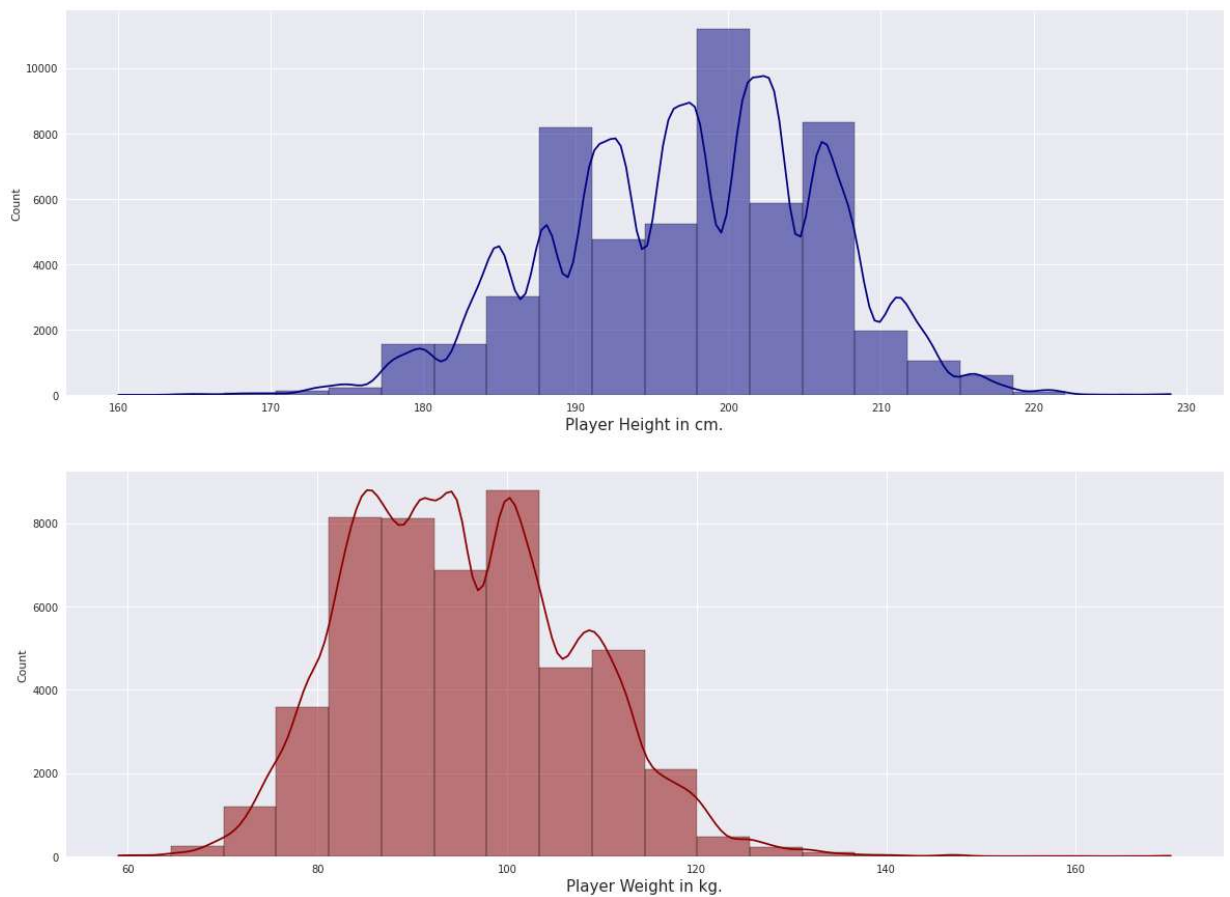
fig.suptitle("Height and Weight Distribution of the Players", fontsize=25, y=0.93)

sns.histplot(x = data["height_cm"], kde=True, ax=ax[0], color="navy", bins=20);
ax[0].set_xlabel("Player Height in cm.", fontsize=15);

sns.histplot(x = data["weight_kg"], kde=True, ax=ax[1], color="darkred", bins=20);
ax[1].set_xlabel("Player Weight in kg.", fontsize=15);

```

## Height and Weight Distribution of the Players



## Player High Schools

In [10]:

```
values = data['high_school'].value_counts().tolist()[0:215]
names = list(dict(data['high_school'].value_counts()).keys())[0:15]

fig = go.Bar(x = names,
             y = values,
             marker = dict(color = 'rgba(0, 101, 58, 0.85)',
                           line=dict(color='rgb(255, 194, 32)',width=1.5)))

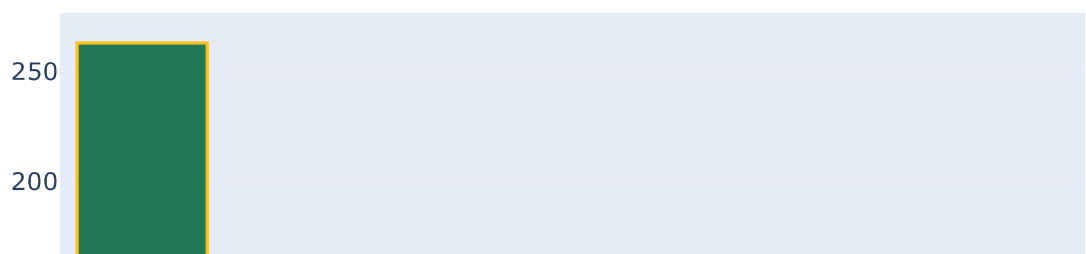
layout = go.Layout()

fig = go.Figure(data = fig, layout = layout)

fig.update_layout(title_text='High Schools of the Players')

fig.show()
```

## High Schools of the Players



## Player Nationality

```
In [11]: values = data['nationality'].value_counts().tolist()[ :20]
names = list(dict(data['nationality'].value_counts()).keys()[ :20])

fig = go.Bar(x = names,
             y = values,
             marker = dict(color = 'rgba(255, 194, 32, 0.85)',
                           line=dict(color='rgb(134, 0, 56)',width=1.75)))

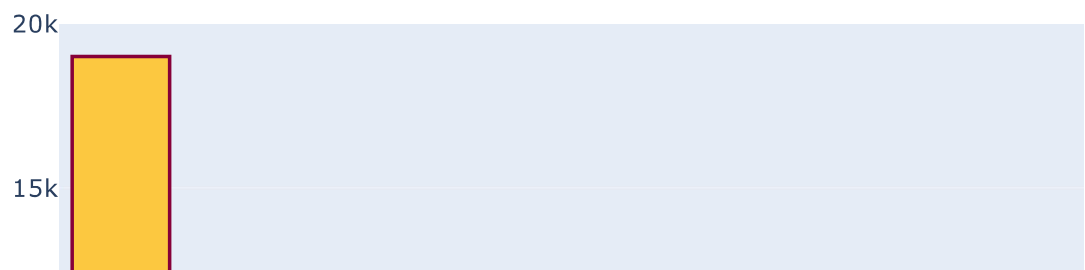
layout = go.Layout()

fig = go.Figure(data = fig, layout = layout)

fig.update_layout(title_text='Nationality of the Players')

fig.show()
```

### Nationality of the Players



In [12]:

```
plt.style.use("seaborn")
fig, ax = plt.subplots(1,2, figsize=(20,7))

fig.suptitle("Field Goals Made vs. Field Goals Attempted", fontsize=25, y=1.0)

sns.histplot(x = data["FGM"], kde=True, ax=ax[0], color="navy", bins=20);
ax[0].set_xlabel("Field Goals Made", fontsize=15);

sns.histplot(x = data["FGA"], kde=True, ax=ax[1], color="darkred", bins=20);
ax[1].set_xlabel("Field Goals Attempt", fontsize=15);
```

Field Goals Made vs. Field Goals Attempted

