WEEKLY PROJECT 19

Baseball Dataset

Importing libraries and dataset

In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
df=pd.read_csv("Baseball_data.csv")
df.head()
```

Out[2]:

	name	handedness	height	weight	avg	HR
0	Tom Brown	R	73	170	0.000	0
1	Denny Lemaster	R	73	182	0.130	4
2	Joe Nolan	L	71	175	0.263	27
3	Denny Doyle	L	69	175	0.250	16
4	Jose Cardenal	R	70	150	0.275	138

1)Top 5 players having height more than average

In [3]:
```
l=len(df['name'])
c=0
for i in range(l):
    c+=df['height'][i]
a=c/l
avg=round(a,1)
r=[]
n=[]
for i in range(l):
    if(df['height'][i]>avg):
        r.append(df['height'][i])
        n.append(df['name'][i])
df_=pd.DataFrame(r,columns=['h'])
df_['na']=n
x=df_.sort_values(by='h',ascending=False).head(5)['h'].values
y=df_.sort_values(by='h',ascending=False).head(5)['na'].values
df1=pd.DataFrame(x,columns=['Height'])
df1['Player_Name']=y
df1.head()
```

Out[3]:

	Height	Player_Name
0	80	Stefan Wever
1	80	J.R. Richard
2	80	Steve Ellsworth
3	80	Mike Smithson
4	79	Walt Bond

2)Players having least weight and left handed

In [10]:
```
l=len(df['name'])
a=[]
b=[]
c=[]
for i in range(l):
    if(df['handedness'][i]=='L'):
        a.append(df['weight'][i])
        b.append(df['name'][i])
        c.append(df['handedness'][i])
df_=pd.DataFrame(b,columns=['na'])
df_['w']=a
df_['hand']=c
x=df_.sort_values(by='w',ascending=True).head(5)['na'].values
y=df_.sort_values(by='w',ascending=True).head(5)['w'].values
z=df_.sort_values(by='w',ascending=True).head(5)['hand'].values
df1=pd.DataFrame(x,columns=['Player_Name'])
df1['Weight']=y
df1['Handedness']=z
df1.head(2)
```

Out[10]:

	Player_Name	Weight	Handedness
0	Curt Ford	150	L
1	Vic Davalillo	150	L

In [ ]:
```
```