

Betül Şeyma Kaya CS240 Project 214993915

What is the most valuable player in the basketball on account of what is worthwhile. Are points or assists important in this matter?

Firstly, I read the excel files basketball\_players and basketball\_awards\_players. I show my columns which I am using.

```
players = pd.read_csv("basketball_players.csv")  
awards = pd.read_csv("basketball_awards_players.csv")
```

I use only most valuable player in basketball\_awards\_players. For this reason, I use loc code and I take only most valuable players.

```
MVP=awards.loc[awards['award']== 'Most Valuable Player']  
#take only award=most valuable players
```

Next, I use describe function. This function calculate automatically standard deviation, mean, max ,min %25,%50%75 values.

```
players.describe()
```

After, I link two values with merge code.

```
players=pd.merge(players, MVP)
```

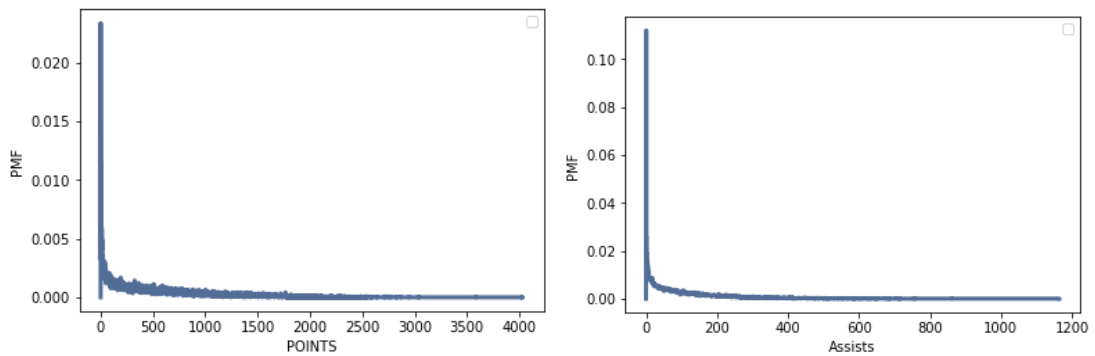
I have defined the values I will use. They are points and assists and I write them. Then, I calculate mean of points and mean of assists. Because I compare the situations of the most valuable players according to general average.

```
mean_points=(points.mean())  
print (mean_points)  
#mean of points  
****
```

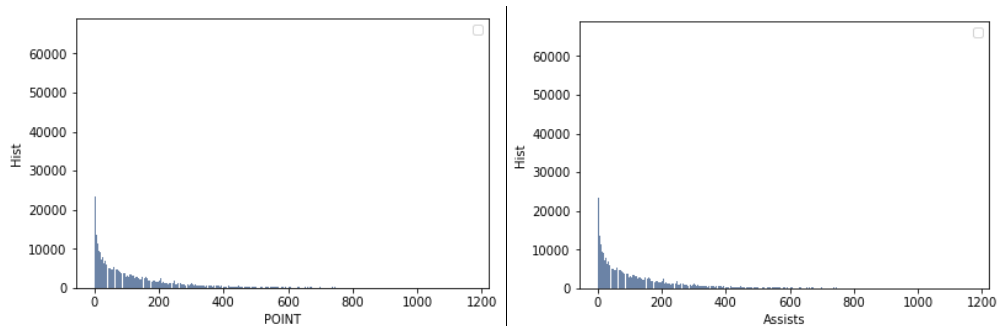
```
mean_assists=(assists.mean())  
print (mean_assists)  
#mean of assists
```

Finally I have drawn the table or graphs. Firstly, I plot pmf of point and pmf of assists.

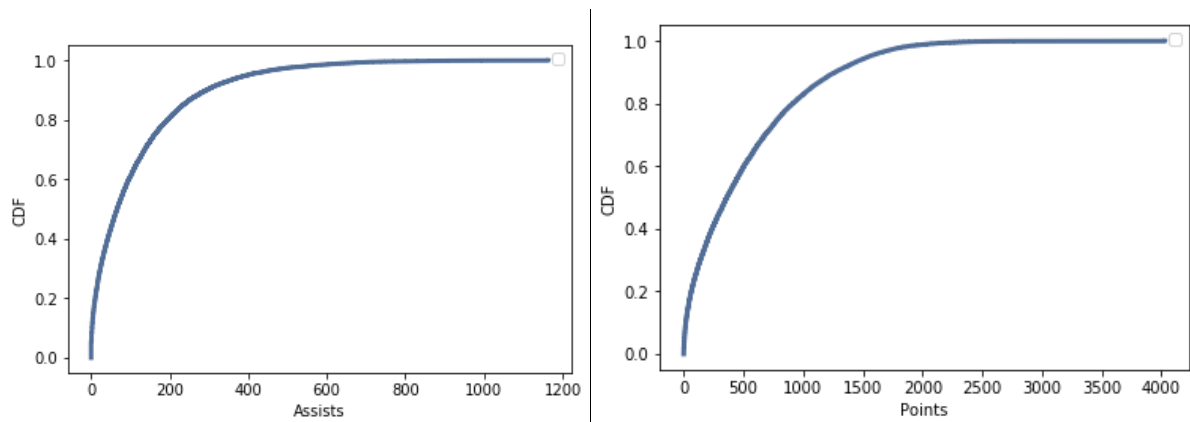
```
point_pmf=thinkstats2.Pmf(players.points)
thinkplot.Pmf(point_pmf)
thinkplot.Config(xlabel='POINTS', ylabel='PMF')
```



Secondly, I plot histogram of plots and histogram of assists.



Finally, I plot cdf of points, cdf of assists.



```
points_cdf=thinkstats2.Cdf(players.points)
```

```
thinkplot.Cdf(points_cdf)
```

```
thinkplot.Config(xlabel='Points', ylabel='CDF')
```

I calculate p value of cdf points and cdf assists

```
value_p=points_cdf[0]
```

```
value_p
```

```
#p value of cdf points
```

```
value_p=assists_cdf[0]
```

```
value_p
```

```
#p value of cdf assists
```

A large p-value ( $> 0.05$ ) indicates weak evidence against the null hypothesis, so you fail to reject the null hypothesis.

A small p-value (typically  $\leq 0.05$ ) indicates strong evidence against the null hypothesis, so you reject the null hypothesis.