

```
In [50]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

from statsmodels.distributions.empirical_distribution import ECDF # Empirical CDF
```

Cricket

```
In [2]: sehwag = pd.read_csv("sehwag.csv")
dravid = pd.read_csv("dravid.csv")
```

```
In [3]: sehwag
```

Out[3]:

	Runs	Mins	BF	4s	6s	SR	Pos	Dismissal	Inns	Unnamed: 9	Opposition	Ground	Start Date	Unnamed: 13
0	1	5	2	0	0	50.00	7	lbw	1	NaN	v Pakistan	Mohali	1 Apr 1999	ODI # 1427
1	19	18	24	0	1	79.16	6	caught	1	NaN	v Zimbabwe	Rajkot	14 Dec 2000	ODI # 1660
2	58	62	54	8	0	107.40	6	bowled	1	NaN	v Australia	Bengaluru	25 Mar 2001	ODI # 1696
3	2	7	7	0	0	28.57	6	caught	2	NaN	v Zimbabwe	Bulawayo	27 Jun 2001	ODI # 1730
4	11	19	16	1	0	68.75	6	not out	2	NaN	v West Indies	Bulawayo	30 Jun 2001	ODI # 1731
...
240	15	21	15	2	0	100.00	2	caught	1	NaN	v Sri Lanka	Hambantota	24 Jul 2012	ODI # 3292
241	3	6	6	0	0	50.00	2	caught	2	NaN	v Sri Lanka	Colombo (RPS)	28 Jul 2012	ODI # 3293
242	34	46	29	6	0	117.24	2	caught	2	NaN	v Sri Lanka	Colombo (RPS)	31 Jul 2012	ODI # 3294
243	4	20	11	1	0	36.36	2	bowled	1	NaN	v Pakistan	Chennai	30 Dec 2012	ODI # 3314
244	31	70	43	3	0	72.09	2	lbw	2	NaN	v Pakistan	Kolkata	3 Jan 2013	ODI # 3315

245 rows × 14 columns

```
In [4]: sehwag.shape
```

Out[4]: (245, 14)

```
In [5]: sehwag["Runs"].describe()
```

Out[5]:

count	245.000000
mean	33.767347
std	34.809419
min	0.000000
25%	8.000000
50%	23.000000
75%	46.000000