

```
In [16]: #Jaypal Mahale  
#743 G
```

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

```
In [17]: df=pd.read_csv('Dataset 6.csv')  
df
```

Out[17]:

	<b>batsman</b>	<b>total_runs</b>	<b>out</b>	<b>numberofballs</b>	<b>average</b>	<b>strikerate</b>
<b>0</b>	V Kohli	5426	152	4111	35.697368	131.987351
<b>1</b>	SK Raina	5386	160	3916	33.662500	137.538304
<b>2</b>	RG Sharma	4902	161	3742	30.447205	130.999466
<b>3</b>	DA Warner	4717	114	3292	41.377193	143.286756
<b>4</b>	S Dhawan	4601	137	3665	33.583942	125.538881
<b>5</b>	CH Gayle	4525	110	2972	41.136364	152.254374
<b>6</b>	MS Dhoni	4450	118	3206	37.711864	138.802246
<b>7</b>	RV Uthappa	4420	156	3381	28.333333	130.730553
<b>8</b>	AB de Villiers	4414	104	2902	42.442308	152.101999
<b>9</b>	G Gambhir	4219	134	3400	31.485075	124.088235
<b>10</b>	AM Rahane	3834	117	3133	32.769231	122.374721
<b>11</b>	KD Karthik	3669	138	2813	26.586957	130.430146
<b>12</b>	SR Watson	3590	115	2566	31.217391	139.906469
<b>13</b>	AT Rayudu	3313	114	2616	29.061404	126.643731
<b>14</b>	YK Pathan	3222	110	2240	29.290909	143.839286
<b>15</b>	BB McCullum	2886	106	2185	27.226415	132.082380
<b>16</b>	PA Patel	2864	127	2352	22.551181	121.768708
<b>17</b>	MK Pandey	2855	95	2352	30.052632	121.386054
<b>18</b>	KA Pollard	2772	102	1879	27.176471	147.525279
<b>19</b>	Yuvraj Singh	2755	109	2121	25.275229	129.891561
<b>20</b>	V Sehwag	2728	98	1746	27.836735	156.242841
<b>21</b>	M Vijay	2589	93	2104	27.838710	123.051331
<b>22</b>	SE Marsh	2489	65	1866	38.292308	133.386924
<b>23</b>	JH Kallis	2427	89	2219	27.269663	109.373592
<b>24</b>	DR Smith	2385	81	1758	29.444444	135.665529
<b>25</b>	SR Tendulkar	2334	71	1943	32.873239	120.123520
<b>26</b>	SV Samson	2215	79	1693	28.037975	130.832841
<b>27</b>	R Dravid	2174	77	1877	28.233766	115.823122
<b>28</b>	AC Gilchrist	2069	76	1492	27.223684	138.672922

In [18]:

```
df.count()
df.describe()
df.head()
```

Out[18]:

	batsman	total_runs	out	numberofballs	average	strikerate
0	V Kohli	5426	152	4111	35.697368	131.987351
1	SK Raina	5386	160	3916	33.662500	137.538304
2	RG Sharma	4902	161	3742	30.447205	130.999466
3	DA Warner	4717	114	3292	41.377193	143.286756
4	S Dhawan	4601	137	3665	33.583942	125.538881

In [19]: `gdf=df.groupby('out')`  
`gdf`

Out[19]: `<pandas.core.groupby.generic.DataFrameGroupBy object at 0x000021F1468C640>`

In [20]: `df.sort_values(by='total_runs',ascending=False)`

Out[20]:

	<b>batsman</b>	<b>total_runs</b>	<b>out</b>	<b>numberofballs</b>	<b>average</b>	<b>strikerate</b>
<b>0</b>	V Kohli	5426	152	4111	35.697368	131.987351
<b>1</b>	SK Raina	5386	160	3916	33.662500	137.538304
<b>2</b>	RG Sharma	4902	161	3742	30.447205	130.999466
<b>3</b>	DA Warner	4717	114	3292	41.377193	143.286756
<b>4</b>	S Dhawan	4601	137	3665	33.583942	125.538881
<b>5</b>	CH Gayle	4525	110	2972	41.136364	152.254374
<b>6</b>	MS Dhoni	4450	118	3206	37.711864	138.802246
<b>7</b>	RV Uthappa	4420	156	3381	28.333333	130.730553
<b>8</b>	AB de Villiers	4414	104	2902	42.442308	152.101999
<b>9</b>	G Gambhir	4219	134	3400	31.485075	124.088235
<b>10</b>	AM Rahane	3834	117	3133	32.769231	122.374721
<b>11</b>	KD Karthik	3669	138	2813	26.586957	130.430146
<b>12</b>	SR Watson	3590	115	2566	31.217391	139.906469
<b>13</b>	AT Rayudu	3313	114	2616	29.061404	126.643731
<b>14</b>	YK Pathan	3222	110	2240	29.290909	143.839286
<b>15</b>	BB McCullum	2886	106	2185	27.226415	132.082380
<b>16</b>	PA Patel	2864	127	2352	22.551181	121.768708
<b>17</b>	MK Pandey	2855	95	2352	30.052632	121.386054
<b>18</b>	KA Pollard	2772	102	1879	27.176471	147.525279
<b>19</b>	Yuvraj Singh	2755	109	2121	25.275229	129.891561
<b>20</b>	V Sehwag	2728	98	1746	27.836735	156.242841
<b>21</b>	M Vijay	2589	93	2104	27.838710	123.051331
<b>22</b>	SE Marsh	2489	65	1866	38.292308	133.386924
<b>23</b>	JH Kallis	2427	89	2219	27.269663	109.373592
<b>24</b>	DR Smith	2385	81	1758	29.444444	135.665529
<b>25</b>	SR Tendulkar	2334	71	1943	32.873239	120.123520
<b>26</b>	SV Samson	2215	79	1693	28.037975	130.832841
<b>27</b>	R Dravid	2174	77	1877	28.233766	115.823122
<b>28</b>	AC Gilchrist	2069	76	1492	27.223684	138.672922

In [21]: `print(df.mean(numeric_only=True))`

```
total_runs      3456.206897
out             110.620690
numberofballs   2604.896552
average         31.177086
strikerate      132.632728
dtype: float64
```

```
In [22]: print(df.median(numeric_only=True))
```

```
total_runs      3222.000000
out             110.000000
numberofballs   2352.000000
average         29.444444
strikerate      130.999466
dtype: float64
```

```
In [23]: df_sample=df.sample(n=3,replace=False,random_state=1)
df_sample
```

```
Out[23]:
```

	<b>batsman</b>	<b>total_runs</b>	<b>out</b>	<b>numberofballs</b>	<b>average</b>	<b>strikerate</b>
<b>14</b>	YK Pathan	3222	110	2240	29.290909	143.839286
<b>21</b>	M Vijay	2589	93	2104	27.838710	123.051331
<b>18</b>	KA Pollard	2772	102	1879	27.176471	147.525279

```
In [25]: df.drop(['strikerate'],axis=1)
```

Out[25]:

	<b>batsman</b>	<b>total_runs</b>	<b>out</b>	<b>numberofballs</b>	<b>average</b>
<b>0</b>	V Kohli	5426	152	4111	35.697368
<b>1</b>	SK Raina	5386	160	3916	33.662500
<b>2</b>	RG Sharma	4902	161	3742	30.447205
<b>3</b>	DA Warner	4717	114	3292	41.377193
<b>4</b>	S Dhawan	4601	137	3665	33.583942
<b>5</b>	CH Gayle	4525	110	2972	41.136364
<b>6</b>	MS Dhoni	4450	118	3206	37.711864
<b>7</b>	RV Uthappa	4420	156	3381	28.333333
<b>8</b>	AB de Villiers	4414	104	2902	42.442308
<b>9</b>	G Gambhir	4219	134	3400	31.485075
<b>10</b>	AM Rahane	3834	117	3133	32.769231
<b>11</b>	KD Karthik	3669	138	2813	26.586957
<b>12</b>	SR Watson	3590	115	2566	31.217391
<b>13</b>	AT Rayudu	3313	114	2616	29.061404
<b>14</b>	YK Pathan	3222	110	2240	29.290909
<b>15</b>	BB McCullum	2886	106	2185	27.226415
<b>16</b>	PA Patel	2864	127	2352	22.551181
<b>17</b>	MK Pandey	2855	95	2352	30.052632
<b>18</b>	KA Pollard	2772	102	1879	27.176471
<b>19</b>	Yuvraj Singh	2755	109	2121	25.275229
<b>20</b>	V Sehwag	2728	98	1746	27.836735
<b>21</b>	M Vijay	2589	93	2104	27.838710
<b>22</b>	SE Marsh	2489	65	1866	38.292308
<b>23</b>	JH Kallis	2427	89	2219	27.269663
<b>24</b>	DR Smith	2385	81	1758	29.444444
<b>25</b>	SR Tendulkar	2334	71	1943	32.873239
<b>26</b>	SV Samson	2215	79	1693	28.037975
<b>27</b>	R Dravid	2174	77	1877	28.233766
<b>28</b>	AC Gilchrist	2069	76	1492	27.223684

In [26]: `df.set_index('average',drop=False)`

Out[26]:

	batsman	total_runs	out	numberofballs	average	strikerate
average						
35.697368	V Kohli	5426	152	4111	35.697368	131.987351
33.662500	SK Raina	5386	160	3916	33.662500	137.538304
30.447205	RG Sharma	4902	161	3742	30.447205	130.999466
41.377193	DA Warner	4717	114	3292	41.377193	143.286756
33.583942	S Dhawan	4601	137	3665	33.583942	125.538881
41.136364	CH Gayle	4525	110	2972	41.136364	152.254374
37.711864	MS Dhoni	4450	118	3206	37.711864	138.802246
28.333333	RV Uthappa	4420	156	3381	28.333333	130.730553
42.442308	AB de Villiers	4414	104	2902	42.442308	152.101999
31.485075	G Gambhir	4219	134	3400	31.485075	124.088235
32.769231	AM Rahane	3834	117	3133	32.769231	122.374721
26.586957	KD Karthik	3669	138	2813	26.586957	130.430146
31.217391	SR Watson	3590	115	2566	31.217391	139.906469
29.061404	AT Rayudu	3313	114	2616	29.061404	126.643731
29.290909	YK Pathan	3222	110	2240	29.290909	143.839286
27.226415	BB McCullum	2886	106	2185	27.226415	132.082380
22.551181	PA Patel	2864	127	2352	22.551181	121.768708
30.052632	MK Pandey	2855	95	2352	30.052632	121.386054
27.176471	KA Pollard	2772	102	1879	27.176471	147.525279
25.275229	Yuvraj Singh	2755	109	2121	25.275229	129.891561
27.836735	V Sehwag	2728	98	1746	27.836735	156.242841
27.838710	M Vijay	2589	93	2104	27.838710	123.051331
38.292308	SE Marsh	2489	65	1866	38.292308	133.386924
27.269663	JH Kallis	2427	89	2219	27.269663	109.373592
29.444444	DR Smith	2385	81	1758	29.444444	135.665529
32.873239	SR Tendulkar	2334	71	1943	32.873239	120.123520
28.037975	SV Samson	2215	79	1693	28.037975	130.832841
28.233766	R Dravid	2174	77	1877	28.233766	115.823122
27.223684	AC Gilchrist	2069	76	1492	27.223684	138.672922

In [27]: `df.sort_index(axis=0,ascending=False)`

Out[27]:

	<b>batsman</b>	<b>total_runs</b>	<b>out</b>	<b>numberofballs</b>	<b>average</b>	<b>strikerate</b>
<b>28</b>	AC Gilchrist	2069	76	1492	27.223684	138.672922
<b>27</b>	R Dravid	2174	77	1877	28.233766	115.823122
<b>26</b>	SV Samson	2215	79	1693	28.037975	130.832841
<b>25</b>	SR Tendulkar	2334	71	1943	32.873239	120.123520
<b>24</b>	DR Smith	2385	81	1758	29.444444	135.665529
<b>23</b>	JH Kallis	2427	89	2219	27.269663	109.373592
<b>22</b>	SE Marsh	2489	65	1866	38.292308	133.386924
<b>21</b>	M Vijay	2589	93	2104	27.838710	123.051331
<b>20</b>	V Sehwag	2728	98	1746	27.836735	156.242841
<b>19</b>	Yuvraj Singh	2755	109	2121	25.275229	129.891561
<b>18</b>	KA Pollard	2772	102	1879	27.176471	147.525279
<b>17</b>	MK Pandey	2855	95	2352	30.052632	121.386054
<b>16</b>	PA Patel	2864	127	2352	22.551181	121.768708
<b>15</b>	BB McCullum	2886	106	2185	27.226415	132.082380
<b>14</b>	YK Pathan	3222	110	2240	29.290909	143.839286
<b>13</b>	AT Rayudu	3313	114	2616	29.061404	126.643731
<b>12</b>	SR Watson	3590	115	2566	31.217391	139.906469
<b>11</b>	KD Karthik	3669	138	2813	26.586957	130.430146
<b>10</b>	AM Rahane	3834	117	3133	32.769231	122.374721
<b>9</b>	G Gambhir	4219	134	3400	31.485075	124.088235
<b>8</b>	AB de Villiers	4414	104	2902	42.442308	152.101999
<b>7</b>	RV Uthappa	4420	156	3381	28.333333	130.730553
<b>6</b>	MS Dhoni	4450	118	3206	37.711864	138.802246
<b>5</b>	CH Gayle	4525	110	2972	41.136364	152.254374
<b>4</b>	S Dhawan	4601	137	3665	33.583942	125.538881
<b>3</b>	DA Warner	4717	114	3292	41.377193	143.286756
<b>2</b>	RG Sharma	4902	161	3742	30.447205	130.999466
<b>1</b>	SK Raina	5386	160	3916	33.662500	137.538304
<b>0</b>	V Kohli	5426	152	4111	35.697368	131.987351

In [28]: `df['batsman'].value_counts()`



```
Out[28]: V Kohli          1
          BB McCullum     1
          R Dravid        1
          SV Samson       1
          SR Tendulkar    1
          DR Smith        1
          JH Kallis       1
          SE Marsh        1
          M Vijay         1
          V Sehwag        1
          Yuvraj Singh    1
          KA Pollard      1
          MK Pandey       1
          PA Patel        1
          YK Pathan       1
          SK Raina        1
          AT Rayudu       1
          SR Watson       1
          KD Karthik      1
          AM Rahane       1
          G Gambhir       1
          AB de Villiers  1
          RV Uthappa      1
          MS Dhoni        1
          CH Gayle        1
          S Dhawan        1
          DA Warner       1
          RG Sharma       1
          AC Gilchrist    1
          Name: batsman, dtype: int64
```

```
In [29]: df['out'].var()
```

```
Out[29]: 725.9581280788176
```

```
In [ ]:
```