```
In [ ]:  Name:Chetan Subhash Sapkal
         Roll No:457
         PRN:202201040152
         D3
```

```
In [23]:  import pandas as pd
          df=pd.read_csv('players.csv')
          print(df)
```

```
           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
..             ...         ...  ...            ...         ...          ...
511        ND Doshi           0    1             13    0.000000     0.000000
512         J Denly           0    1              1    0.000000     0.000000
513         S Ladda           0    2              9    0.000000     0.000000
514  V Pratap Singh           0    1              1    0.000000     0.000000
515       S Kaushik           0    1              1    0.000000     0.000000

[516 rows x 6 columns]
```

```
In [7]:  print(df[['batsman','average']])
```

```
           batsman     average
0          V Kohli   35.697368
1         SK Raina   33.662500
2        RG Sharma   30.447205
3         DA Warner   41.377193
4         S Dhawan   33.583942
..             ...         ...
511        ND Doshi    0.000000
512         J Denly    0.000000
513         S Ladda    0.000000
514  V Pratap Singh    0.000000
515       S Kaushik    0.000000

[516 rows x 2 columns]
```

```
In [8]:  print(df.describe())
```

```
          total_runs         out  numberofballs     average   strikerate
count     516.000000  516.000000     516.000000  482.000000   516.000000
mean      430.625969   17.063953     335.645349   16.355074   105.433442
std       882.275431   28.124511     663.593679   11.146222    39.633938
min         0.000000    0.000000       1.000000    0.000000     0.000000
25%        15.000000    2.000000      17.000000    8.000000    84.362069
50%        74.000000    5.500000      68.000000   14.607143   111.651584
75%       340.500000   18.000000     290.000000   23.657609   130.499036
max      5426.000000  161.000000    4111.000000   88.000000   250.000000
```

```
In [21]:  print("Average runs by batsmen are",df['total_runs'].mean())
          print("Average strikerate of batsmen is",df['strikerate'].mean())
```

```
Average runs by batsmen are 430.6259689922481
Average strikerate of batsmen is 105.4334420702907
```

```
In [20]: print("Median of total runs scored is",df['total_runs'].median())
         print("Median of strikerate is",df['strikerate'].median())
         print("Median of average is",df['average'].median())
```

```
Median of total runs scored is 74.0
Median of strikerate is 111.6515837
Median of average is 14.607142855
```

```
In [22]: print("Max runs scored are",df["total_runs"].max())
         print("Max average is",df["average"].max())
```

```
Max runs scored are 5426
Max average is 88.0
```

```
In [25]: print("The players with runs scored higher than 1000 are")
         print(df.loc[df['total_runs']>1000])
```

```
The players with runs scored higher than 1000 are
         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
..           ...         ...  ...            ...        ...          ...
64     HH Pandya        1082   38            689  28.473684   157.039187
65   LMP Simmons        1079   27            851  39.962963   126.792009
66   LRPL Taylor        1017   41            818  24.804878   124.327628
67       M Vohra        1013   42            767  24.119048   132.073012
68  KP Pietersen        1001   28            740  35.750000   135.270270

[69 rows x 6 columns]
```

```
In [24]: print("The players with average higher than 30 are")
         print(df.loc[df['average']>30])
```

The players with average higher than 30 are

	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
5	CH Gayle	4525	110	2972	41.136364	152.254374
6	MS Dhoni	4450	118	3206	37.711864	138.802246
8	AB de Villiers	4414	104	2902	42.442308	152.101999
9	G Gambhir	4219	134	3400	31.485075	124.088235
10	AM Rahane	3834	117	3133	32.769231	122.374721
12	SR Watson	3590	115	2566	31.217391	139.906469
17	MK Pandey	2855	95	2352	30.052632	121.386054
22	SE Marsh	2489	65	1866	38.292308	133.386924
25	SR Tendulkar	2334	71	1943	32.873239	120.123520
29	SPD Smith	2038	59	1570	34.542373	129.808917
30	JP Duminy	2031	49	1633	41.448980	124.372321
31	KL Rahul	1988	49	1428	40.571429	139.215686
32	MEK Hussey	1977	52	1608	38.019231	122.947761
34	F du Plessis	1861	56	1459	33.232143	127.553119
35	DA Miller	1858	54	1327	34.407407	140.015072
38	RR Pant	1767	49	1071	36.061224	164.985994
46	Q de Kock	1478	46	1109	32.130435	133.273219
48	S Badrinath	1441	48	1210	30.020833	119.090909
49	AD Russell	1415	41	752	34.512195	188.164894
51	BJ Hodge	1400	46	1117	30.434783	125.335721
52	JC Buttler	1396	37	919	37.729730	151.904244
55	KS Williamson	1309	34	964	38.500000	135.788382
56	CA Lynn	1288	37	909	34.810811	141.694169
57	SS Tiwary	1276	42	1064	30.380952	119.924812
61	ML Hayden	1107	27	802	41.000000	138.029925
65	LMP Simmons	1079	27	851	39.962963	126.792009
68	KP Pietersen	1001	28	740	35.750000	135.270270
72	A Symonds	974	26	748	37.461538	130.213904
97	HM Amla	577	13	407	44.384615	141.769042
98	V Shankar	563	17	416	33.117647	135.336538
105	S Gill	510	14	377	36.428571	135.278515
106	OA Shah	506	13	386	38.923077	131.088083
108	MP Stoinis	476	15	362	31.733333	131.491713
109	J Bairstow	459	8	284	57.375000	161.619718
138	M Ali	302	10	177	30.200000	170.621469
158	TM Head	205	6	148	34.166667	138.513513
159	PD Collingwood	203	4	156	50.750000	130.128205
166	Bipul Sharma	187	6	121	31.166667	154.545455
170	JJ Roy	183	5	135	36.600000	135.555556
172	AC Voges	181	6	142	30.166667	127.464789
180	MN van Wyk	167	3	132	55.666667	126.515152
183	R Parag	161	4	126	40.250000	127.777778
200	Kamran Akmal	128	4	78	32.000000	164.102564
222	DJ Harris	111	3	101	37.000000	109.900990
243	Iqbal Abdulla	88	1	83	88.000000	106.024096
272	A Flintoff	62	2	53	31.000000	116.981132
314	AN Ahmed	36	1	26	36.000000	138.461538
317	B Sumanth	35	1	37	35.000000	94.594595
322	M Santner	33	1	23	33.000000	143.478261
323	D Salunkhe	33	1	24	33.000000	137.500000

In [26]:
```python
print(df.tail())
```

```
          batsman  total_runs  out  numberofballs  average  strikerate
511       ND Doshi           0    1             13      0.0         0.0
512        J Denly           0    1              1      0.0         0.0
513        S Ladda           0    2              9      0.0         0.0
514  V Pratap Singh          0    1              1      0.0         0.0
515      S Kaushik           0    1              1      0.0         0.0
```

In [27]:
```python
print(df.head())
```

```
     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 [18]:
```python
print(df.groupby('out').count())
```

```
     batsman  total_runs  numberofballs  average  strikerate
out
0         34          34             34        0          34
1         76          76             76       76          76
2         57          57             57       57          57
3         44          44             44       44          44
4         21          21             21       21          21
..       ...         ...            ...      ...         ...
138        1           1              1        1           1
152        1           1              1        1           1
156        1           1              1        1           1
160        1           1              1        1           1
161        1           1              1        1           1

[89 rows x 5 columns]
```

In [45]:
```python
print("Correlation between total runs and strikerate is",df['total_runs'].corr(d
print("Correlation between total runs and average is",df['total_runs'].corr(df['
```

```
Correlation between total runs and strikerate is 0.3080241594515927
Correlation between total runs and average is 0.556466521126907
```

In [43]:
```python
print("Covariance between total runs and strikerate is",df['total_runs'].cov(df[
```

```
Covariance between total runs and strikerate is 10771.004242830573
```

In [48]:
```python
print("Best batsmen are")
print(df.loc[(df['total_runs']>1000) & (df['strikerate']>130) & (df['average']>3
```

```
Best batsmen are
          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
5          CH Gayle       4525   110            2972   41.136364   152.254374
6          MS Dhoni       4450   118            3206   37.711864   138.802246
8     AB de Villiers      4414   104            2902   42.442308   152.101999
12        SR Watson       3590   115            2566   31.217391   139.906469
22         SE Marsh       2489    65            1866   38.292308   133.386924
31         KL Rahul       1988    49            1428   40.571429   139.215686
35         DA Miller      1858    54            1327   34.407407   140.015072
38          RR Pant       1767    49            1071   36.061224   164.985994
46        Q de Kock       1478    46            1109   32.130435   133.273219
49        AD Russell      1415    41             752   34.512195   188.164894
52        JC Buttler      1396    37             919   37.729730   151.904244
55     KS Williamson      1309    34             964   38.500000   135.788382
56          CA Lynn       1288    37             909   34.810811   141.694169
61         ML Hayden      1107    27             802   41.000000   138.029925
68       KP Pietersen     1001    28             740   35.750000   135.270270
```

In [3]:
```python
print('batsmen with min runs are')
print(df.loc[df['total_runs']==0])
```

```
batsmen with min runs are
          batsman  total_runs   out   numberofballs   average   strikerate
500        K Ahmed          0     1               1      0.0          0.0
501        C Nanda          0     1               1      0.0          0.0
502      YA Abdulla         0     0               1      NaN          0.0
503        U Kaul          0     0               1      NaN          0.0
504         P Raj          0     0               1      NaN          0.0
505     S Lamichhane        0     1               1      0.0          0.0
506      RR Bhatkal         0     2               2      0.0          0.0
507       IC Pandey         0     1               2      0.0          0.0
508        L Ablish         0     0               2      NaN          0.0
509     Abdur Razzak        0     0               2      NaN          0.0
510      Sunny Gupta        0     1               1      0.0          0.0
511        ND Doshi         0     1              13      0.0          0.0
512        J Denly         0     1               1      0.0          0.0
513        S Ladda         0     2               9      0.0          0.0
514    V Pratap Singh       0     1               1      0.0          0.0
515      S Kaushik         0     1               1      0.0          0.0
```

In [5]:
```python
print("average runs made by batsmen are",df['average'].mean())
```

average runs made by batsmen are 16.35507443897925

In [12]:
```python
print("batsmen with most strikerate are")
print(df.groupby('strikerate').get_group(df['strikerate'].max()))
```

```
batsmen with most strikerate are
          batsman  total_runs   out   numberofballs   average   strikerate
443      B Stanlake          5     0               2      NaN        250.0
449    KMDN Kulasekara       5     0               2      NaN        250.0
```

In [11]:
```python
print(df.iloc[::-1])
```

	batsman	total_runs	out	numberofballs	average	strikerate
515	S Kaushik	0	1	1	0.000000	0.000000
514	V Pratap Singh	0	1	1	0.000000	0.000000
513	S Ladda	0	2	9	0.000000	0.000000
512	J Denly	0	1	1	0.000000	0.000000
511	ND Doshi	0	1	13	0.000000	0.000000
..	...	...	...	...	...	...
4	S Dhawan	4601	137	3665	33.583942	125.538881
3	DA Warner	4717	114	3292	41.377193	143.286756
2	RG Sharma	4902	161	3742	30.447205	130.999466
1	SK Raina	5386	160	3916	33.662500	137.538304
0	V Kohli	5426	152	4111	35.697368	131.987351

[516 rows x 6 columns]

In [13]:
```python
print('batsmen who played highest no. of balls')
print(df.groupby('numberofballs').get_group(df['numberofballs'].max()))
```

batsmen who played highest no. of balls

	batsman	total_runs	out	numberofballs	average	strikerate
0	V Kohli	5426	152	4111	35.697368	131.987351

In [19]:
```python
print("total number of runs scored by players",df['total_runs'].sum())
```

total number of runs scored by players 222203

In [ ]: