Cricket_Test_match_dataset Analysis by Pandas Library

In [1]:

```
#import library and read the dataset

import pandas as pd
df = pd.read_csv(r'Cricket_Test_match_dataset.csv')
df
```

Out[1]:

	Player	Span	Mat	Inns	NO	Runs	HS	Ave	BF	SR	100	50	0	4s	6s
0	DG Bradman (AUS)	1928-1948	52	80	10	6996	334	99.94	9800.0	58.60	29	13	7	626	6
1	AC Voges (AUS)	2015-2016	20	31	7	1485	269*	61.87	2667.0	55.68	5	4	2	186	5
2	RG Pollock (SA)	1963-1970	23	41	4	2256	274	60.97	1707.0	54.48	7	11	1	246	11
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00	10	5	2	104	1
4	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00	10	5	2	104	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
62	KD Walters (AUS)	1965-1981	74	125	14	5357	250	48.26	8662.0	49.16	15	33	4	525	23
63	GC Smith (ICC/SA)	2002-2014	117	205	13	9265	277	48.25	15525.0	59.67	27	38	11	1165	24
64	WH Ponsford (AUS)	1924-1934	29	48	4	2122	266	48.22	3118.0	44.77	7	6	1	119	0
65	SJ McCabe (AUS)	1930-1938	39	62	5	2748	232	48.21	3217.0	60.02	6	13	4	241	5
66	DR Jardine (ENG)	1928-1934	22	33	6	1296	127	48.00	2110.0	25.59	1	10	2	53	0

67 rows × 15 columns

In [2]:

```
#rename multiple columns in a list

df = df.rename(columns = {'Mat':'Matches', 'NO':'Not_Outs', 'HS':'Highest_Inns_Score', 'BF'
df
```

Out[2]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
0	DG Bradman (AUS)	1928-1948	52	80	10	6996	334	99.94	9800.0	58.6
1	AC Voges (AUS)	2015-2016	20	31	7	1485	269*	61.87	2667.0	55.6
2	RG Pollock (SA)	1963-1970	23	41	4	2256	274	60.97	1707.0	54.4
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.0
4	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.0
...	...	...	...	...	...	...	...	...	...	..
62	KD Walters (AUS)	1965-1981	74	125	14	5357	250	48.26	8662.0	49.1
63	GC Smith (ICC/SA)	2002-2014	117	205	13	9265	277	48.25	15525.0	59.6
64	WH Ponsford (AUS)	1924-1934	29	48	4	2122	266	48.22	3118.0	44.7
65	SJ McCabe (AUS)	1930-1938	39	62	5	2748	232	48.21	3217.0	60.0
66	DR Jardine (ENG)	1928-1934	22	33	6	1296	127	48.00	2110.0	25.5

67 rows × 15 columns

In [3]:

```
df.head()
```

Out[3]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
0	DG Bradman (AUS)	1928-1948	52	80	10	6996	334	99.94	9800.0	58.60
1	AC Voges (AUS)	2015-2016	20	31	7	1485	269*	61.87	2667.0	55.68
2	RG Pollock (SA)	1963-1970	23	41	4	2256	274	60.97	1707.0	54.48
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
4	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00

In [4]:

```
#check null values
df.isnull().any()
```

Out[4]:

```
Player                 False
Span                   False
Matches                False
Inns                   False
Not_Outs               False
Runs                   False
Highest_Inns_Score     False
Ave                    False
Balls_Faced             True
Batting_Strike_Rate     True
100                    False
50                     False
0                      False
4s                     False
6s                     False
dtype: bool
```

In [5]:

```
df[df['Balls_Faced'].isna()==1]
```

Out[5]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
10	ED Weekes (WI)	1948-1958	48	81	5	4455	207	58.61	NaN	NaN
18	CL Walcott (WI)	1948-1960	44	74	7	3798	220	56.68	NaN	NaN
57	Hon.FS Jackson (ENG)	1893-1905	20	33	4	1415	144*	48.79	NaN	NaN

In [6]:

```
df['Balls_Faced']=df['Balls_Faced'].fillna(0)
df['Batting_Strike_Rate']=df['Batting_Strike_Rate'].fillna(0)
df
```

Out[6]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
0	DG Bradman (AUS)	1928-1948	52	80	10	6996	334	99.94	9800.0	58.60
1	AC Voges (AUS)	2015-2016	20	31	7	1485	269*	61.87	2667.0	55.68
2	RG Pollock (SA)	1963-1970	23	41	4	2256	274	60.97	1707.0	54.48
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
4	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
...	...	...	...	...	...	...	...	...	...	...
62	KD Walters (AUS)	1965-1981	74	125	14	5357	250	48.26	8662.0	49.16
63	GC Smith (ICC/SA)	2002-2014	117	205	13	9265	277	48.25	15525.0	59.67
64	WH Ponsford (AUS)	1924-1934	29	48	4	2122	266	48.22	3118.0	44.77
65	SJ McCabe (AUS)	1930-1938	39	62	5	2748	232	48.21	3217.0	60.02
66	DR Jardine (ENG)	1928-1934	22	33	6	1296	127	48.00	2110.0	25.59

67 rows × 15 columns

In [7]:

```
df[df['Player']=='ED Weekes (WI)']
df[df['Player']=='CL Walcott (WI)']
df[df['Player']=='Hon.FS Jackson (ENG)']
```

Out[7]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
57	Hon.FS Jackson (ENG)	1893-1905	20	33	4	1415	144*	48.79	0.0	0.0

In [8]:

```
#drop duplicates

df.duplicated()
```

Out[8]:

```
0     False
1     False
2     False
3     False
4      True
      ...  
62    False
63    False
64    False
65    False
66    False
Length: 67, dtype: bool
```

In [9]:

```
df[df['Player'].duplicated()==1] #find duplicated values
```

Out[9]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
4	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
13	GS Sobers (WI)	1954-1974	93	160	21	8032	365*	57.78	4063.0	53.58
17	JB Hobbs (ENG)	1908-1930	61	102	7	5410	211	56.94	5363.0	46.22
60	V Kohli (IND)	2011-2024	116	197	12	9017	254*	48.74	16150.0	55.83

In [10]:

```
df[df['Player'].isin(['GA Headley (WI)','GS Sobers (WI)','JB Hobbs (ENG)','V Kohli (IND)'])
```

Out[10]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
4	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
12	GS Sobers (WI)	1954-1974	93	160	21	8032	365*	57.78	4063.0	53.58
13	GS Sobers (WI)	1954-1974	93	160	21	8032	365*	57.78	4063.0	53.58
16	JB Hobbs (ENG)	1908-1930	61	102	7	5410	211	56.94	5363.0	46.22
17	JB Hobbs (ENG)	1908-1930	61	102	7	5410	211	56.94	5363.0	46.22
59	V Kohli (IND)	2011-2024	116	197	12	9017	254*	48.74	16150.0	55.83
60	V Kohli (IND)	2011-2024	116	197	12	9017	254*	48.74	16150.0	55.83

```
In [11]: df = df.drop_duplicates() # delete duplicates
```

```
In [12]: df[df['Player'].isin(['GA Headley (WI)','GS Sobers (WI)','JB Hobbs (ENG)','V Kohli (IND)'])
```

Out[12]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
12	GS Sobers (WI)	1954-1974	93	160	21	8032	365*	57.78	4063.0	53.58
16	JB Hobbs (ENG)	1908-1930	61	102	7	5410	211	56.94	5363.0	46.22
59	V Kohli (IND)	2011-2024	116	197	12	9017	254*	48.74	16150.0	55.83

```
In [13]: #split up Span into Start and End date
         df['Span'].str.split()
```

```
Out[13]: 0      [1928-1948]
         1      [2015-2016]
         2      [1963-1970]
         3      [1930-1954]
         5      [1924-1935]
                   ...
         62     [1965-1981]
         63     [2002-2014]
         64     [1924-1934]
         65     [1930-1938]
         66     [1928-1934]
         Name: Span, Length: 63, dtype: object
```

```
In [14]: df['Span'].str.split(pat = '-')
```

```
Out[14]: 0      [1928, 1948]
         1      [2015, 2016]
         2      [1963, 1970]
         3      [1930, 1954]
         5      [1924, 1935]
                    ...
         62     [1965, 1981]
         63     [2002, 2014]
         64     [1924, 1934]
         65     [1930, 1938]
         66     [1928, 1934]
         Name: Span, Length: 63, dtype: object
```

```
In [15]: df['Span'].str.split(pat = '-').str[1]
```

```
Out[15]: 0      1948
         1      2016
         2      1970
         3      1954
         5      1935
                ...
         62     1981
         63     2014
         64     1934
         65     1938
         66     1934
         Name: Span, Length: 63, dtype: object
```

In [16]:
```
df['Rookie_year'] = df['Span'].str.split(pat='-').str[0]
```

```
C:\Users\verma\AppData\Local\Temp\ipykernel_17604\742699055.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_gu
ide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/st
able/user_guide/indexing.html#returning-a-view-versus-a-copy)
  df['Rookie_year'] = df['Span'].str.split(pat='-').str[0]
```

In [17]:
```
df['Final_year'] = df['Span'].str.split(pat='-').str[1]
```

```
C:\Users\verma\AppData\Local\Temp\ipykernel_17604\3985314040.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_gu
ide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/st
able/user_guide/indexing.html#returning-a-view-versus-a-copy)
  df['Final_year'] = df['Span'].str.split(pat='-').str[1]
```

In [18]:
```
df
```

Out[18]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
0	DG Bradman (AUS)	1928-1948	52	80	10	6996	334	99.94	9800.0	58.60
1	AC Voges (AUS)	2015-2016	20	31	7	1485	269*	61.87	2667.0	55.68
2	RG Pollock (SA)	1963-1970	23	41	4	2256	274	60.97	1707.0	54.48
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
5	H Sutcliffe (ENG)	1924-1935	54	84	9	4555	194	60.73	6558.0	34.59
...	...	...	...	...	...	...	...	...	...	..
62	KD Walters (AUS)	1965-1981	74	125	14	5357	250	48.26	8662.0	49.16
63	GC Smith (ICC/SA)	2002-2014	117	205	13	9265	277	48.25	15525.0	59.67
64	WH Ponsford (AUS)	1924-1934	29	48	4	2122	266	48.22	3118.0	44.77
65	SJ McCabe (AUS)	1930-1938	39	62	5	2748	232	48.21	3217.0	60.02
66	DR Jardine (ENG)	1928-1934	22	33	6	1296	127	48.00	2110.0	25.59

63 rows × 17 columns

In [19]:
```
#drop Span column

df.drop(['Span'],axis  = 1)
```

Out[19]:

	Player	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate	100
0	DG Bradman (AUS)	52	80	10	6996	334	99.94	9800.0	58.60	29
1	AC Voges (AUS)	20	31	7	1485	269*	61.87	2667.0	55.68	5
2	RG Pollock (SA)	23	41	4	2256	274	60.97	1707.0	54.48	7
3	GA Headley (WI)	22	40	4	2190	270*	60.83	416.0	56.00	10
5	H Sutcliffe (ENG)	54	84	9	4555	194	60.73	6558.0	34.59	16
...	...	...	...	...	...	...	...	...	...	...
62	KD Walters (AUS)	74	125	14	5357	250	48.26	8662.0	49.16	15
63	GC Smith (ICC/SA)	117	205	13	9265	277	48.25	15525.0	59.67	27
64	WH Ponsford (AUS)	29	48	4	2122	266	48.22	3118.0	44.77	7
65	SJ McCabe (AUS)	39	62	5	2748	232	48.21	3217.0	60.02	6
66	DR Jardine (ENG)	22	33	6	1296	127	48.00	2110.0	25.59	1

63 rows × 16 columns

In [20]:
```
df.head()
```

Out[20]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
0	DG Bradman (AUS)	1928-1948	52	80	10	6996	334	99.94	9800.0	58.60
1	AC Voges (AUS)	2015-2016	20	31	7	1485	269*	61.87	2667.0	55.68
2	RG Pollock (SA)	1963-1970	23	41	4	2256	274	60.97	1707.0	54.48
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
5	H Sutcliffe (ENG)	1924-1935	54	84	9	4555	194	60.73	6558.0	34.59

In [21]:
```
#Question -> Split up the Country from the player
```

In [22]:
```
df['Player'].str.split(pat='(')
```

Out[22]:
```
0      [DG Bradman , AUS)]
1        [AC Voges , AUS)]
2       [RG Pollock , SA)]
3       [GA Headley , WI)]
5     [H Sutcliffe , ENG)]
              ...
62     [KD Walters , AUS)]
63    [GC Smith , ICC/SA)]
64    [WH Ponsford , AUS)]
65      [SJ McCabe , AUS)]
66     [DR Jardine , ENG)]
Name: Player, Length: 63, dtype: object
```

In [23]:
```
df['Country'] = df['Player'].str.split(pat='(').str[1]
```

```
C:\Users\verma\AppData\Local\Temp\ipykernel_17604\1102775048.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)
  df['Country'] = df['Player'].str.split(pat='(').str[1]
```

In [24]:
```
df['Country'] = df['Country'].str.split(pat = ')').str[0]
df['Country']
```

```
C:\Users\verma\AppData\Local\Temp\ipykernel_17604\32286923.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)
  df['Country'] = df['Country'].str.split(pat = ')').str[0]
```

Out[24]:
```
0        AUS
1        AUS
2         SA
3         WI
5        ENG
       ...
62       AUS
63    ICC/SA
64       AUS
65       AUS
66       ENG
Name: Country, Length: 63, dtype: object
```

In [25]:
```
df.head()
```

Out[25]:

	Player	Span	Matches	Inns	Not_Outs	Runs	Highest_Inns_Score	Ave	Balls_Faced	Batting_Strike_Rate
0	DG Bradman (AUS)	1928-1948	52	80	10	6996	334	99.94	9800.0	58.60
1	AC Voges (AUS)	2015-2016	20	31	7	1485	269*	61.87	2667.0	55.68
2	RG Pollock (SA)	1963-1970	23	41	4	2256	274	60.97	1707.0	54.48
3	GA Headley (WI)	1930-1954	22	40	4	2190	270*	60.83	416.0	56.00
5	H Sutcliffe (ENG)	1924-1935	54	84	9	4555	194	60.73	6558.0	34.59

In [26]:
```
#change datatypes
df.dtypes
```

Out[26]:
```
Player                  object
Span                    object
Matches                  int64
Inns                     int64
Not_Outs                 int64
Runs                     int64
Highest_Inns_Score      object
Ave                    float64
Balls_Faced            float64
Batting_Strike_Rate    float64
100                      int64
50                       int64
0                        int64
4s                       int64
6s                       int64
Rookie_year             object
Final_year              object
Country                 object
dtype: object
```

In [27]:
```
df['Highest_Inns_Score'].str.split(pat='*').str[0]
```

Out[27]:
```
0      334
1      269
2      274
3      270
5      194
      ...
62     250
63     277
64     266
65     232
66     127
Name: Highest_Inns_Score, Length: 63, dtype: object
```

```
In [28]: df['Highest_Inns_Score'] = df['Highest_Inns_Score'].str.split(pat='*').str[0]
```

```
C:\Users\verma\AppData\Local\Temp\ipykernel_17604\805221489.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_gu
ide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/st
able/user_guide/indexing.html#returning-a-view-versus-a-copy)
  df['Highest_Inns_Score'] = df['Highest_Inns_Score'].str.split(pat='*').str[0]
```

```
In [29]: df['Highest_Inns_Score'].astype('int')
```

```
Out[29]: 0     334
         1     269
         2     274
         3     270
         5     194
              ...
         62    250
         63    277
         64    266
         65    232
         66    127
         Name: Highest_Inns_Score, Length: 63, dtype: int32
```

```
In [30]: df['Highest_Inns_Score'] = df['Highest_Inns_Score'].astype('int')
```

```
C:\Users\verma\AppData\Local\Temp\ipykernel_17604\2413117690.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_gu
ide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/st
able/user_guide/indexing.html#returning-a-view-versus-a-copy)
  df['Highest_Inns_Score'] = df['Highest_Inns_Score'].astype('int')
```

```
In [31]: df.dtypes
```

```
Out[31]: Player                  object
         Span                    object
         Matches                  int64
         Inns                     int64
         Not_Outs                 int64
         Runs                     int64
         Highest_Inns_Score       int32
         Ave                    float64
         Balls_Faced            float64
         Batting_Strike_Rate    float64
         100                      int64
         50                       int64
         0                        int64
         4s                       int64
         6s                       int64
         Rookie_year             object
         Final_year              object
         Country                 object
         dtype: object
```

```
In [32]: #question check to fix Rookie year and final year
```

```
In [33]: df = df.astype({'Rookie_year':'int','Final_year':'int'})
```

In [34]:
```
df.dtypes
```

Out[34]:
```
Player                  object
Span                    object
Matches                  int64
Inns                     int64
Not_Outs                 int64
Runs                     int64
Highest_Inns_Score       int32
Ave                    float64
Balls_Faced            float64
Batting_Strike_Rate    float64
100                      int64
50                       int64
0                        int64
4s                       int64
6s                       int64
Rookie_year              int32
Final_year               int32
Country                 object
dtype: object
```

In [35]:
```
df.isnull().any()
```

Out[35]:
```
Player                 False
Span                   False
Matches                False
Inns                   False
Not_Outs               False
Runs                   False
Highest_Inns_Score     False
Ave                    False
Balls_Faced            False
Batting_Strike_Rate    False
100                    False
50                     False
0                      False
4s                     False
6s                     False
Rookie_year            False
Final_year             False
Country                False
dtype: bool
```

In [36]:
```
pd.set_option('display.max_rows',None) # show full dataset
df
```

14	Sangakkara (SL)	2000-2015	134	233	17	12400	319	57.40	22882.0
15	SPD Smith (AUS)	2010-2024	109	195	25	9685	239	56.97	18100.0
16	JB Hobbs (ENG)	1908-1930	61	102	7	5410	211	56.94	5363.0
18	CL Walcott (WI)	1948-1960	44	74	7	3798	220	56.68	0.0
19	L Hutton (ENG)	1937-1955	79	138	15	6971	364	56.67	2844.0
20	JH Kallis (ICC/SA)	1995-2013	166	280	40	13289	224	55.37	28903.0
21	GE Tyldesley (ENG)	1921-1929	14	20	2	990	122	55.00	178.0
22	KS Williamson (NZ)	2010-2024	102	180	17	8881	251	54.48	17267.0

In [37]:
```
##question build out Career_length column
#int convert rooki/final year and subtract
```

In [38]:
```
df['career_lentgh']=df['Final_year'] - df['Rookie_year']
df
```

12	GS Sobers (WI)	1954-1974	93	160	21	8032	365	57.78	4063.0
14	KC Sangakkara (SL)	2000-2015	134	233	17	12400	319	57.40	22882.0
15	SPD Smith (AUS)	2010-2024	109	195	25	9685	239	56.97	18100.0
16	JB Hobbs (ENG)	1908-1930	61	102	7	5410	211	56.94	5363.0
18	CL Walcott (WI)	1948-1960	44	74	7	3798	220	56.68	0.0
19	L Hutton (ENG)	1937-1955	79	138	15	6971	364	56.67	2844.0
20	JH Kallis (ICC/SA)	1995-2013	166	280	40	13289	224	55.37	28903.0
21	GE Tyldesley	1921-	14	20	2	990	122	55.00	178.0

In [39]:
```
#Question 1-> Cricketers in this Df what is the average career length
```

In [40]:
```
df['career_lentgh'].mean()
```

Out[40]:
```
12.968253968253968
```

In [41]:
```
#question Avg Batting_strike_Rate for crickters who playes over 10 years
```

In [42]:
```
df[df['career_lentgh']>10]['Batting_Strike_Rate'].mean()
```

Out[42]:
```
47.967727272727274
```

In [43]:
```
#question Find number of crickters who played before 1960
```

In [44]:
```
df[df['Rookie_year']<1960]['Player'].count()
```

Out[44]:
```
23
```

In [45]:
```
#question Max Highest Inns Score by country
```

In [46]:

```
df.groupby('Country')['Highest_Inns_Score'].max().to_frame('Highincountry').reset_index().s
```

Out[46]:

	Country	Highincountry
5	ICC/WI	400
0	AUS	380
10	SL	374
11	WI	365
1	ENG	364
3	ICC/PAK	329
2	ICC/IND	319
8	PAK	313
9	SA	278
4	ICC/SA	277
6	IND	254
7	NZ	251
12	ZIM	232

In [47]:

```
#question  hundreds,Fifties,ducks(0) avg by country
```

In [48]:

```
df.groupby('Country')[['100','50','0']].mean()
```

Out[48]:

	100	50	0
Country			
AUS	20.625000	28.375000	8.562500
ENG	12.846154	21.230769	4.307692
ICC/IND	29.500000	47.500000	12.000000
ICC/PAK	25.000000	46.000000	15.000000
ICC/SA	36.000000	48.000000	13.500000
ICC/WI	34.000000	48.000000	17.000000
IND	24.200000	30.800000	9.000000
NZ	32.000000	35.000000	11.000000
PAK	17.600000	23.800000	7.800000
SA	9.800000	20.200000	3.400000
SL	28.666667	44.000000	12.333333
WI	16.625000	25.625000	7.250000
ZIM	12.000000	27.000000	5.000000