In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

In [2]:
```
df=pd.read_csv("C:\\Users\\AYUSH SONY\\Desktop\\Border Gavaskar Trophy Analysis\\BGTres
```

In [3]:
```
df.head()
```

Out[3]:

	Test Series Year	Test Number	Innings	Venue	Host	Highest Scorer	Team	Runs by highest scorer	best bowler	wickets by best bowler	team total	Wir
0	1996-97	1	1	Delhi	India	Michael Slater	Australia	44	Anil Kumble	4.0	182/10	I
1	1996-97	1	2	Delhi	India	Nayan Mongia	India	152	Paul Reiffel	3.0	361/10	I
2	1996-97	1	3	Delhi	India	Steve Waugh	Australia	67	Anil Kumble	5.0	234/10	I
3	1996-97	1	4	Delhi	India	Mohd. Azharuddin	India	21	Paul Reiffel	2.0	56/3	I
4	1997-98	1	1	Chennai	India	Navjot Sidhu	India	62	Shane Warne	4.0	257/10	I

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

Out[4]:
```
(214, 18)
```

In [5]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 214 entries, 0 to 213
Data columns (total 18 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   Test Series Year        214 non-null    object 
 1   Test Number             214 non-null    int64  
 2   Innings                 214 non-null    int64  
 3   Venue                   214 non-null    object 
 4   Host                    214 non-null    object 
 5   Highest Scorer          214 non-null    object 
 6   Team                    214 non-null    object 
 7   Runs by highest scorer  214 non-null    int64  
 8   best bowler             214 non-null    object 
 9   wickets by best bowler  213 non-null    float64
 10  team total              214 non-null    object 
 11  Winner                  214 non-null    object 
 12  Win Margin              214 non-null    object 
 13  MOTM                    214 non-null    object 
 14  Ind captain             214 non-null    object 
 15  Aus captain             214 non-null    object 
```

```
 16  MOTS                    214 non-null    object
 17  Series Win              214 non-null    object
dtypes: float64(1), int64(3), object(14)
memory usage: 18.5+ KB
```

In [6]:
```
# there are no missing values in the dataset...
```

In [7]:
```
df.nunique()
```

Out[7]:
```
Test Series Year          16
Test Number                4
Innings                    4
Venue                     18
Host                       2
Highest Scorer            66
Team                       2
Runs by highest scorer   128
best bowler               58
wickets by best bowler     9
team total               186
Winner                     3
Win Margin                35
MOTM                      39
Ind captain               10
Aus captain                8
MOTS                      14
Series Win                 3
dtype: int64
```

In [11]:
```
df.duplicated().sum()  # There are not any duplicate values in the dataset..
```

Out[11]:
```
0
```

In [12]:
```
df.isnull().sum()
```

Out[12]:
```
Test Series Year         0
Test Number              0
Innings                  0
Venue                    0
Host                     0
Highest Scorer           0
Team                     0
Runs by highest scorer   0
best bowler              0
wickets by best bowler   1
team total               0
Winner                   0
Win Margin               0
MOTM                     0
Ind captain              0
Aus captain              0
MOTS                     0
Series Win               0
dtype: int64
```

In [13]:
```
df.describe()
```

Out[13]:

	Test Number	Innings	Runs by highest scorer	wickets by best bowler
count	214.000000	214.000000	214.000000	213.000000
mean	2.336449	2.429907	94.429907	3.915493

	Test Number	Innings	Runs by highest scorer	wickets by best bowler
std	1.095822	1.097262	56.356446	1.649012
min	1.000000	1.000000	4.000000	0.000000
25%	1.000000	1.000000	55.000000	3.000000
50%	2.000000	2.000000	78.500000	4.000000
75%	3.000000	3.000000	122.750000	5.000000
max	4.000000	4.000000	329.000000	8.000000

In [23]:
```
df.head(10)
```

Out[23]:

	Test Series Year	Test Number	Innings	Venue	Host	Highest Scorer	Team	Runs by highest scorer	best bowler	wickets by best bowler	team total	V
0	1996-97	1	1	Delhi	India	Michael Slater	Australia	44	Anil Kumble	4.0	182/10	
1	1996-97	1	2	Delhi	India	Nayan Mongia	India	152	Paul Reiffel	3.0	361/10	
2	1996-97	1	3	Delhi	India	Steve Waugh	Australia	67	Anil Kumble	5.0	234/10	
3	1996-97	1	4	Delhi	India	Mohd. Azharuddin	India	21	Paul Reiffel	2.0	56/3	
4	1997-98	1	1	Chennai	India	Navjot Sidhu	India	62	Shane Warne	4.0	257/10	
5	1997-98	1	2	Chennai	India	Ian Healy	Australia	90	Anil Kumble	4.0	328/10	
6	1997-98	1	3	Chennai	India	Sachin Tendulkar	India	155	Gavin Robertson	1.0	418/4	
7	1997-98	1	4	Chennai	India	Shane Warne	Australia	35	Anil Kumble	4.0	168/10	
8	1997-98	2	1	Kolkata	India	Steve Waugh	Australia	80	Anil Kumble	3.0	233/10	
9	1997-98	2	2	Kolkata	India	Mohd. Azharuddin	India	163	Gavin Robertson	2.0	633/5	

In [28]:
```
for i in df.columns:
    print(i)
    print(df[i].unique())
```

```
Test Series Year
['1996-97' '1997-98' '1999-00' '2000-01' '2003-04' '2004-05' '2007-08'
 '2008-09' '2010-11' '2011-12' '2012-13' '2014-15' '2016-17' '2018-19'
 '2020-21' '2022-23']
Test Number
[1 2 3 4]
Innings
```

```
[1 2 3 4]
Venue
['Delhi' 'Chennai' 'Kolkata' 'Bengaluru' 'Adelaide' 'Melbourne' 'Sydney'
 'Mumbai' 'Brisbane' 'Nagpur' 'Perth' 'Mohali' 'Hyderabad' 'Pune' 'Ranchi'
 'Dharamsala' 'Indore' 'Ahmedabad']
Host
['India' 'Australia']
Highest Scorer
['Michael Slater' 'Nayan Mongia' 'Steve Waugh' 'Mohd. Azharuddin'
 'Navjot Sidhu' 'Ian Healy' 'Sachin Tendulkar' 'Shane Warne' 'Mark Taylor'
 'Mark Waugh' 'Greg Blewett' 'Sourav Ganguly' 'Adam Gilchrist'
 'Justin Langer' 'VVS Laxman' 'Matthew Hayden' 'Rahul Dravid'
 'Ricky Ponting' 'Virendra Sehwag' 'Simon Katich' 'Michael Clarke'
 'Parthiv Patel' 'Damien Martyn' 'Mohd. Kaif' 'Andrew Symonds'
 'Michael Hussey' 'Anil Kumble' 'Zaheer Khan' 'Shane Watson'
 'Gautam Gambhir' 'Marcus North' 'Cheteshwar Pujara' 'Ed Cowan' 'MS Dhoni'
 'Virat Kohli' 'David Warner' 'Moises Henriques' 'Mitchell Starc'
 'Shikhar Dhawan' 'Phillip Hughes' 'Peter Siddle' 'Murali Vijay'
 'Steve Smith' 'Chris Rogers' 'Shaun Marsh' 'Virat Kohli ' 'Matt Renshaw'
 'KL Rahul' 'Peter Handscomb' 'Ravindra Jadeja ' 'Glenn Maxwell'
 'Travis Head' 'Marcus Harris' 'Usman Khawaja' 'Ajinkya Rahane'
 'Mayank Agarwal' 'Pat Cummins' 'Tim Paine' 'Joe Burns'
 'Marnus Labuschagne' 'Cameron Green' 'Shubman Gill' 'Rishabh Pant'
 'Shardul Thakur' 'Rohit Sharma' 'Axar Patel']
Team
['Australia' 'India']
Runs by highest scorer
[ 44 152  67  21  62  90 155  35  80 163  45 177 153 102 150  61  88  43
  91 116  55  52 223 167  76 122  65  28 110  59 281 203 126  57  66 121
 144  99 242 233  72 195 257  92  53 241 125 151  46  60  58 104  12 114
  31  69  24 124  73  42 162 154 145  93  79  81 140 146  41  49  78 206
 112  16 109  77  98  56 128 214  68  89  32 329  83 180  75 221 130 224
  19 204 187  34  51  50  82 115 141 133 192 169  54 117 147  71  64 178
 202 111  63 123  70  30 106  22 193   4  74   9  48 131  84  97 108 120
  25 186]
best bowler
['Anil Kumble' 'Paul Reiffel' 'Shane Warne' 'Gavin Robertson' 'Adam Dale'
 'Michael Kasprowicz' 'Venkatesh Prasad' 'Ajit Agarkar' 'Damien Fleming'
 'Javagal Srinath' 'Brett Lee' 'Glenn McGrath' 'Harbhajan Singh'
 'Jason Gillespie' 'Rahul Sanghvi' 'Colin Miller' 'Zaheer Khan'
 'Nathan Bracken' 'Andy Bichel' 'Simon Katich' 'Stuart MacGill'
 'Brad Williams' 'Murali Kartik' 'Michael Clarke' 'Stuart Clark'
 'Mitchell Johnson' 'Andrew Symonds' 'RP Singh' 'Irfan Pathan'
 'Ishant Sharma' 'Amit Mishra' 'Peter Siddle' 'Virendra Sehwag'
 'Jason Krejza' 'Shane Watson' 'Ben Hilfenhaus' 'Pragyan Ojha'
 'Umesh Yadav' 'James Pattinson' 'Ravichandran Ashwin' 'Nathan Lyon'
 'Ravindra Jadeja' 'Glenn Maxwell' 'Karn Sharma' 'Josh Hazlewood'
 'Mohd. Shami' 'Ryan Harris' 'Mitchell Starc' "Steve O'Keefe"
 'Pat Cummins' 'Kuldeep Yadav' 'Jasprit Bumrah' 'Mohd. Siraj'
 'Navdeep Saini' 'T Natarajan' 'Todd Murphy' 'Matthew Kuhnemann'
 'Axar Patel']
wickets by best bowler
[ 4.  3.  5.  2.  1.  6.  0.  7.  8. nan]
team total
['182/10' '361/10' '234/10' '56/3' '257/10' '328/10' '418/4' '168/10'
 '233/10' '633/5' '181/10' '424/10' '400/10' '169/10' '195/2' '441/10'
 '285/10' '239/8' '110/10' '405/10' '238/10' '208/5' '195/10' '150/10'
 '552/5' '261/10' '176/10' '349/10' '219/10' '47/0' '445/10' '171/10'
 '657/7' '212/10' '391/10' '501/10' '264/10' '155/8' '323/10' '409/10'
 '284/3' '73/2' '556/10' '523/10' '196/10' '233/6' '366/10' '558/10'
 '286/10' '97/1' '705/7' '474/10' '211/2' '357/6' '246/10' '228/10'
 '239/10' '235/10' '376/10' '369/10' '19/0' '398/10' '185/10' '329/5'
 '200/10' '104/10' '203/10' '205/10' '93/10' '343/10' '351/7' '161/10'
 '463/10' '532/10' '401/7' '210/10' '330/10' '294/10' '340/10' '526/10'
 '563/10' '269/7' '430/10' '360/10' '228/6' '177/4' '469/10' '268/10'
```

```
 '314/3' '613/7' '577/10' '31/0' '355/10' '295/10' '209/10' '428/10'
 '192/10' '216/9' '478/10' '495/10' '223/10' '207/3' '333/10' '282/10'
 '240/10' '191/10' '659/4' '604/7' '272/10' '167/5' '201/10' '380/10'
 '572/10' '241/10' '50/2' '237/9' '503/10' '131/10' '408/10' '499/10'
 '136/4' '262/10' '164/10' '158/4' '517/7' '444/10' '290/5' '315/10'
 '505/10' '224/10' '130/6' '530/10' '465/10' '318/9' '174/6' '572/7'
 '475/10' '251/6' '252/7' '260/10' '105/10' '107/10' '189/10' '276/10'
 '274/10' '112/10' '451/10' '603/9' '204/6' '300/10' '332/10' '137/10'
 '106/2' '250/10' '307/10' '291/10' '326/10' '283/10' '243/10' '140/10'
 '443/7' '151/10' '106/8' '622/7' '6/0' '244/10' '36/10' '93/2' '70/2'
 '338/10' '312/6' '334/5' '336/10' '329/7' '177/10' '91/10' '263/10'
 '113/10' '118/4' '109/10' '197/10' '163/10' '78/1' '480/10' '571/9'
 '175/2']
Winner
['India' 'Australia' 'Drawn']
Win Margin
['7w' '179r' '219ri' '8w' '285r' '180r' '141ri' '10w' '171r' '2w' '0' '4w'
 '9w' '217r' '342r' '13r' '337r' '122r' '72r' '320r' '172r' '1w' '68ri'
 '37ri' '298r' '135ri' '6w' '48r' '333r' '75r' '31r' '146r' '137r' '3w'
 '132r']
MOTM
['Nayan Mongia' 'Sachin Tendulkar' 'Javagal Srinath' 'Michael Kasprowicz'
 'michael Kasprowicz' 'Steve Waugh' 'steve Waugh' 'Glenn McGrath'
 'Adam Gilchrist' 'VVS Laxman' 'Matthew Hayden' 'Sourav Ganguly'
 'Rahul Dravid' 'Ricky Ponting' 'Michael Clarke' 'Anil Kumble'
 'Damien Martyn' 'Murali Kartik' 'Andrew Symonds' 'Irfan Pathan'
 'Zaheer Khan' 'MS Dhoni' 'Jason Krejza' 'James Pattinson' 'David Warner'
 'Peter Siddle' 'Cheteshwar Pujara' 'Shikhar Dhawan' 'Ravindra Jadeja'
 'Nathan Lyon' 'Steve Smith' 'Ryan Harris' "Steve O'Keefe" 'KL Rahul'
 'Jasprit Bumrah' 'Tim Paine' 'Ajinkya Rahane' 'Rishabh Pant'
 'Virat Kohli']
Ind captain
['Sachin Tendulkar' 'Mohd. Azharuddin' 'mohd. Azharuddin' 'Sourav Ganguly'
 'Rahul Dravid' 'Anil Kumble' 'MS Dhoni' 'Virat Kohli' 'Ajinkya Rahane'
 'Rohit Sharma']
Aus captain
['Mark Taylor' 'Steve Waugh' 'Adam Gilchrist' 'Ricky Ponting'
 'Michael Clarke' 'Steve Smith' 'Tim Paine' 'Pat Cummins']
MOTS
['Nayan Mongia' 'Sachin Tendulkar' 'Harbhajan Singh' 'Rahul Dravid'
 'Damien Martyn' 'Brett Lee' 'Ishant Sharma' 'Michael Clarke'
 'Ravichandran Ashwin' 'Steve Smith' "Ravindra Jadeja'"
 'Cheteshwar Pujara' 'Pat Cummins' 'Ravichandran Ashwin-Ravindra Jadeja']
Series Win
['India' 'Australia' 'Drawn']
```

In [30]:
```
for i in df.columns:
    print(i)
    print(df[i].value_counts().sort_values(ascending=True))
```

```
Test Series Year
1996-97     4
2010-11     8
1999-00    11
1997-98    11
2000-01    12
2022-23    14
2011-12    14
2018-19    15
2007-08    15
2012-13    15
2016-17    15
2003-04    16
2020-21    16
2008-09    16
```

```
2004-05    16
2014-15    16
Name: Test Series Year, dtype: int64
Test Number
4    41
3    53
2    57
1    63
Name: Test Number, dtype: int64
Innings
4    46
3    56
2    56
1    56
Name: Innings, dtype: int64
Venue
Hyderabad      3
Ahmedabad      3
Ranchi         3
Dharamsala     4
Indore         4
Pune           4
Kolkata        7
Mumbai         8
Nagpur        11
Perth         11
Mohali        12
Brisbane      12
Delhi         16
Chennai       16
Bengaluru     20
Sydney        25
Adelaide      27
Melbourne     28
Name: Venue, dtype: int64
Host
Australia    103
India        111
Name: Host, dtype: int64
Highest Scorer
Joe Burns             1
Matt Renshaw          1
Ravindra Jadeja       1
Chris Rogers          1
Mohd. Kaif            1
                     ..
Matthew Hayden       10
Cheteshwar Pujara    11
Steve Smith          12
Virat Kohli          12
Sachin Tendulkar     19
Name: Highest Scorer, Length: 66, dtype: int64
Team
India        106
Australia    108
Name: Team, dtype: int64
Runs by highest scorer
32     1
214    1
206    1
204    1
203    1
      ..
91     4
44     4
```

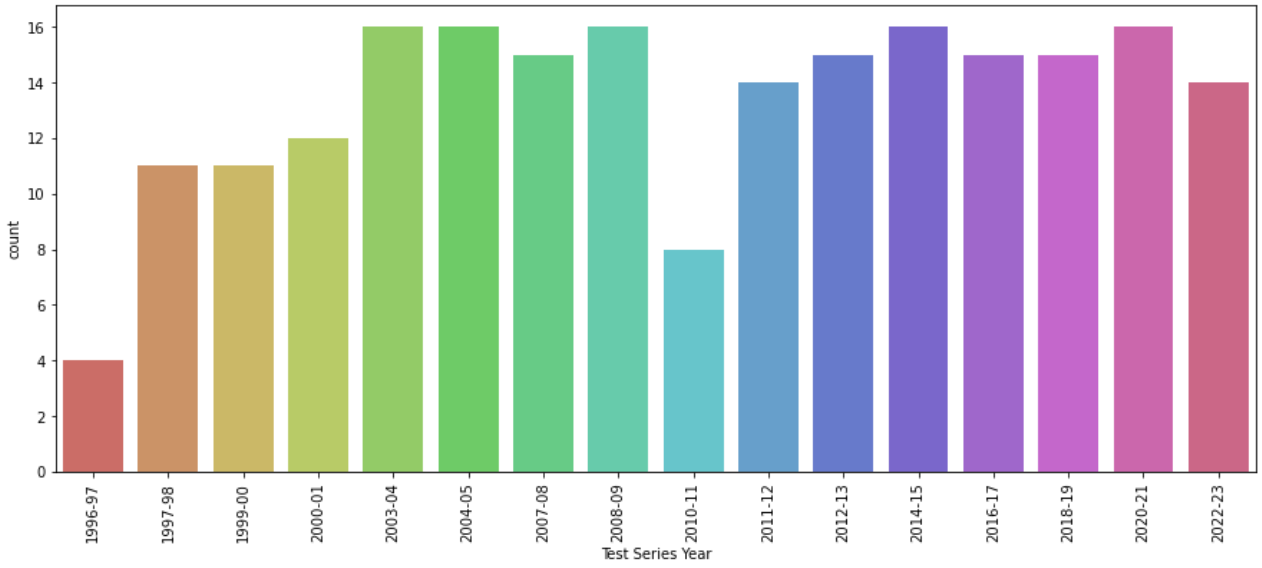
```
55     5
72     6
45     6
Name: Runs by highest scorer, Length: 128, dtype: int64
best bowler
Matthew Kuhnemann       1
Andy Bichel             1
Navdeep Saini           1
Rahul Sanghvi           1
Glenn Maxwell           1
Shane Watson            1
Brad Williams           1
Jason Krejza            1
Michael Clarke          1
Damien Fleming          1
Todd Murphy             1
Pragyan Ojha            1
Colin Miller            1
Murali Kartik           1
Simon Katich            1
Axar Patel              1
Nathan Bracken          1
Adam Dale               1
Irfan Pathan            1
RP Singh                1
Virendra Sehwag         1
Andrew Symonds          1
T Natarajan             1
Venkatesh Prasad        1
Michael Kasprowicz      1
Steve O'Keefe           2
Gavin Robertson         2
Karn Sharma             2
Paul Reiffel            2
Stuart MacGill          2
Mohd. Siraj             2
Kuldeep Yadav           2
Javagal Srinath         2
Ryan Harris             2
Amit Mishra             2
Mitchell Starc          3
Stuart Clark            3
Shane Warne             4
James Pattinson         4
Peter Siddle            4
Mohd. Shami             4
Umesh Yadav             4
Jasprit Bumrah          4
Brett Lee               5
Ajit Agarkar            5
Ishant Sharma           6
Josh Hazlewood          6
Ben Hilfenhaus          6
Glenn McGrath           6
Zaheer Khan             7
Pat Cummins             7
Jason Gillespie         7
Mitchell Johnson       10
Ravindra Jadeja        12
Harbhajan Singh        13
Ravichandran Ashwin    15
Nathan Lyon            16
Anil Kumble            18
Name: best bowler, dtype: int64
wickets by best bowler
```

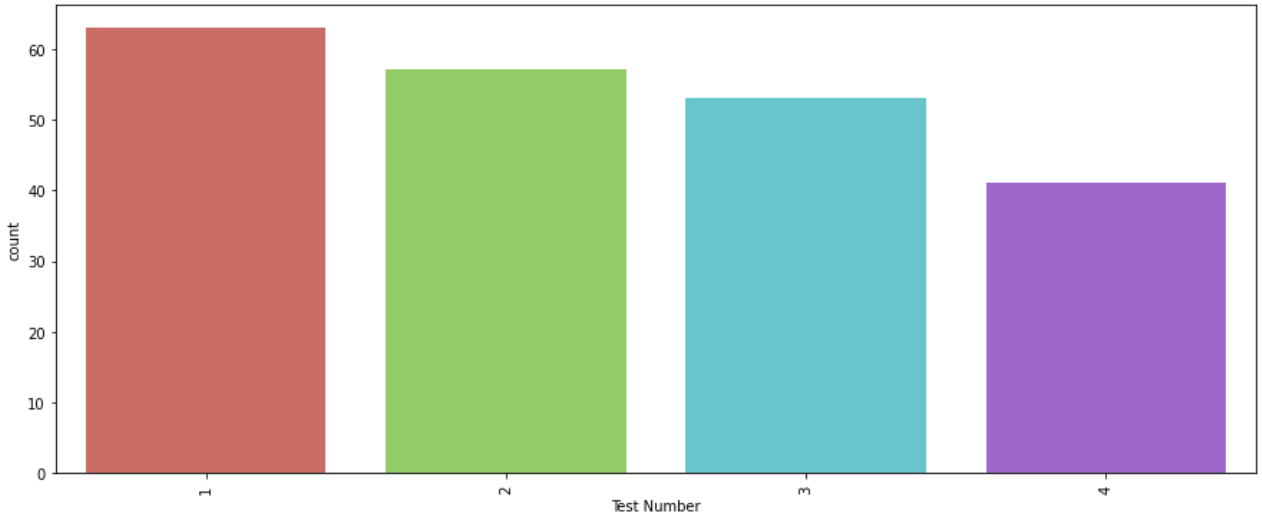
```
0.0     4
8.0     5
7.0     8
1.0    15
6.0    17
2.0    18
3.0    43
5.0    44
4.0    59
Name: wickets by best bowler, dtype: int64
team total
228/6     1
274/10    1
532/10    1
475/10    1
633/5     1
         ..
441/10    2
405/10    2
400/10    3
369/10    3
195/10    3
Name: team total, Length: 186, dtype: int64
Winner
Drawn        44
Australia    77
India        93
Name: Winner, dtype: int64
Win Margin
68ri      3
135ri     3
37ri      3
132r      3
141ri     3
219ri     3
13r       4
2w        4
75r       4
342r      4
3w        4
320r      4
180r      4
1w        4
10w       4
285r      4
172r      4
48r       4
333r      4
179r      4
337r      4
298r      4
217r      4
31r       4
137r      4
72r       4
171r      4
146r      4
4w        8
122r      8
7w        8
9w        8
6w       12
8w       20
0        44
Name: Win Margin, dtype: int64
```

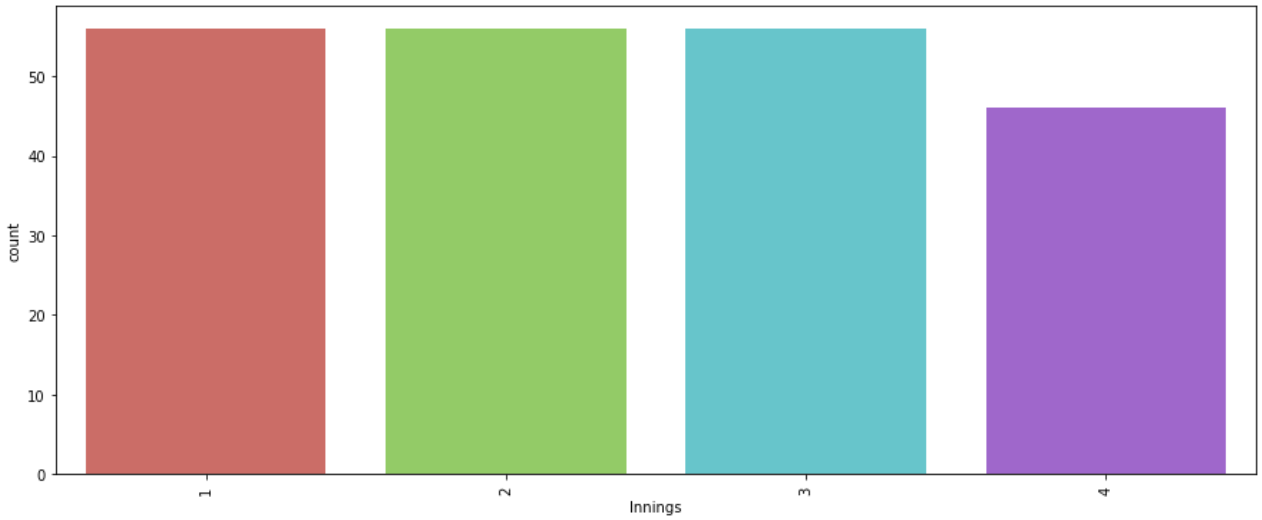
```
MOTM
Steve Waugh            2
michael Kasprowicz     2
Michael Kasprowicz     2
steve Waugh            2
David Warner           3
Virat Kohli            3
Javagal Srinath        3
Glenn McGrath          3
Shikhar Dhawan         4
Ricky Ponting          4
Damien Martyn          4
Jasprit Bumrah         4
Steve O'Keefe          4
Irfan Pathan           4
Murali Kartik          4
James Pattinson        4
Nayan Mongia           4
Rishabh Pant           4
Rahul Dravid           4
Adam Gilchrist         4
Andrew Symonds         4
Ajinkya Rahane         4
Ryan Harris            4
Peter Siddle           4
Jason Krejza           4
Sourav Ganguly         4
Tim Paine              4
Anil Kumble            4
KL Rahul               4
Michael Clarke         7
VVS Laxman             8
MS Dhoni               8
Matthew Hayden         8
Zaheer Khan            8
Steve Smith           12
Nathan Lyon           12
Cheteshwar Pujara     13
Ravindra Jadeja       15
Sachin Tendulkar      19
Name: MOTM, dtype: int64
Ind captain
mohd. Azharuddin     1
Rahul Dravid         8
Mohd. Azharuddin    10
Rohit Sharma        14
Sachin Tendulkar    15
Ajinkya Rahane      16
Anil Kumble         23
Sourav Ganguly      36
Virat Kohli         38
MS Dhoni            53
Name: Ind captain, dtype: int64
Aus captain
Pat Cummins        7
Adam Gilchrist    12
Mark Taylor       15
Tim Paine         31
Michael Clarke    33
Steve Smith       34
Steve Waugh       39
Ricky Ponting     43
Name: Aus captain, dtype: int64
MOTS
Nayan Mongia                          4
```

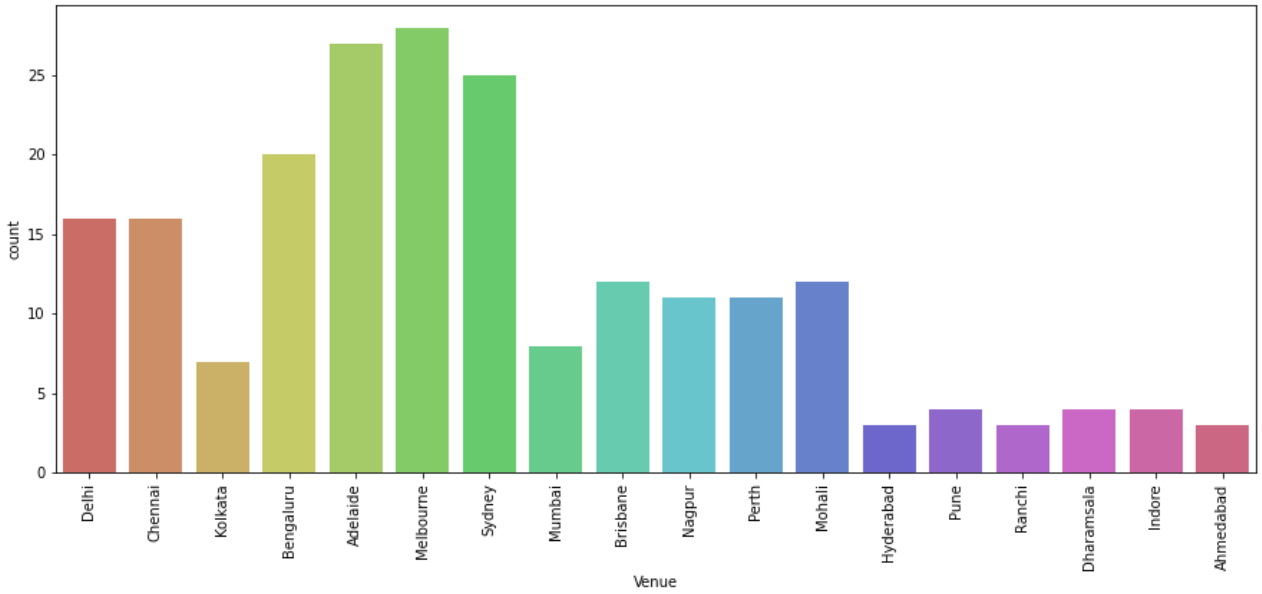
```
Harbhajan Singh                    12
Ravichandran Ashwin-Ravindra Jadeja    14
Michael Clarke                     14
Brett Lee                          15
Ravichandran Ashwin                15
Cheteshwar Pujara                  15
Ravindra Jadeja'                   15
Steve Smith                        16
Pat Cummins                        16
Damien Martyn                      16
Rahul Dravid                       16
Ishant Sharma                      16
Sachin Tendulkar                   30
Name: MOTS, dtype: int64
Series Win
Drawn         16
Australia     72
India        126
Name: Series Win, dtype: int64
```

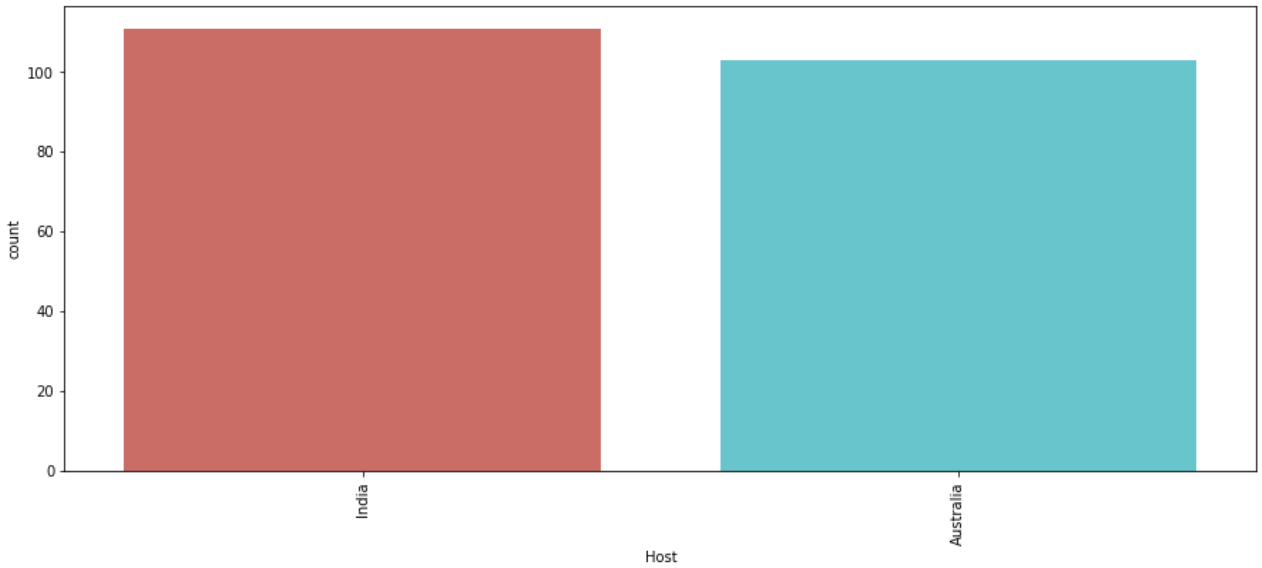
In [36]:
```
for i in df.columns:
    plt.figure(figsize=(15,6))
    sns.countplot(df[i], data=df, palette='hls')
    plt.xticks(rotation=90)
    plt.show()
```

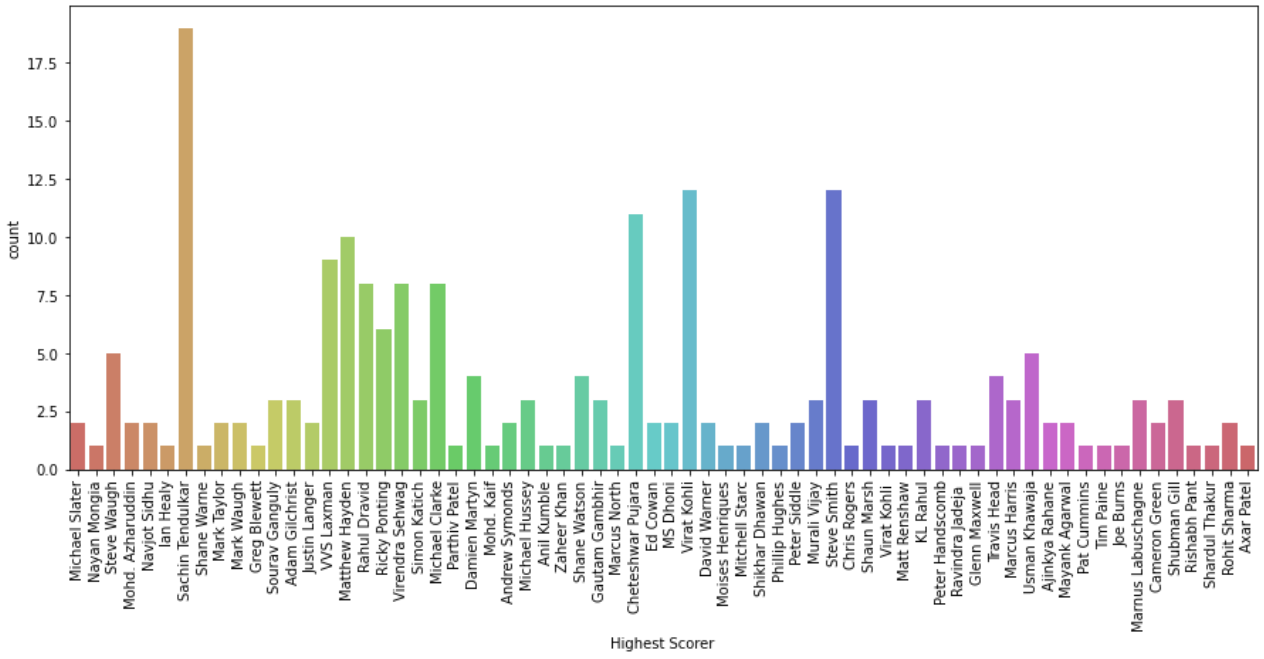


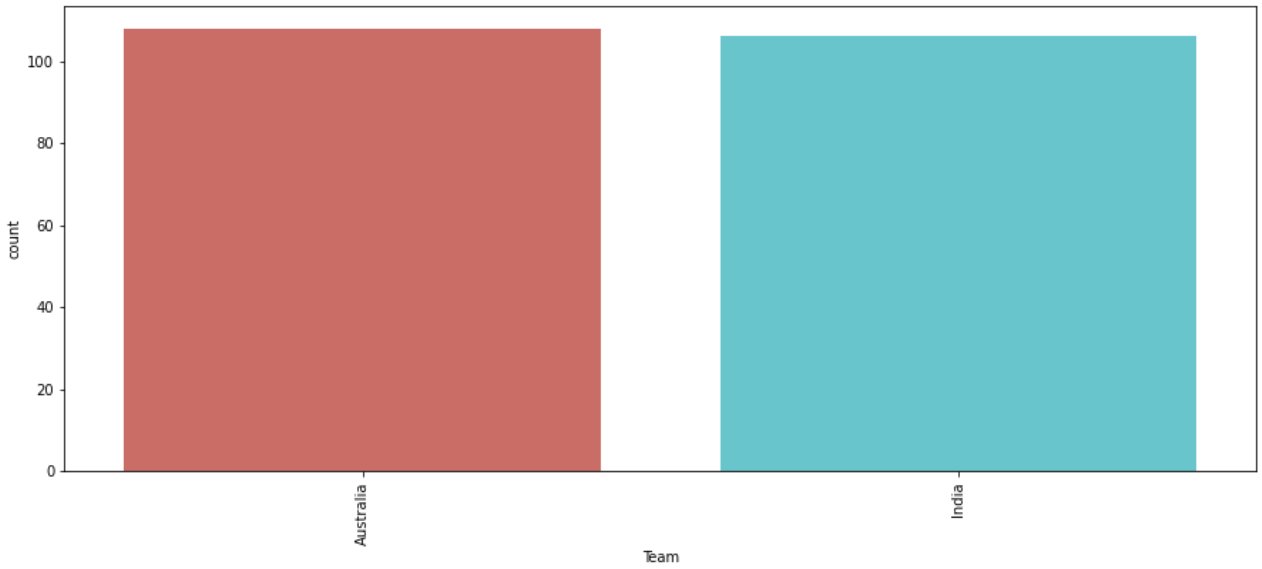


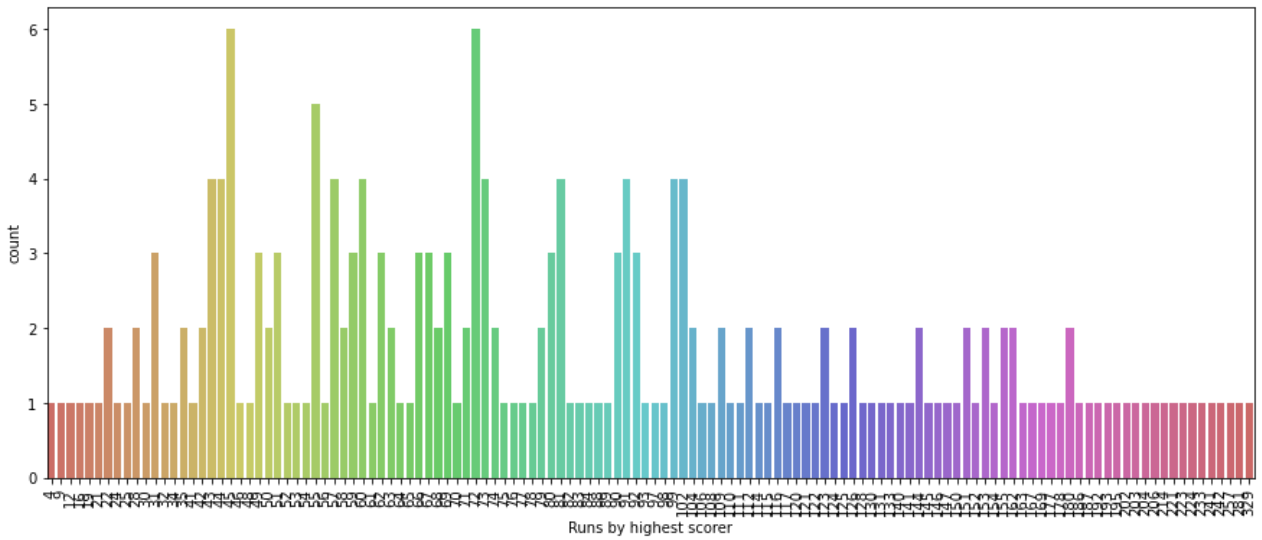

count
0
10
20
30
40
50
1
2
3
4
Innings

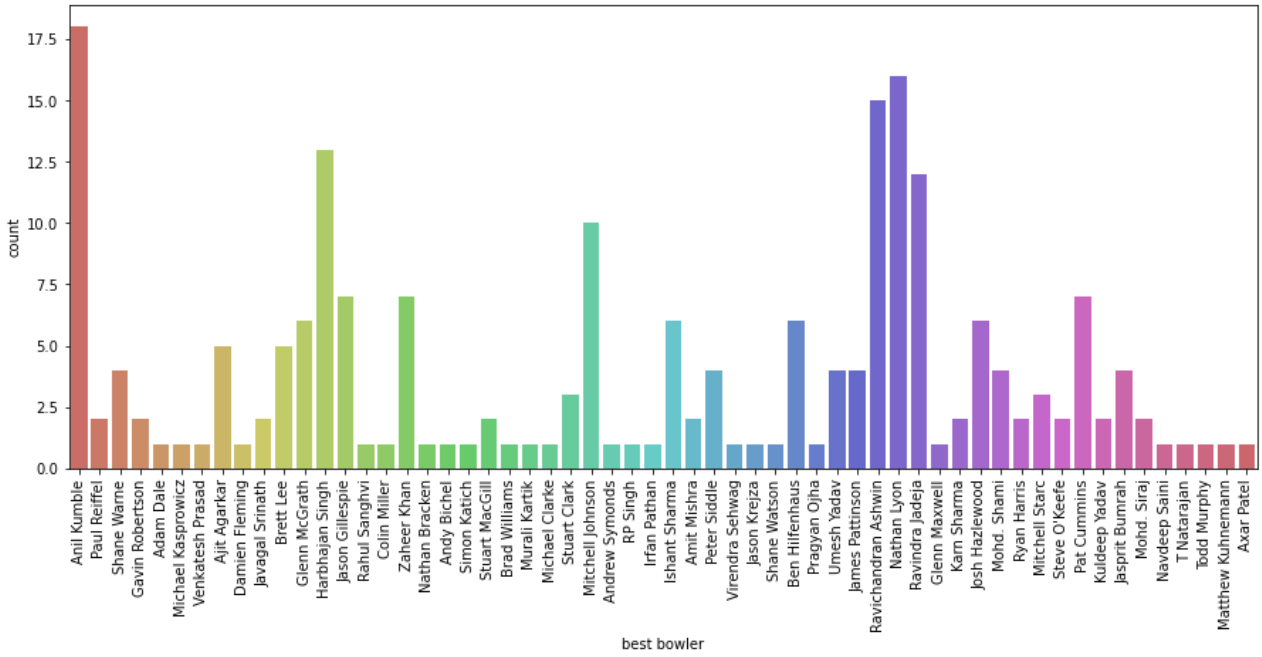

count
0
5
10
15
20
25
Delhi
Chennai
Kolkata
Bengaluru
Adelaide
Melbourne
Sydney
Mumbai
Brisbane
Nagpur
Perth
Mohali
Hyderabad
Pune
Ranchi
Dharamsala
Indore
Ahmedabad
Venue

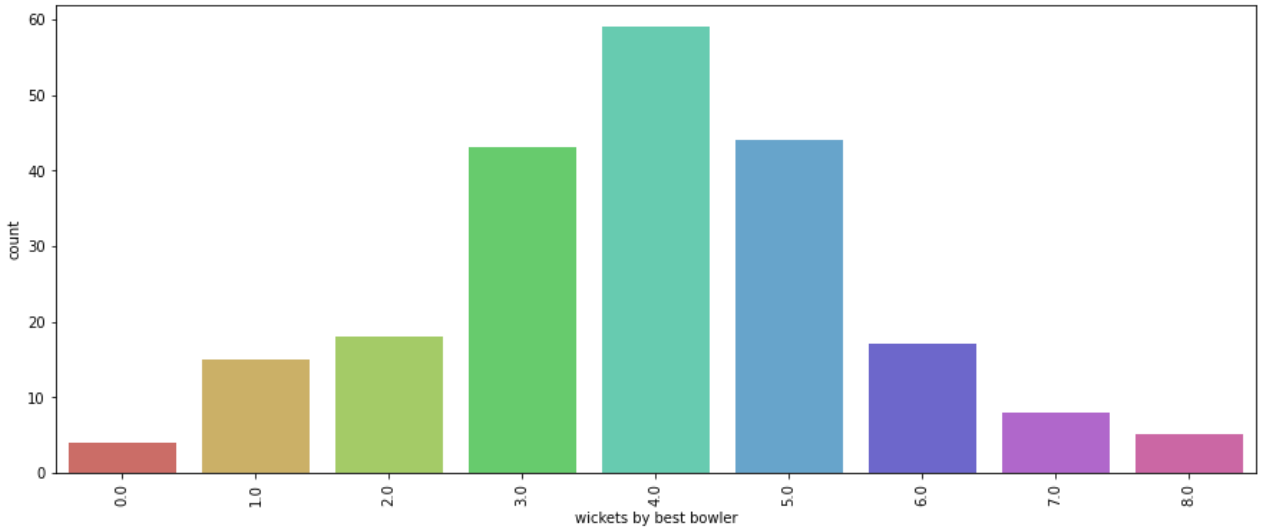

count
0
20
40
60
80
100
India
Australia
Host

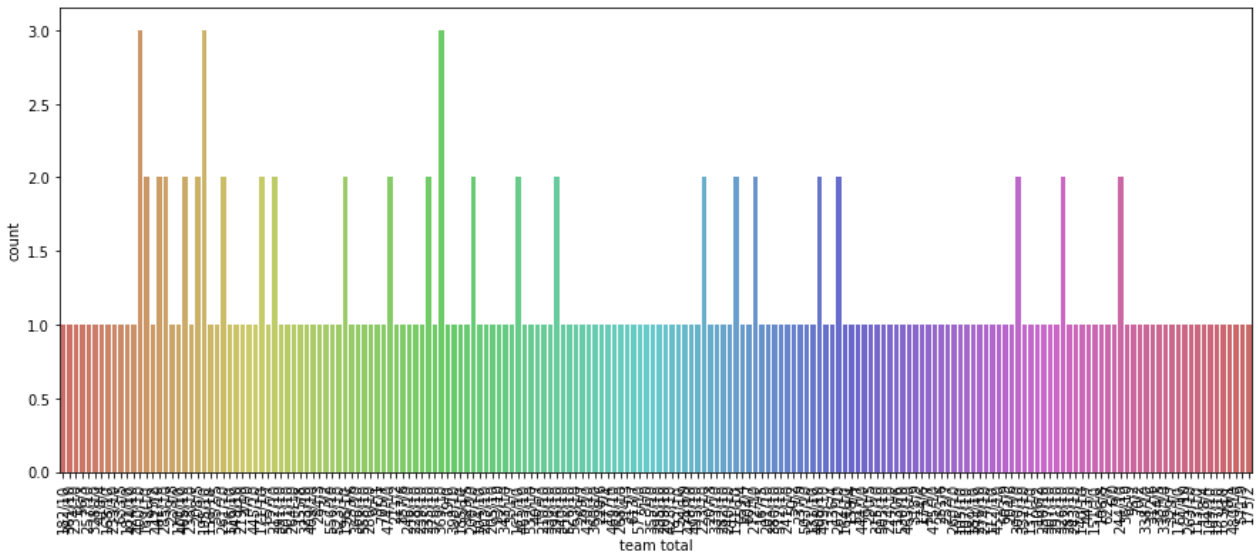

count
Highest Scorer
Michael Slater
Nayan Mongia
Steve Waugh
Mohd. Azharuddin
Navjot Sidhu
Ian Healy
Sachin Tendulkar
Shane Warne
Mark Taylor
Mark Waugh
Greg Blewett
Sourav Ganguly
Adam Gilchrist
Justin Langer
VVS Laxman
Matthew Hayden
Rahul Dravid
Ricky Ponting
Virendra Sehwag
Simon Katich
Michael Clarke
Parthiv Patel
Damien Martyn
Mohd. Kaif
Andrew Symonds
Michael Hussey
Anil Kumble
Zaheer Khan
Shane Watson
Gautam Gambhir
Marcus North
Cheteshwar Pujara
Ed Cowan
MS Dhoni
Virat Kohli
David Warner
Moises Henriques
Mitchell Starc
Shikhar Dhawan
Phillip Hughes
Peter Siddle
Murali Vijay
Steve Smith
Chris Rogers
Shaun Marsh
Virat Kohli
Matt Renshaw
KL Rahul
Peter Handscomb
Ravindra Jadeja
Glenn Maxwell
Travis Head
Marcus Harris
Usman Khawaja
Ajinkya Rahane
Mayank Agarwal
Pat Cummins
Tim Paine
Joe Burns
Marnus Labuschagne
Cameron Green
Shubman Gill
Rishabh Pant
Shardul Thakur
Rohit Sharma
Axar Patel

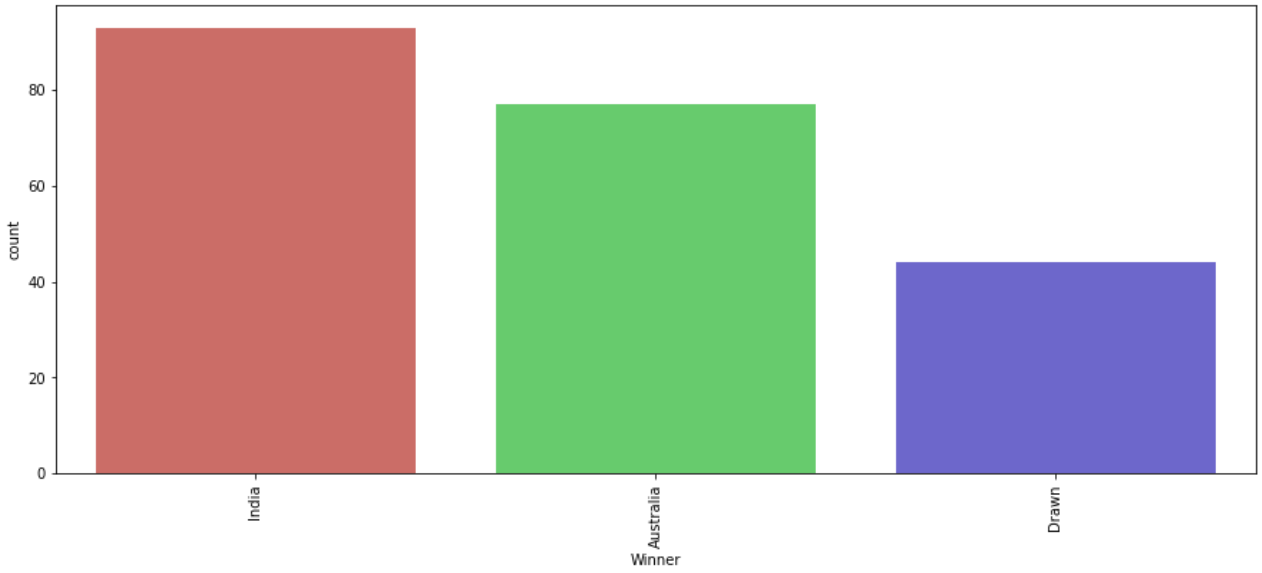

count
Team
Australia
India

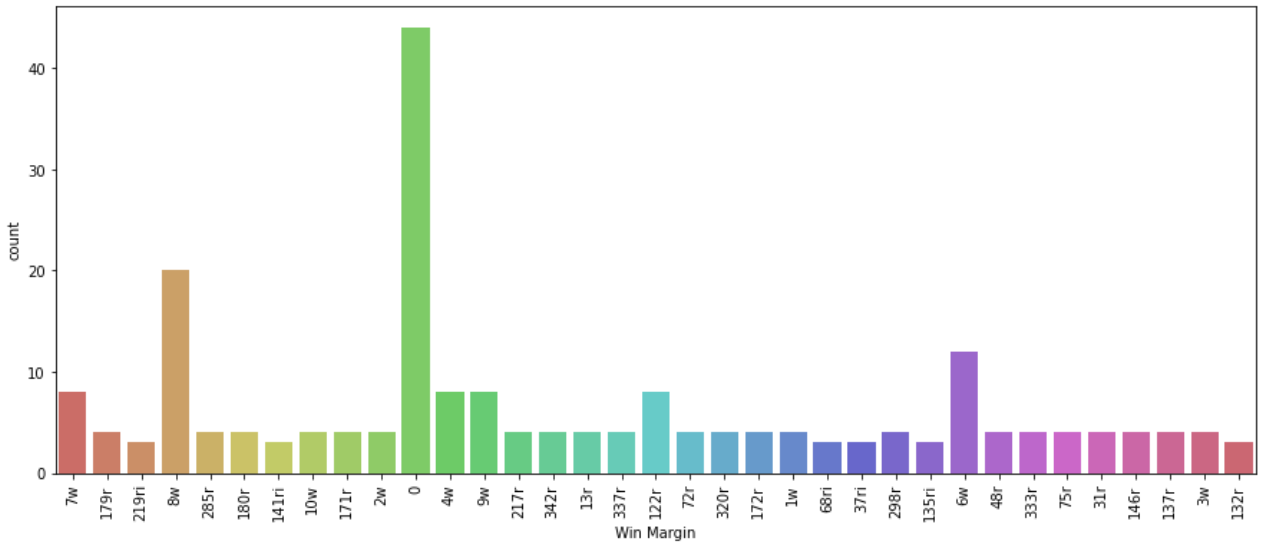

count
Runs by highest scorer

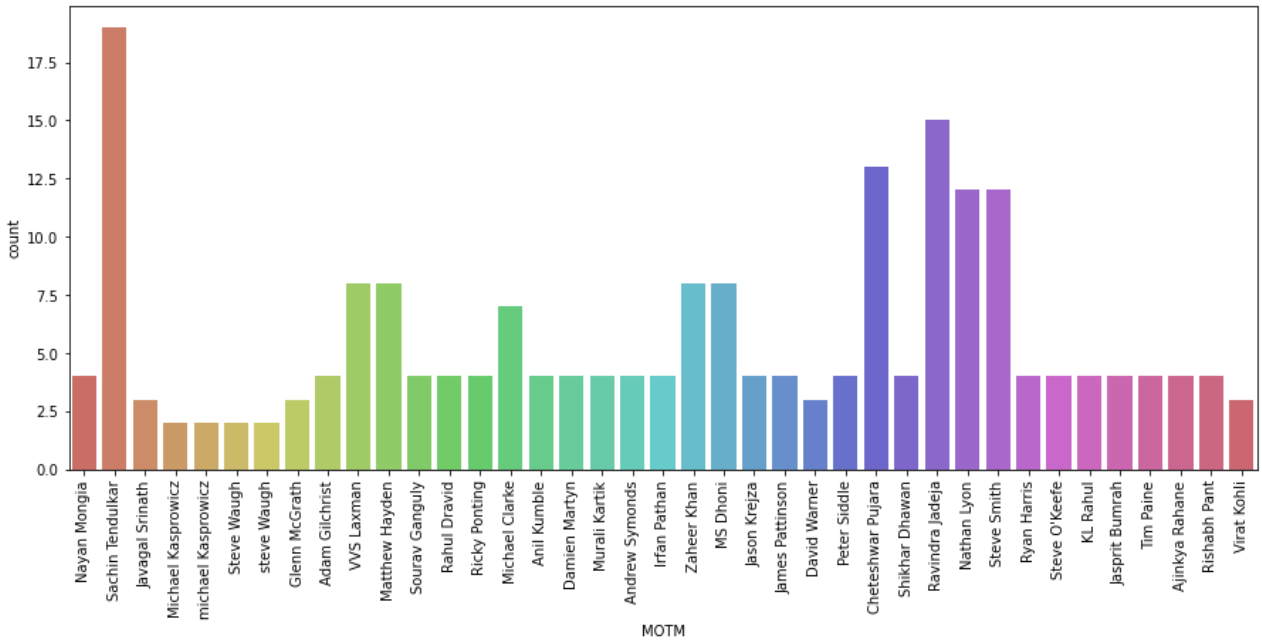

count
best bowler
Anil Kumble
Paul Reiffel
Shane Warne
Gavin Robertson
Adam Dale
Michael Kasprowicz
Venkatesh Prasad
Ajit Agarkar
Damien Fleming
Javagal Srinath
Brett Lee
Glenn McGrath
Harbhajan Singh
Jason Gillespie
Rahul Sanghvi
Colin Miller
Zaheer Khan
Nathan Bracken
Andy Bichel
Simon Katich
Stuart MacGill
Brad Williams
Murali Kartik
Michael Clarke
Stuart Clark
Mitchell Johnson
Andrew Symonds
RP Singh
Irfan Pathan
Ishant Sharma
Amit Mishra
Peter Siddle
Virendra Sehwag
Jason Krejza
Shane Watson
Ben Hilfenhaus
Pragyan Ojha
Umesh Yadav
James Pattinson
Ravichandran Ashwin
Nathan Lyon
Ravindra Jadeja
Glenn Maxwell
Karn Sharma
Josh Hazlewood
Mohd. Shami
Ryan Harris
Mitchell Starc
Steve O'Keefe
Pat Cummins
Kuldeep Yadav
Jasprit Bumrah
Mohd. Siraj
Navdeep Saini
T Natarajan
Todd Murphy
Matthew Kuhnemann
Axar Patel

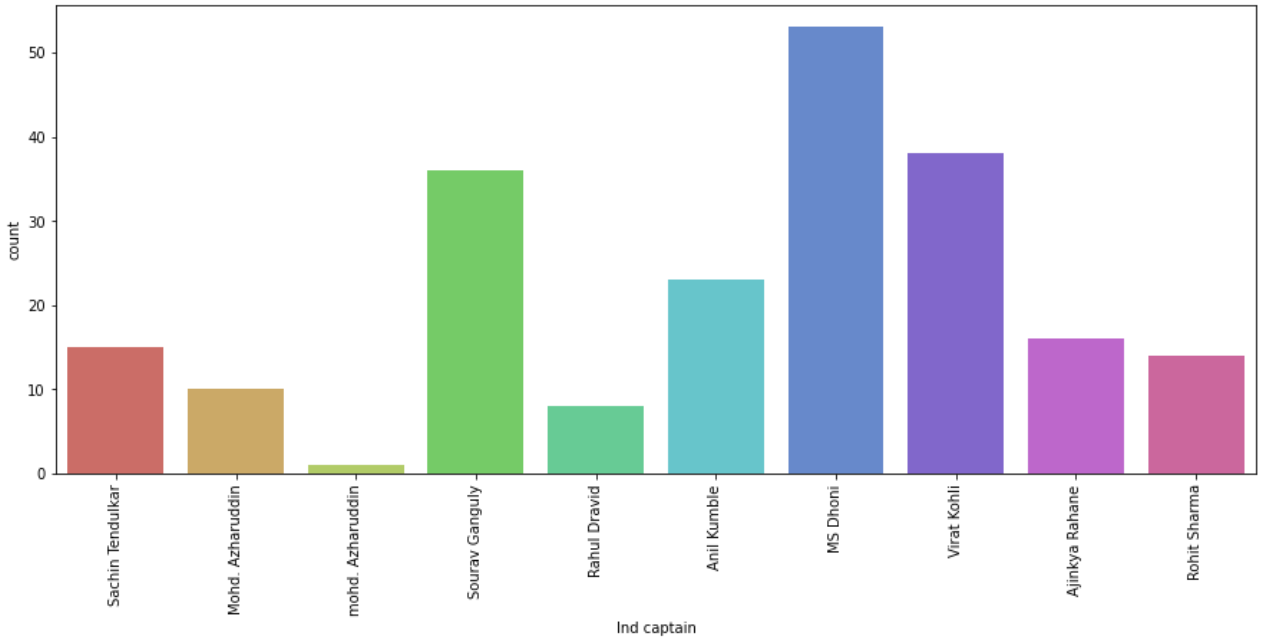

count
wickets by best bowler

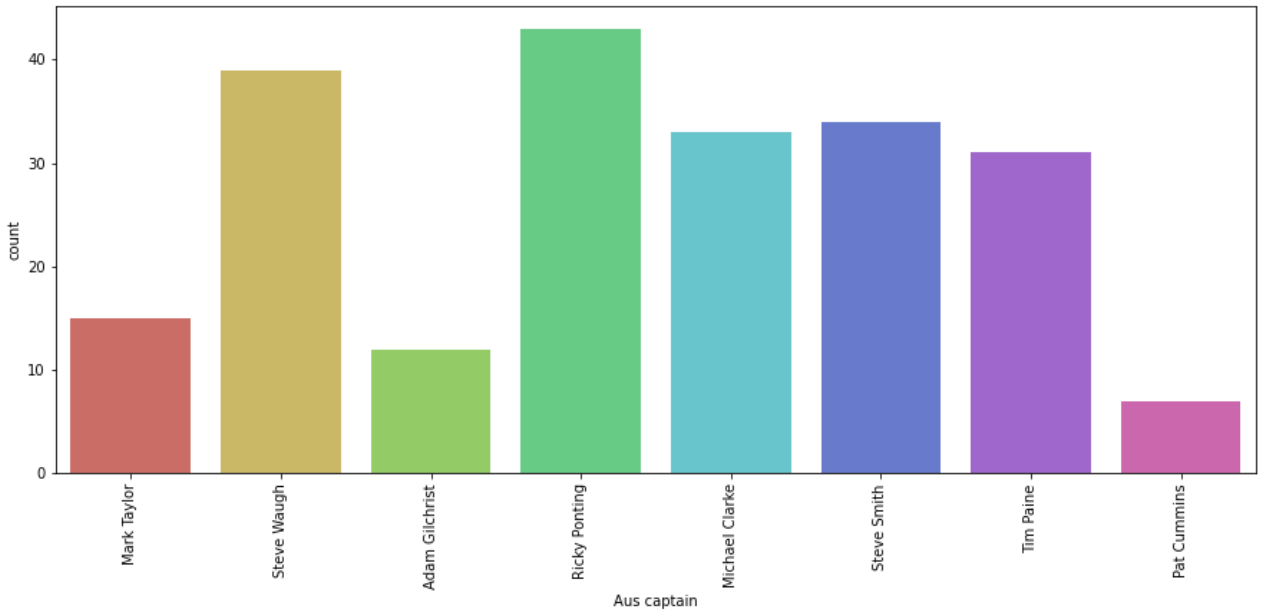

count
team total

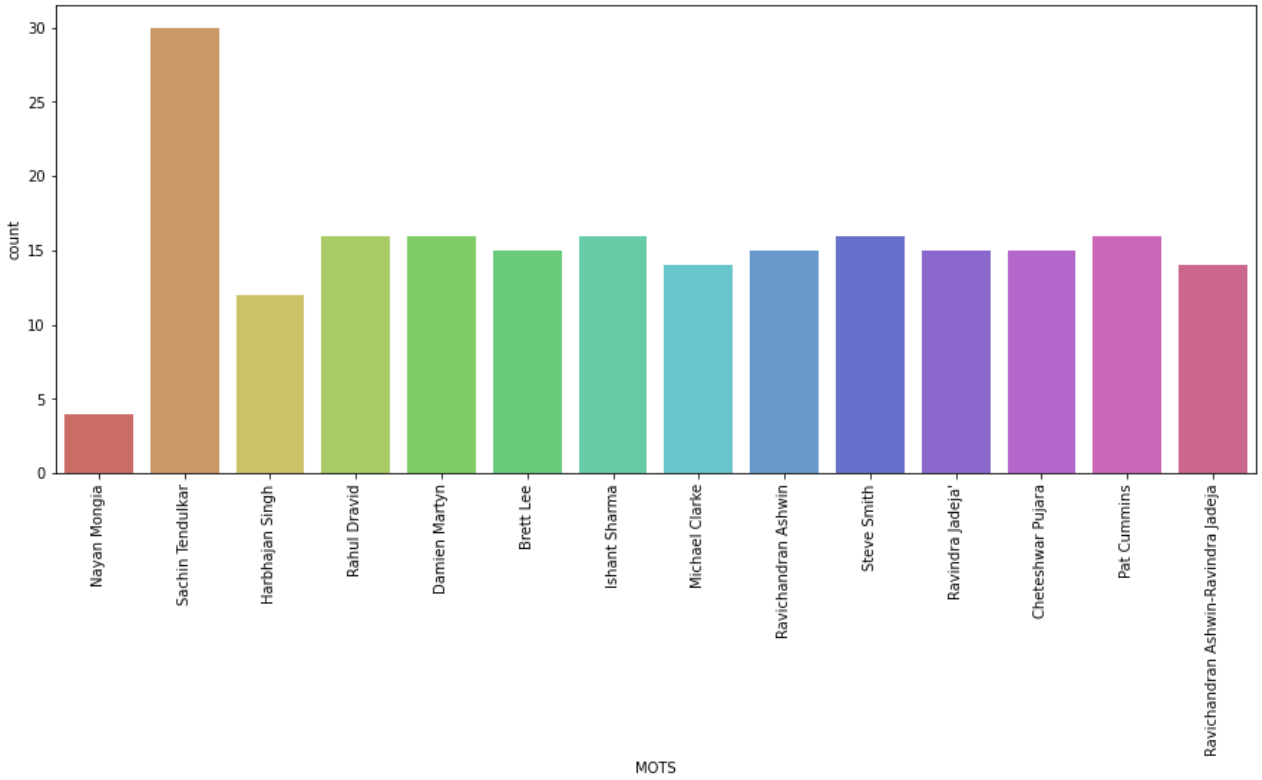

count
0
20
40
60
80
India
Australia
Drawn
Winner

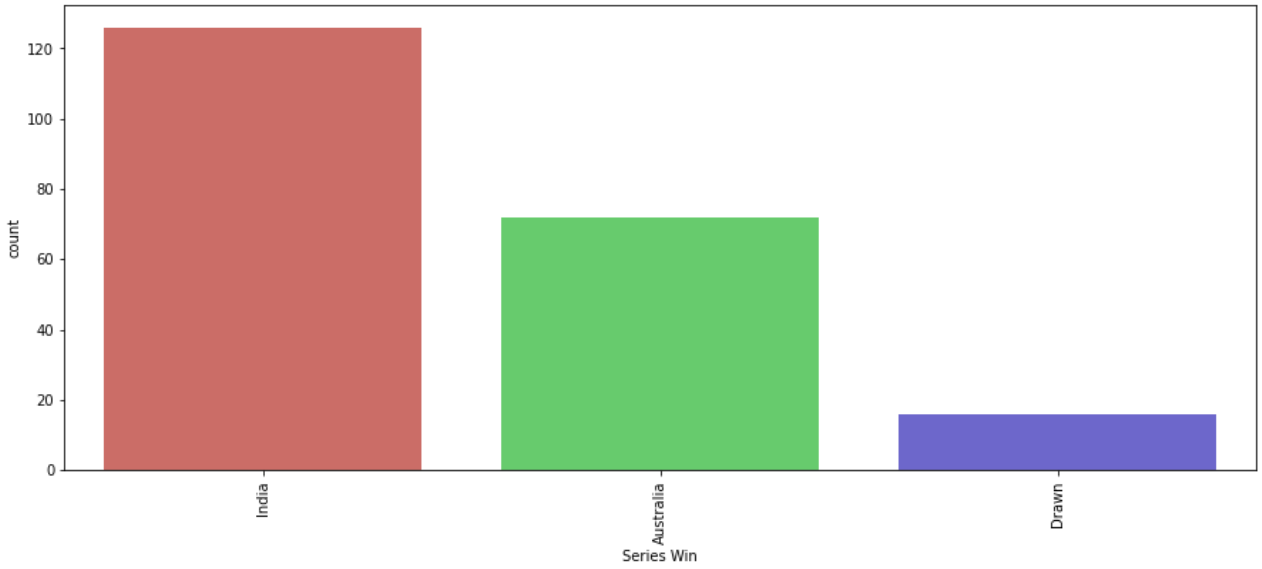

count
0
10
20
30
40
7w
179r
219ri
8w
285r
180r
141ri
10w
171r
2w
0
4w
9w
217r
342r
13r
337r
122r
72r
320r
172r
1w
68ri
37ri
298r
135ri
6w
48r
333r
75r
31r
146r
137r
3w
132r
Win Margin


count
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Nayan Mongia
Sachin Tendulkar
Javagal Srinath
Michael Kasprowicz
michael Kasprowicz
Steve Waugh
steve Waugh
Glenn McGrath
Adam Gilchrist
VVS Laxman
Matthew Hayden
Sourav Ganguly
Rahul Dravid
Ricky Ponting
Michael Clarke
Anil Kumble
Damien Martyn
Murali Kartik
Andrew Symonds
Irfan Pathan
Zaheer Khan
MS Dhoni
Jason Krejza
James Pattinson
David Warner
Peter Siddle
Cheteshwar Pujara
Shikhar Dhawan
Ravindra Jadeja
Nathan Lyon
Steve Smith
Ryan Harris
Steve O'Keefe
KL Rahul
Jasprit Bumrah
Tim Paine
Ajinkya Rahane
Rishabh Pant
Virat Kohli
MOTM


count
0
10
20
30
40
50
Sachin Tendulkar
Mohd. Azharuddin
mohd. Azharuddin
Sourav Ganguly
Rahul Dravid
Anil Kumble
MS Dhoni
Virat Kohli
Ajinkya Rahane
Rohit Sharma
Ind captain


count
0
10
20
30
40
Mark Taylor
Steve Waugh
Adam Gilchrist
Ricky Ponting
Michael Clarke
Steve Smith
Tim Paine
Pat Cummins
Aus captain





In [37]:
```
new_data=df[['Test Series Year', 'Test Number', 'Innings', 'Venue', 'Host', 'Team', 'Wi
 'Ind captain', 'Aus captain', 'MOTS', 'Series Win']]
```

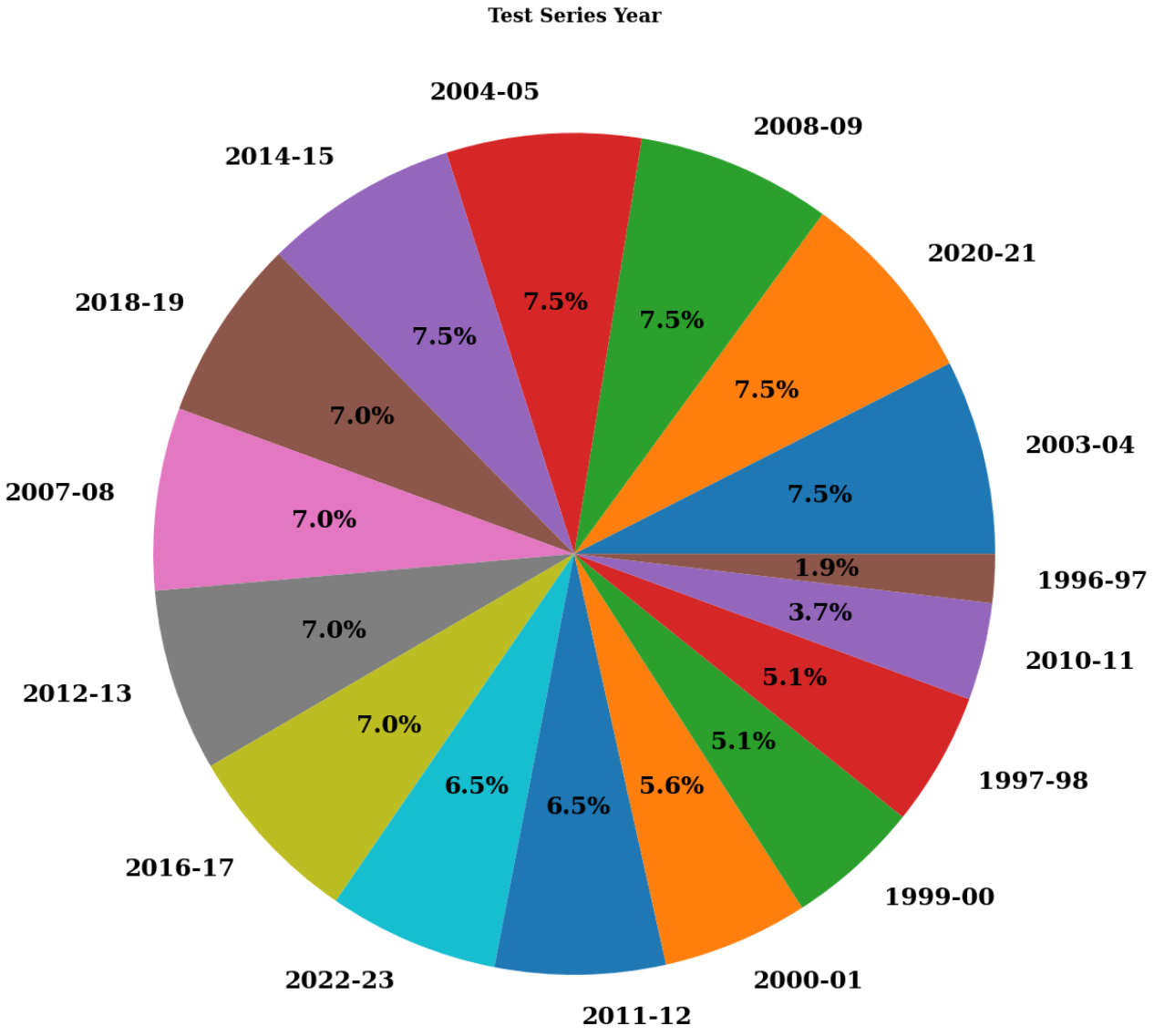
In [38]:
```
new_data.head()
```

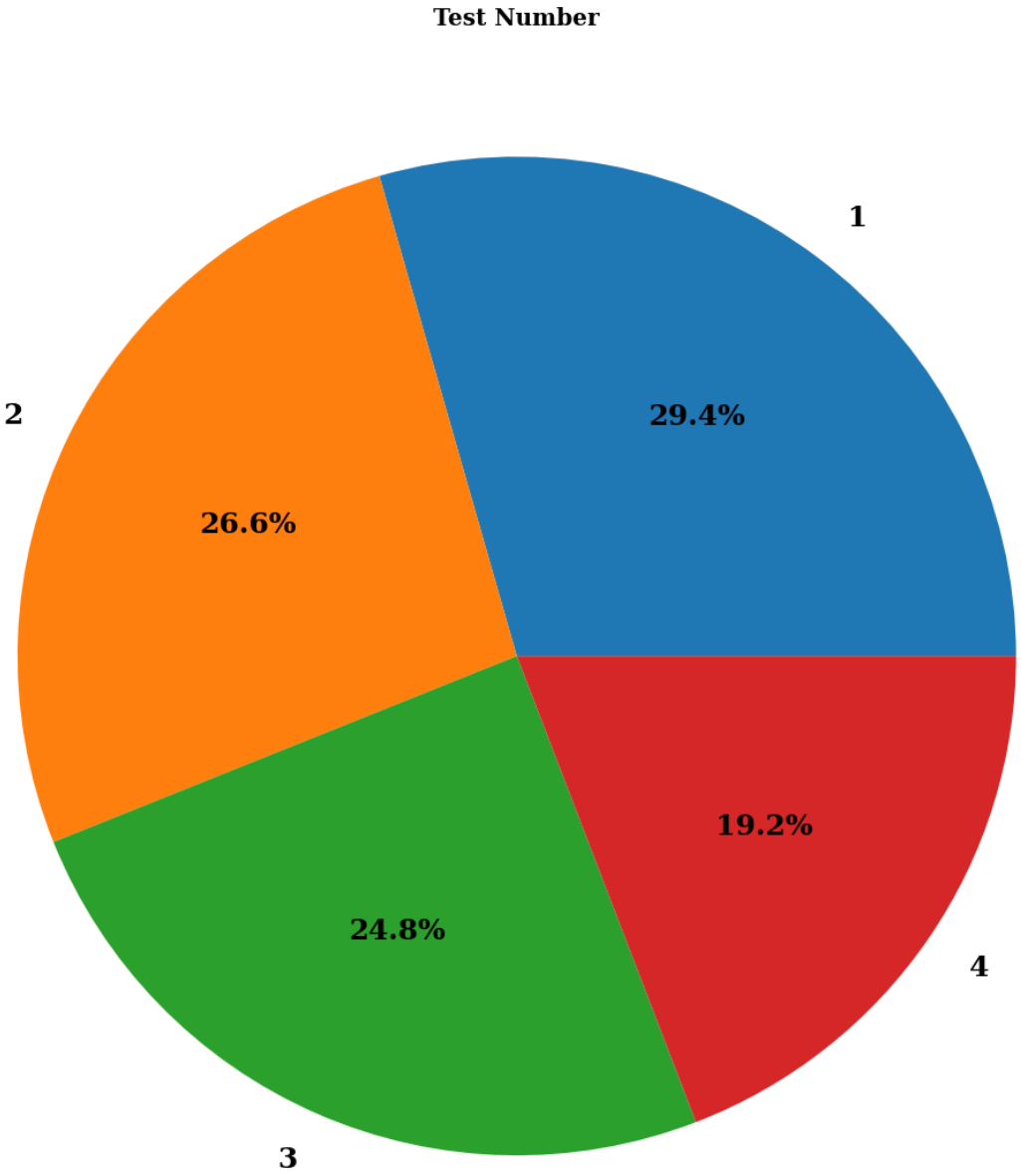
Out[38]:

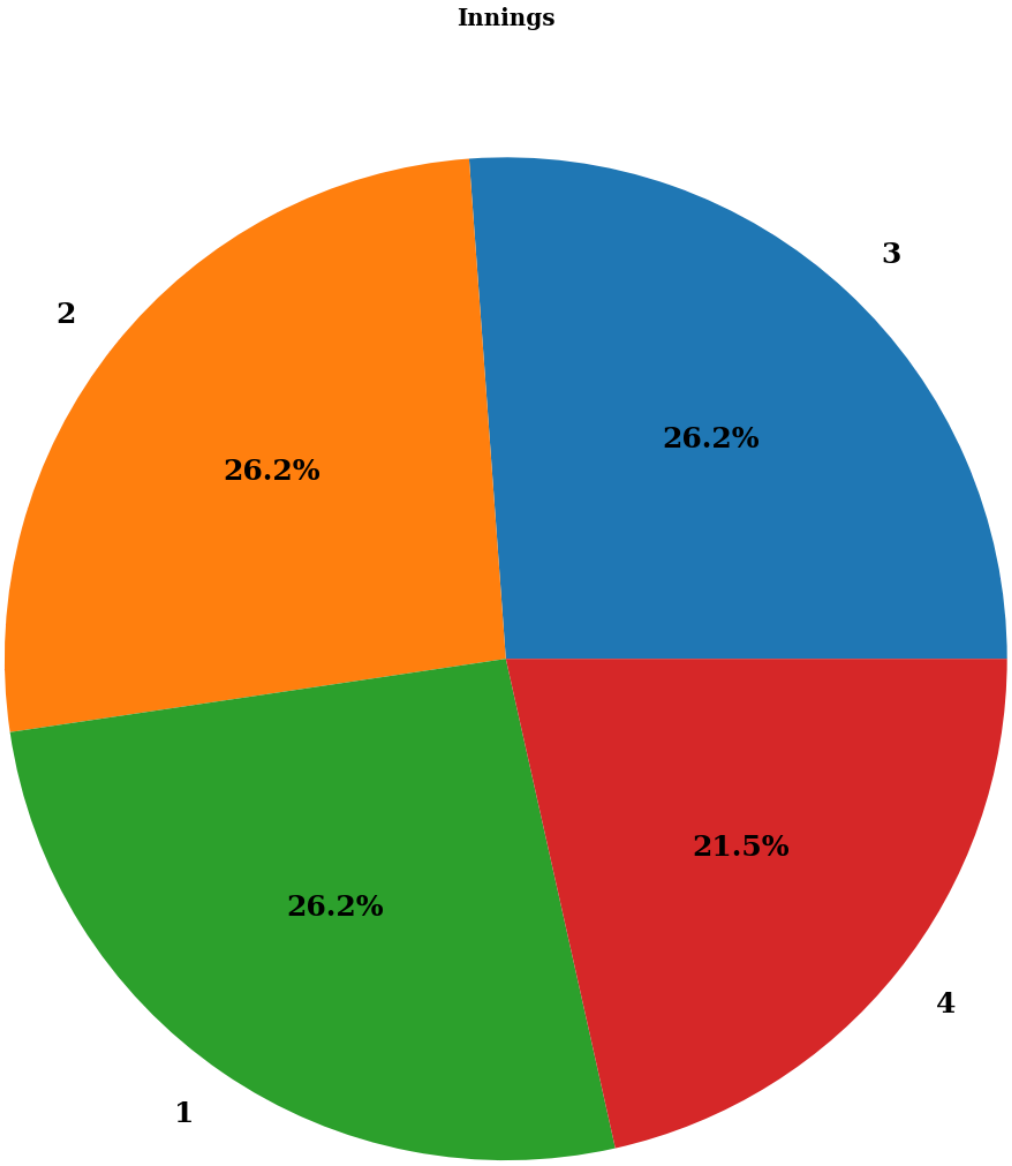
	Test Series Year	Test Number	Innings	Venue	Host	Team	Winner	Ind captain	Aus captain	MOTS	Series Win
0	1996-97	1	1	Delhi	India	Australia	India	Sachin Tendulkar	Mark Taylor	Nayan Mongia	India
1	1996-97	1	2	Delhi	India	India	India	Sachin Tendulkar	Mark Taylor	Nayan Mongia	India
2	1996-	1	3	Delhi	India	Australia	India	Sachin	Mark	Nayan	India

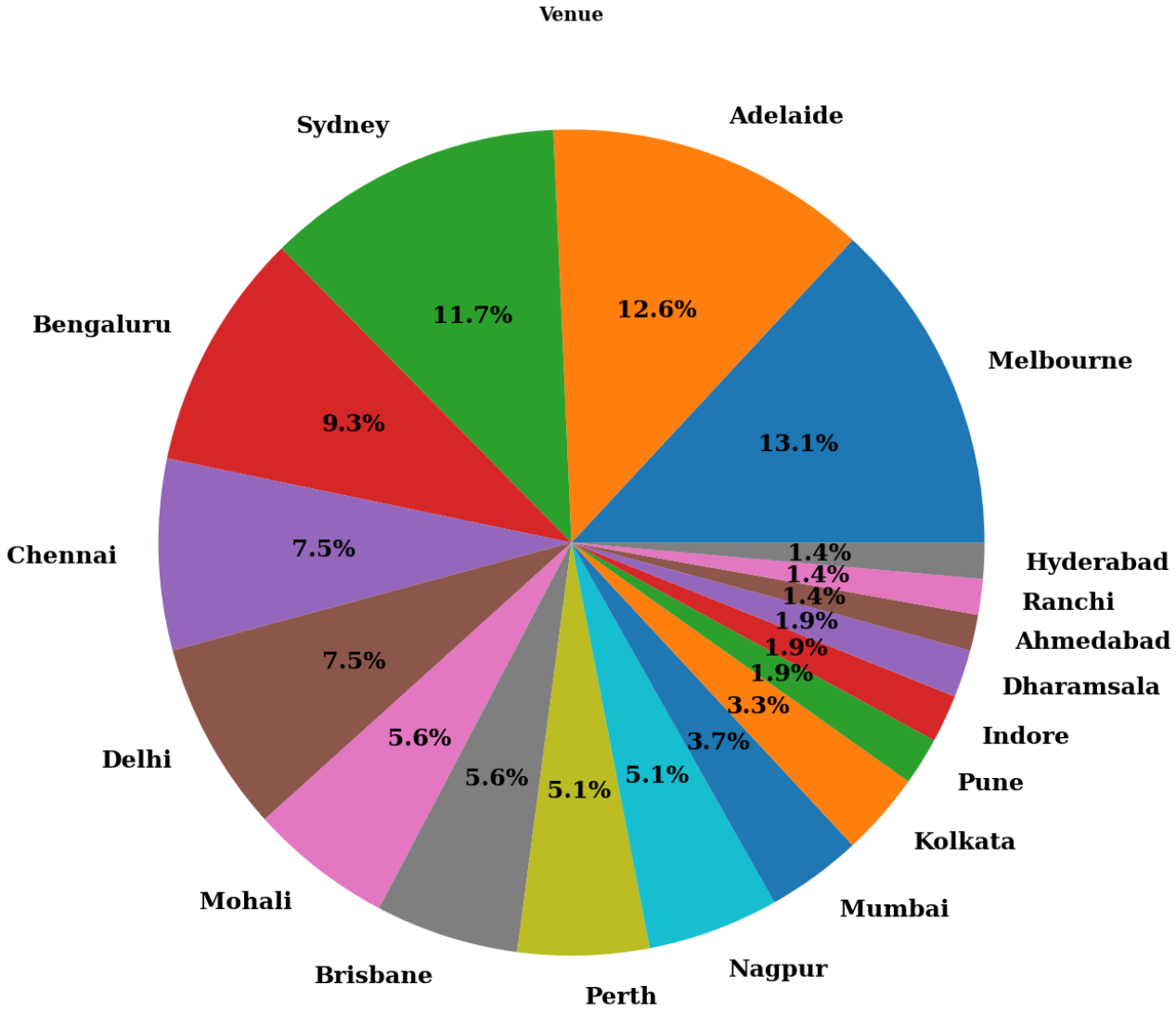
	Test Series Year	Test Number	Innings	Venue	Host	Team	Winner	Ind captain	Aus captain	MOTS	Series Win
	97							Tendulkar	Taylor	Mongia	
3	1996-97	1	4	Delhi	India	India	India	Sachin Tendulkar	Mark Taylor	Nayan Mongia	India
4	1997-98	1	1	Chennai	India	India	India	Mohd. Azharuddin	Mark Taylor	Sachin Tendulkar	India

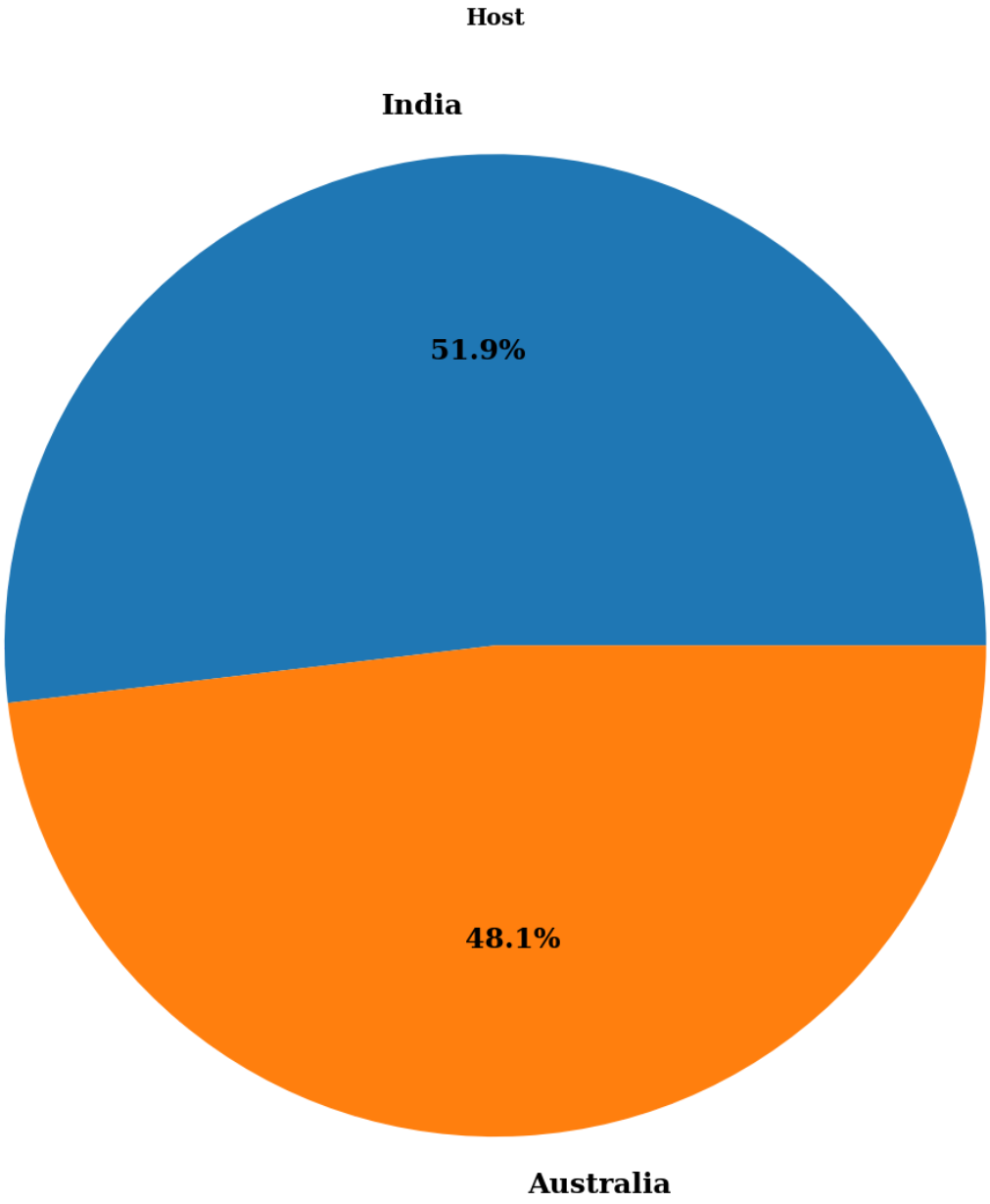
In [40]:
```
for i in new_data:
    plt.figure(figsize=(30,20))
    plt.pie(df[i].value_counts(), labels=df[i].value_counts().index, autopct='%1.1f%%',
 'color': 'black',
'weight': 'bold',
'family': 'serif' })
    hfont = {'fontname':'serif', 'weight': 'bold'}
    plt.title(i, size=20, **hfont)
    plt.show()
```

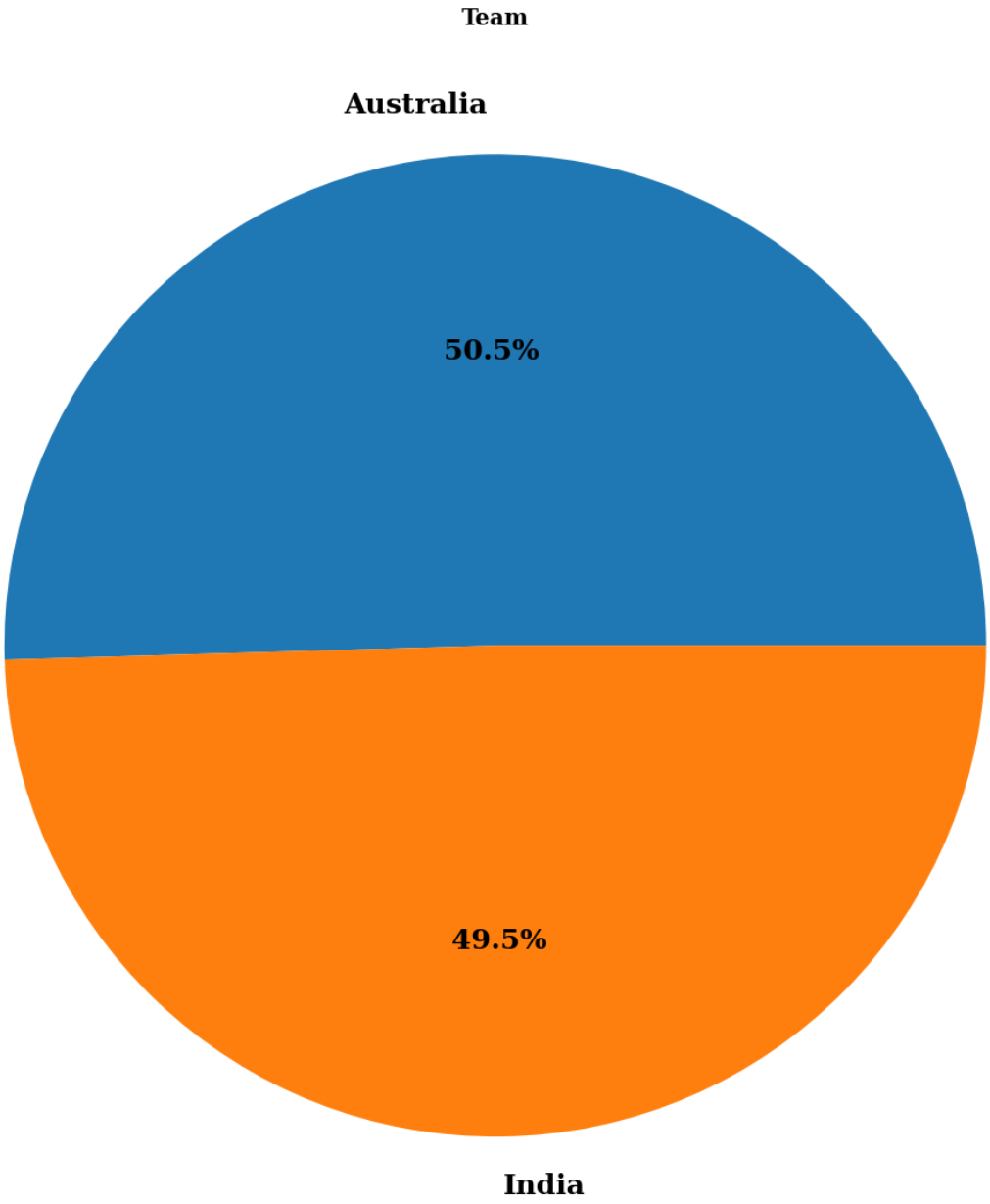


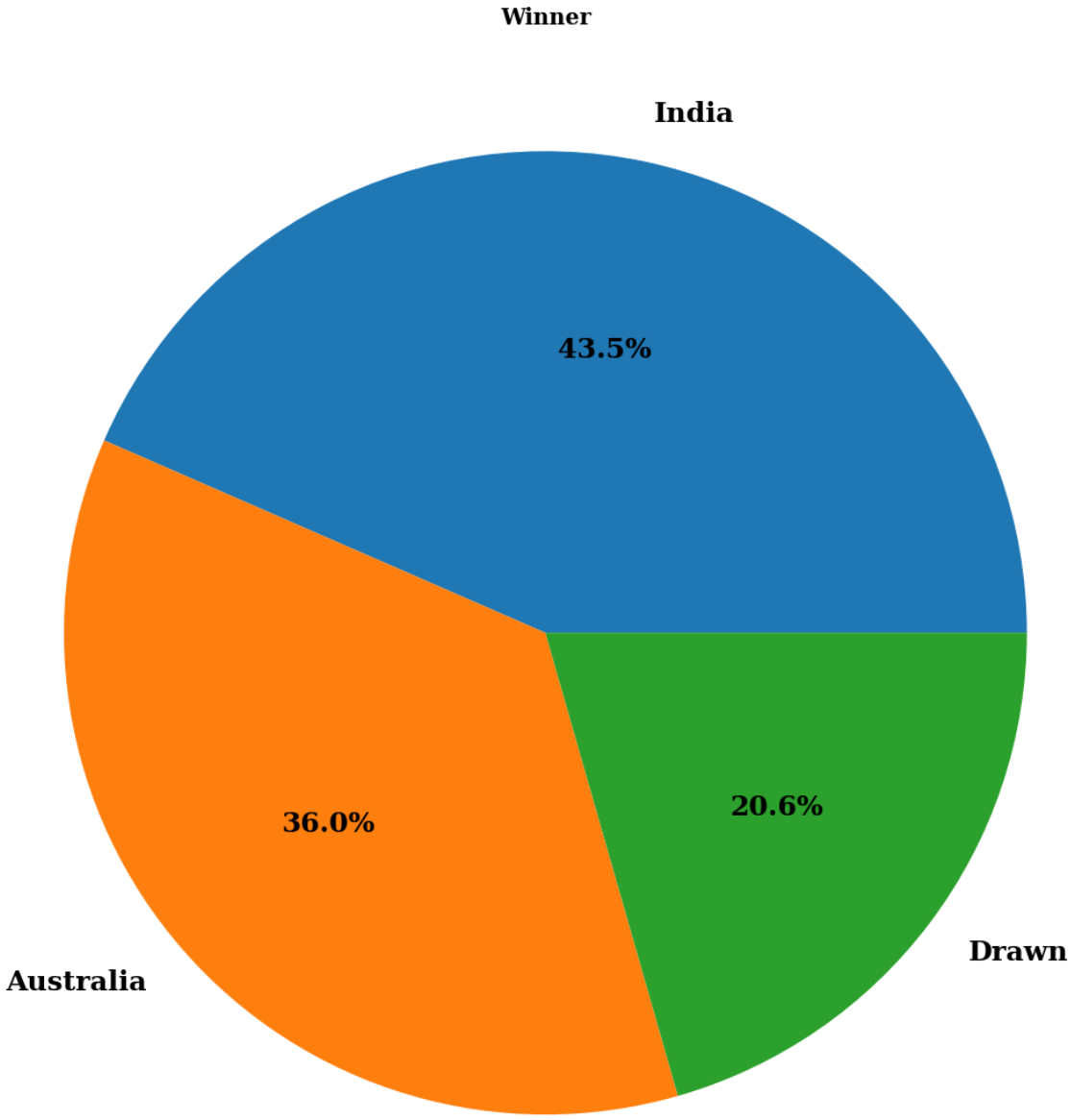

Test Number
1
29.4%
2
26.6%
3
24.8%
4
19.2%

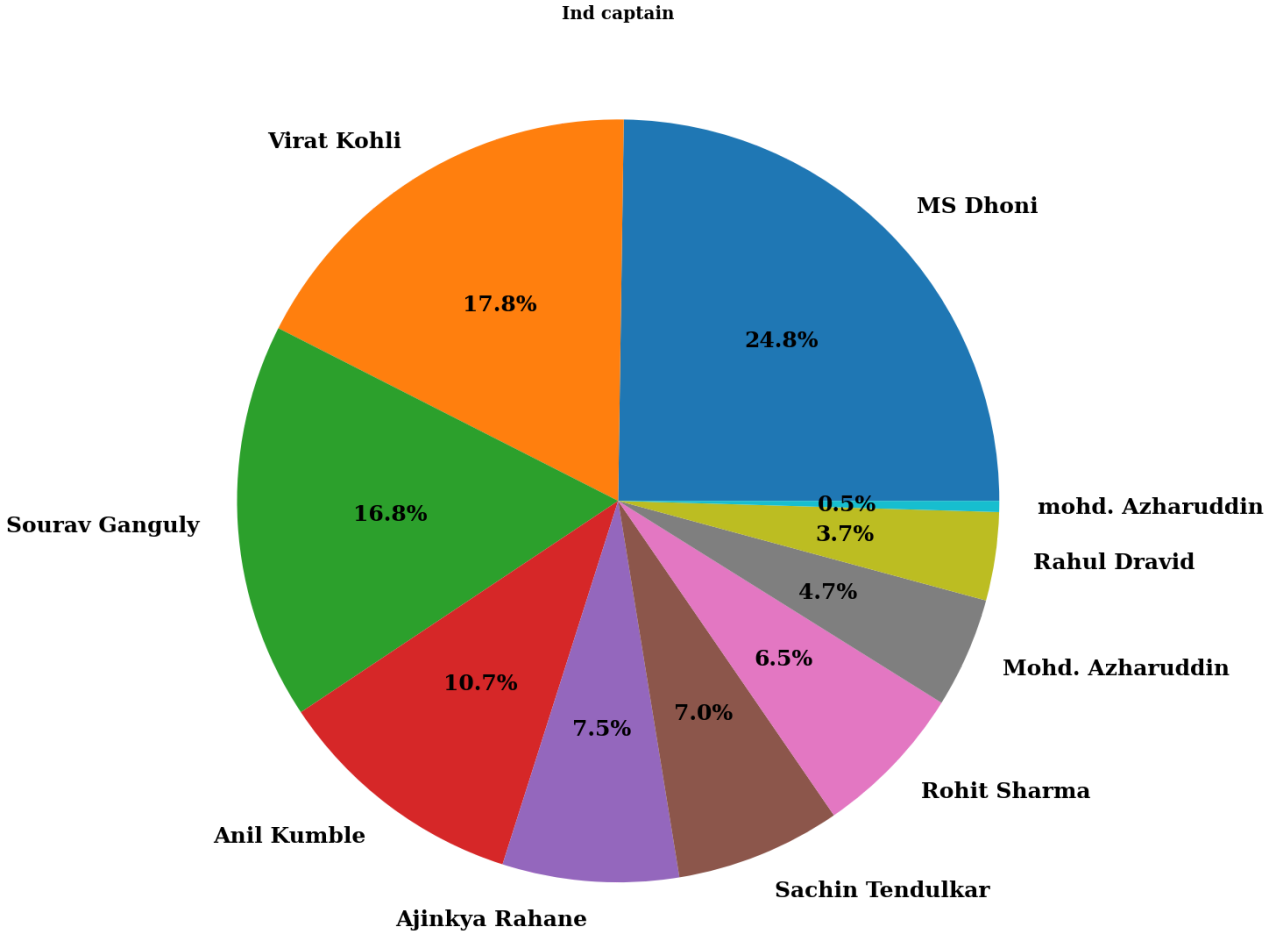

Innings
3
26.2%
2
26.2%
21.5%
4
26.2%
1

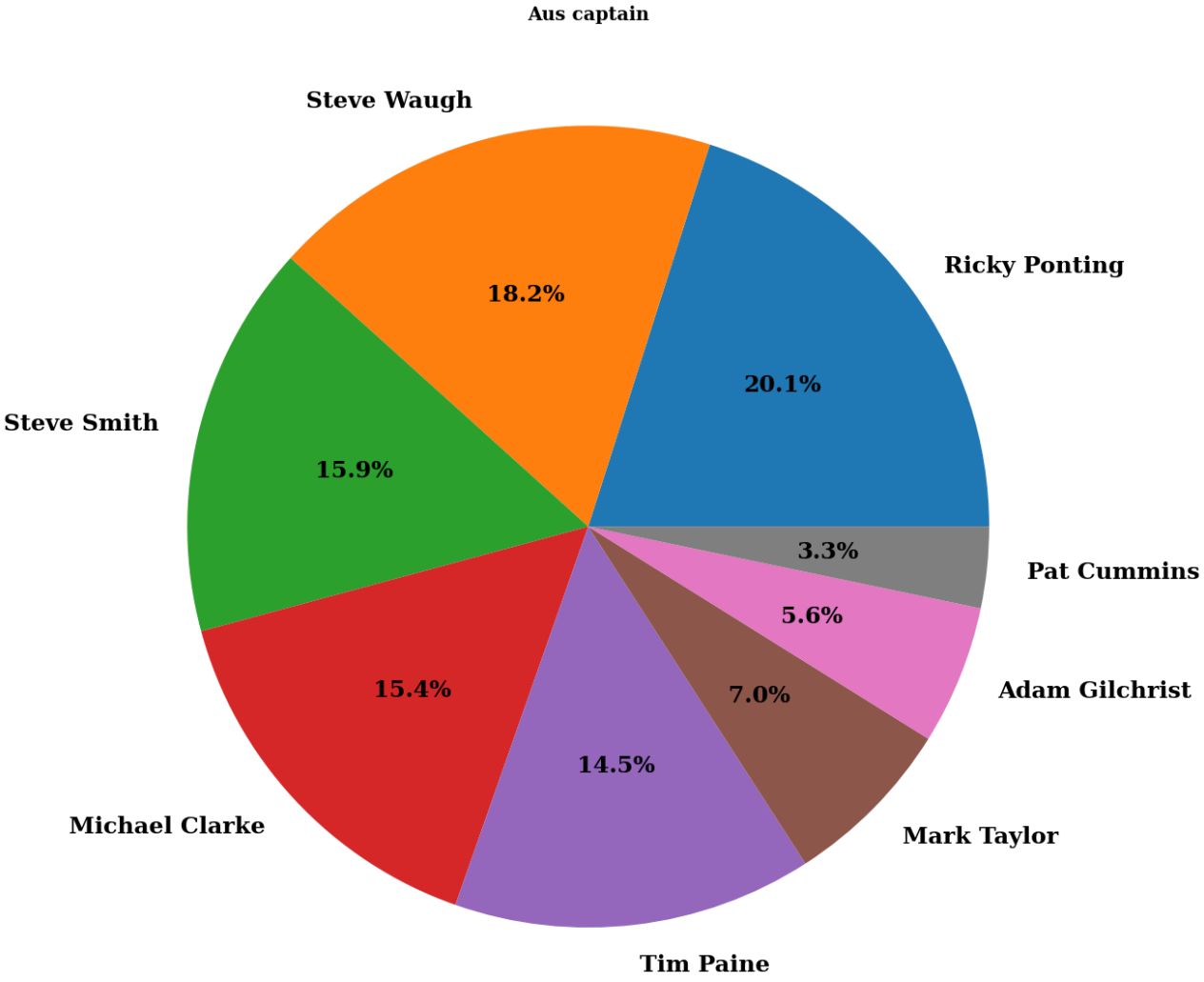

Venue
Sydney
Adelaide
Bengaluru
Melbourne
Chennai
Hyderabad
Ranchi
Ahmedabad
Dharamsala
Indore
Pune
Kolkata
Mumbai
Nagpur
Perth
Brisbane
Mohali
Delhi
11.7%
12.6%
13.1%
9.3%
7.5%
1.4%
1.4%
1.4%
1.9%
1.9%
1.9%
3.3%
3.7%
5.1%
5.1%
5.6%
5.6%
7.5%

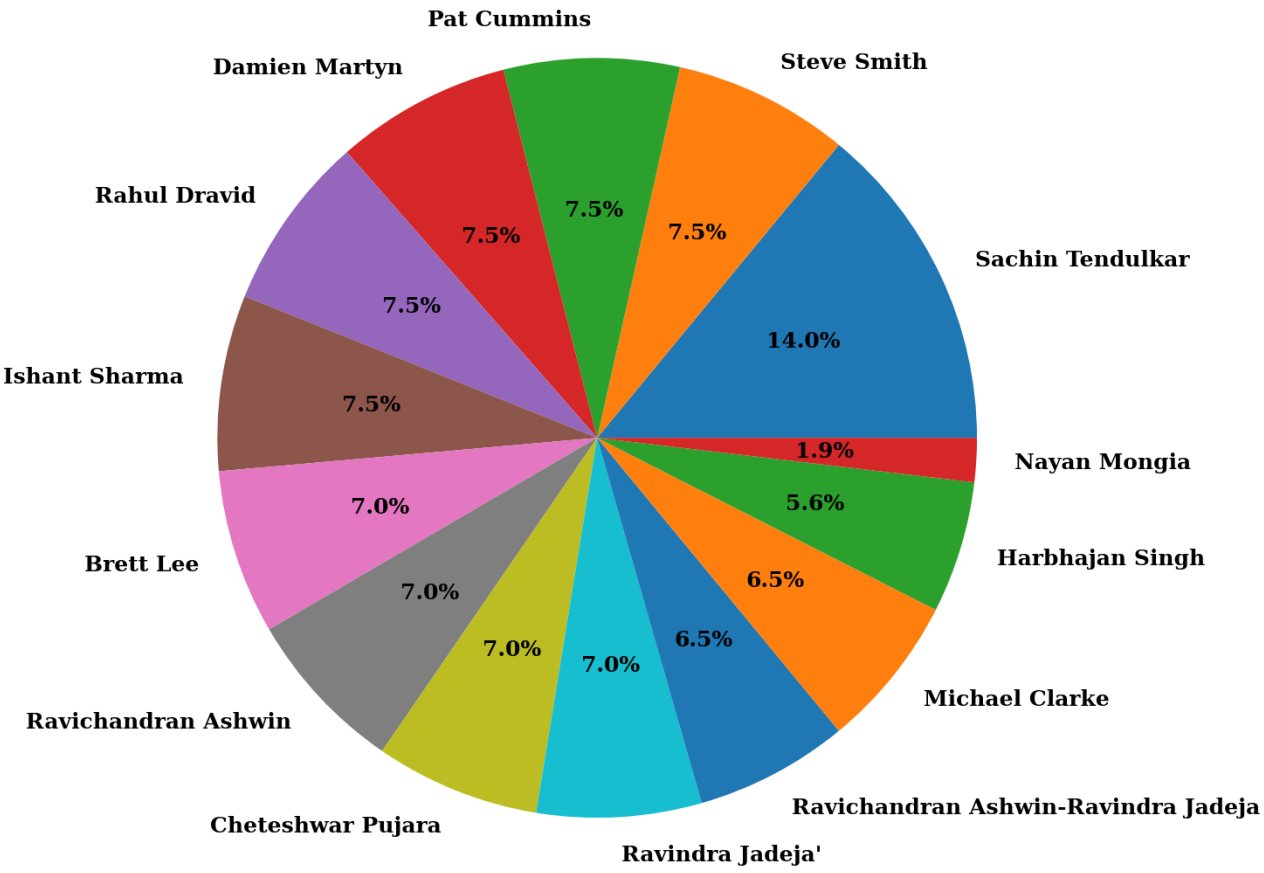

Host
India
51.9%
48.1%
Australia

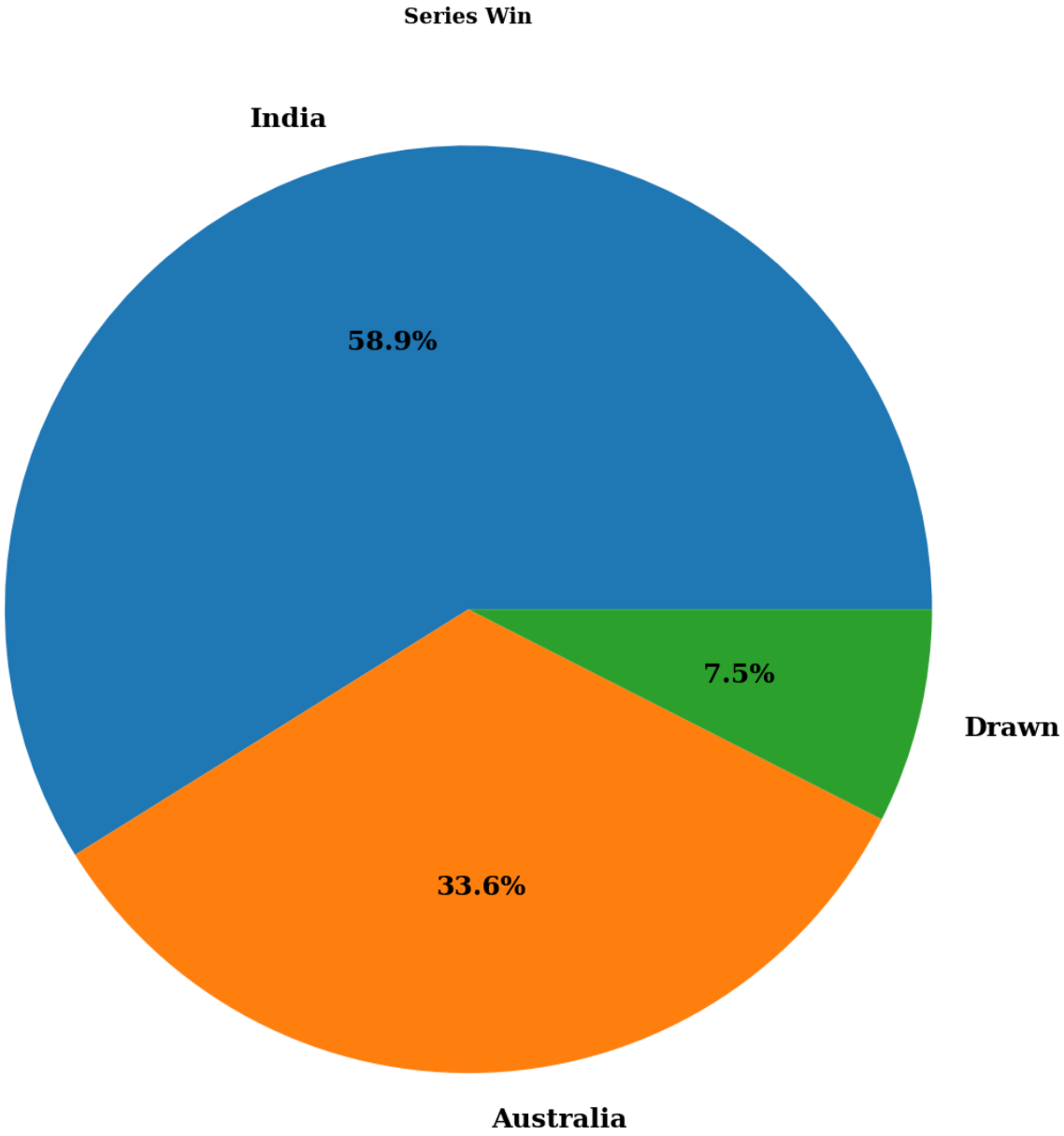

Team
Australia
50.5%
49.5%
India


Winner
India
43.5%
Australia
36.0%
Drawn
20.6%

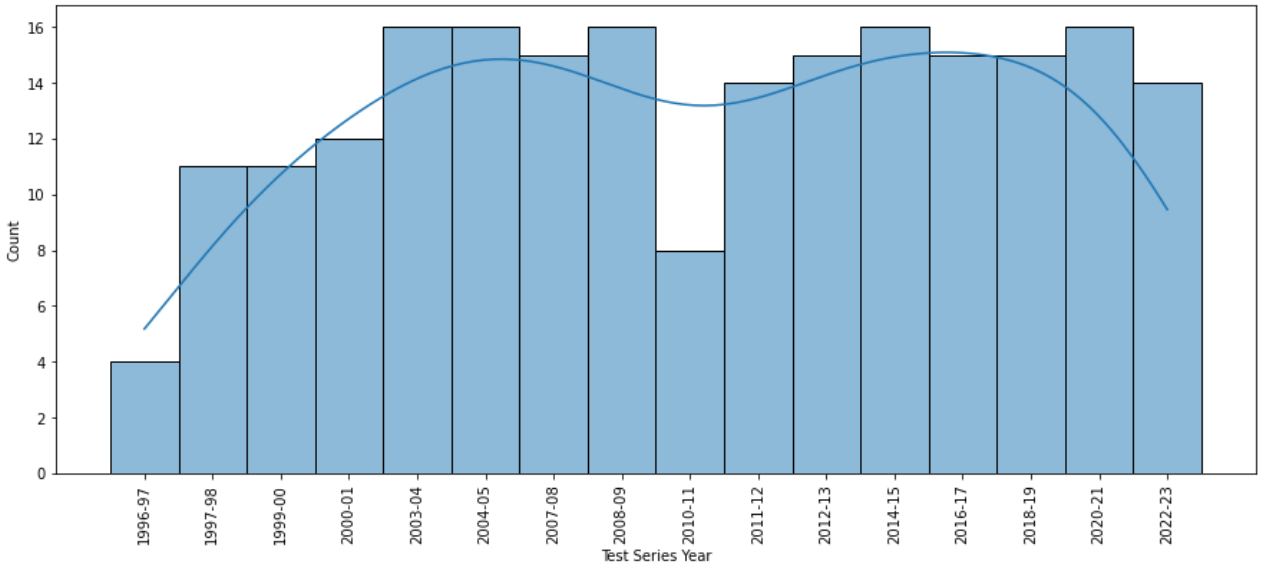

Ind captain
Virat Kohli
MS Dhoni
17.8%
24.8%
Sourav Ganguly
16.8%
0.5%
mohd. Azharuddin
3.7%
Rahul Dravid
4.7%
Mohd. Azharuddin
6.5%
10.7%
7.0%
7.5%
Rohit Sharma
Anil Kumble
Sachin Tendulkar
Ajinkya Rahane

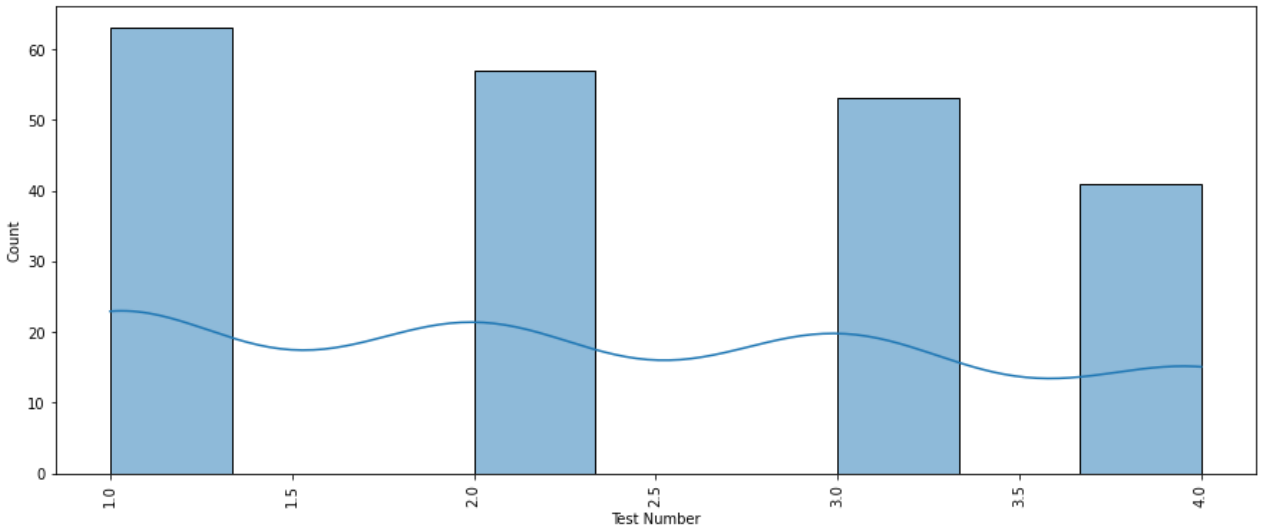

Aus captain
Steve Waugh
18.2%
Ricky Ponting
20.1%
Steve Smith
15.9%
3.3%
Pat Cummins
5.6%
Adam Gilchrist
15.4%
7.0%
14.5%
Michael Clarke
Mark Taylor
Tim Paine

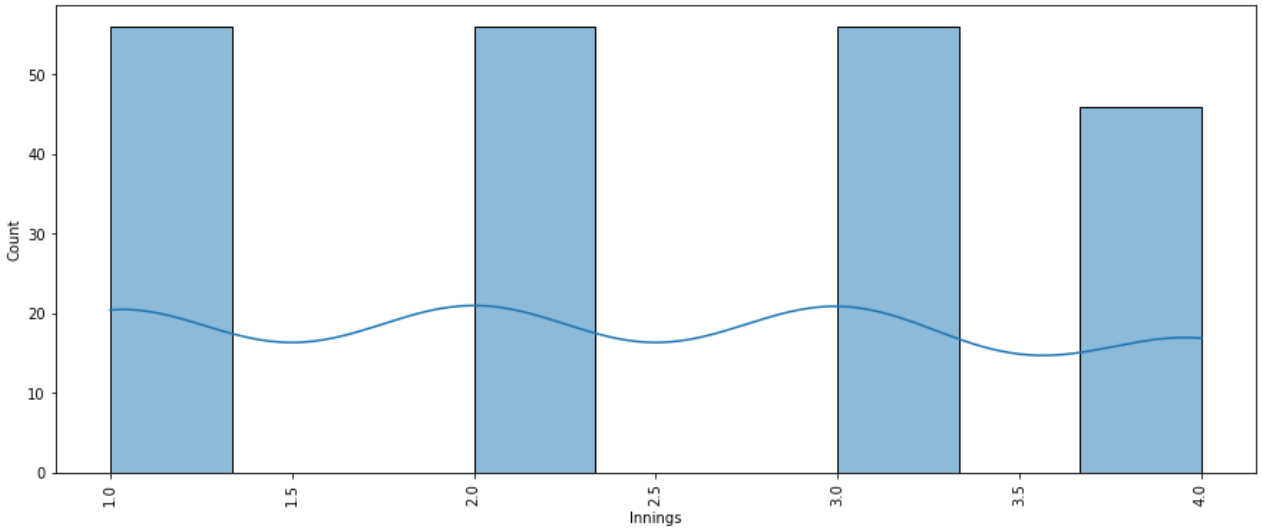

MOTS
Pat Cummins
Damien Martyn
Steve Smith
Rahul Dravid
Sachin Tendulkar
Ishant Sharma
Nayan Mongia
Harbhajan Singh
Brett Lee
Michael Clarke
Ravichandran Ashwin
Ravichandran Ashwin-Ravindra Jadeja
Cheteshwar Pujara
Ravindra Jadeja'
7.5%
7.5%
7.5%
7.5%
14.0%
7.5%
1.9%
5.6%
7.0%
6.5%
7.0%
6.5%
7.0%
7.0%

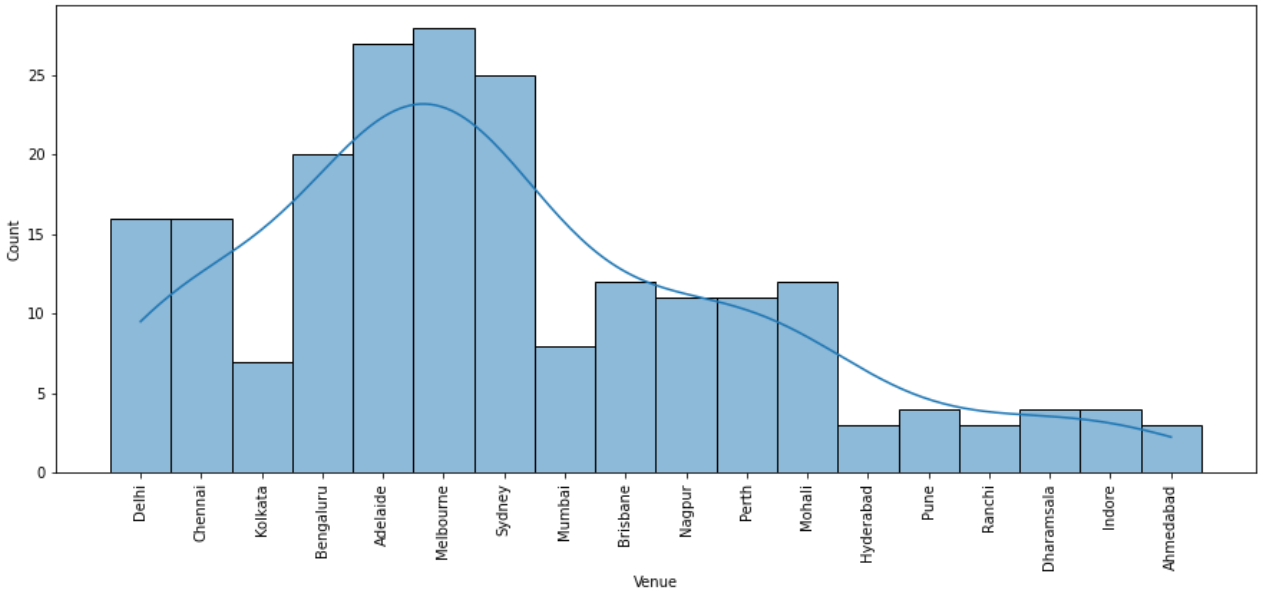


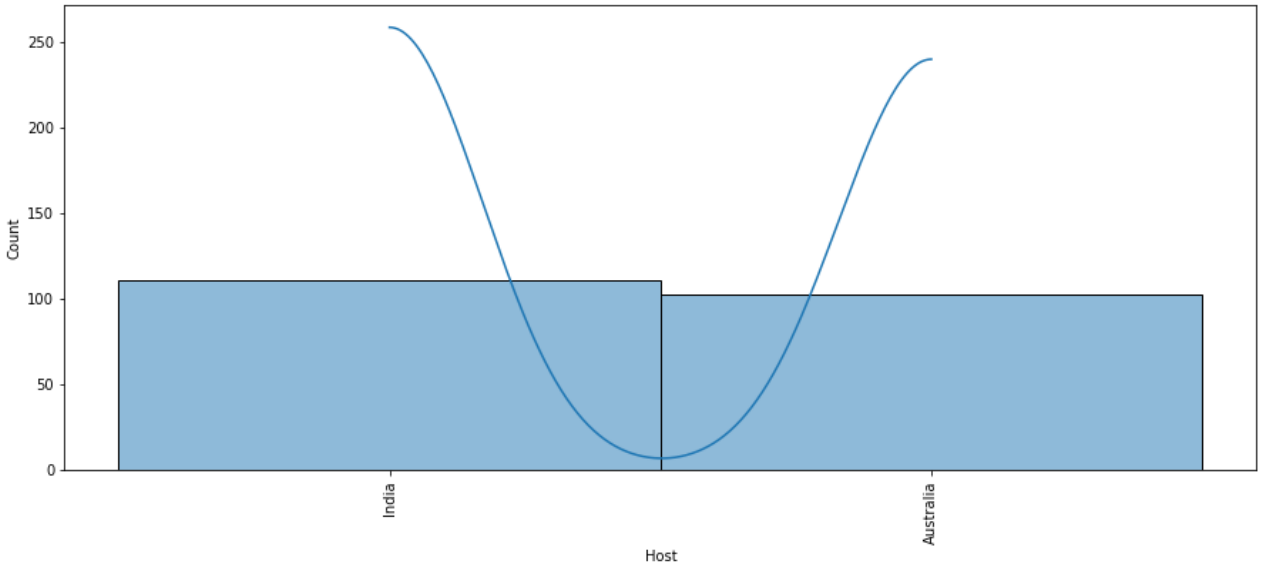
In [41]:
```
for i in df.columns:
    plt.figure(figsize=(15,6))
    sns.histplot(df[i], kde = True, palette = 'hls')
    plt.xticks(rotation = 90)
    plt.show()
```

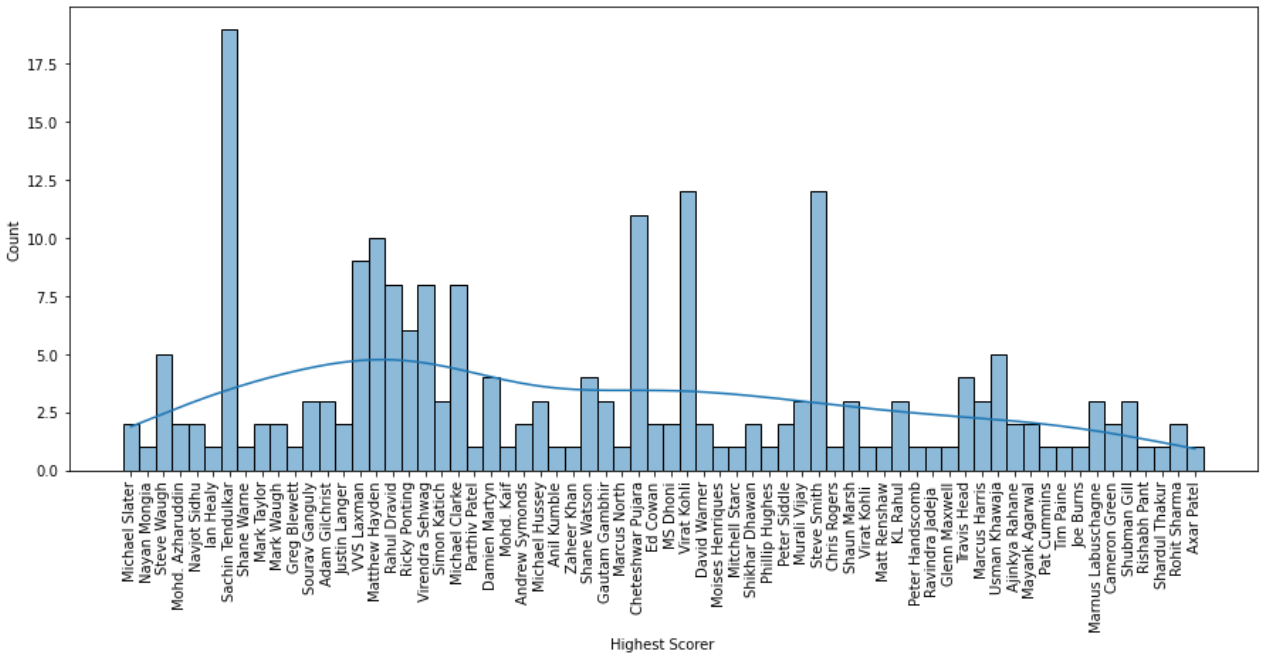

Count
Test Series Year
1996-97
1997-98
1999-00
2000-01
2003-04
2004-05
2007-08
2008-09
2010-11
2011-12
2012-13
2014-15
2016-17
2018-19
2020-21
2022-23

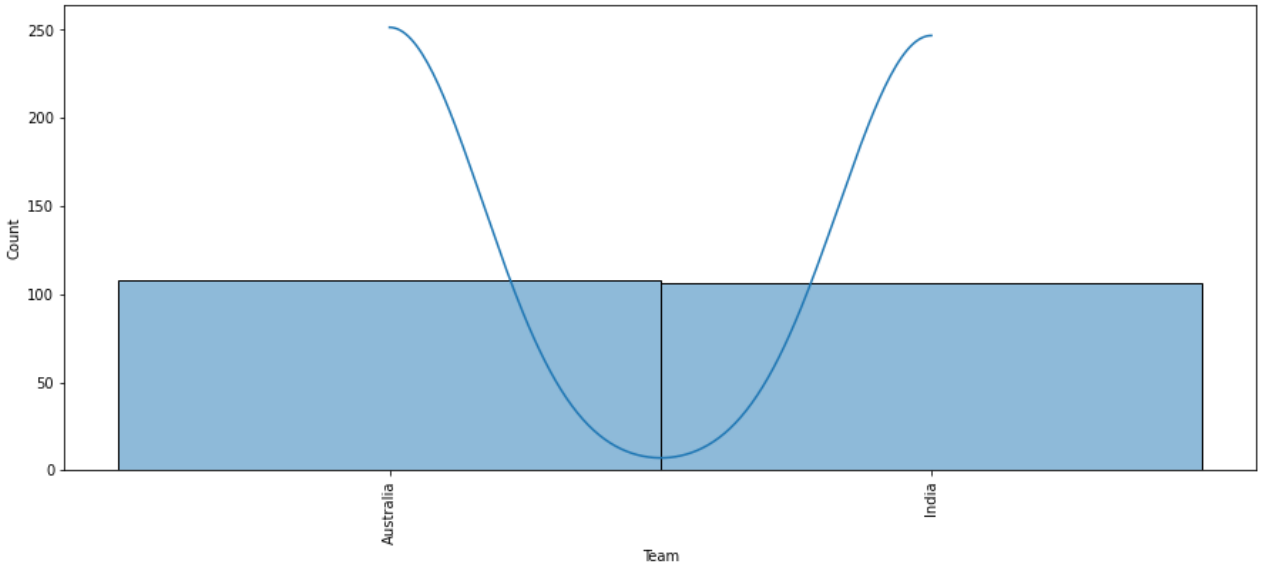

Count
Test Number

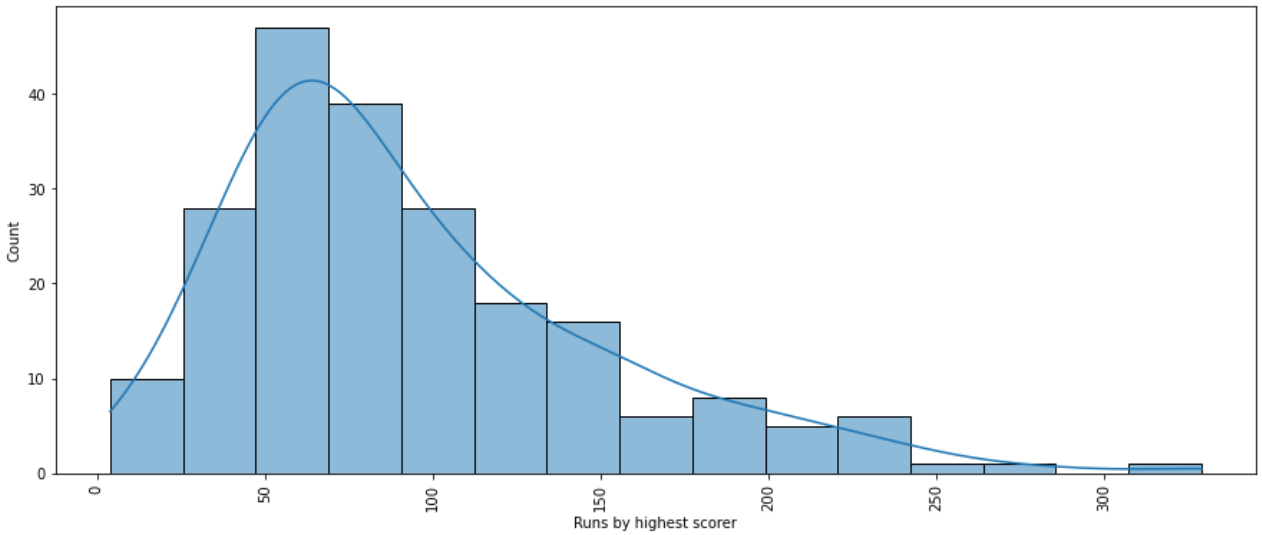

Count
Innings

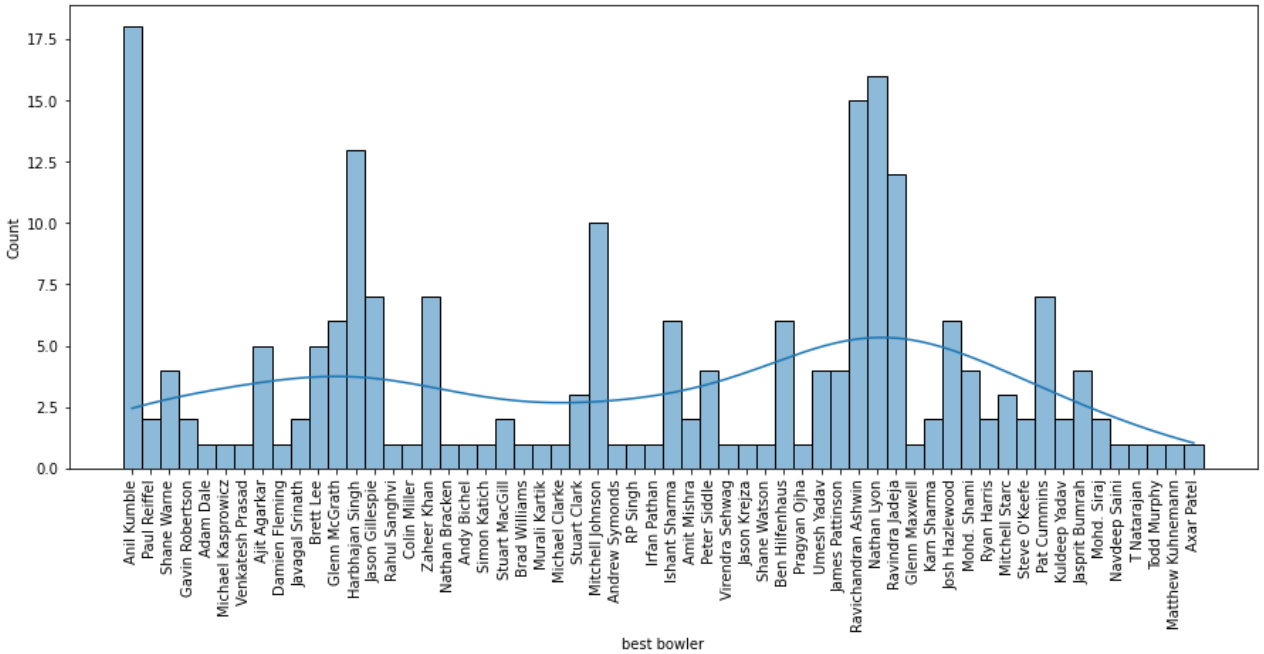

Count
Venue
Delhi
Chennai
Kolkata
Bengaluru
Adelaide
Melbourne
Sydney
Mumbai
Brisbane
Nagpur
Perth
Mohali
Hyderabad
Pune
Ranchi
Dharamsala
Indore
Ahmedabad

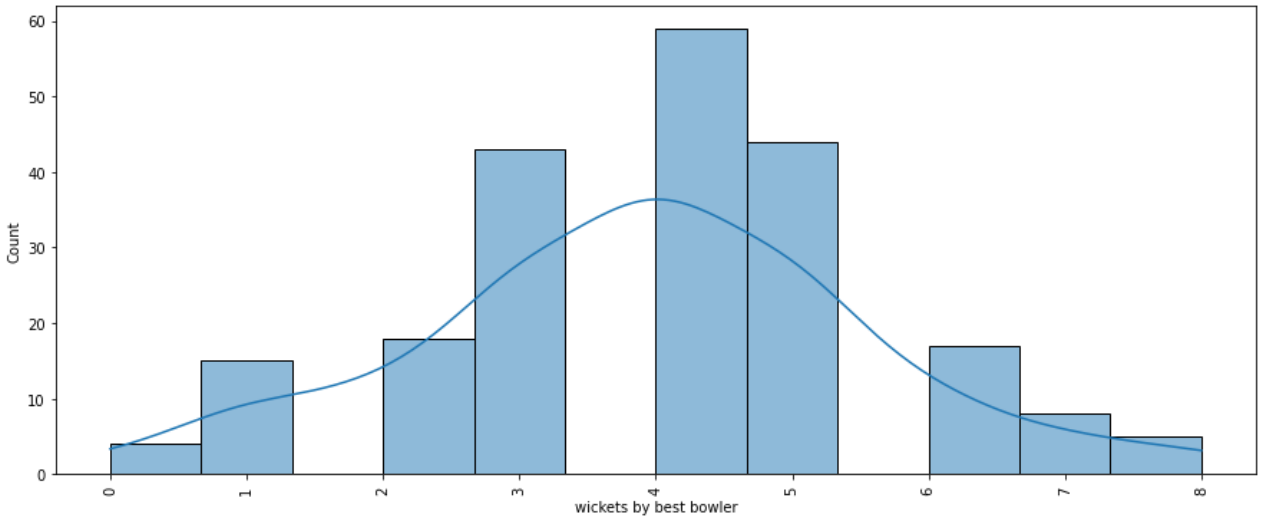

Count
Host
India
Australia

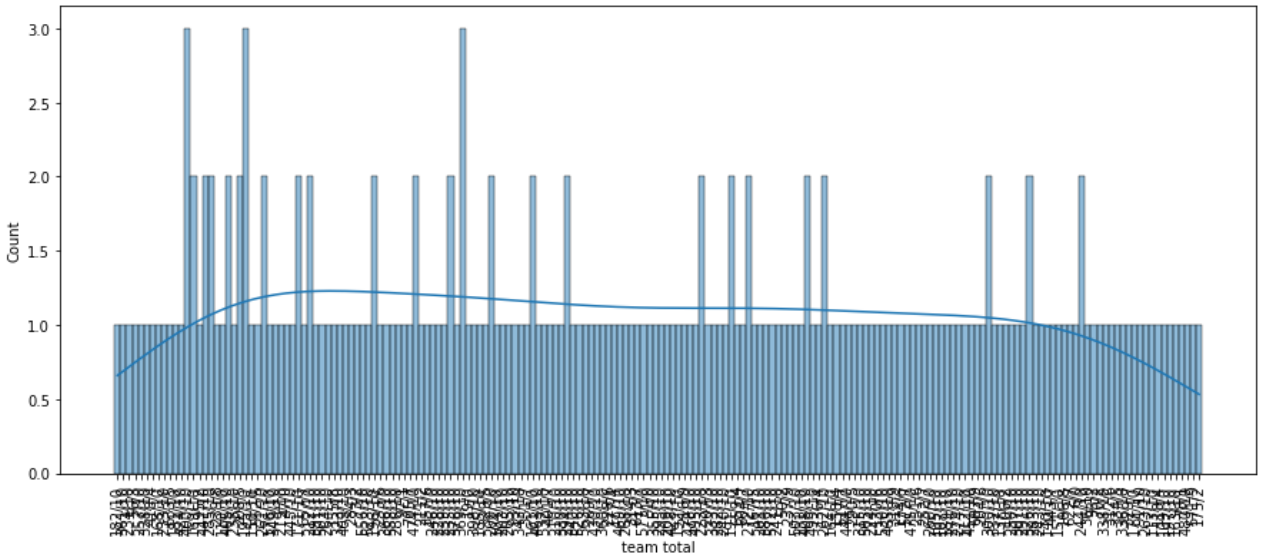

Count
Highest Scorer
Michael Slater
Nayan Mongia
Steve Waugh
Mohd. Azharuddin
Navjot Sidhu
Ian Healy
Sachin Tendulkar
Shane Warne
Mark Taylor
Mark Waugh
Greg Blewett
Sourav Ganguly
Adam Gilchrist
Justin Langer
VVS Laxman
Matthew Hayden
Rahul Dravid
Ricky Ponting
Virendra Sehwag
Simon Katich
Michael Clarke
Parthiv Patel
Damien Martyn
Mohd. Kaif
Andrew Symonds
Michael Hussey
Anil Kumble
Zaheer Khan
Shane Watson
Gautam Gambhir
Marcus North
Cheteshwar Pujara
Ed Cowan
MS Dhoni
Virat Kohli
David Warner
Moises Henriques
Mitchell Starc
Shikhar Dhawan
Phillip Hughes
Peter Siddle
Murali Vijay
Steve Smith
Chris Rogers
Shaun Marsh
Virat Kohli
Matt Renshaw
KL Rahul
Peter Handscomb
Ravindra Jadeja
Glenn Maxwell
Travis Head
Marcus Harris
Usman Khawaja
Ajinkya Rahane
Mayank Agarwal
Pat Cummins
Tim Paine
Joe Burns
Marnus Labuschagne
Cameron Green
Shubman Gill
Rishabh Pant
Shardul Thakur
Rohit Sharma
Axar Patel

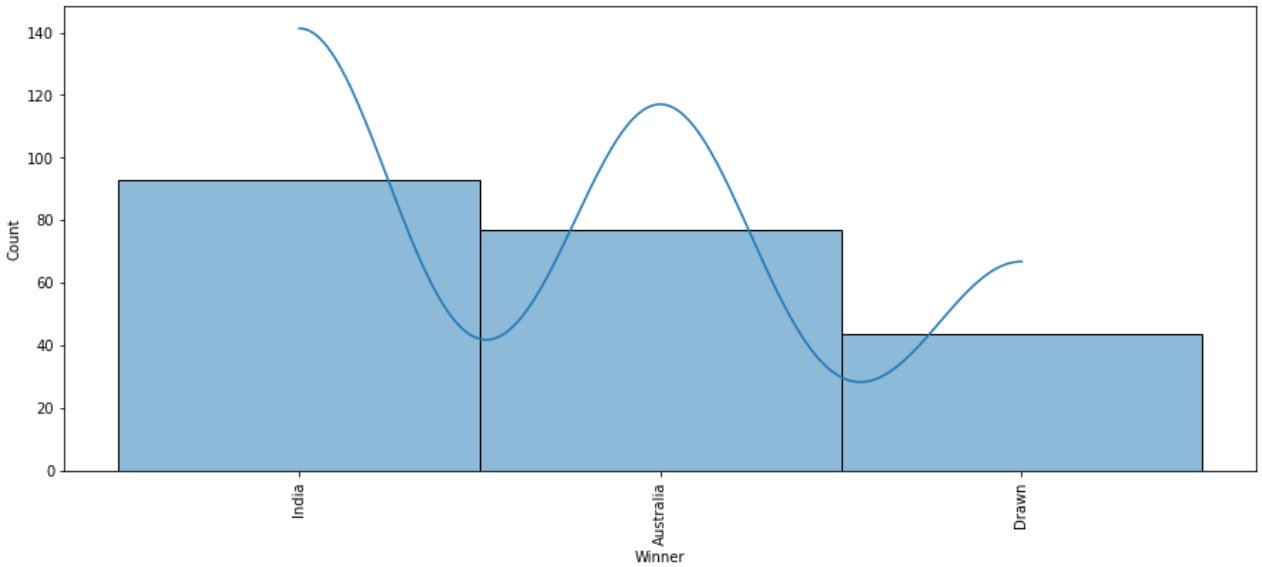

Count
Australia
India
Team

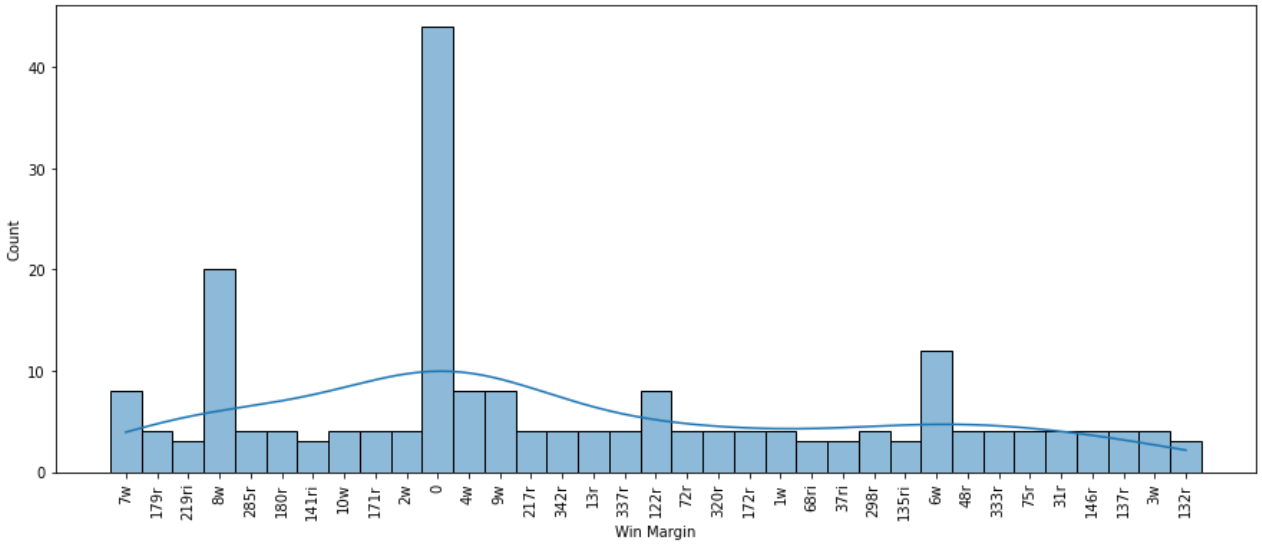

Count
Runs by highest scorer

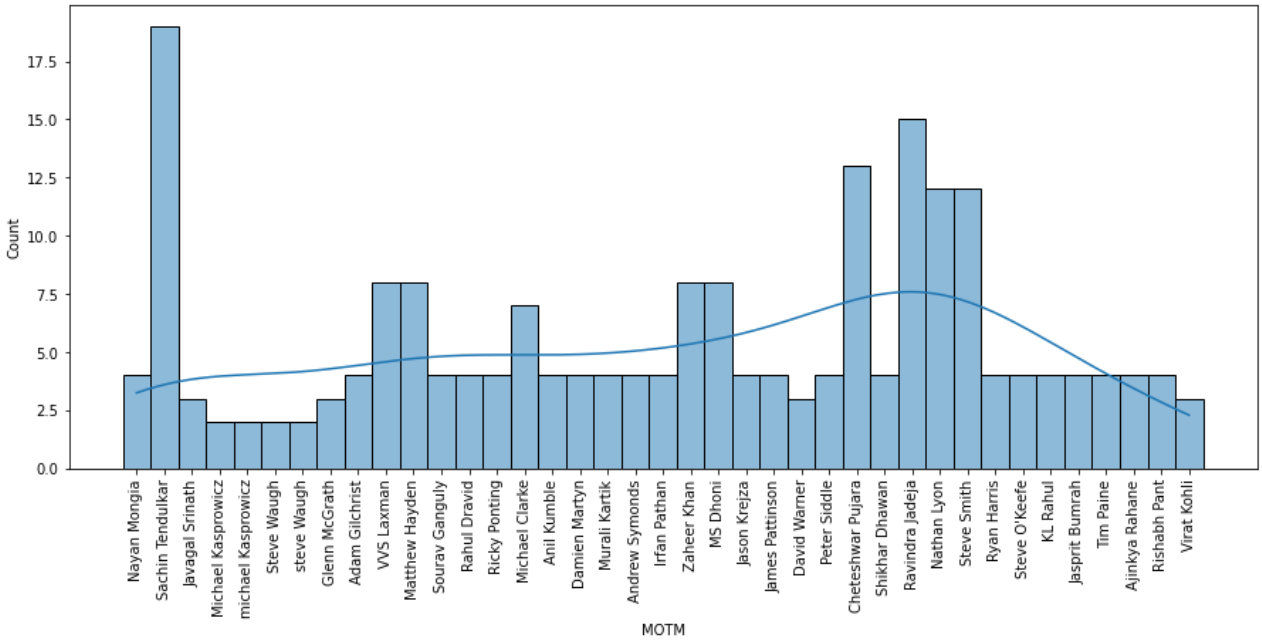

Count
Anil Kumble
Paul Reiffel
Shane Warne
Gavin Robertson
Adam Dale
Michael Kasprowicz
Venkatesh Prasad
Ajit Agarkar
Damien Fleming
Javagal Srinath
Brett Lee
Glenn McGrath
Harbhajan Singh
Jason Gillespie
Rahul Sanghvi
Colin Miller
Zaheer Khan
Nathan Bracken
Andy Bichel
Simon Katich
Stuart MacGill
Brad Williams
Murali Kartik
Michael Clarke
Stuart Clark
Mitchell Johnson
Andrew Symonds
RP Singh
Irfan Pathan
Ishant Sharma
Amit Mishra
Peter Siddle
Virendra Sehwag
Jason Krejza
Shane Watson
Ben Hilfenhaus
Pragyan Ojha
Umesh Yadav
James Pattinson
Ravichandran Ashwin
Nathan Lyon
Ravindra Jadeja
Glenn Maxwell
Karn Sharma
Josh Hazlewood
Mohd. Shami
Ryan Harris
Mitchell Starc
Steve O'Keefe
Pat Cummins
Kuldeep Yadav
Jasprit Bumrah
Mohd. Siraj
Navdeep Saini
T Natarajan
Todd Murphy
Matthew Kuhnemann
Axar Patel
best bowler

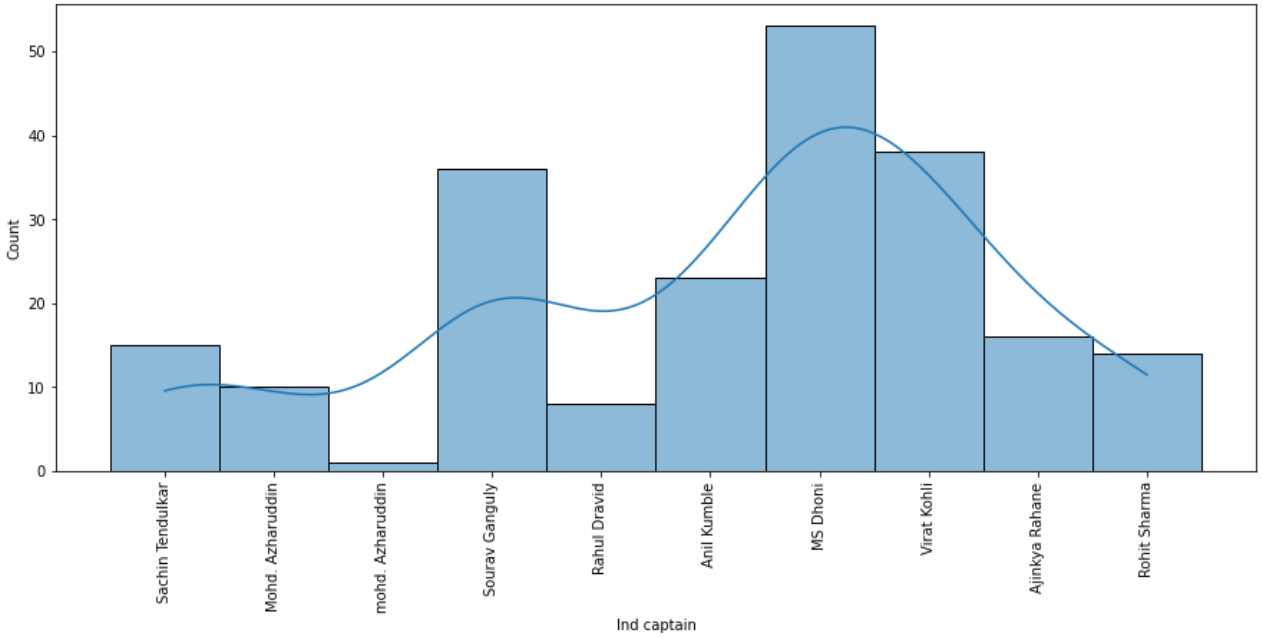

Count
wickets by best bowler

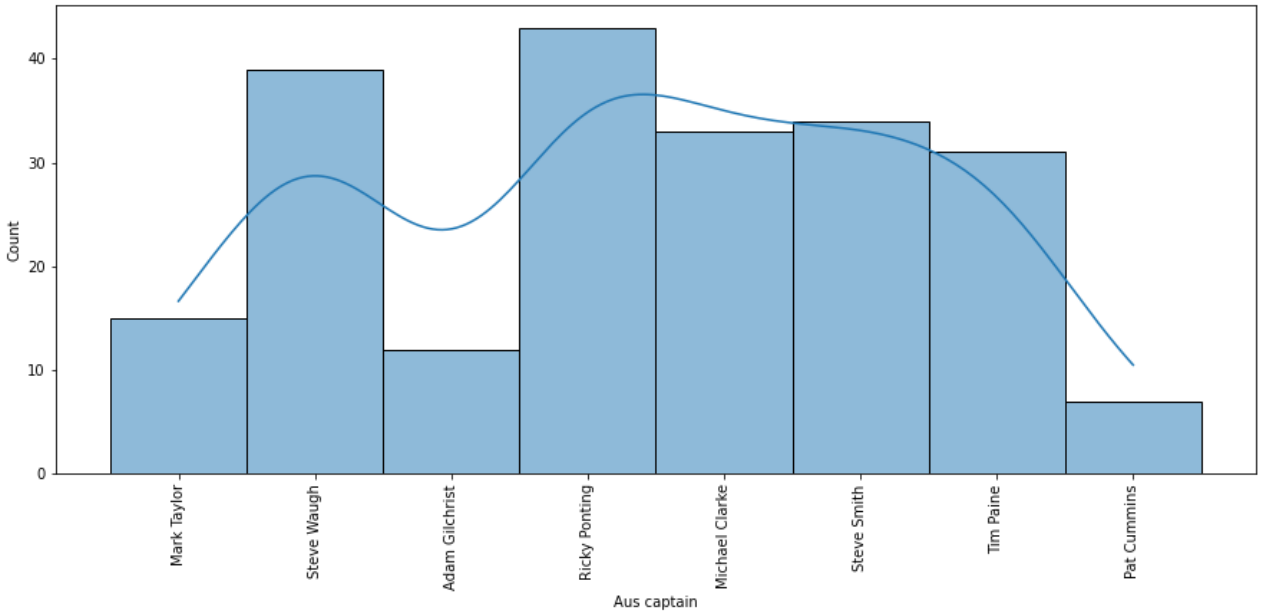

Count
team total

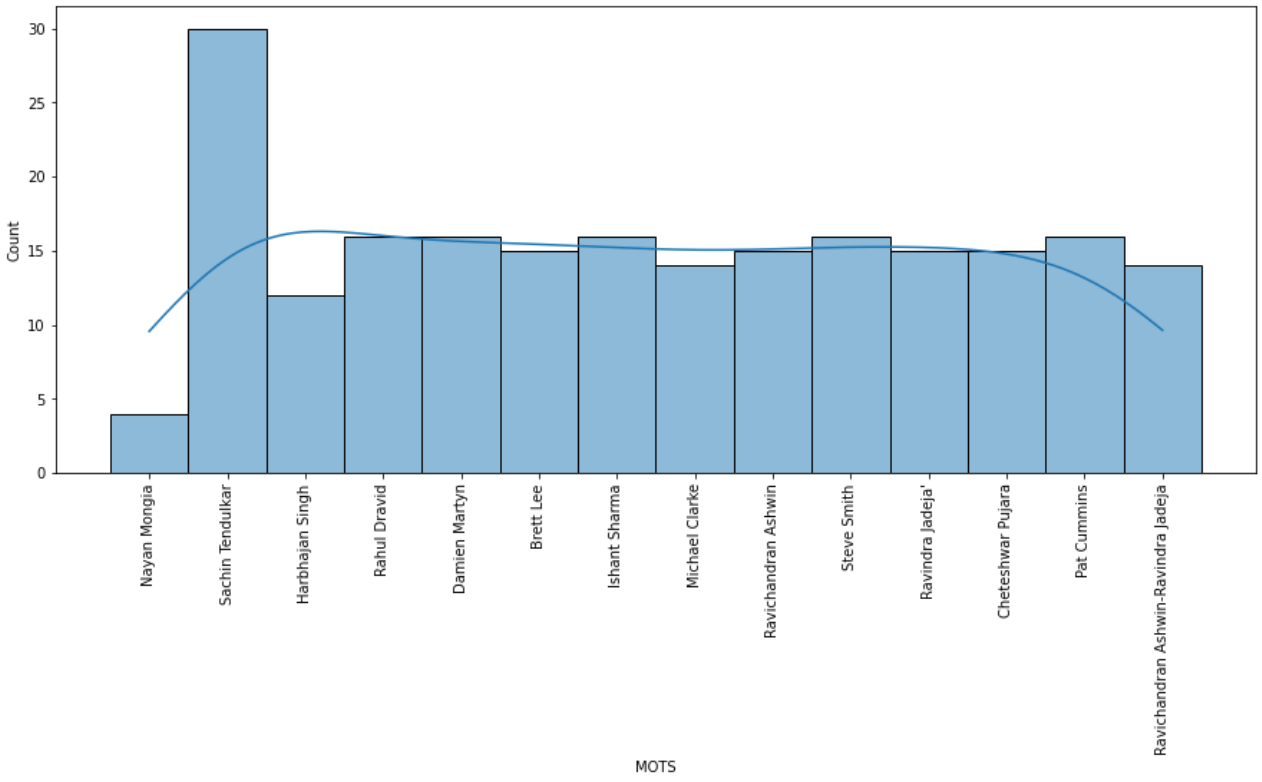

Count
India
Australia
Drawn
Winner

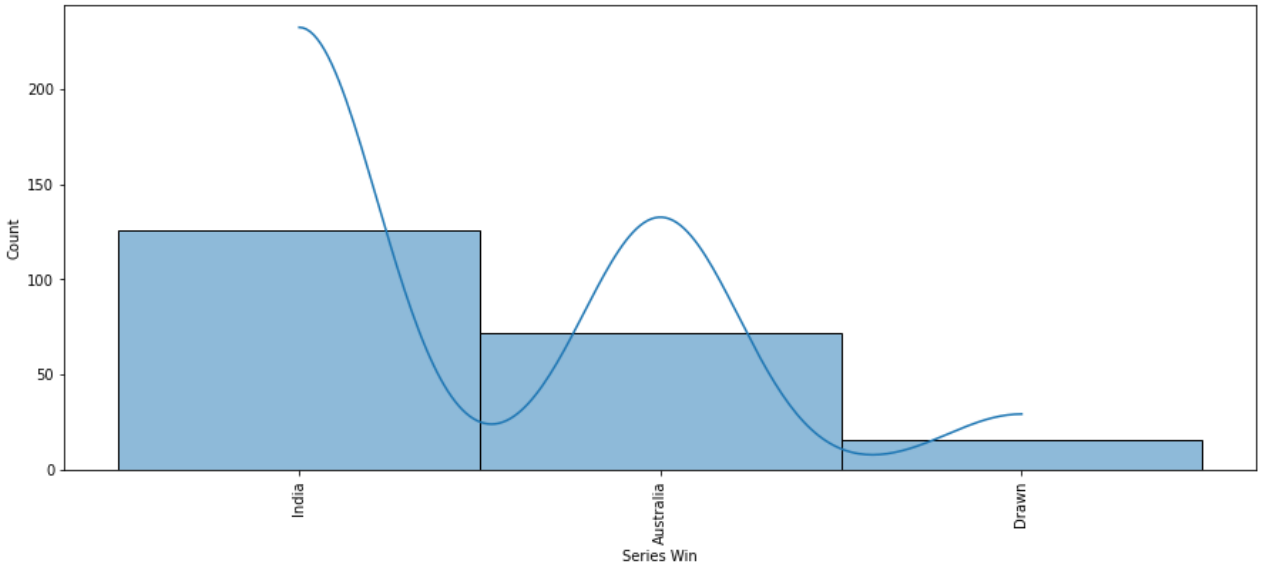

Count
Win Margin
7w
179r
219ri
8w
285r
180r
141ri
10w
171r
2w
0
4w
9w
217r
342r
13r
337r
122r
72r
320r
172r
1w
68ri
37ri
298r
135ri
6w
48r
333r
75r
31r
146r
137r
3w
132r


Count
MOTM
Nayan Mongia
Sachin Tendulkar
Javagal Srinath
Michael Kasprowicz
michael Kasprowicz
Steve Waugh
steve Waugh
Glenn McGrath
Adam Gilchrist
VVS Laxman
Matthew Hayden
Sourav Ganguly
Rahul Dravid
Ricky Ponting
Michael Clarke
Anil Kumble
Damien Martyn
Murali Kartik
Andrew Symonds
Irfan Pathan
Zaheer Khan
MS Dhoni
Jason Krejza
James Pattinson
David Warner
Peter Siddle
Cheteshwar Pujara
Shikhar Dhawan
Ravindra Jadeja
Nathan Lyon
Steve Smith
Ryan Harris
Steve O'Keefe
KL Rahul
Jasprit Bumrah
Tim Paine
Ajinkya Rahane
Rishabh Pant
Virat Kohli

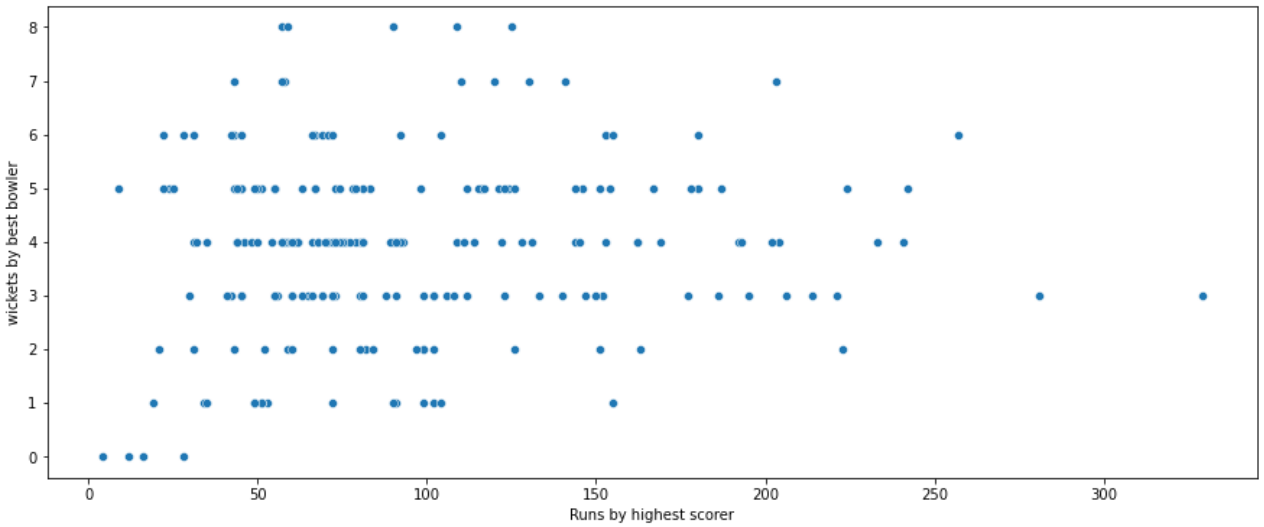

Count
Ind captain
Sachin Tendulkar
Mohd. Azharuddin
mohd. Azharuddin
Sourav Ganguly
Rahul Dravid
Anil Kumble
MS Dhoni
Virat Kohli
Ajinkya Rahane
Rohit Sharma


Count
0
10
20
30
40
Mark Taylor
Steve Waugh
Adam Gilchrist
Ricky Ponting
Michael Clarke
Steve Smith
Tim Paine
Pat Cummins
Aus captain

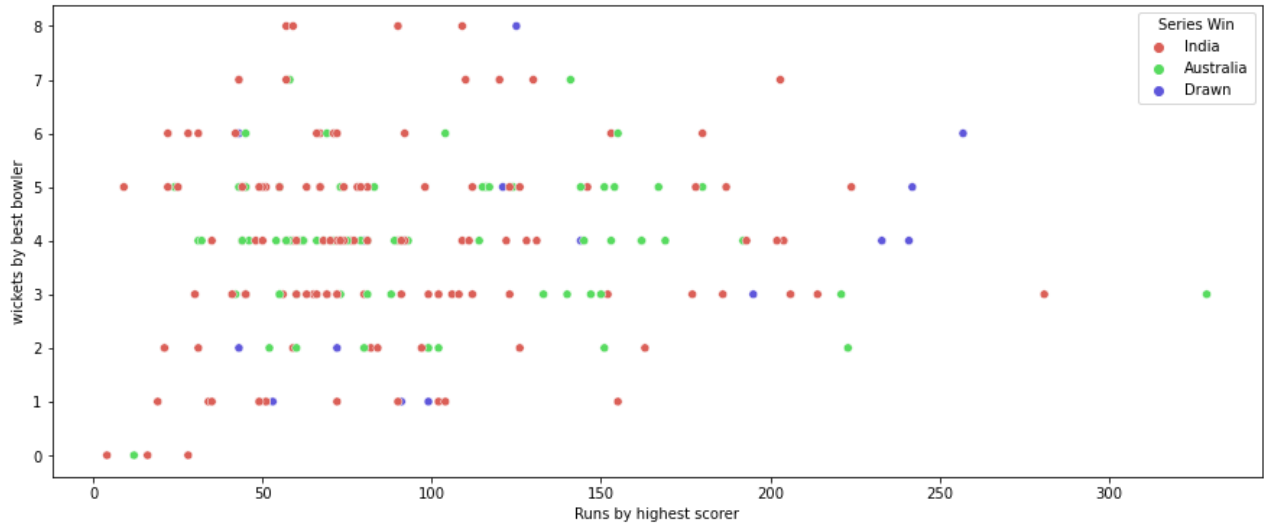

Count
0
5
10
15
20
25
30
Nayan Mongia
Sachin Tendulkar
Harbhajan Singh
Rahul Dravid
Damien Martyn
Brett Lee
Ishant Sharma
Michael Clarke
Ravichandran Ashwin
Steve Smith
Ravindra Jadeja'
Cheteshwar Pujara
Pat Cummins
Ravichandran Ashwin-Ravindra Jadeja
MOTS



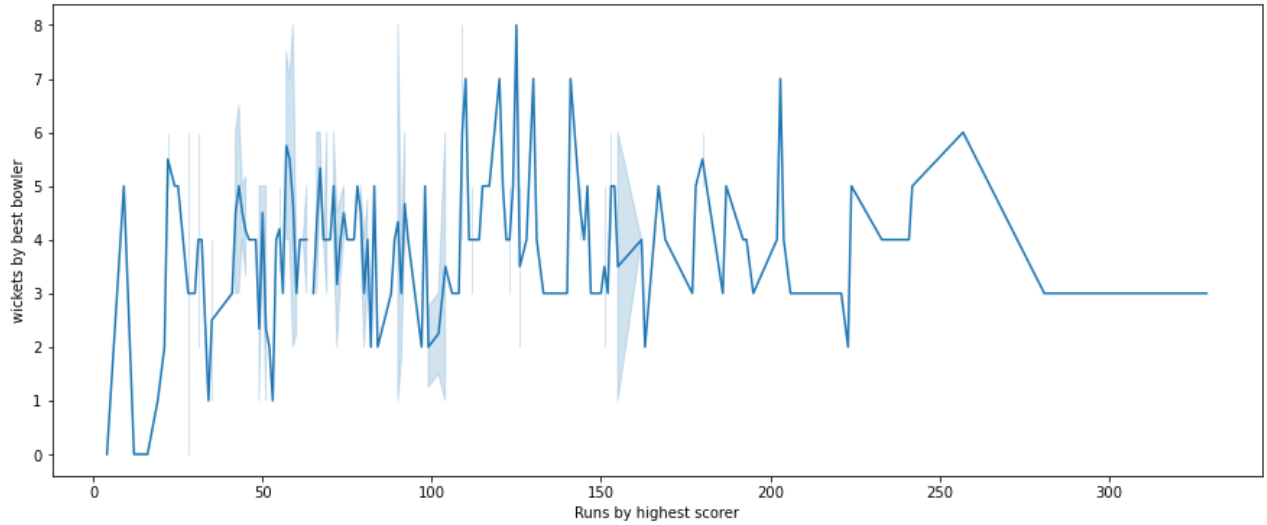
In [42]:
```
plt.figure(figsize=(15,6))
sns.scatterplot(x = df['Runs by highest scorer'], y = df['wickets by best bowler'], pal
plt.show()
```



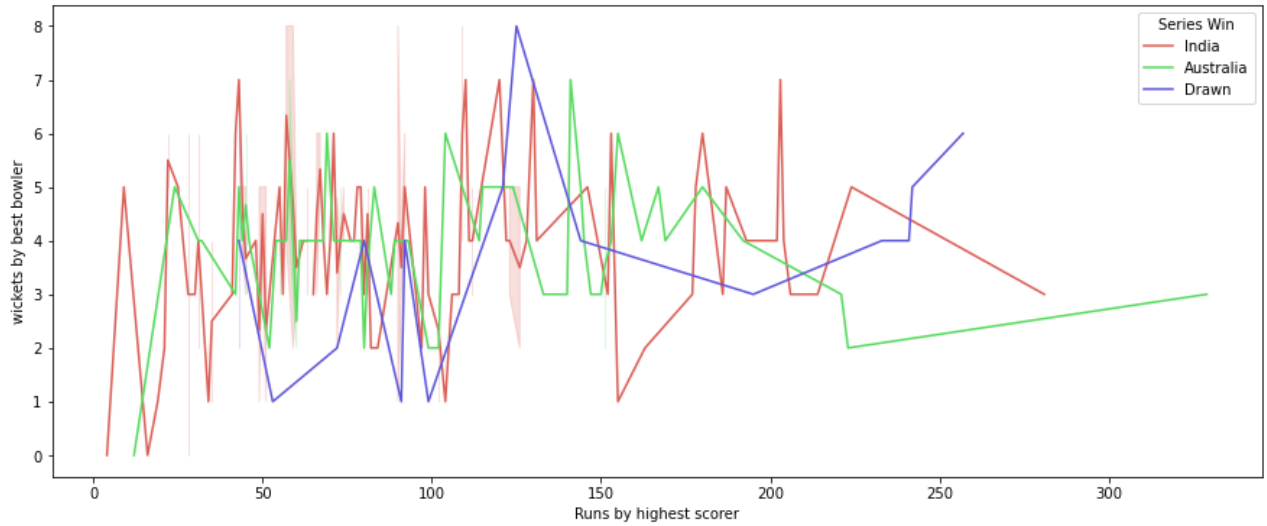
In [44]:
```
plt.figure(figsize=(15,6))
sns.scatterplot(x = df['Runs by highest scorer'], y = df['wickets by best bowler'], hue
plt.show()
```



In [43]:
```
plt.figure(figsize=(15,6))
sns.lineplot(x = df['Runs by highest scorer'], y = df['wickets by best bowler'], palett
plt.show()
```



In [45]:
```
plt.figure(figsize=(15,6))
sns.lineplot(x = df['Runs by highest scorer'], y = df['wickets by best bowler'], hue=df
plt.show()
```



In [52]:
```
df1 = df.groupby(["Test Series Year","Test Number"])["Winner","Host"].first()
```

In [53]:
```
df1.head()
```

Out[53]:

		Winner	Host
Test Series Year	**Test Number**		
1996-97	**1**	India	India
1997-98	**1**	India	India
	2	India	India
	3	Australia	India
1999-00	**1**	Australia	Australia

In [54]:
```
df1=df1.reset_index()
```

In [60]:
```
df1
```

Out[60]:

	Test Series Year	Test Number	Winner	Host
0	1996-97	1	India	India
1	1997-98	1	India	India
2	1997-98	2	India	India
3	1997-98	3	Australia	India
4	1999-00	1	Australia	Australia
5	1999-00	2	Australia	Australia
6	1999-00	3	Australia	Australia
7	2000-01	1	Australia	India
8	2000-01	2	India	India
9	2000-01	3	India	India
10	2003-04	1	Drawn	Australia
11	2003-04	2	India	Australia
12	2003-04	3	Australia	Australia
13	2003-04	4	Drawn	Australia
14	2004-05	1	Australia	India
15	2004-05	2	Drawn	India
16	2004-05	3	Australia	India
17	2004-05	4	India	India
18	2007-08	1	Australia	Australia
19	2007-08	2	Australia	Australia

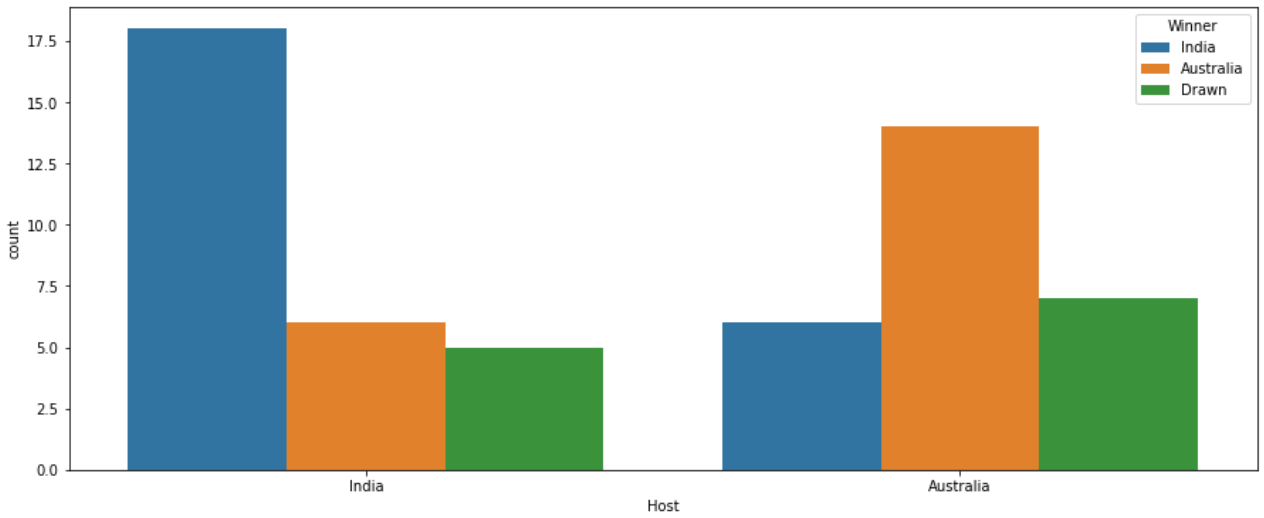
	Test Series Year	Test Number	Winner	Host
20	2007-08	3	India	Australia
21	2007-08	4	Drawn	Australia
22	2008-09	1	Drawn	India
23	2008-09	2	India	India
24	2008-09	3	Drawn	India
25	2008-09	4	India	India
26	2010-11	1	India	India
27	2010-11	2	India	India
28	2011-12	1	Australia	Australia
29	2011-12	2	Australia	Australia
30	2011-12	3	Australia	Australia
31	2011-12	4	Australia	Australia
32	2012-13	1	India	India
33	2012-13	2	India	India
34	2012-13	3	India	India
35	2012-13	4	India	India
36	2014-15	1	Australia	Australia
37	2014-15	2	Australia	Australia
38	2014-15	3	Drawn	Australia
39	2014-15	4	Drawn	Australia
40	2016-17	1	Australia	India
41	2016-17	2	India	India
42	2016-17	3	Drawn	India
43	2016-17	4	India	India
44	2018-19	1	India	Australia
45	2018-19	2	Australia	Australia
46	2018-19	3	India	Australia
47	2018-19	4	Drawn	Australia
48	2020-21	1	Australia	Australia
49	2020-21	2	India	Australia
50	2020-21	3	Drawn	Australia
51	2020-21	4	India	Australia
52	2022-23	1	India	India

	Test Series Year	Test Number	Winner	Host
53	2022-23	2	India	India
54	2022-23	3	Australia	India
55	2022-23	4	Drawn	India

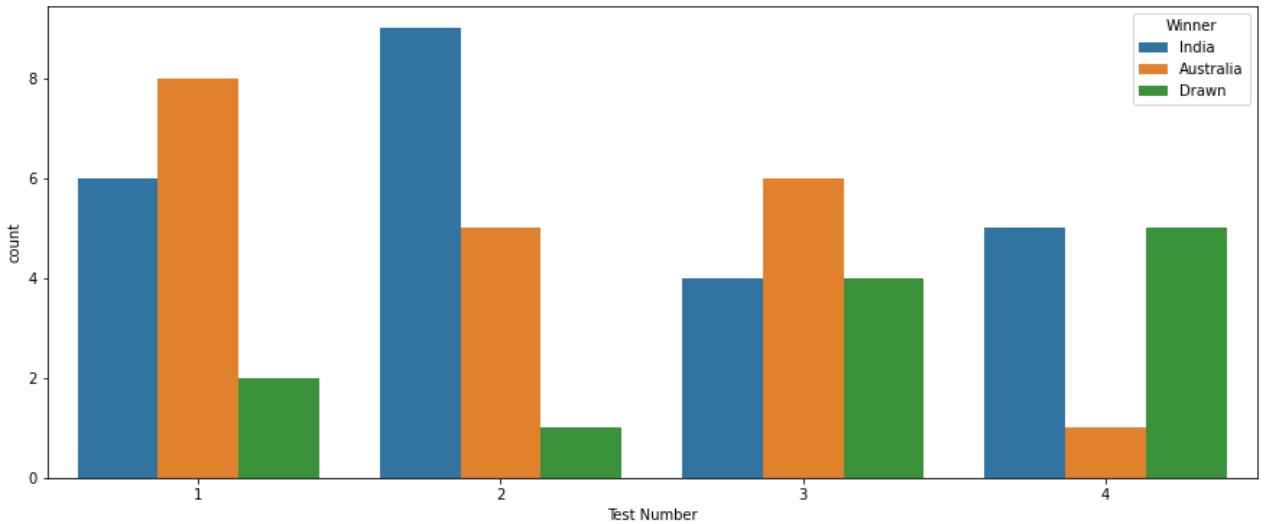
In [58]:
```
df1.shape
```

Out[58]: (56, 4)

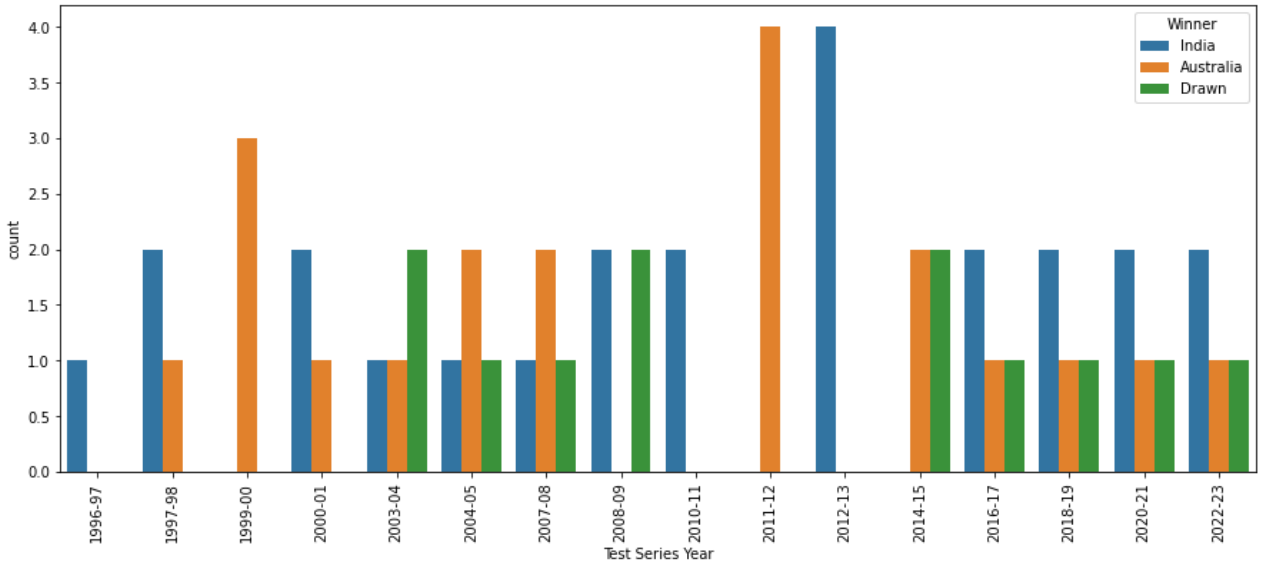
In [61]:
```
plt.figure(figsize=(15,6))
sns.countplot(x = df1["Host"], hue = df1["Winner"], data=df1)
plt.show()
```



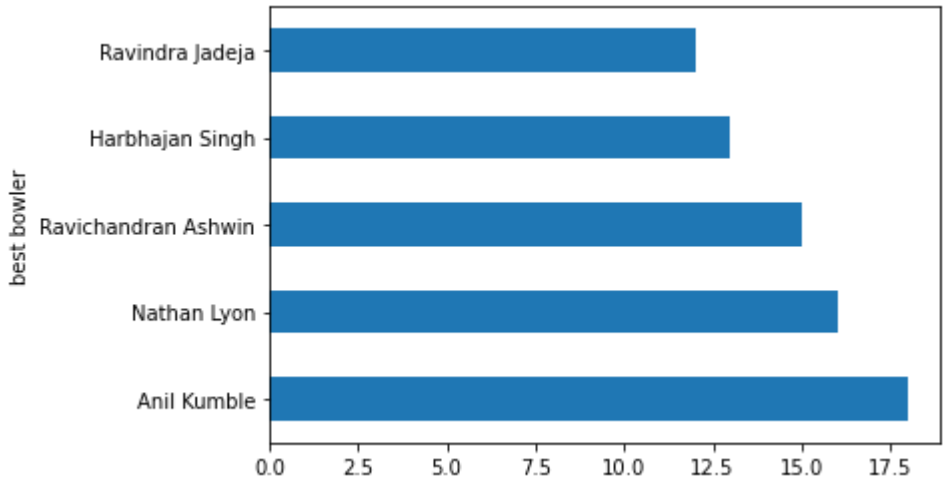
In [62]:
```
plt.figure(figsize=(15,6))
sns.countplot(x = df1["Test Number"], hue = df1["Winner"], data=df1)
plt.show()
```



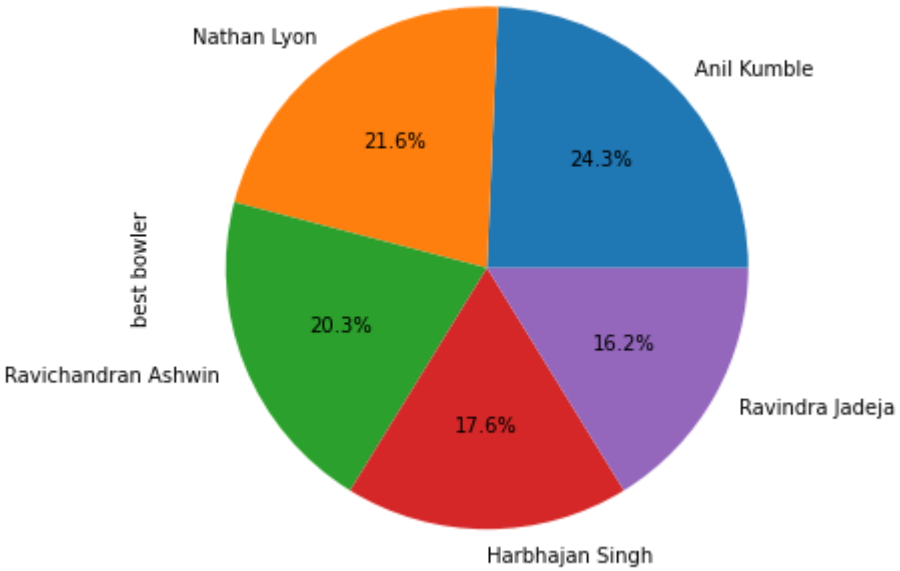
In [65]:
```
plt.figure(figsize=(15,6))
sns.countplot(x = df1["Test Series Year"], hue = df1["Winner"], data=df1)
plt.xticks(rotation=90)
plt.show()
```



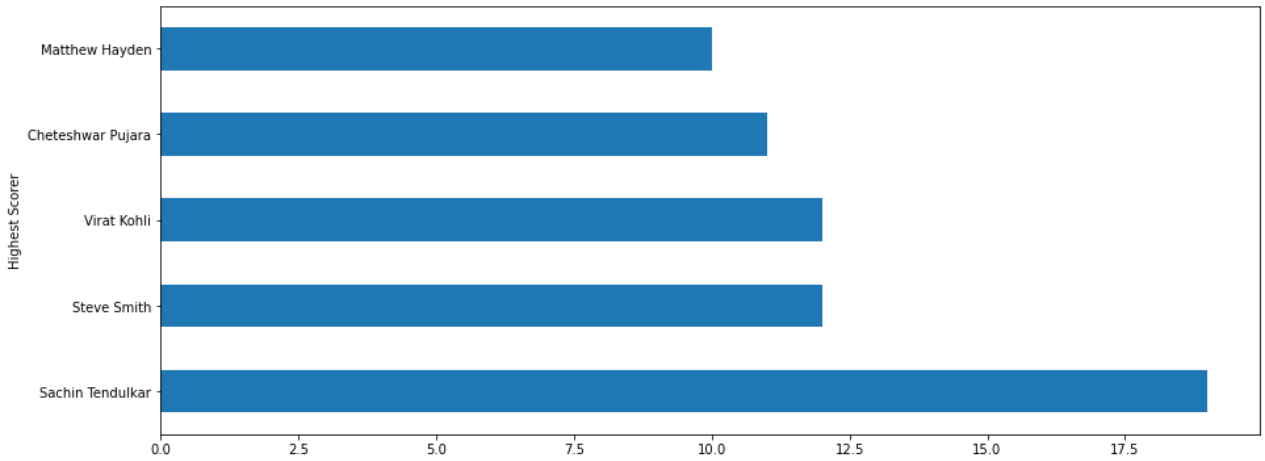
In [73]:
```
df.groupby("best bowler")["best bowler"].count().sort_values(ascending = False).head().
plt.show()
```



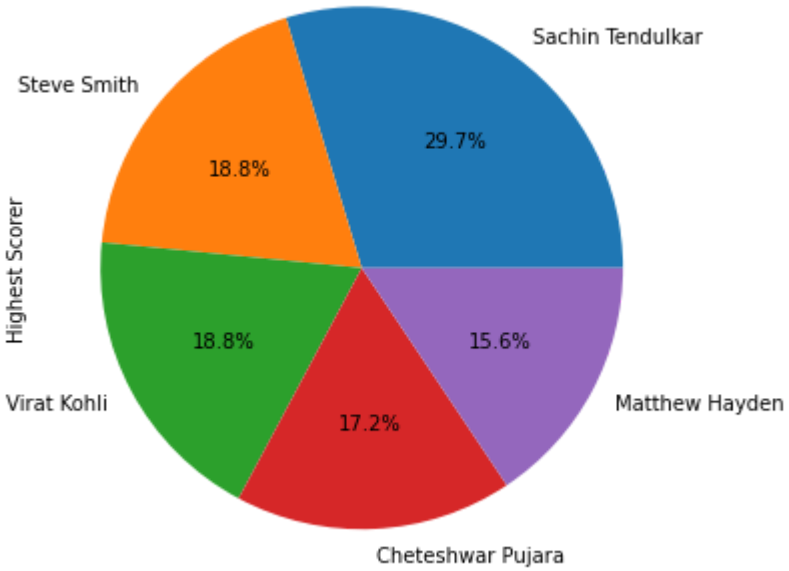
In [77]:
```
plt.figure(figsize=(15,6))
df.groupby("best bowler")["best bowler"].count().sort_values(ascending = False).head().
plt.show()
```



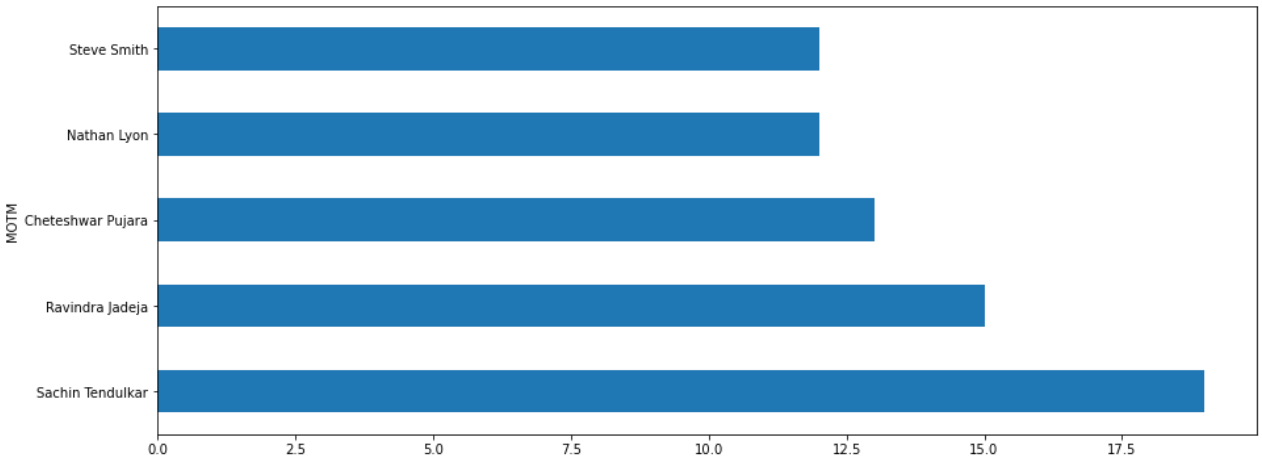
In [75]:
```
plt.figure(figsize=(15,6))
df.groupby("Highest Scorer")["Highest Scorer"].count().sort_values(ascending = False).h
plt.show()
```



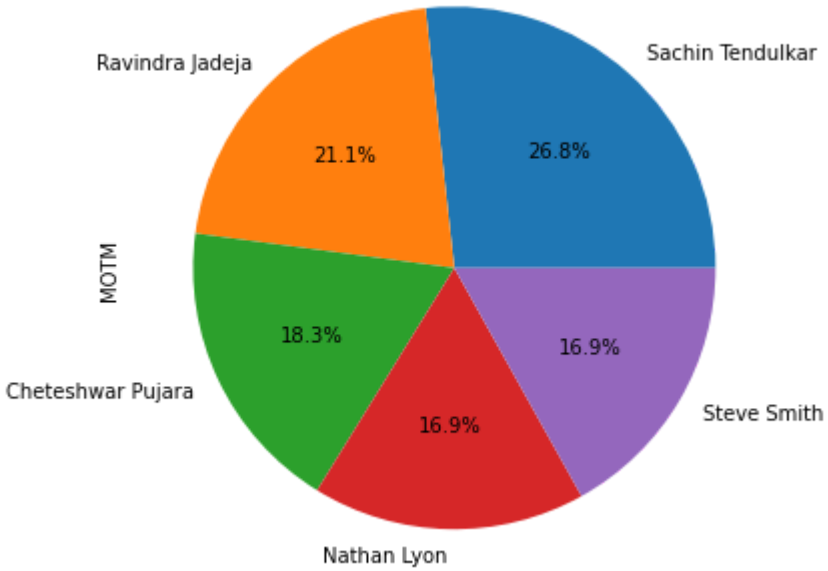
In [76]:
```
plt.figure(figsize=(15,6))
df.groupby("Highest Scorer")["Highest Scorer"].count().sort_values(ascending = False).h
plt.show()
```



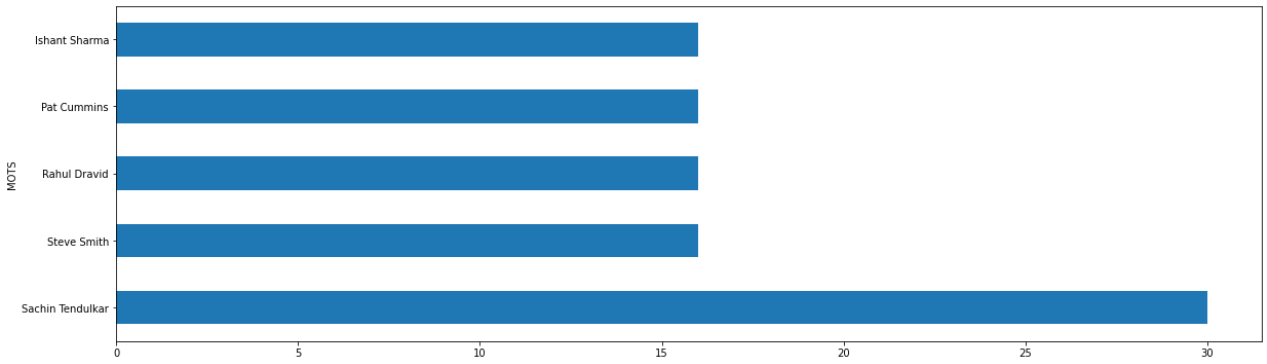
In [78]:
```
plt.figure(figsize=(15,6))
df.groupby("MOTM")["MOTM"].count().sort_values(ascending = False).head().plot(kind="bar
plt.show()
```



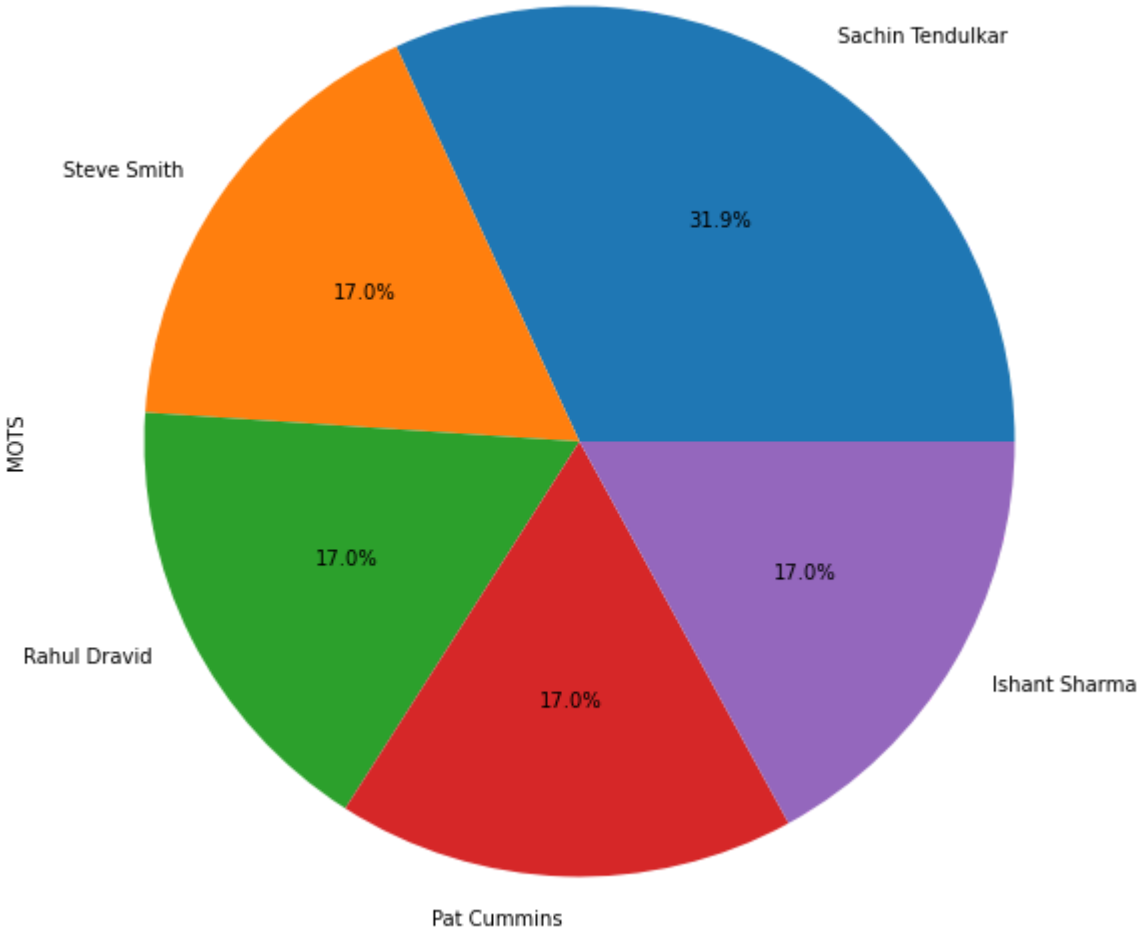
In [79]:
```
plt.figure(figsize=(15,6))
df.groupby("MOTM")["MOTM"].count().sort_values(ascending = False).head().plot(kind="pie
plt.show()
```



In [81]:
```
plt.figure(figsize=(20, 6))
df.groupby("MOTS")["MOTS"].count().sort_values(ascending = False).head().plot(kind="bar
plt.show()
```



In [82]:
```
plt.figure(figsize=(20, 10))
df.groupby("MOTS")["MOTS"].count().sort_values(ascending = False).head().plot(kind="pie
plt.show()
```



In [83]:
```
df2=df.groupby(["Test Series Year","Test Number"])["Winner","Ind captain","Aus captain"
```

In [84]:
```
df2.head()
```

Out[84]:

Test Series Year	Test Number	Winner	Ind captain	Aus captain
1996-97	**1**	India	Sachin Tendulkar	Mark Taylor
1997-98	**1**	India	Mohd. Azharuddin	Mark Taylor
	2	India	Mohd. Azharuddin	Mark Taylor
	3	Australia	Mohd. Azharuddin	Mark Taylor
1999-00	**1**	Australia	Sachin Tendulkar	Steve Waugh

In [85]:
```
df2 = df2.reset_index()
```

In [86]:
```
df2.head()
```

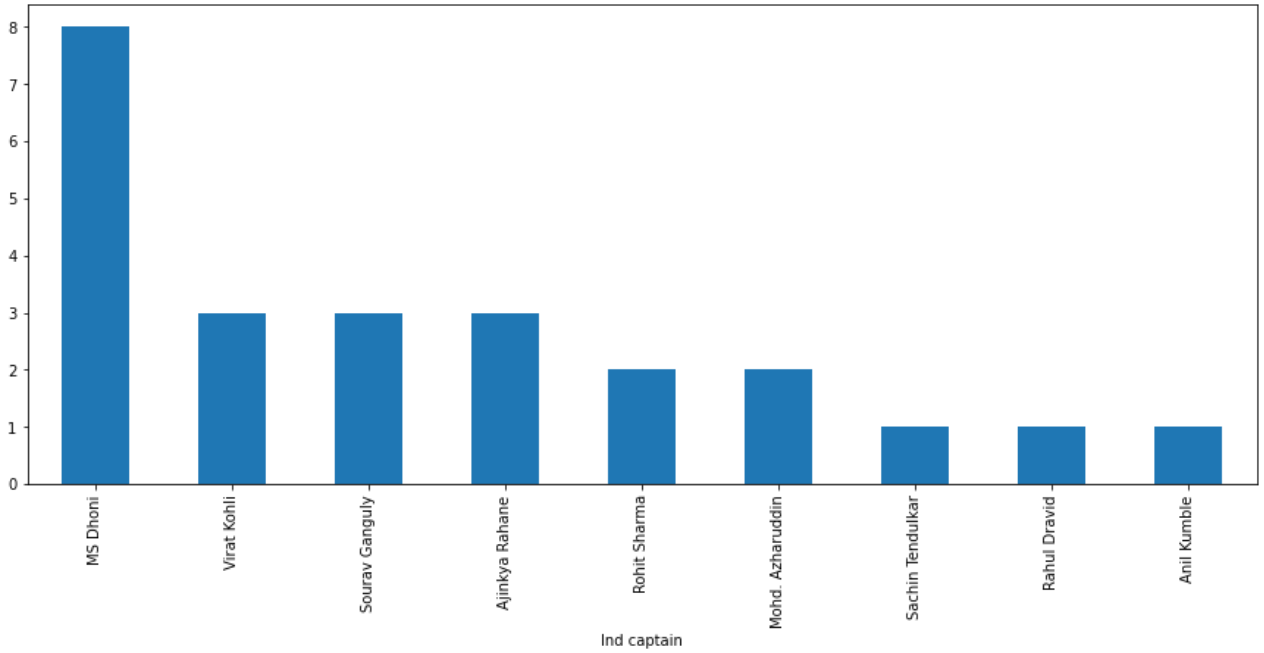
Out[86]:

	Test Series Year	Test Number	Winner	Ind captain	Aus captain
0	1996-97	1	India	Sachin Tendulkar	Mark Taylor
1	1997-98	1	India	Mohd. Azharuddin	Mark Taylor
2	1997-98	2	India	Mohd. Azharuddin	Mark Taylor
3	1997-98	3	Australia	Mohd. Azharuddin	Mark Taylor
4	1999-00	1	Australia	Sachin Tendulkar	Steve Waugh

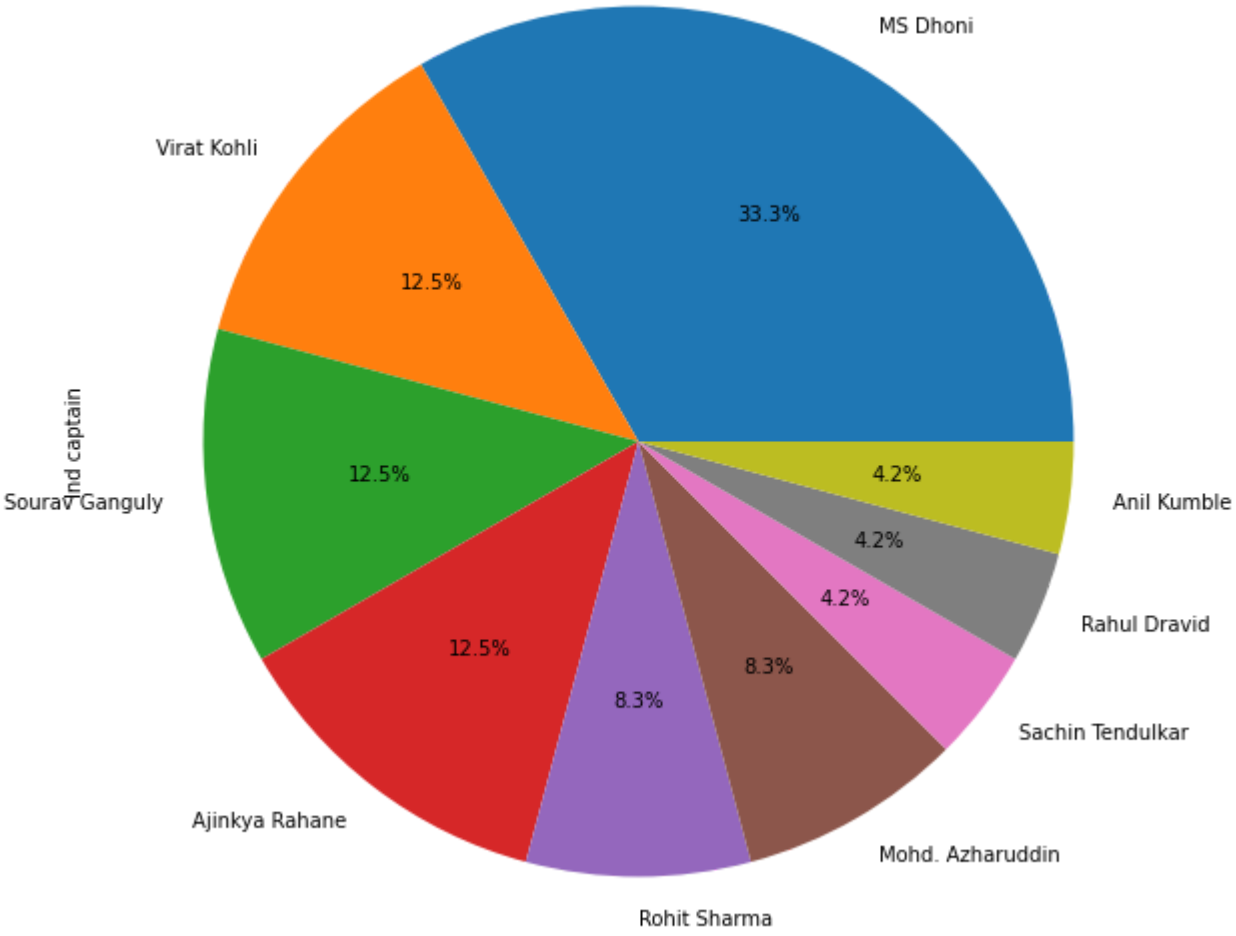
In [87]:
```
df2.shape
```

Out[87]:
```
(56, 5)
```

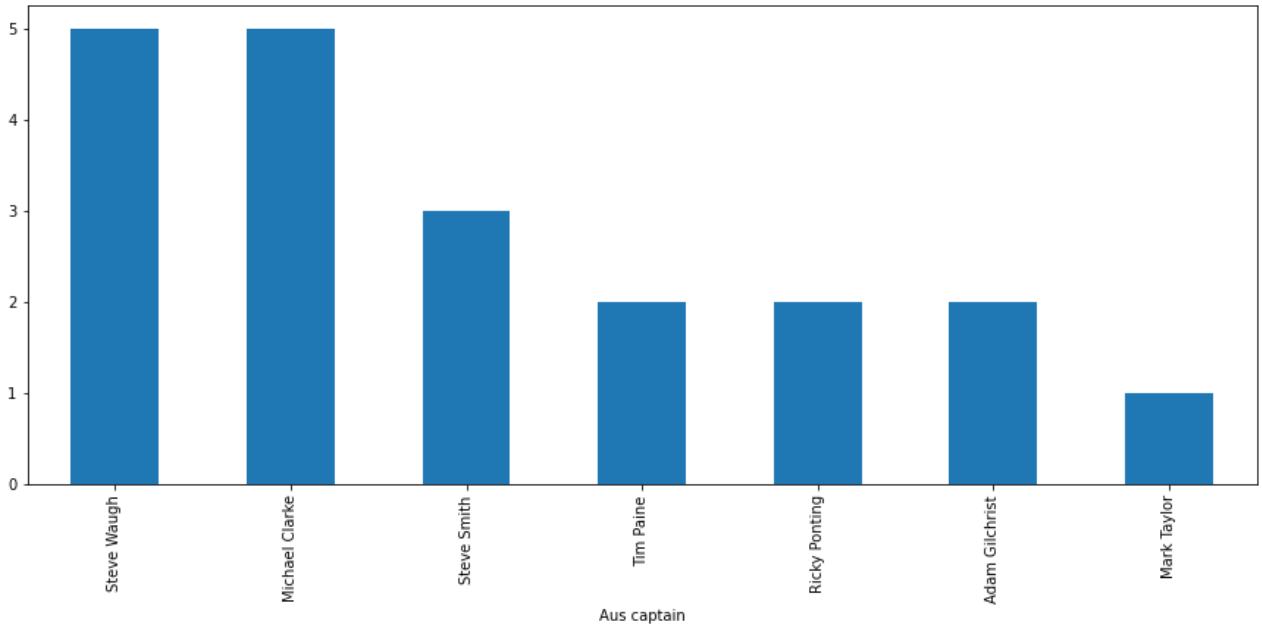
In [88]:
```
ind_W = df2[df2["Winner"]=="India"]
plt.figure(figsize=(15,6))
ind_W.groupby("Ind captain")["Ind captain"].count().sort_values(ascending=False).plot(k
plt.show()
```



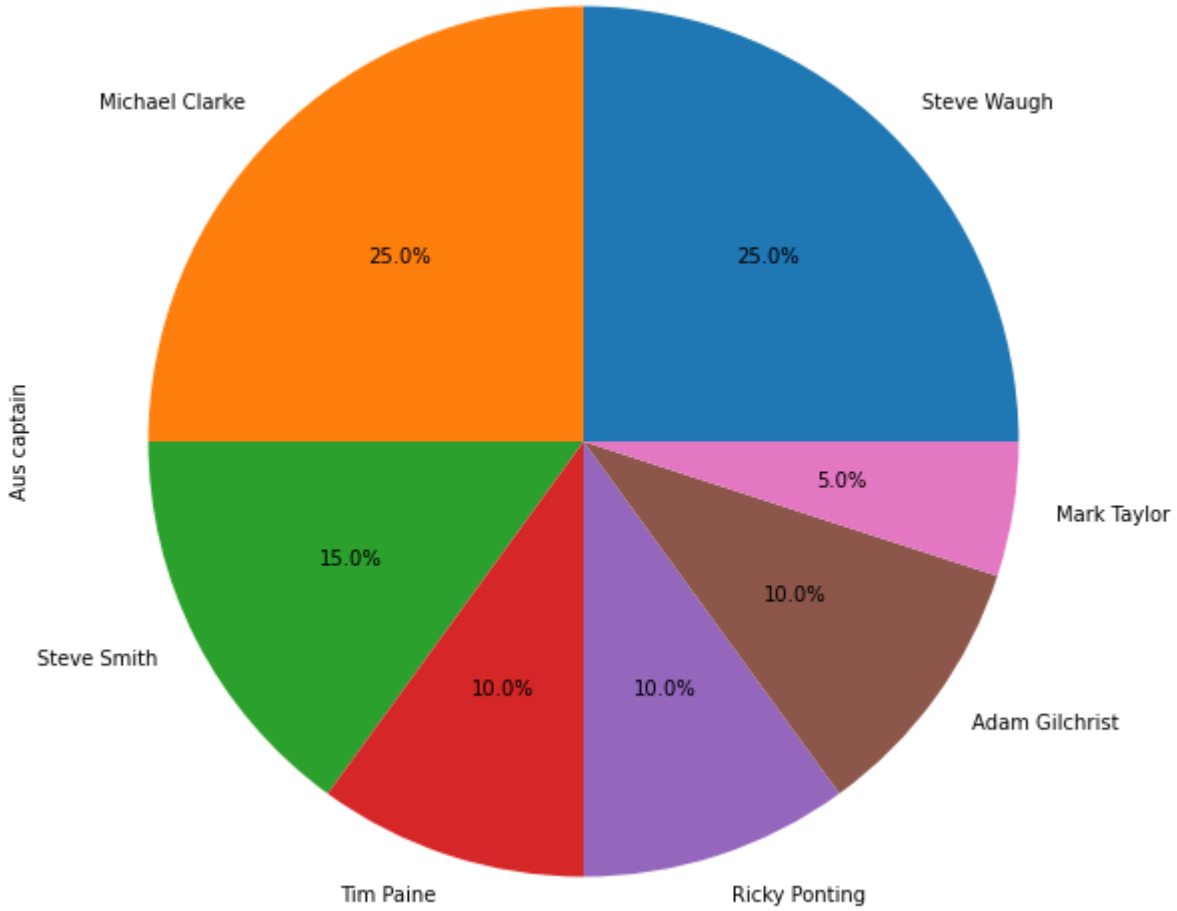
In [89]:
```
plt.figure(figsize=(30, 10))
ind_W.groupby("Ind captain")["Ind captain"].count().sort_values(ascending=False).plot(k
plt.show()
```



In [90]:
```
aus_W = df2[df2["Winner"]=="Australia"]
plt.figure(figsize=(15,6))
aus_W.groupby("Aus captain")["Aus captain"].count().sort_values(ascending=False).plot(k
plt.show()
```



In [92]:
```
plt.figure(figsize=(30, 10))
aus_W.groupby("Aus captain")["Aus captain"].count().sort_values(ascending=False).plot(k
plt.show()
```



In [93]:
```
Runs=[]
Wics=[]
run_wic=df['team total'].str.split("/").tolist()
print(run_wic)
```

```
[['182', '10'], ['361', '10'], ['234', '10'], ['56', '3'], ['257', '10'], ['328', '10'],
['418', '4'], ['168', '10'], ['233', '10'], ['633', '5'], ['181', '10'], ['424', '10'],
['400', '10'], ['169', '10'], ['195', '2'], ['441', '10'], ['285', '10'], ['239', '8'],
['110', '10'], ['405', '10'], ['238', '10'], ['208', '5'], ['195', '10'], ['150', '10'],
['552', '5'], ['261', '10'], ['176', '10'], ['349', '10'], ['219', '10'], ['47', '0'],
['445', '10'], ['171', '10'], ['657', '7'], ['212', '10'], ['391', '10'], ['501', '10'],
['264', '10'], ['155', '8'], ['323', '10'], ['409', '10'], ['284', '3'], ['73', '2'],
['556', '10'], ['523', '10'], ['196', '10'], ['233', '6'], ['366', '10'], ['558', '10'],
['286', '10'], ['97', '1'], ['705', '7'], ['474', '10'], ['211', '2'], ['357', '6'], ['4
74', '10'], ['246', '10'], ['228', '10'], ['239', '10'], ['235', '10'], ['376', '10'],
['369', '10'], ['19', '0'], ['398', '10'], ['185', '10'], ['329', '5'], ['200', '10'],
['104', '10'], ['203', '10'], ['205', '10'], ['93', '10'], ['343', '10'], ['196', '10'],
['351', '7'], ['161', '10'], ['463', '10'], ['532', '10'], ['401', '7'], ['210', '10'],
['330', '10'], ['212', '10'], ['294', '10'], ['340', '10'], ['526', '10'], ['563', '1
0'], ['269', '7'], ['430', '10'], ['360', '10'], ['228', '6'], ['177', '4'], ['469', '1
0'], ['268', '10'], ['314', '3'], ['195', '10'], ['613', '7'], ['577', '10'], ['208',
'5'], ['31', '0'], ['441', '10'], ['355', '10'], ['295', '10'], ['209', '10'], ['428',
'10'], ['405', '10'], ['192', '10'], ['216', '9'], ['478', '10'], ['495', '10'], ['223',
'10'], ['207', '3'], ['333', '10'], ['282', '10'], ['240', '10'], ['169', '10'], ['191',
'10'], ['659', '4'], ['400', '10'], ['161', '10'], ['369', '10'], ['171', '10'], ['604',
'7'], ['272', '10'], ['167', '5'], ['201', '10'], ['380', '10'], ['572', '10'], ['241',
'10'], ['50', '2'], ['237', '9'], ['503', '10'], ['131', '10'], ['408', '10'], ['499',
'10'], ['223', '10'], ['136', '4'], ['262', '10'], ['272', '10'], ['164', '10'], ['158',
'4'], ['517', '7'], ['444', '10'], ['290', '5'], ['315', '10'], ['408', '10'], ['505',
'10'], ['224', '10'], ['130', '6'], ['530', '10'], ['465', '10'], ['318', '9'], ['174',
'6'], ['572', '7'], ['475', '10'], ['251', '6'], ['252', '7'], ['260', '10'], ['105', '1
0'], ['285', '10'], ['107', '10'], ['189', '10'], ['276', '10'], ['274', '10'], ['112',
'10'], ['451', '10'], ['603', '9'], ['204', '6'], ['300', '10'], ['332', '10'], ['137',
'10'], ['106', '2'], ['250', '10'], ['235', '10'], ['307', '10'], ['291', '10'], ['32
6', '10'], ['283', '10'], ['243', '10'], ['140', '10'], ['443', '7'], ['151', '10'],
['106', '8'], ['261', '10'], ['622', '7'], ['300', '10'], ['6', '0'], ['244', '10'],
['191', '10'], ['36', '10'], ['93', '2'], ['195', '10'], ['326', '10'], ['200', '10'],
['70', '2'], ['338', '10'], ['244', '10'], ['312', '6'], ['334', '5'], ['369', '10'],
['336', '10'], ['294', '10'], ['329', '7'], ['177', '10'], ['400', '10'], ['91', '10'],
['263', '10'], ['262', '10'], ['113', '10'], ['118', '4'], ['109', '10'], ['197', '1
0'], ['163', '10'], ['78', '1'], ['480', '10'], ['571', '9'], ['175', '2']]
```

In [94]:
```
for runs,wics in run_wic:
    Runs.append(runs)
    Wics.append(wics)
```

In [95]:
```
print(Runs)
```

```
['182', '361', '234', '56', '257', '328', '418', '168', '233', '633', '181', '424', '40
0', '169', '195', '441', '285', '239', '110', '405', '238', '208', '195', '150', '552',
'261', '176', '349', '219', '47', '445', '171', '657', '212', '391', '501', '264', '15
5', '323', '409', '284', '73', '556', '523', '196', '233', '366', '558', '286', '97', '7
05', '474', '211', '357', '474', '246', '228', '239', '235', '376', '369', '19', '398',
'185', '329', '200', '104', '203', '205', '93', '343', '196', '351', '161', '463', '53
2', '401', '210', '330', '212', '294', '340', '526', '563', '269', '430', '360', '228',
'177', '469', '268', '314', '195', '613', '577', '208', '31', '441', '355', '295', '20
9', '428', '405', '192', '216', '478', '495', '223', '207', '333', '282', '240', '169',
'191', '659', '400', '161', '369', '171', '604', '272', '167', '201', '380', '572', '24
1', '50', '237', '503', '131', '408', '499', '223', '136', '262', '272', '164', '158',
'517', '444', '290', '315', '408', '505', '224', '130', '530', '465', '318', '174', '57
2', '475', '251', '252', '260', '105', '285', '107', '189', '276', '274', '112', '451',
'603', '204', '300', '332', '137', '106', '250', '235', '307', '291', '326', '283', '24
3', '140', '443', '151', '106', '261', '622', '300', '6', '244', '191', '36', '93', '19
```

```
5', '326', '200', '70', '338', '244', '312', '334', '369', '336', '294', '329', '177',
'400', '91', '263', '262', '113', '118', '109', '197', '163', '78', '480', '571', '175']
```

In [96]:
```
print(Wics)
```

```
['10', '10', '10', '3', '10', '10', '4', '10', '10', '5', '10', '10', '10', '10', '2',
'10', '10', '8', '10', '10', '10', '5', '10', '10', '5', '10', '10', '10', '10', '0', '1
0', '10', '7', '10', '10', '10', '10', '8', '10', '10', '3', '2', '10', '10', '10', '6',
'10', '10', '10', '1', '7', '10', '2', '6', '10', '10', '10', '10', '10', '10', '10',
'0', '10', '10', '5', '10', '10', '10', '10', '10', '10', '10', '7', '10', '10', '10',
'7', '10', '10', '10', '10', '10', '10', '10', '7', '10', '10', '6', '4', '10', '10',
'3', '10', '7', '10', '5', '0', '10', '10', '10', '10', '10', '10', '10', '9', '10', '1
0', '10', '3', '10', '10', '10', '10', '10', '4', '10', '10', '10', '10', '7', '10',
'5', '10', '10', '10', '10', '2', '9', '10', '10', '10', '10', '10', '4', '10', '10', '1
0', '4', '7', '10', '5', '10', '10', '10', '10', '6', '10', '10', '9', '6', '7', '10',
'6', '7', '10', '10', '10', '10', '10', '10', '10', '10', '10', '9', '6', '10', '10', '1
0', '2', '10', '10', '10', '10', '10', '10', '10', '10', '7', '10', '8', '10', '7', '1
0', '0', '10', '10', '10', '2', '10', '10', '10', '2', '10', '10', '6', '5', '10', '10',
'10', '7', '10', '10', '10', '10', '10', '10', '4', '10', '10', '10', '1', '10', '9',
'2']
```

In [97]:
```
df["RUNS"]=Runs
df["WICKETS"]=Wics
df["RUNS"]=df["RUNS"].astype('int32')
df["WICKETS"]=df["WICKETS"].astype('int32')
```

In [98]:
```
df.head()
```

Out[98]:

	Test Series Year	Test Number	Innings	Venue	Host	Highest Scorer	Team	Runs by highest scorer	best bowler	wickets by best bowler	team total	Wi
0	1996-97	1	1	Delhi	India	Michael Slater	Australia	44	Anil Kumble	4.0	182/10	I
1	1996-97	1	2	Delhi	India	Nayan Mongia	India	152	Paul Reiffel	3.0	361/10	I
2	1996-97	1	3	Delhi	India	Steve Waugh	Australia	67	Anil Kumble	5.0	234/10	I
3	1996-97	1	4	Delhi	India	Mohd. Azharuddin	India	21	Paul Reiffel	2.0	56/3	I
4	1997-98	1	1	Chennai	India	Navjot Sidhu	India	62	Shane Warne	4.0	257/10	I

In [99]:
```
df.shape
```

Out[99]:
```
(214, 20)
```

In [100…
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 214 entries, 0 to 213
Data columns (total 20 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   Test Series Year       214 non-null    object
 1   Test Number            214 non-null    int64
```

```
 2   Innings                 214 non-null    int64
 3   Venue                   214 non-null    object
 4   Host                    214 non-null    object
 5   Highest Scorer          214 non-null    object
 6   Team                    214 non-null    object
 7   Runs by highest scorer  214 non-null    int64
 8   best bowler             214 non-null    object
 9   wickets by best bowler  213 non-null    float64
 10  team total              214 non-null    object
 11  Winner                  214 non-null    object
 12  Win Margin              214 non-null    object
 13  MOTM                    214 non-null    object
 14  Ind captain             214 non-null    object
 15  Aus captain             214 non-null    object
 16  MOTS                    214 non-null    object
 17  Series Win              214 non-null    object
 18  RUNS                    214 non-null    int32
 19  WICKETS                 214 non-null    int32
dtypes: float64(1), int32(2), int64(3), object(14)
memory usage: 20.1+ KB
```

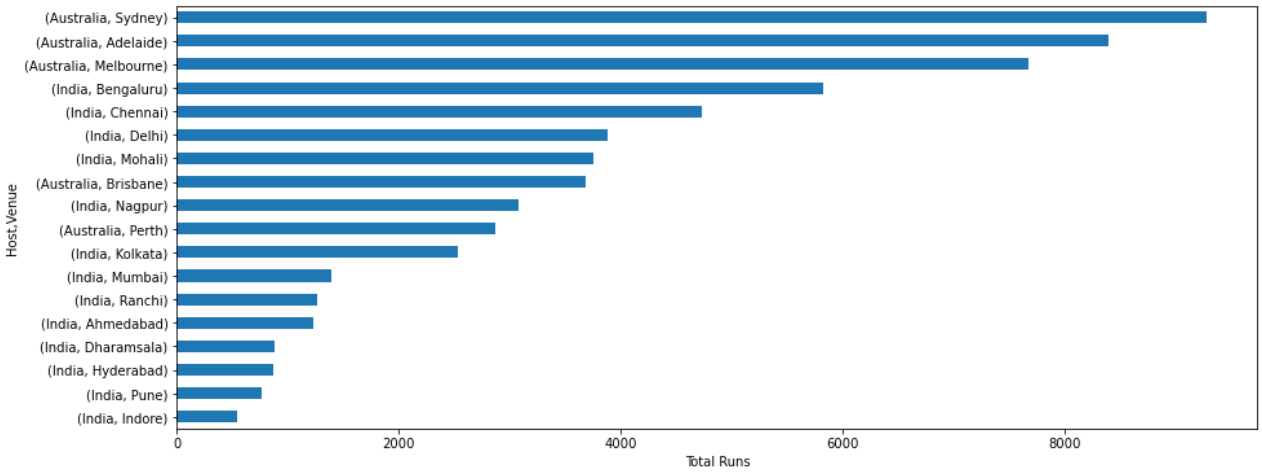
In [101…
```
df.groupby("Team")["RUNS"].sum()
```

Out[101…
```
Team
Australia    31870
India        30737
Name: RUNS, dtype: int32
```

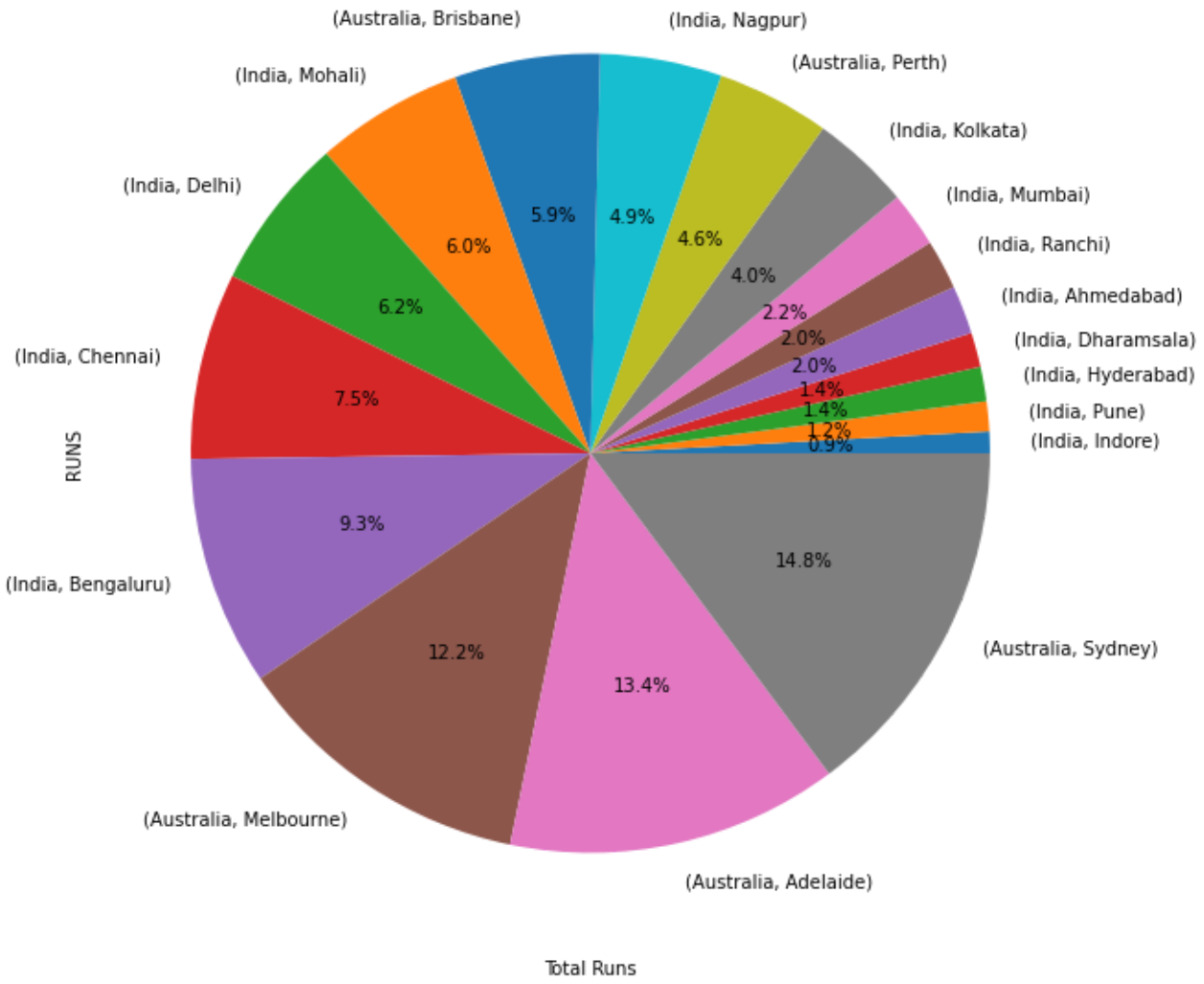
In [102…
```
df.groupby("Team")["WICKETS"].sum()
```

Out[102…
```
Team
Australia    927
India        905
Name: WICKETS, dtype: int32
```

In [103…
```
plt.figure(figsize=(15,6))
df.groupby(["Host","Venue"])["RUNS"].sum().sort_values().plot(kind="barh")
plt.xlabel("Total Runs")
plt.show()
```

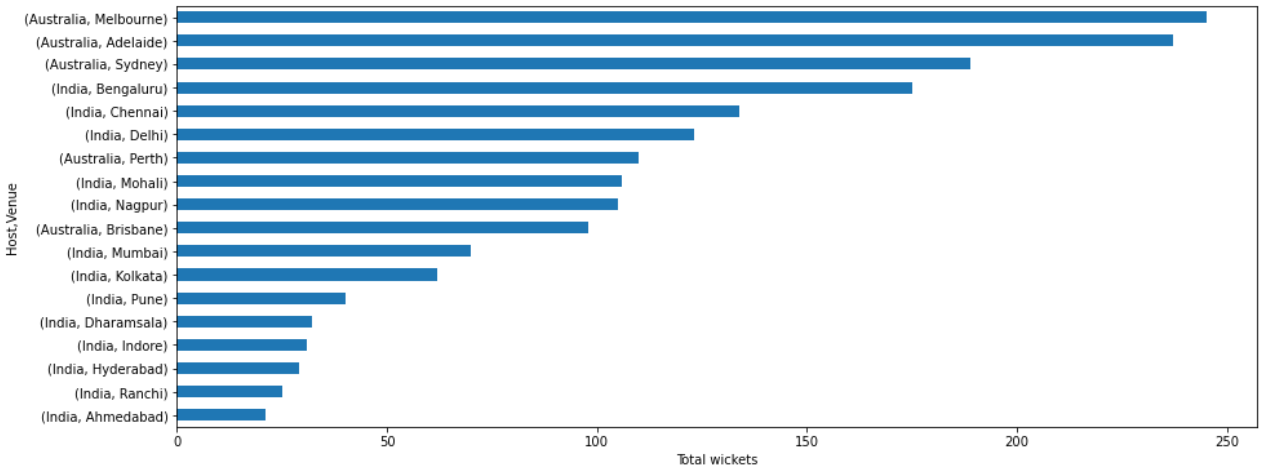


In [104…
```
plt.figure(figsize=(30,10))
df.groupby(["Host","Venue"])["RUNS"].sum().sort_values().plot(kind="pie", autopct='%1.1
plt.xlabel("Total Runs")
plt.show()
```



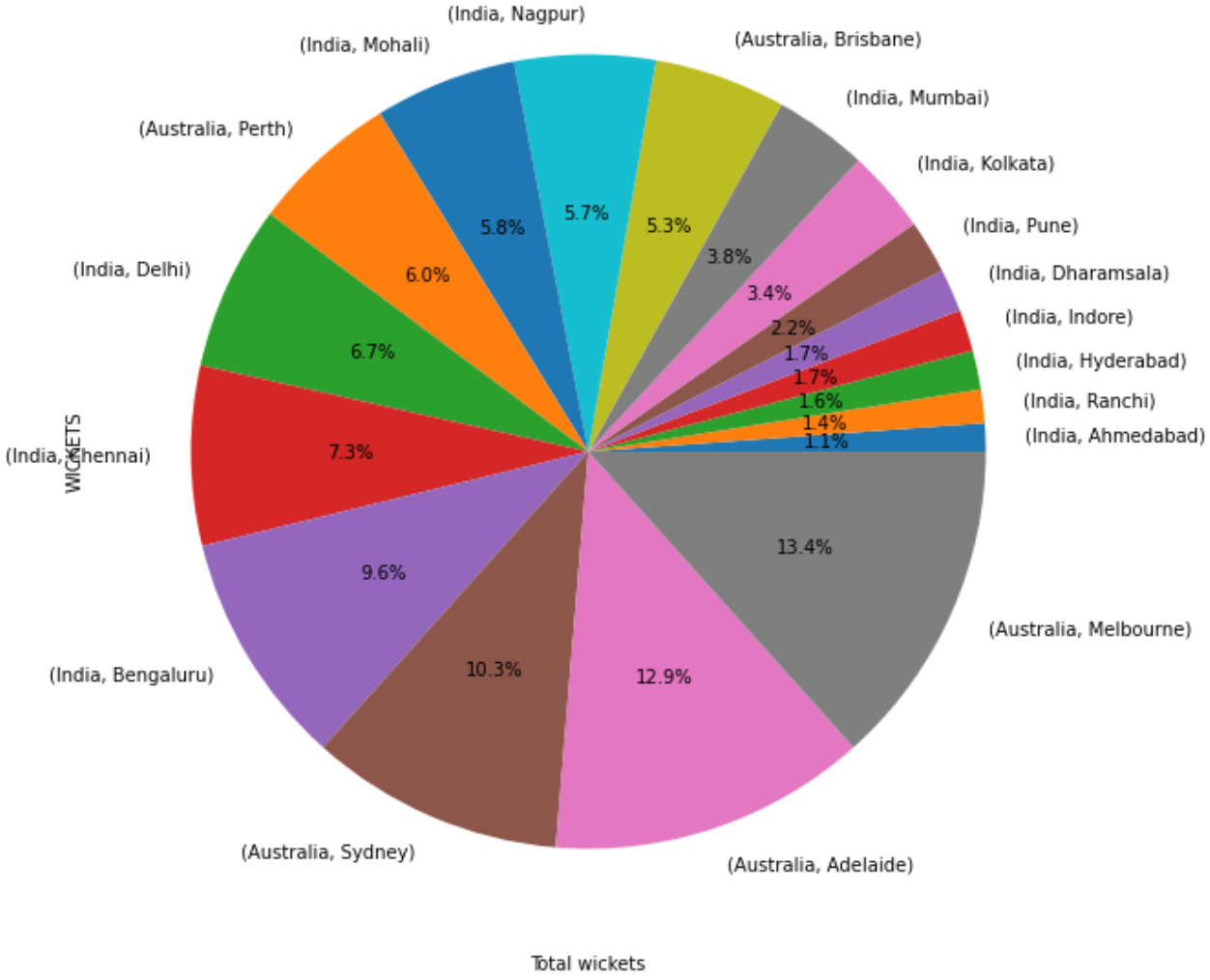
In [105…

```
plt.figure(figsize=(15,6))
df.groupby(["Host","Venue"])["WICKETS"].sum().sort_values().plot(kind="barh")
plt.xlabel("Total wickets")
plt.show()
```



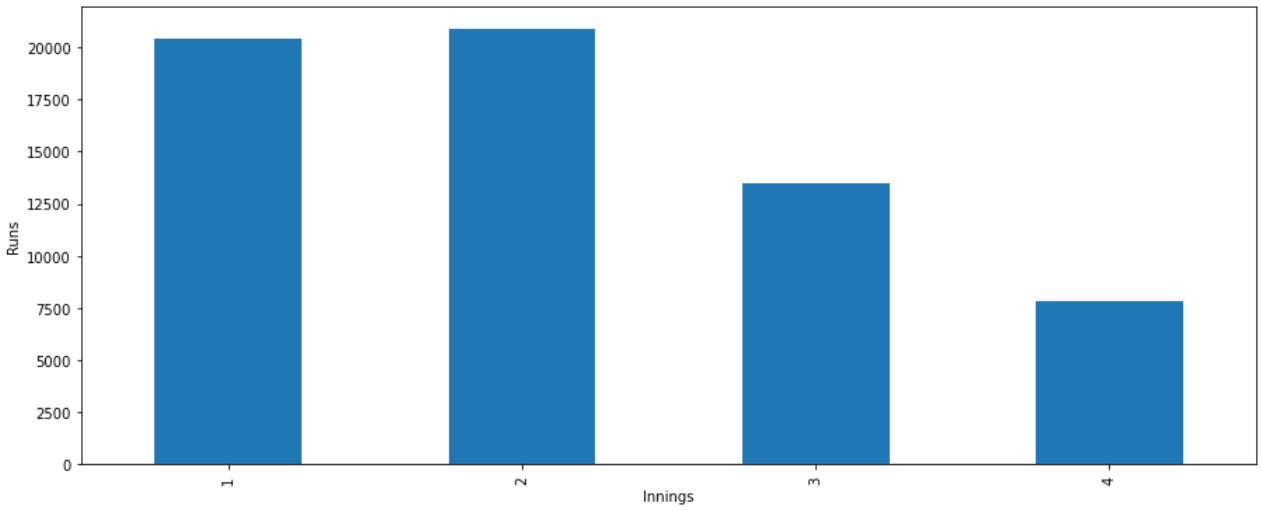
In [106…

```
plt.figure(figsize=(30, 10))
df.groupby(["Host","Venue"])["WICKETS"].sum().sort_values().plot(kind='pie', autopct='%
plt.xlabel("Total wickets")
plt.show()
```



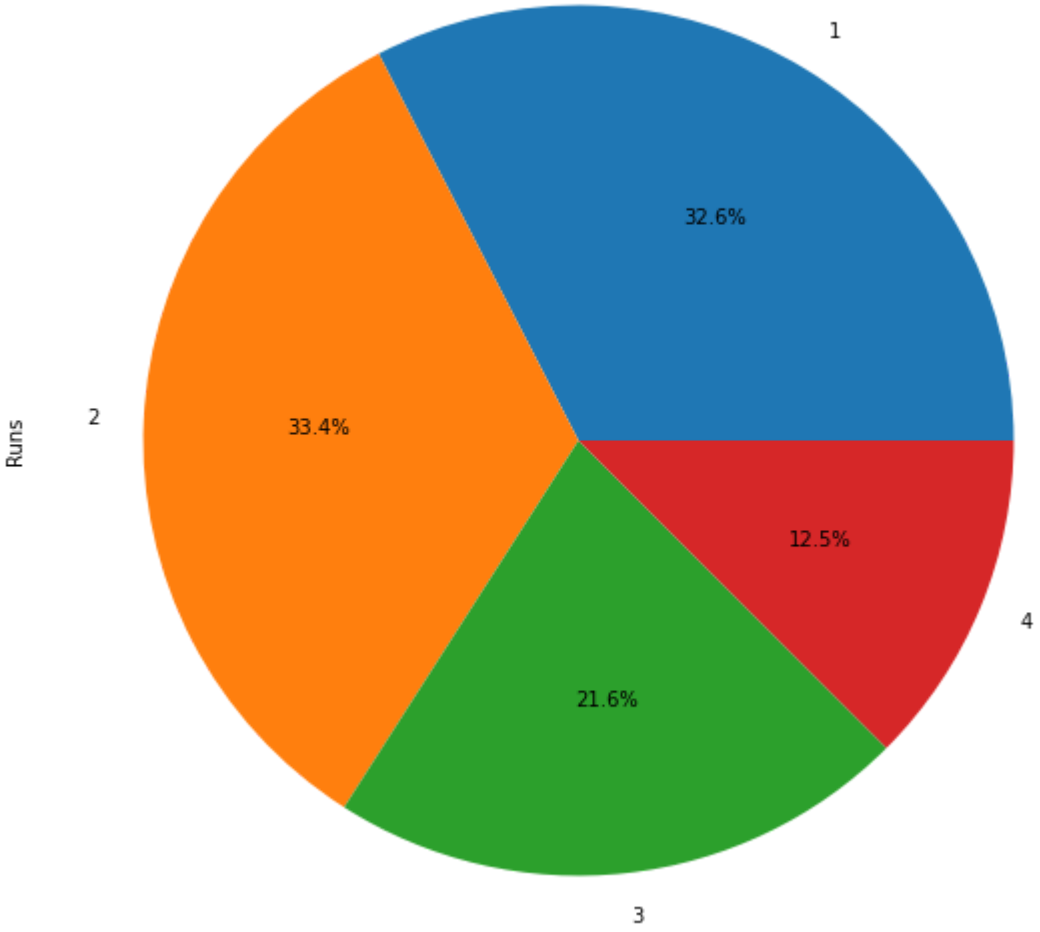
In [107…

```
plt.figure(figsize=(15,6))
df.groupby("Innings")["RUNS"].sum().plot(kind="bar")
plt.ylabel("Runs")
plt.show()
```



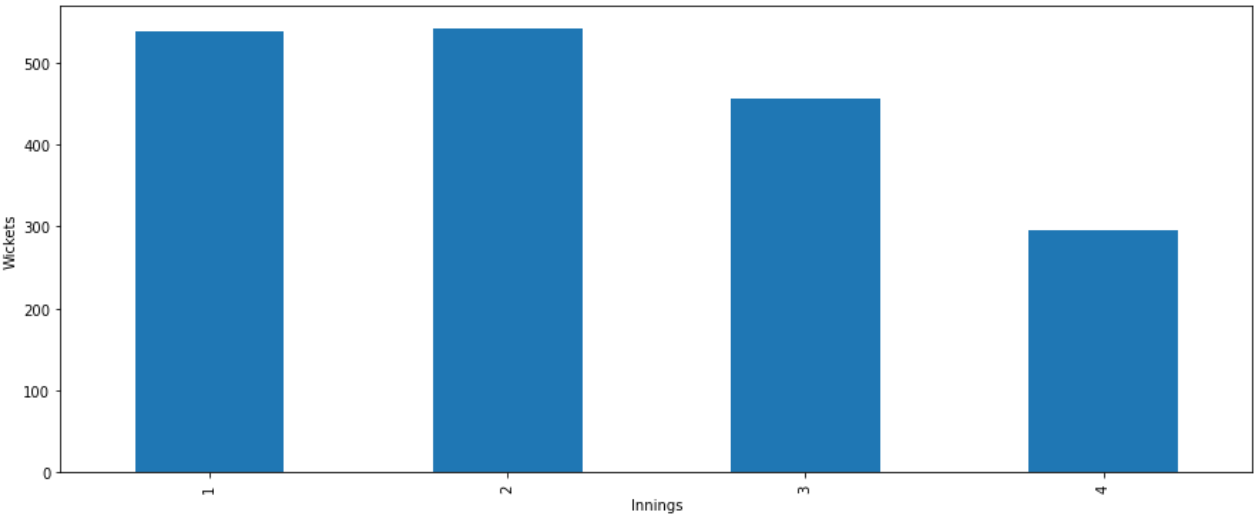
In [108…

```
plt.figure(figsize=(30, 10))
df.groupby("Innings")["RUNS"].sum().plot(kind="pie", autopct='%1.1f%%')
plt.ylabel("Runs")
plt.show()
```



In [109…

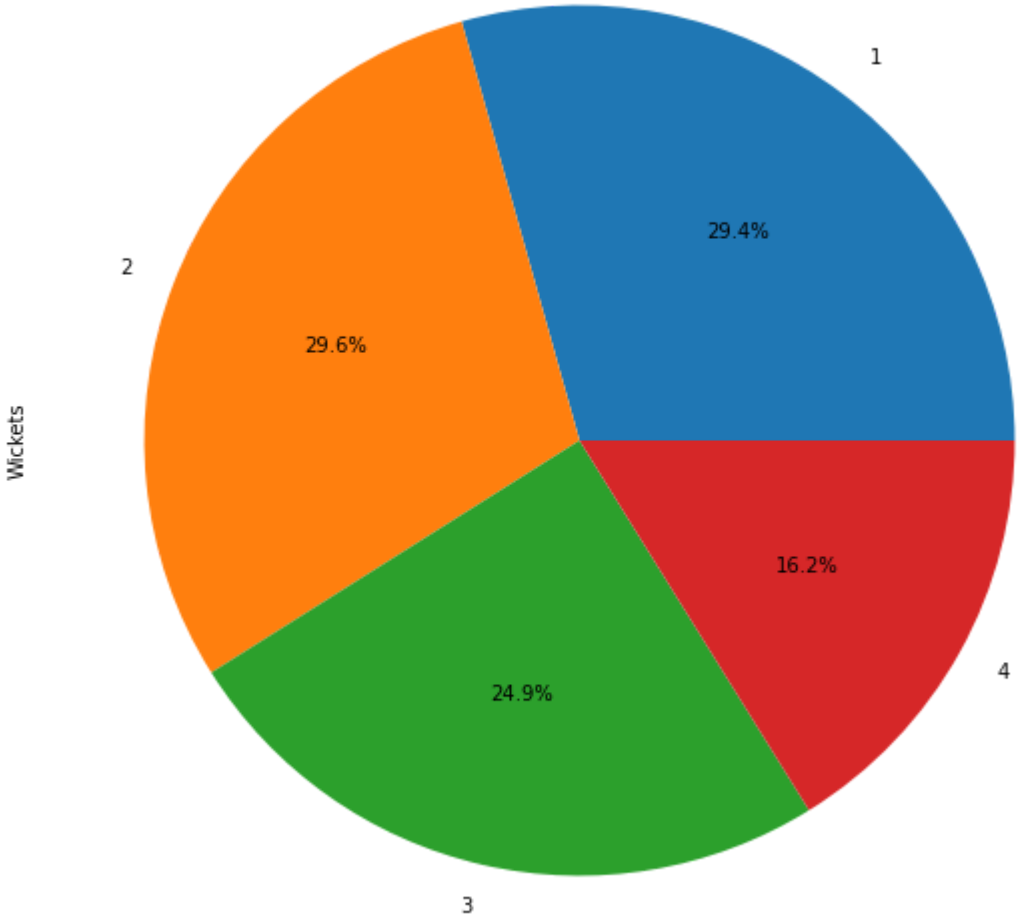
```
plt.figure(figsize=(15,6))
df.groupby("Innings")["WICKETS"].sum().plot(kind="bar")
plt.ylabel("Wickets")
plt.show()
```



In [110…

```
plt.figure(figsize=(30, 10))
df.groupby("Innings")["WICKETS"].sum().plot(kind="pie", autopct='%1.1f%%')
```

```
plt.ylabel("Wickets")
plt.show()
```



In [ ]: