In [1]:

```
import pandas as pd
```

In [2]:

```
df = pd.read_csv('BGTresults.csv')
```

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	Winner	Win Margin	MOTM	Ind captain	c
0	1996-97	1	1	Delhi	India	Michael Slater	Australia	44	Anil Kumble	4.0	182/10	India	7w	Nayan Mongia	Sachin Tendulkar	
1	1996-97	1	2	Delhi	India	Nayan Mongia	India	152	Paul Reiffel	3.0	361/10	India	7w	Nayan Mongia	Sachin Tendulkar	
2	1996-97	1	3	Delhi	India	Steve Waugh	Australia	67	Anil Kumble	5.0	234/10	India	7w	Nayan Mongia	Sachin Tendulkar	
3	1996-97	1	4	Delhi	India	Mohd. Azharuddin	India	21	Paul Reiffel	2.0	56/3	India	7w	Nayan Mongia	Sachin Tendulkar	
4	1997-98	1	1	Chennai	India	Navjot Sidhu	India	62	Shane Warne	4.0	257/10	India	179r	Sachin Tendulkar	Mohd. Azharuddin	

In [4]:

```
df.tail()
```

Out[4]:

	Test Series Year	Test Number	Innings	Venue	Host	Highest Scorer	Team	Runs by highest scorer	best bowler	wickets by best bowler	team total	Winner	Win Margin	MOTM	cap
195	2020-21	3	4	Sydney	Australia	Rishabh Pant	India	97	Josh Hazlewood	2.0	334/5	Drawn	0	Steve Smith	Ajin Rah
196	2020-21	4	1	Brisbane	Australia	Marnus Labuschagne	Australia	108	T Natarajan	3.0	369/10	India	3w	Rishabh Pant	Ajin Rah
197	2020-21	4	2	Brisbane	Australia	Shardul Thakur	India	67	Josh Hazlewood	5.0	336/10	India	3w	Rishabh Pant	Ajin Rah
198	2020-21	4	3	Brisbane	Australia	Steve Smith	Australia	55	Mohd. Siraj	5.0	294/10	India	3w	Rishabh Pant	Ajin Rah
199	2020-21	4	4	Brisbane	Australia	Shubman Gill	India	91	Pat Cummins	4.0	329/7	India	3w	Rishabh Pant	Ajin Rah

In [5]:

```
df.shape
```

Out[5]:

```
(200, 18)
```

In [6]:

```
df.columns
```

Out[6]:

```
Index(['Test Series Year', 'Test Number', 'Innings', 'Venue', 'Host',
       'Highest Scorer', 'Team', 'Runs by highest scorer', 'best bowler',
       'wickets by best bowler', 'team total', 'Winner', 'Win Margin', 'MOTM',
       'Ind captain', 'Aus captain', 'MOTS', 'Series Win'],
      dtype='object')
```

In [7]:

```
df.duplicated().sum()
```

Out[7]:

```
0
```

In [8]:

```
df.isnull().sum()
```

Out[8]:

```
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 [9]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 18 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   Test Series Year        200 non-null    object 
 1   Test Number             200 non-null    int64  
 2   Innings                 200 non-null    int64  
 3   Venue                   200 non-null    object 
 4   Host                    200 non-null    object 
 5   Highest Scorer          200 non-null    object 
 6   Team                    200 non-null    object 
 7   Runs by highest scorer  200 non-null    int64  
 8   best bowler             200 non-null    object 
 9   wickets by best bowler  199 non-null    float64
 10  team total              200 non-null    object 
 11  Winner                  200 non-null    object 
 12  Win Margin              200 non-null    object 
 13  MOTM                    200 non-null    object 
 14  Ind captain             200 non-null    object 
 15  Aus captain             200 non-null    object 
 16  MOTS                    200 non-null    object 
 17  Series Win              200 non-null    object 
dtypes: float64(1), int64(3), object(14)
memory usage: 28.2+ KB
```

In [10]:

```
df.describe()
```

Out[10]:

	Test Number	Innings	Runs by highest scorer	wickets by best bowler
count	200.000000	200.000000	200.000000	199.000000
mean	2.325000	2.440000	95.695000	3.874372
std	1.097908	1.101119	56.528393	1.604695
min	1.000000	1.000000	4.000000	0.000000
25%	1.000000	1.000000	56.750000	3.000000
50%	2.000000	2.000000	79.500000	4.000000
75%	3.000000	3.000000	123.250000	5.000000
max	4.000000	4.000000	329.000000	8.000000

In [11]:

```
median_b = df['wickets by best bowler'].median()
df['wickets by best bowler'] = df['wickets by best bowler'].fillna(median_b)
```

In [12]:

```
df
```

Out[12]:

	Test Series Year	Test Number	Innings	Venue	Host	Highest Scorer	Team	Runs by highest scorer	best bowler	wickets by best bowler	team total	Winner	Win Margin	MOTM	
0	1996-97	1	1	Delhi	India	Michael Slater	Australia	44	Anil Kumble	4.0	182/10	India	7w	Nayan Mongia	T
1	1996-97	1	2	Delhi	India	Nayan Mongia	India	152	Paul Reiffel	3.0	361/10	India	7w	Nayan Mongia	T
2	1996-97	1	3	Delhi	India	Steve Waugh	Australia	67	Anil Kumble	5.0	234/10	India	7w	Nayan Mongia	T
3	1996-97	1	4	Delhi	India	Mohd. Azharuddin	India	21	Paul Reiffel	2.0	56/3	India	7w	Nayan Mongia	T
4	1997-98	1	1	Chennai	India	Navjot Sidhu	India	62	Shane Warne	4.0	257/10	India	179r	Sachin Tendulkar	Az
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
195	2020-21	3	4	Sydney	Australia	Rishabh Pant	India	97	Josh Hazlewood	2.0	334/5	Drawn	0	Steve Smith	
196	2020-21	4	1	Brisbane	Australia	Marnus Labuschagne	Australia	108	T Natarajan	3.0	369/10	India	3w	Rishabh Pant	
197	2020-21	4	2	Brisbane	Australia	Shardul Thakur	India	67	Josh Hazlewood	5.0	336/10	India	3w	Rishabh Pant	
198	2020-21	4	3	Brisbane	Australia	Steve Smith	Australia	55	Mohd. Siraj	5.0	294/10	India	3w	Rishabh Pant	
199	2020-21	4	4	Brisbane	Australia	Shubman Gill	India	91	Pat Cummins	4.0	329/7	India	3w	Rishabh Pant	

200 rows × 18 columns

In [13]:

```
df.nunique()
```

Out[13]:

```
Test Series Year          15
Test Number                4
Innings                    4
Venue                     16
Host                       2
Highest Scorer            64
Team                       2
Runs by highest scorer   125
best bowler               55
wickets by best bowler     9
team total               174
Winner                     3
Win Margin                34
MOTM                      38
Ind captain                9
Aus captain                7
MOTS                      13
Series Win                 3
dtype: int64
```

In [14]:

```
import matplotlib.pyplot as plt
import seaborn as sns
```

In [15]:

```
import warnings
warnings.filterwarnings('ignore')
```

In [16]:

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

```
 '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']
Ind captain
['Sachin Tendulkar' 'Mohd. Azharuddin' 'mohd. Azharuddin' 'Sourav Ganguly'
 'Rahul Dravid' 'Anil Kumble' 'MS Dhoni' 'Virat Kohli' 'Ajinkya Rahane']
Aus captain
['Mark Taylor' 'Steve Waugh' 'Adam Gilchrist' 'Ricky Ponting'
 'Michael Clarke' 'Steve Smith' 'Tim Paine']
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']
Series Win
['India' 'Australia' 'Drawn']
```

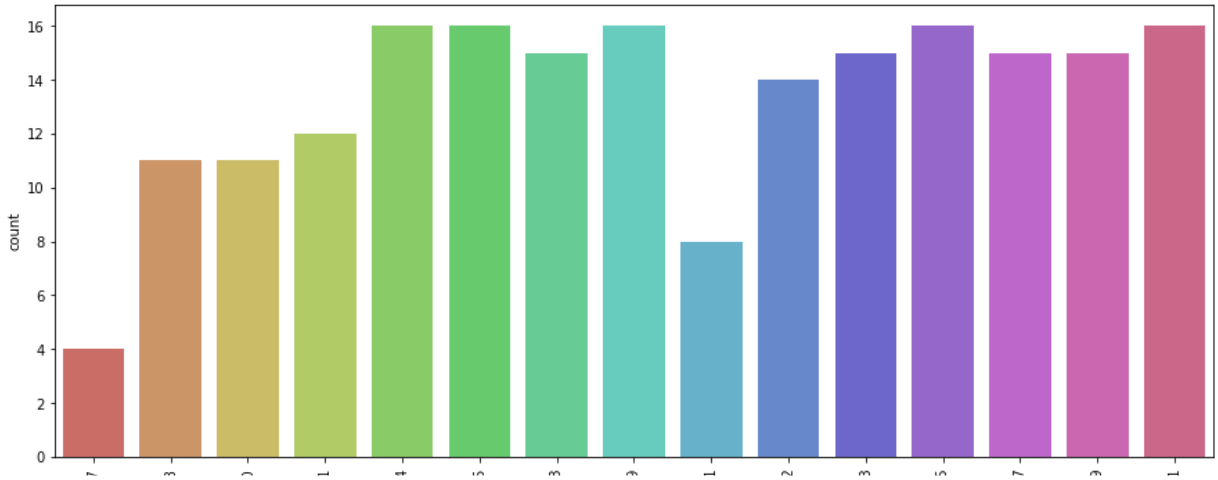
In [17]:

```
for i in df.columns:
    print(i)
    print(df[i].value_counts())
```

```
Name: Test Series Year, dtype: int64
Test Number
1    60
2    53
3    49
4    38
Name: Test Number, dtype: int64
Innings
1    52
2    52
3    52
4    44
Name: Innings, dtype: int64
Venue
Melbourne     28
Adelaide      27
Sydney        25
Bengaluru     20
Chennai       16
```

In [18]:

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

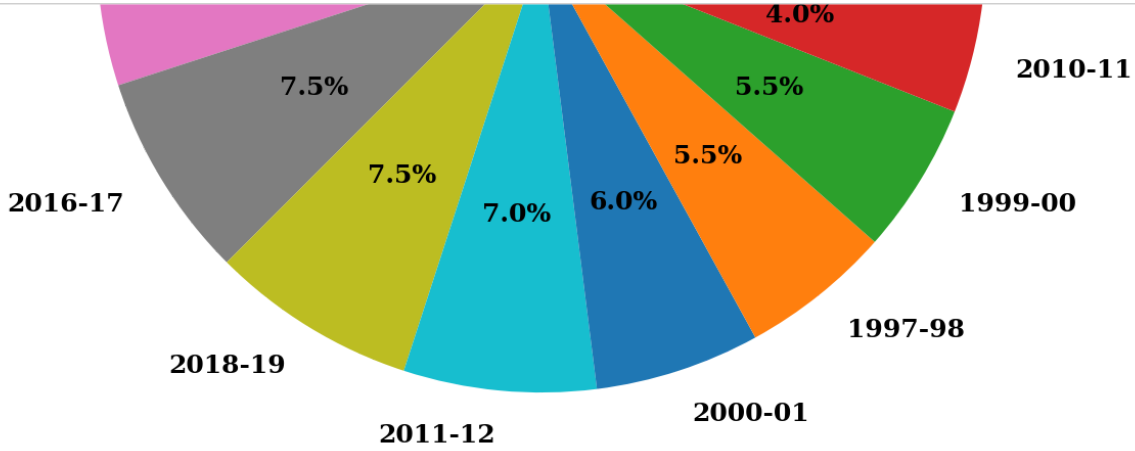


In [19]:

```
df_new = df[['Test Series Year', 'Test Number', 'Innings', 'Venue', 'Host', 'Team', 'Winner',
             'Ind captain', 'Aus captain', 'MOTS', 'Series Win']]
```

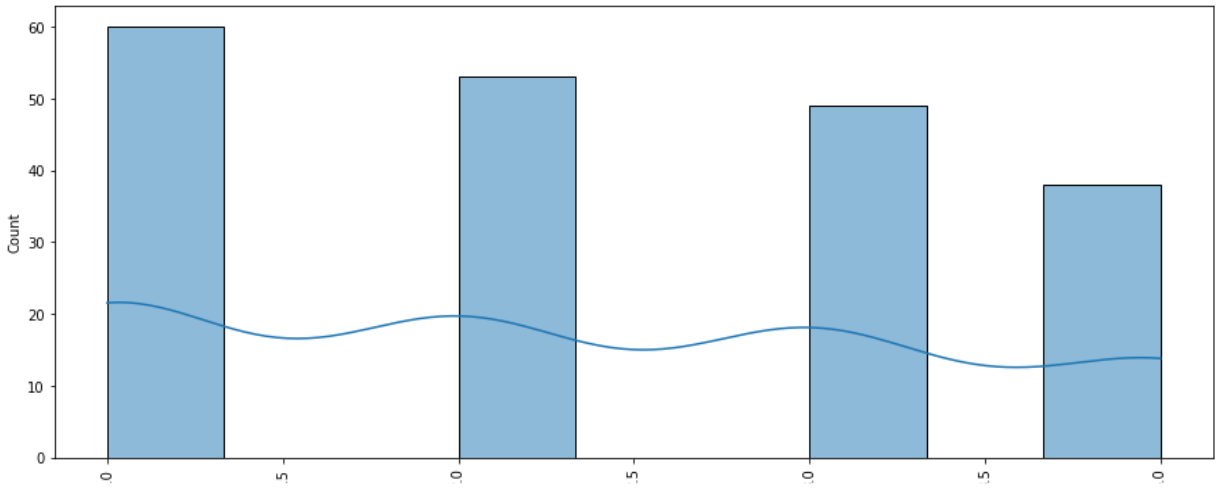
In [20]:

```
for i in df_new:
    plt.figure(figsize=(30,20))
    plt.pie(df[i].value_counts(), labels=df[i].value_counts().index, autopct='%1.1f%%', textprops={ 'fontsize': 25,
                                                                                                    'color': 'black',
                                                                                                    'weight': 'bold',
                                                                                                    'family': 'serif' }
    hfont = {'fontname':'serif', 'weight': 'bold'}
    plt.title(i, size=20, **hfont)
    plt.show()
```



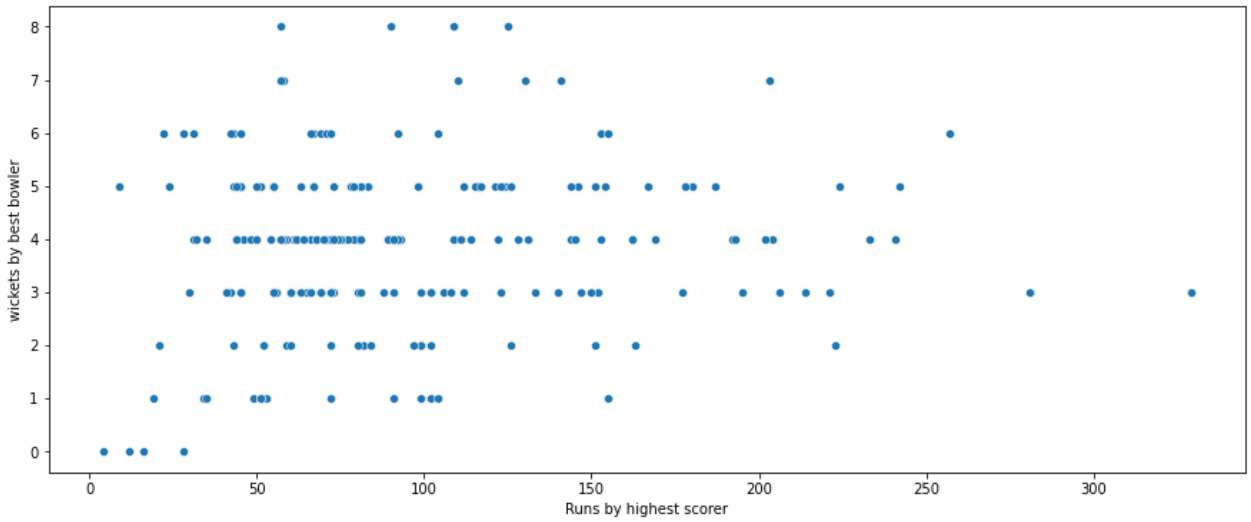
In [21]:

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



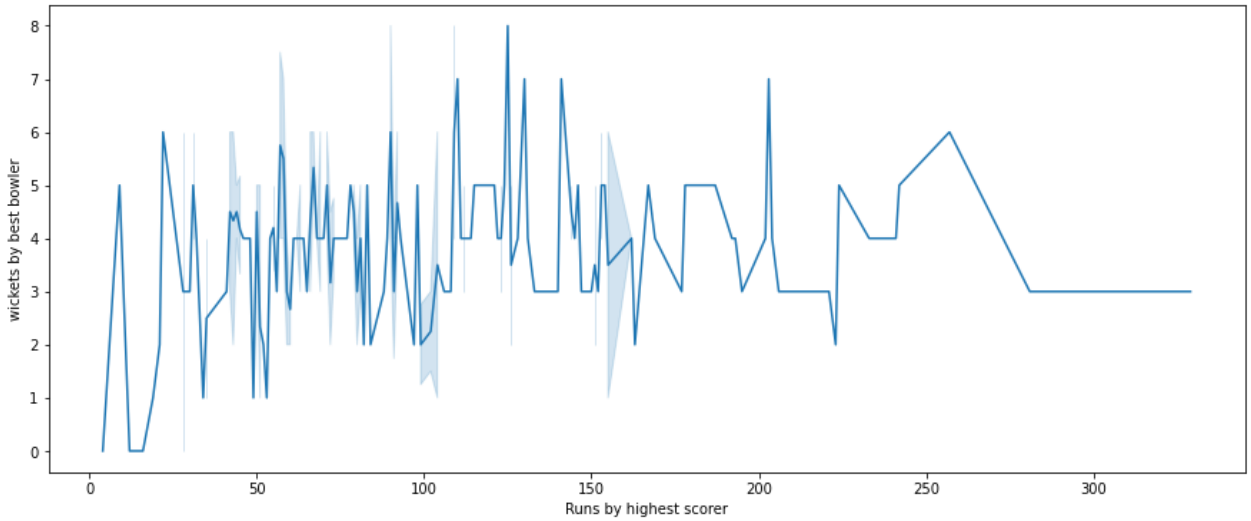
In [22]:

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



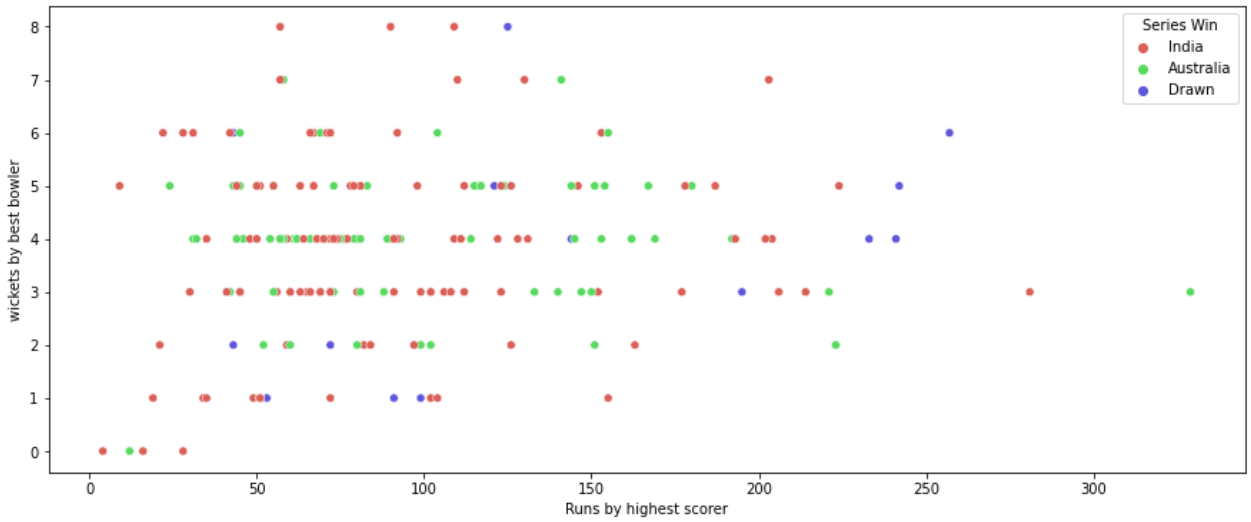
In [23]:

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



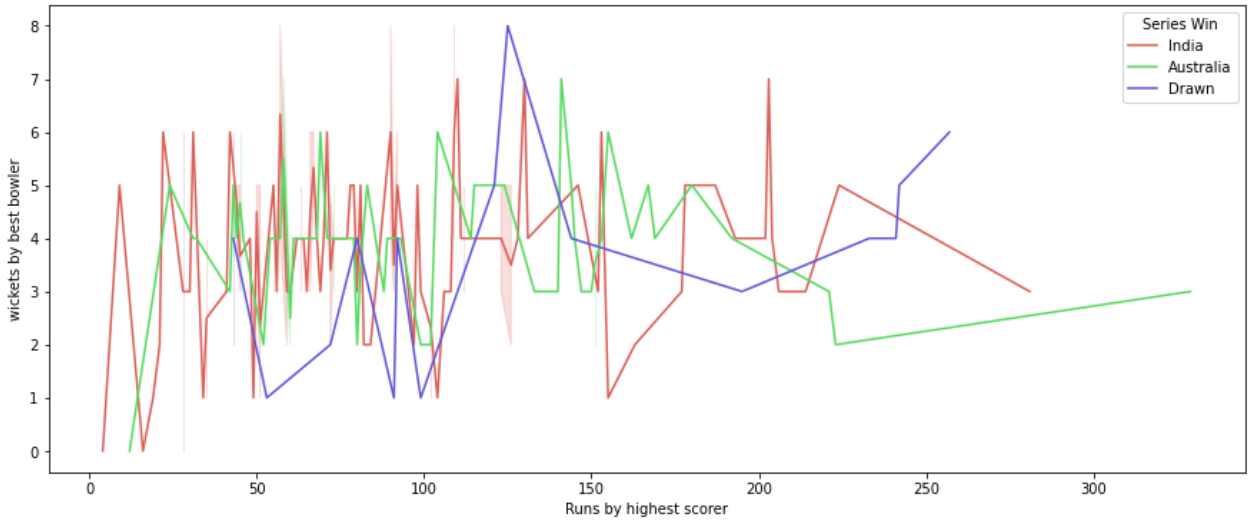
In [24]:

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



In [25]:

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



In [26]:

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

In [27]:

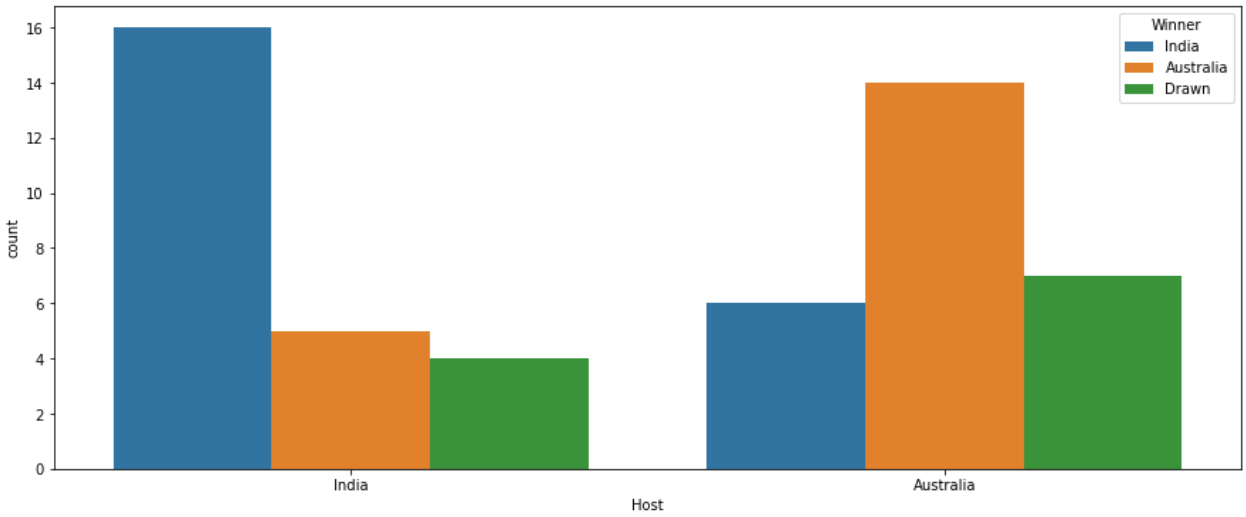
```
df1
```

Out[27]:

	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
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

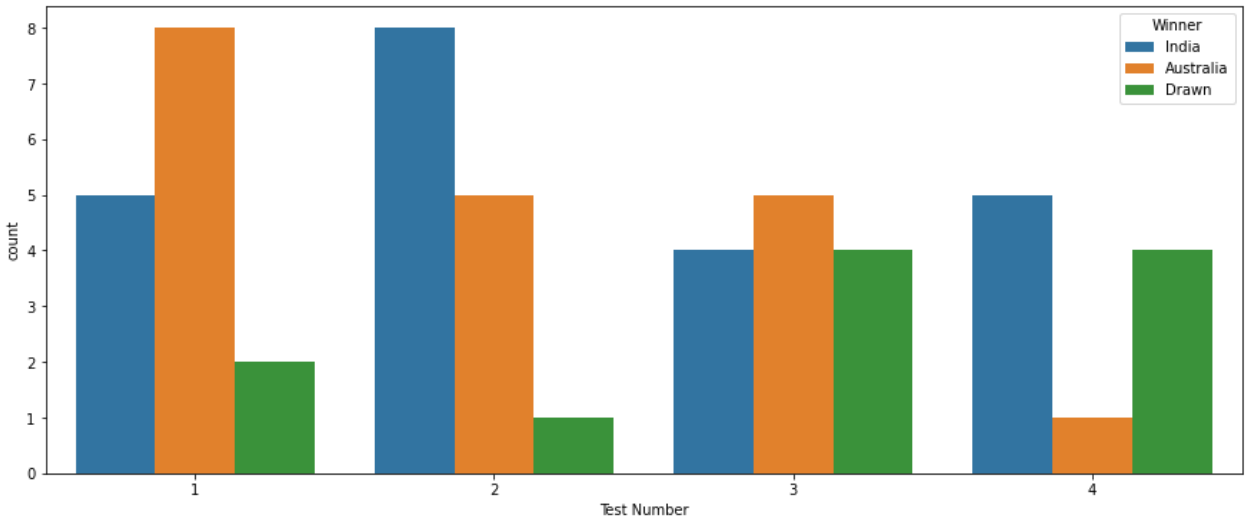
In [28]:

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



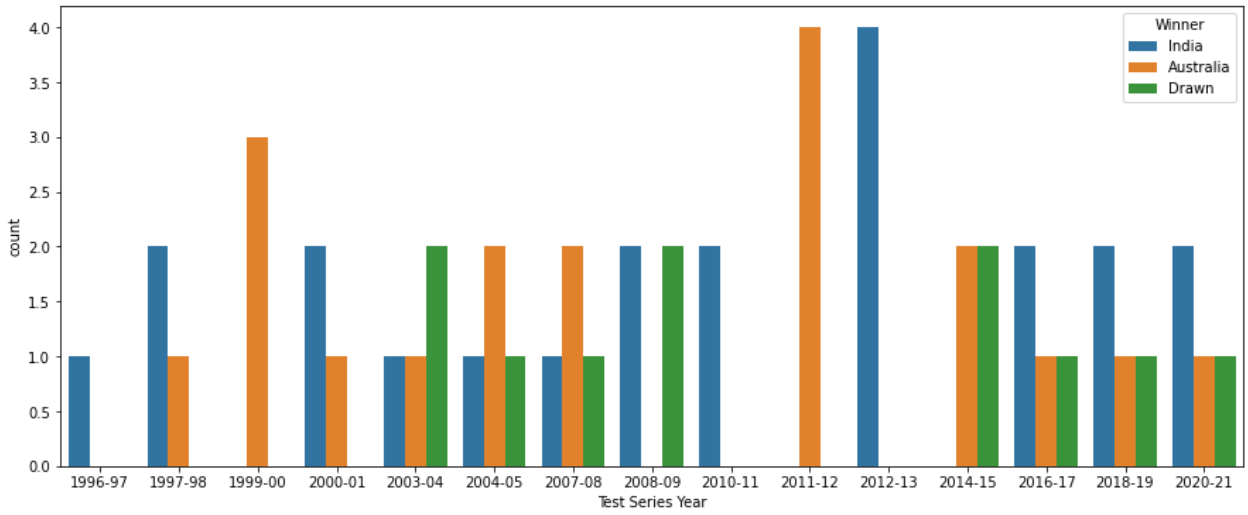
In [29]:

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



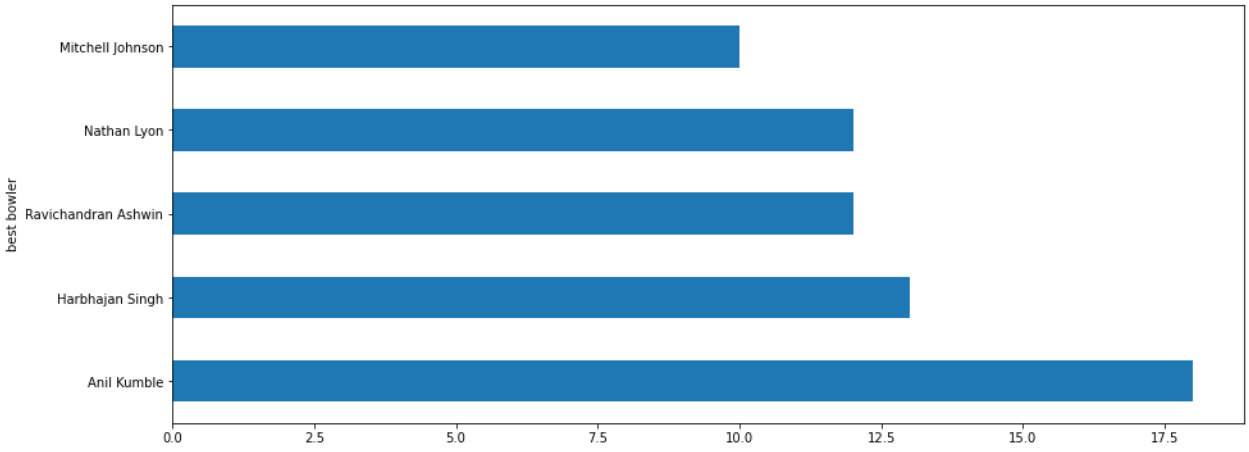
In [30]:

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



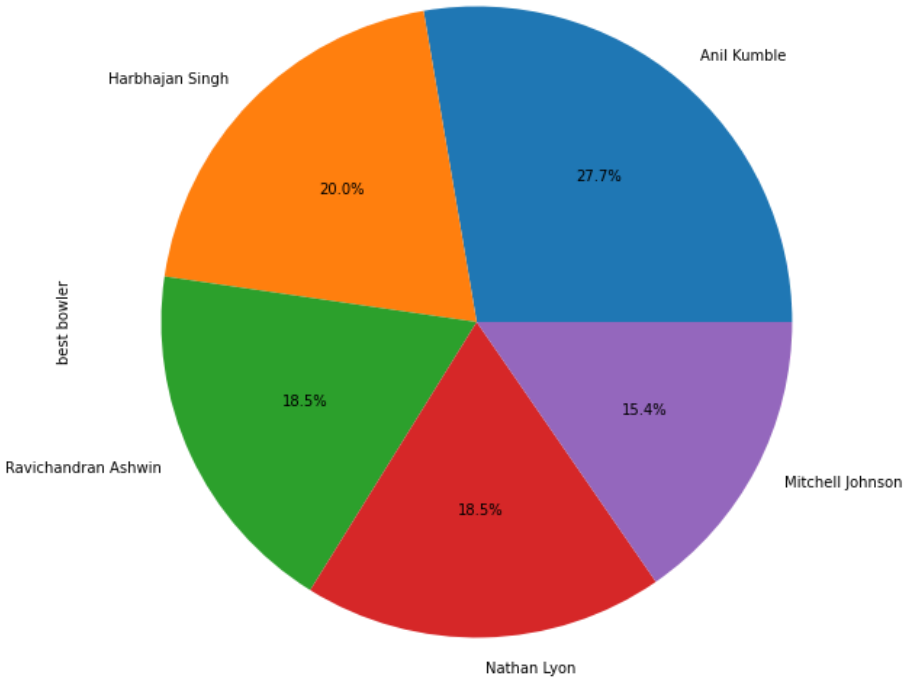
In [31]:

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



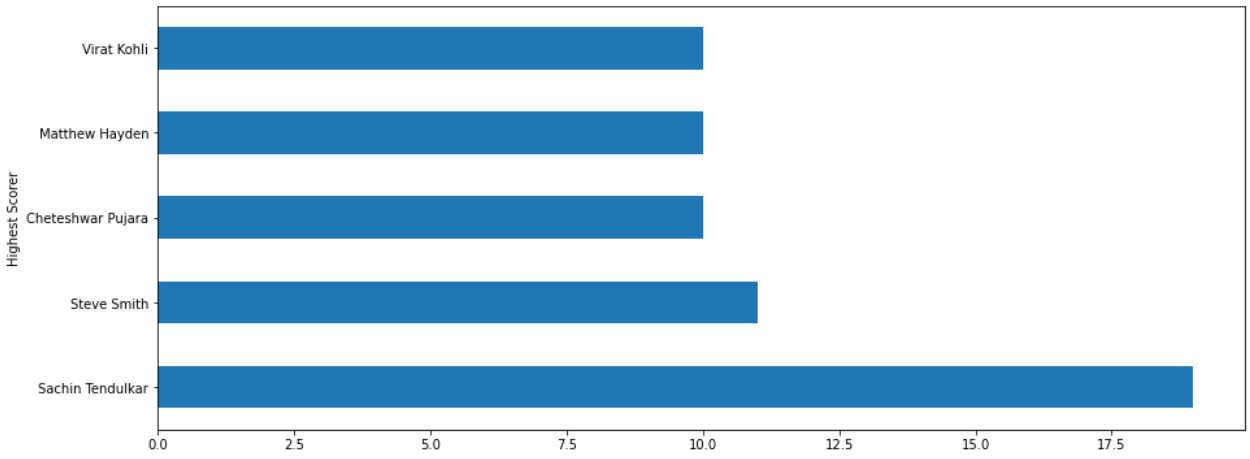
In [32]:

```
plt.figure(figsize=(30,10))
df.groupby("best bowler")["best bowler"].count().sort_values(ascending = False).head().plot(kind="pie", autopct='%1.1f%%')
plt.show()
```



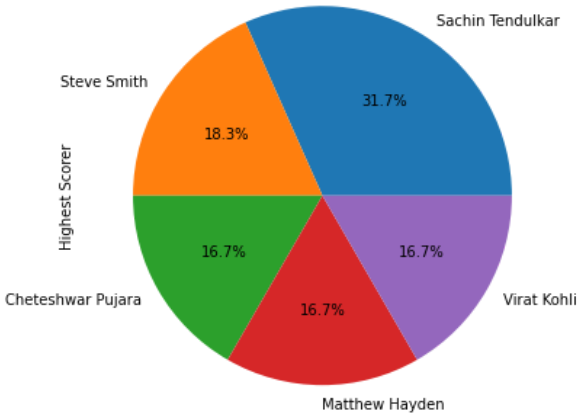
In [33]:

```
plt.figure(figsize=(15,6))
df.groupby("Highest Scorer")["Highest Scorer"].count().sort_values(ascending = False).head().plot(kind="barh")
plt.show()
```



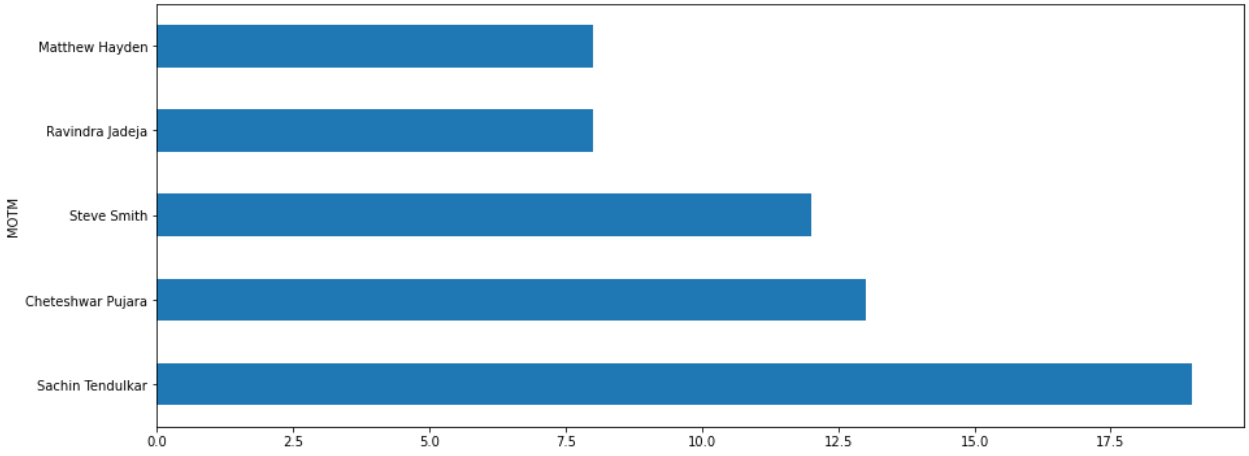
In [34]:

```
plt.figure(figsize=(15,6))
df.groupby("Highest Scorer")["Highest Scorer"].count().sort_values(ascending = False).head().plot(kind="pie", autopct='%1.1f%%
plt.show()
```



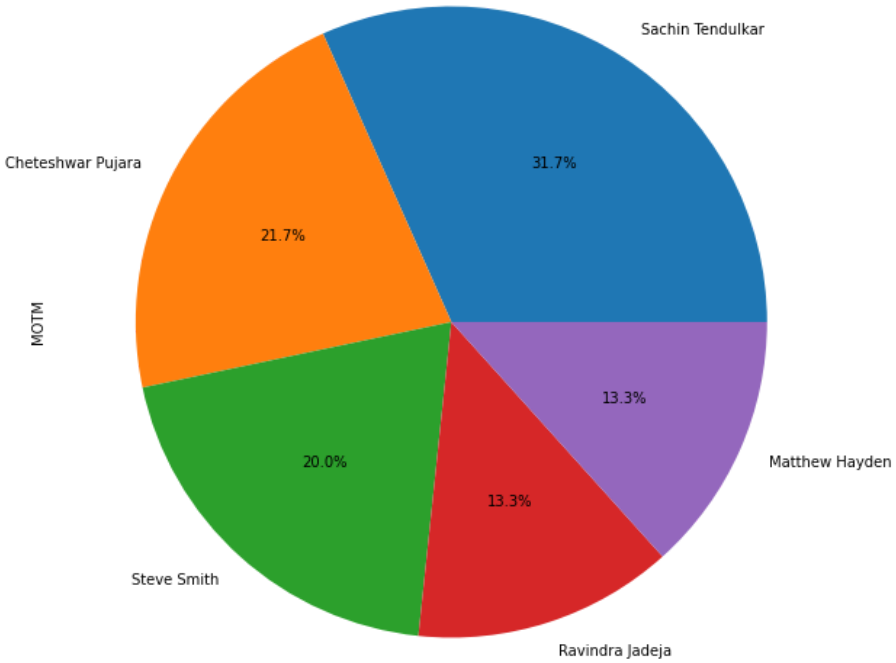
In [35]:

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



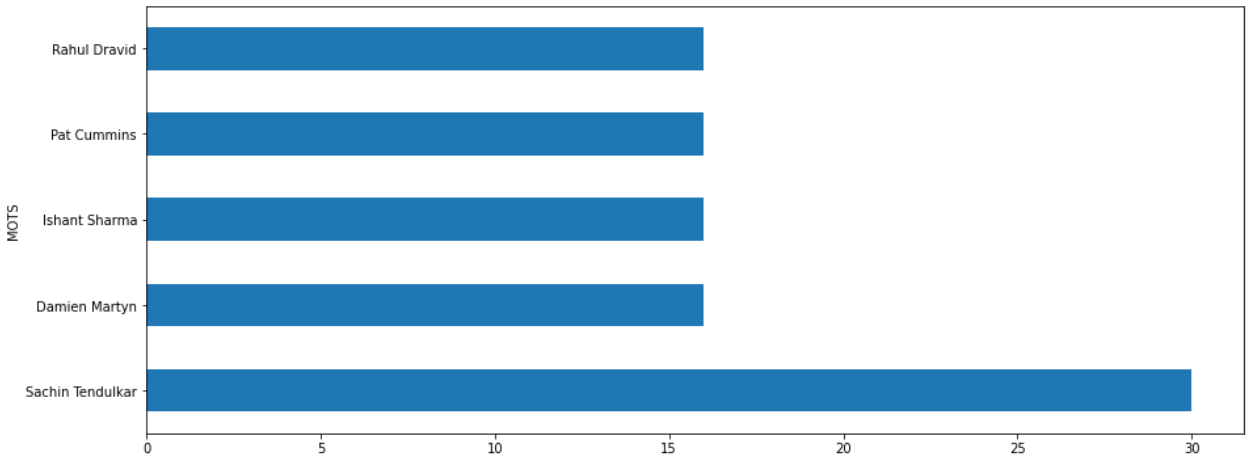
In [36]:

```
plt.figure(figsize=(30, 10))
df.groupby("MOTM")["MOTM"].count().sort_values(ascending = False).head().plot(kind="pie", autopct='%1.1f%%')
plt.show()
```



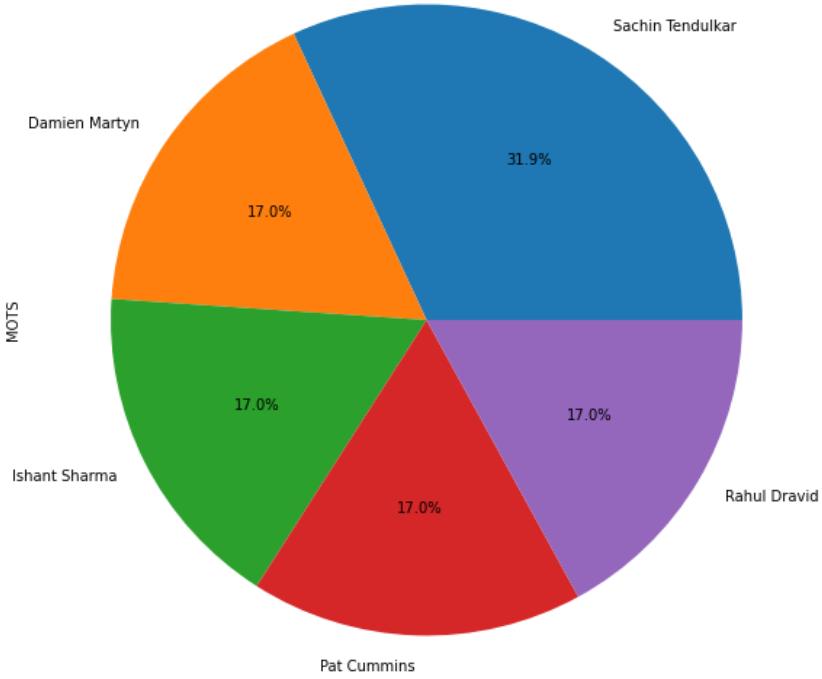
In [37]:

```
plt.figure(figsize=(15, 6))
df.groupby("MOTS")["MOTS"].count().sort_values(ascending = False).head().plot(kind="barh")
plt.show()
```



In [38]:

```
plt.figure(figsize=(30, 10))
df.groupby("MOTS")["MOTS"].count().sort_values(ascending = False).head().plot(kind="pie", autopct='%1.1f%%')
plt.show()
```



In [39]:

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

In [40]:

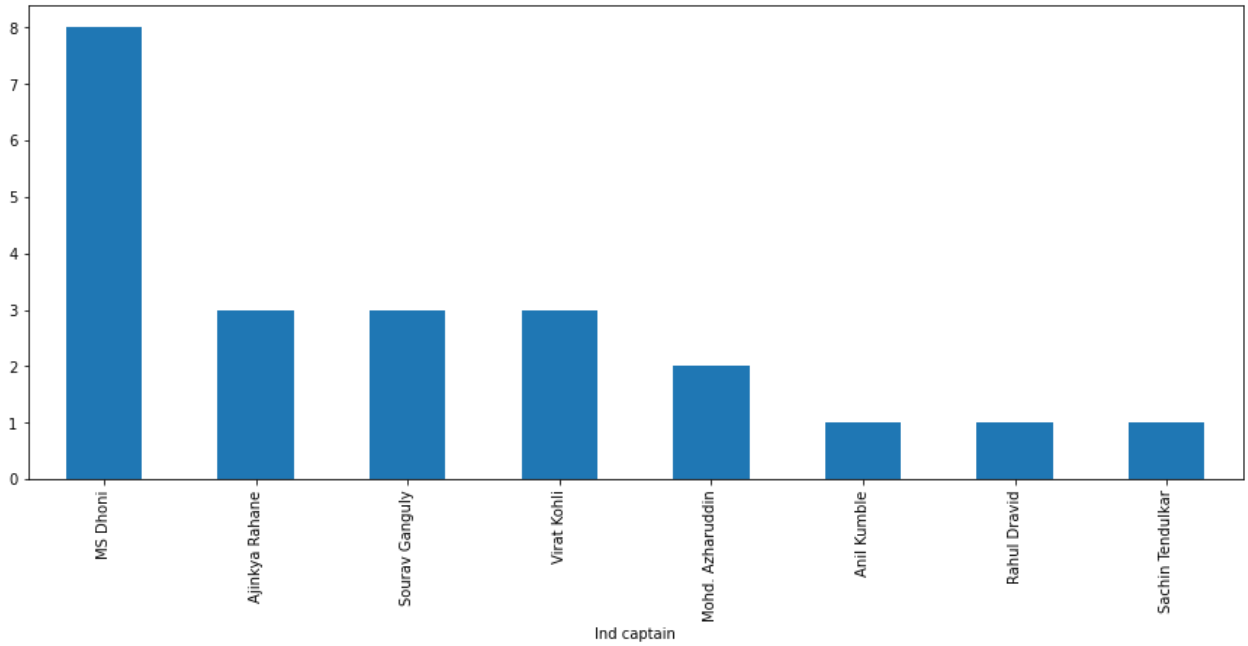
```
df2
```

Out[40]:

	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
5	1999-00	2	Australia	Sachin Tendulkar	Steve Waugh
6	1999-00	3	Australia	Sachin Tendulkar	Steve Waugh
7	2000-01	1	Australia	Sourav Ganguly	Steve Waugh
8	2000-01	2	India	Sourav Ganguly	Steve Waugh
9	2000-01	3	India	Sourav Ganguly	Steve Waugh
10	2003-04	1	Drawn	Sourav Ganguly	Steve Waugh
11	2003-04	2	India	Sourav Ganguly	Steve Waugh
12	2003-04	3	Australia	Sourav Ganguly	Steve Waugh
13	2003-04	4	Drawn	Sourav Ganguly	Steve Waugh
14	2004-05	1	Australia	Sourav Ganguly	Adam Gilchrist
15	2004-05	2	Drawn	Sourav Ganguly	Adam Gilchrist
16	2004-05	3	Australia	Rahul Dravid	Adam Gilchrist
17	2004-05	4	India	Rahul Dravid	Ricky Ponting
18	2007-08	1	Australia	Anil Kumble	Ricky Ponting
19	2007-08	2	Australia	Anil Kumble	Ricky Ponting
20	2007-08	3	India	Anil Kumble	Ricky Ponting
21	2007-08	4	Drawn	Anil Kumble	Ricky Ponting
22	2008-09	1	Drawn	Anil Kumble	Ricky Ponting
23	2008-09	2	India	MS Dhoni	Ricky Ponting
24	2008-09	3	Drawn	Anil Kumble	Ricky Ponting
25	2008-09	4	India	MS Dhoni	Ricky Ponting
26	2010-11	1	India	MS Dhoni	Ricky Ponting
27	2010-11	2	India	MS Dhoni	Ricky Ponting
28	2011-12	1	Australia	MS Dhoni	Michael Clarke
29	2011-12	2	Australia	MS Dhoni	Michael Clarke
30	2011-12	3	Australia	MS Dhoni	Michael Clarke
31	2011-12	4	Australia	MS Dhoni	Michael Clarke
32	2012-13	1	India	MS Dhoni	Michael Clarke
33	2012-13	2	India	MS Dhoni	Michael Clarke
34	2012-13	3	India	MS Dhoni	Michael Clarke
35	2012-13	4	India	MS Dhoni	Michael Clarke
36	2014-15	1	Australia	Virat Kohli	Michael Clarke
37	2014-15	2	Australia	MS Dhoni	Steve Smith
38	2014-15	3	Drawn	MS Dhoni	Steve Smith
39	2014-15	4	Drawn	Virat Kohli	Steve Smith
40	2016-17	1	Australia	Virat Kohli	Steve Smith
41	2016-17	2	India	Virat Kohli	Steve Smith
42	2016-17	3	Drawn	Virat Kohli	Steve Smith
43	2016-17	4	India	Ajinkya Rahane	Steve Smith
44	2018-19	1	India	Virat Kohli	Tim Paine
45	2018-19	2	Australia	Virat Kohli	Tim Paine
46	2018-19	3	India	Virat Kohli	Tim Paine
47	2018-19	4	Drawn	Virat Kohli	Tim Paine
48	2020-21	1	Australia	Virat Kohli	Tim Paine
49	2020-21	2	India	Ajinkya Rahane	Tim Paine
50	2020-21	3	Drawn	Ajinkya Rahane	Tim Paine
51	2020-21	4	India	Ajinkya Rahane	Tim Paine

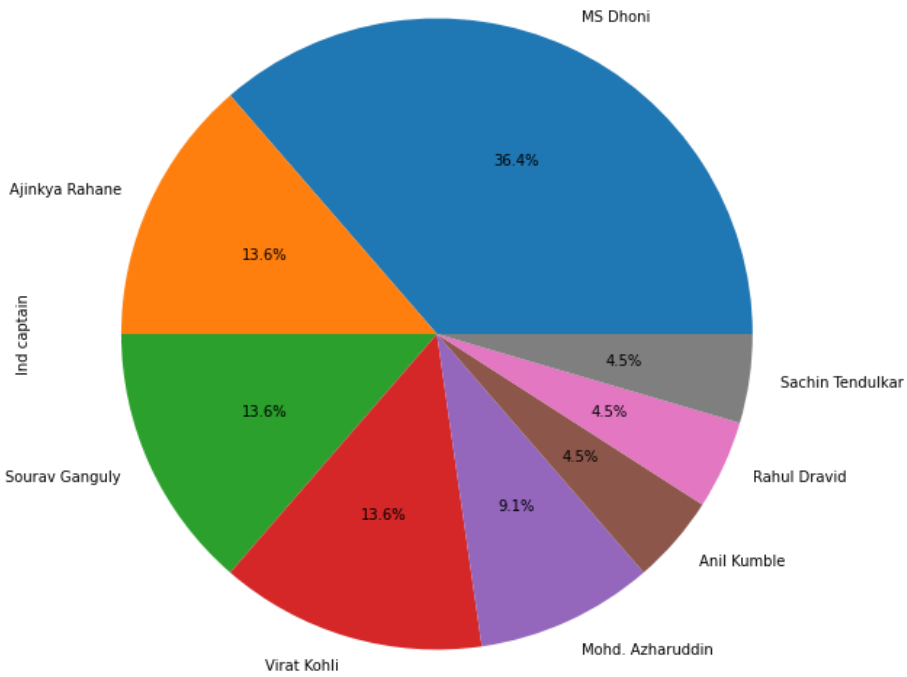
In [41]:

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



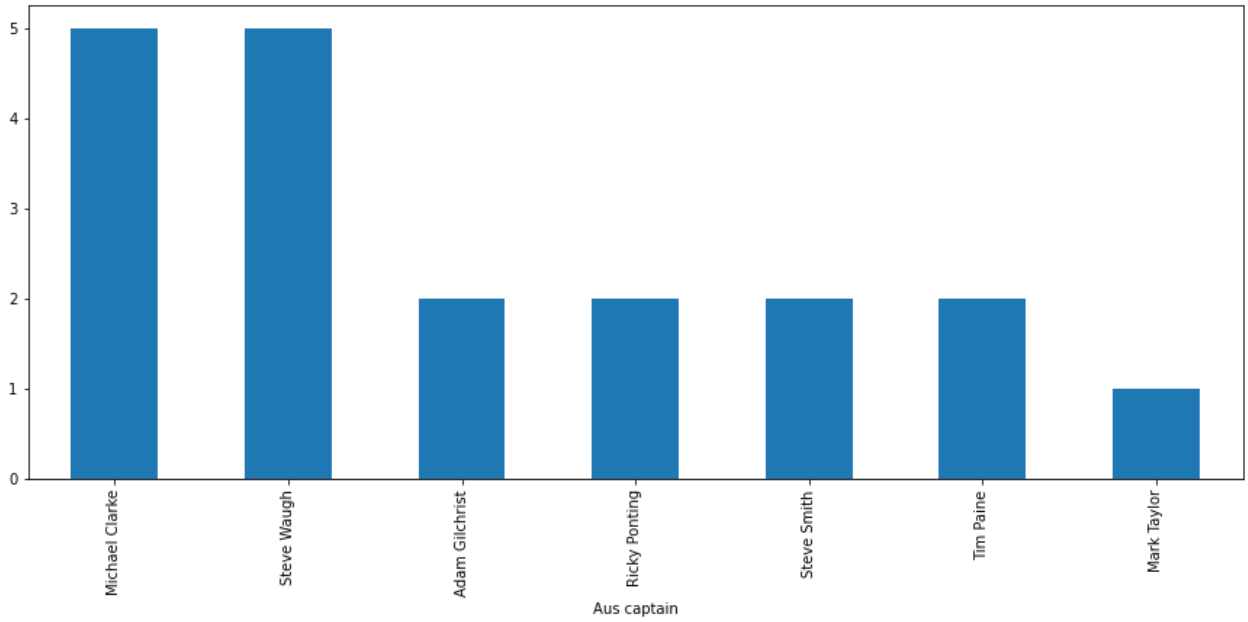
In [42]:

```
plt.figure(figsize=(30, 10))
ind_W.groupby("Ind captain")["Ind captain"].count().sort_values(ascending=False).plot(kind="pie", autopct='%1.1f%%')
plt.show()
```



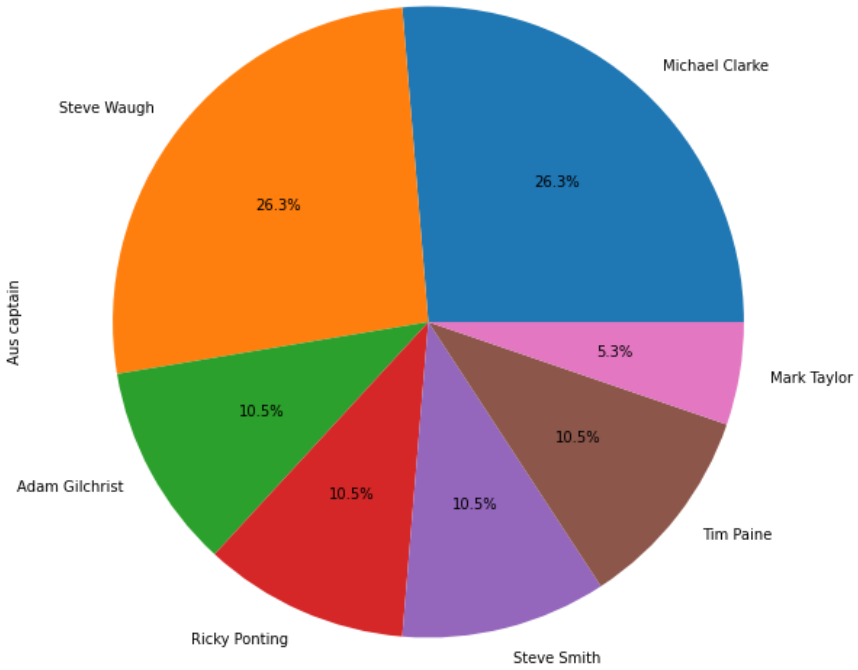
In [43]:

```
Aus_W = df2[df2["Winner"]=="Australia"]
plt.figure(figsize=(15,6))
Aus_W.groupby("Aus captain")["Aus captain"].count().sort_values(ascending=False).plot(kind="bar")
plt.show()
```



In [44]:

```
plt.figure(figsize=(30, 10))
Aus_W.groupby("Aus captain")["Aus captain"].count().sort_values(ascending=False).plot(kind="pie", autopct='%1.1f%%')
plt.show()
```



In [45]:

```
Runs=[]
Wics=[]
run_wic=df['team total'].str.split("/").tolist()
for runs,wics in run_wic:
    Runs.append(runs)
    Wics.append(wics)
```

In [46]:

```
df["RUNS"]=Runs
df["WICKETS"]=Wics
df.drop("team total",axis=1)
df["RUNS"]=df["RUNS"].astype('int32')
df["WICKETS"]=df["WICKETS"].astype('int32')
```

In [47]:

```
df
```

Out[47]:

	Test Series Year	Test Number	Innings	Venue	Host	Highest Scorer	Team	Runs by highest scorer	best bowler	wickets by best bowler	team total	Winner	Win Margin	MOTM	
0	1996-97	1	1	Delhi	India	Michael Slater	Australia	44	Anil Kumble	4.0	182/10	India	7w	Nayan Mongia	T
1	1996-97	1	2	Delhi	India	Nayan Mongia	India	152	Paul Reiffel	3.0	361/10	India	7w	Nayan Mongia	T
2	1996-97	1	3	Delhi	India	Steve Waugh	Australia	67	Anil Kumble	5.0	234/10	India	7w	Nayan Mongia	T
3	1996-97	1	4	Delhi	India	Mohd. Azharuddin	India	21	Paul Reiffel	2.0	56/3	India	7w	Nayan Mongia	T
4	1997-98	1	1	Chennai	India	Navjot Sidhu	India	62	Shane Warne	4.0	257/10	India	179r	Sachin Tendulkar	Az
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
195	2020-21	3	4	Sydney	Australia	Rishabh Pant	India	97	Josh Hazlewood	2.0	334/5	Drawn	0	Steve Smith	
196	2020-21	4	1	Brisbane	Australia	Marnus Labuschagne	Australia	108	T Natarajan	3.0	369/10	India	3w	Rishabh Pant	
197	2020-21	4	2	Brisbane	Australia	Shardul Thakur	India	67	Josh Hazlewood	5.0	336/10	India	3w	Rishabh Pant	
198	2020-21	4	3	Brisbane	Australia	Steve Smith	Australia	55	Mohd. Siraj	5.0	294/10	India	3w	Rishabh Pant	
199	2020-21	4	4	Brisbane	Australia	Shubman Gill	India	91	Pat Cummins	4.0	329/7	India	3w	Rishabh Pant	

200 rows × 20 columns

In [48]:

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

Out[48]:

```
Team
Australia    30296
India        29114
Name: RUNS, dtype: int32
```

In [49]:

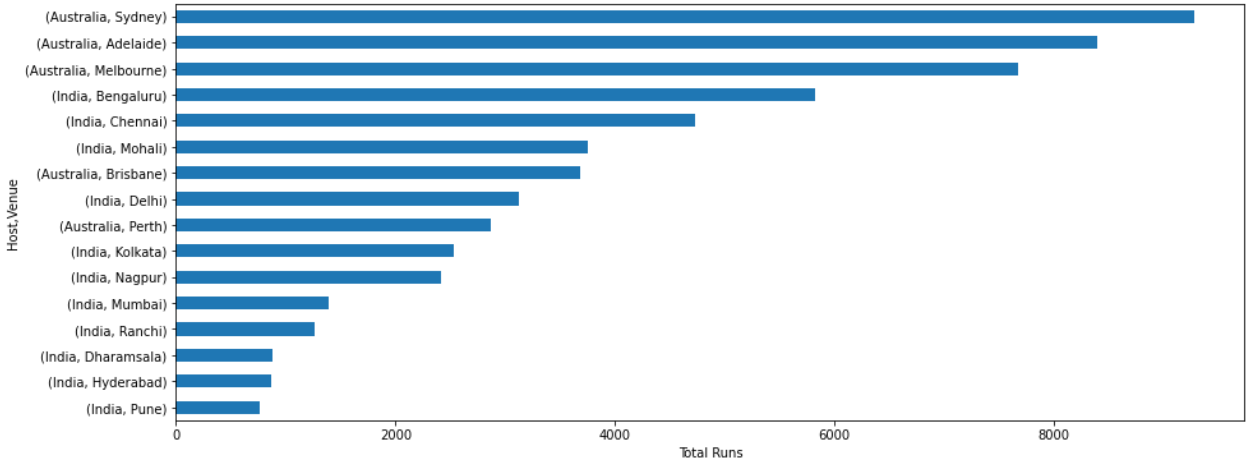
```
df.groupby("Team")["WICKETS"].sum()
```

Out[49]:

```
Team
Australia    864
India        852
Name: WICKETS, dtype: int32
```

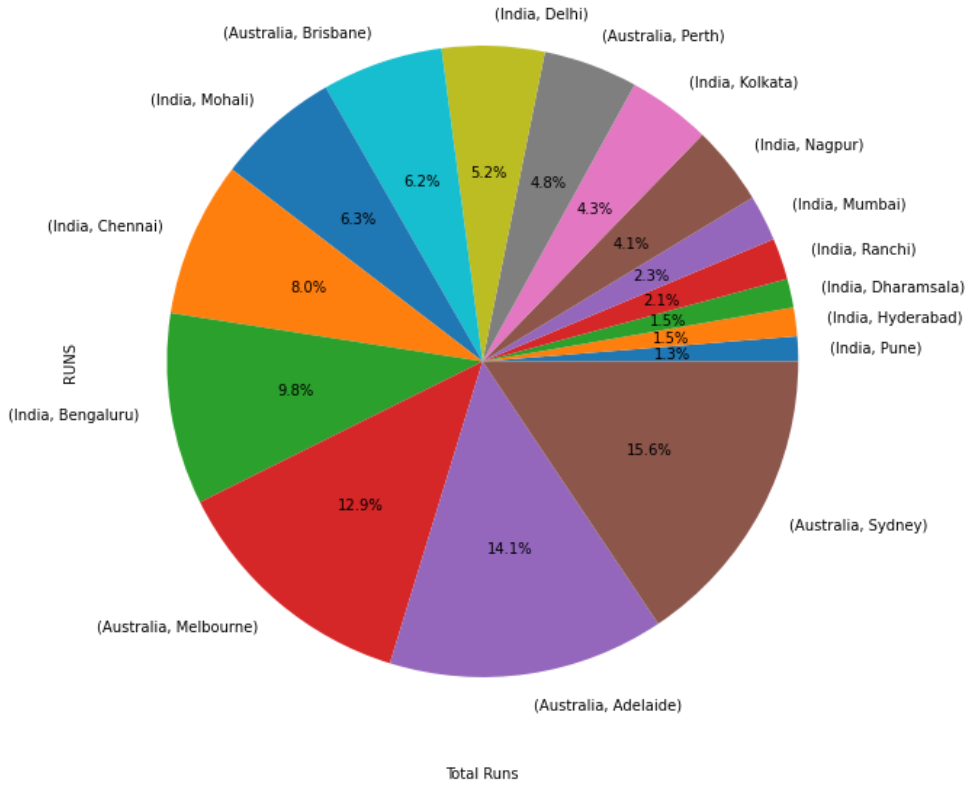
In [50]:

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



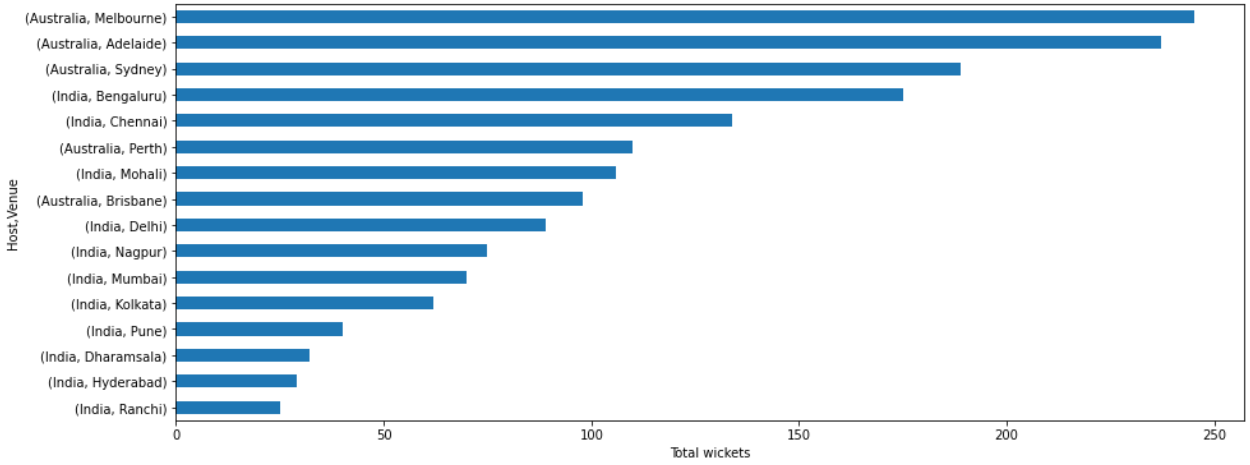
In [51]:

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



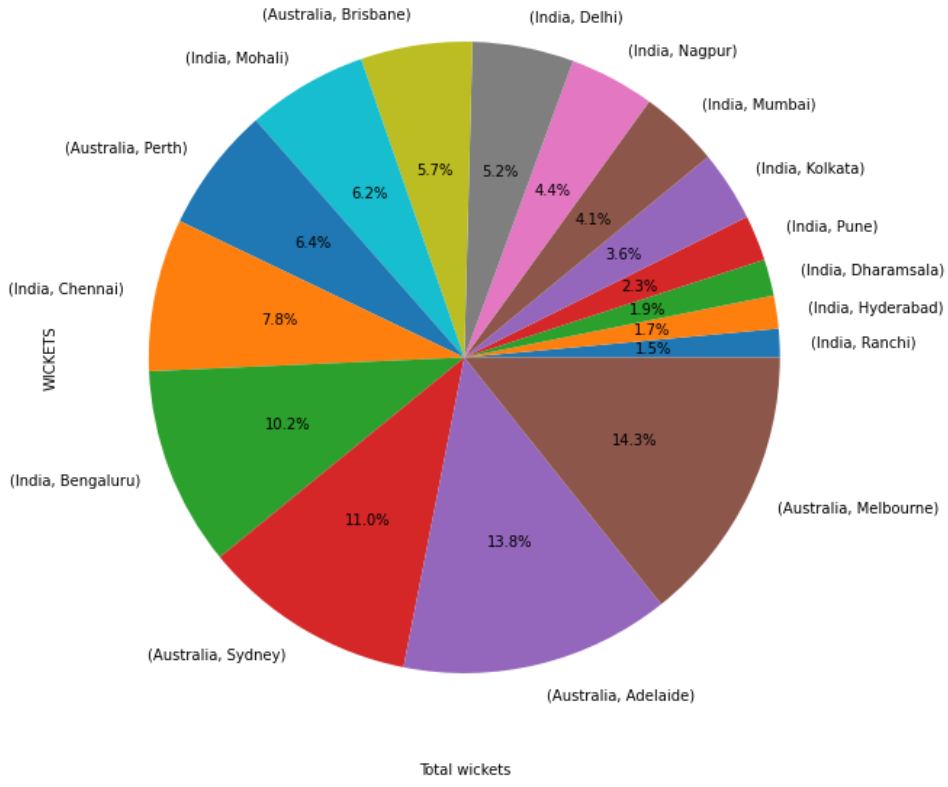
In [52]:

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



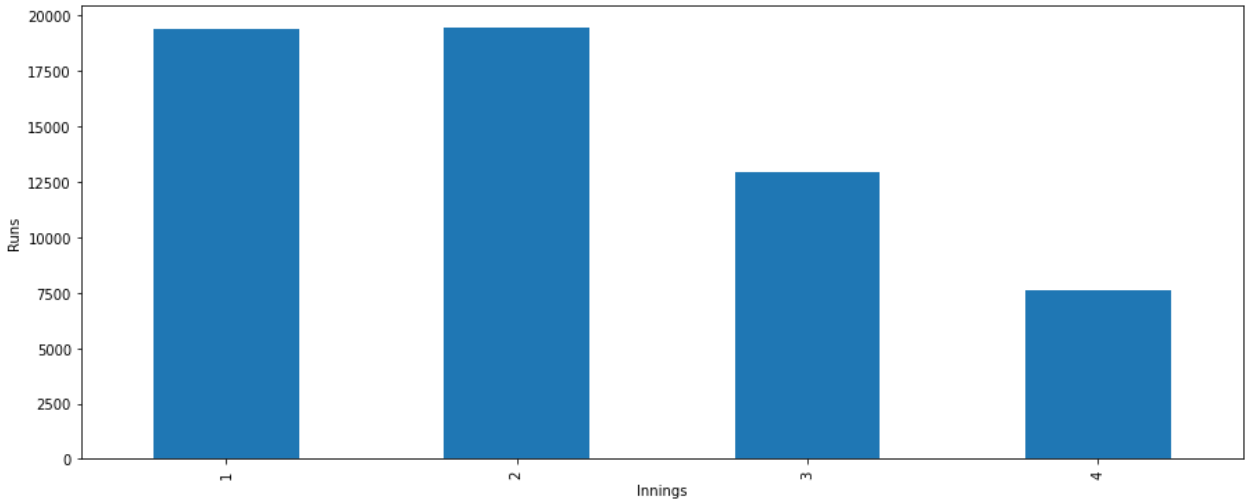
In [53]:

```
plt.figure(figsize=(30, 10))
df.groupby(["Host","Venue"])["WICKETS"].sum().sort_values().plot(kind='pie', autopct='%1.1f%%')
plt.xlabel("Total wickets")
plt.show()
```



In [54]:

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



In [55]:

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