Virat Kohali Centuries (75)

Virat Kohli, the Indian cricket sensation, has achieved a remarkable feat by scoring 75 centuries in international cricket. Known for his exceptional batting skills and unwavering determination, Kohli has left an indelible mark on the cricketing world.

With his aggressive yet calculated approach, Kohli has consistently displayed a hunger for runs and an ability to dominate the opposition. His centuries have come across all formats of the game, including Test matches, One Day Internationals (ODIs), and Twenty20 Internationals (T20Is).

Kohli's century tally is a testament to his consistency and adaptability across different conditions and against various opponents. He has a knack for rising to the occasion in crucial matches and has been a key contributor to India's success on numerous occasions.

Beyond the sheer number of centuries, Kohli's innings have often been a masterclass in batting technique and shot selection. He combines classic stroke play with modern aggression, making him a complete batsman in every sense.

Kohli's records and achievements go beyond the number of centuries. He has amassed numerous accolades and has consistently been ranked among the top batsmen in the world. His passion for the game, coupled with his relentless pursuit of excellence, has made him an inspiration for aspiring cricketers worldwide.

In summary, Virat Kohli's 75 centuries stand as a testament to his exceptional talent, consistency, and dedication to the sport. His achievements have etched his name among the all-time greats of the game and have made him a source of pride for his country, India.

Importing important Liabraries for analysis and visualisations

```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

File path

```python
file = r'd:\rg67266\Downloads\Virat-Kohli-International-Cricket-Centuries.csv'
```

Read the CSV Datasets

```python
df = pd.read_csv(file)
```

```python
df
```

Out[ ]:

	No.	Runs	Against	Position	Innings	Venue	Ground	Date	Result
0	1	107	Sri Lanka	4	2	Eden Gardens, Kolkata	Home	24-Dec-09	Won
1	2	102*	Bangladesh	3	2	Sher-e-Bangla Cricket Stadium, Dhaka	Away	11-Jan-10	Won
2	3	118	Australia	3	2	APCA-VDCA Stadium, Visakhapatnam	Home	20-Oct-10	Won
3	4	105	New Zealand	3	1	Nehru Stadium, Guwahati	Home	28-Nov-10	Won
4	5	100*	Bangladesh	4	1	Sher-e-Bangla Cricket Stadium, Dhaka	Away	19-Feb-11	Won
...	...	...	...	...	...	...	...	...	...
70	71	122*	Afghanistan	2	1	Dubai International Cricket Stadium, Dubai	Neutral	08-Sep-22	Won
71	72	113	Bangladesh	3	1	Zohur Ahmed Chowdhury Stadium, Chittagong	Away	10-Dec-22	Won
72	73	113	Sri Lanka	3	1	ACA Stadium, Guwahati	Home	10-Jan-23	Won
73	74	166*	Sri Lanka	3	1	Greenfield International Stadium, Thiruvananth...	Home	15-Jan-23	Won
74	75	186	Australia	4	2	Narendra Modi Stadium, Ahmedabad	Home	09-Mar-23	Drawn

75 rows × 9 columns

Check the columns and rows

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

Out[ ]: (75, 9)

Columns

```
In [ ]:  df.columns
```

```
Out[ ]:  Index(['No.', 'Runs', 'Against', 'Position', 'Innings', 'Venue', 'Ground',
                'Date', 'Result'],
               dtype='object')
```

Renaming the Columns

```
In [ ]:  df = df.rename(columns={
             'No.' : "Century No"
         })
```

```
In [ ]:  df.columns
```

```
Out[ ]:  Index(['Century No', 'Runs', 'Against', 'Position', 'Innings', 'Venue',
                'Ground', 'Date', 'Result'],
               dtype='object')
```

Checking the null columns or rows

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

```
Out[ ]:  Century No    0
         Runs          0
         Against       0
         Position      0
         Innings       0
         Venue         0
         Ground        0
         Date          0
         Result        0
         dtype: int64
```
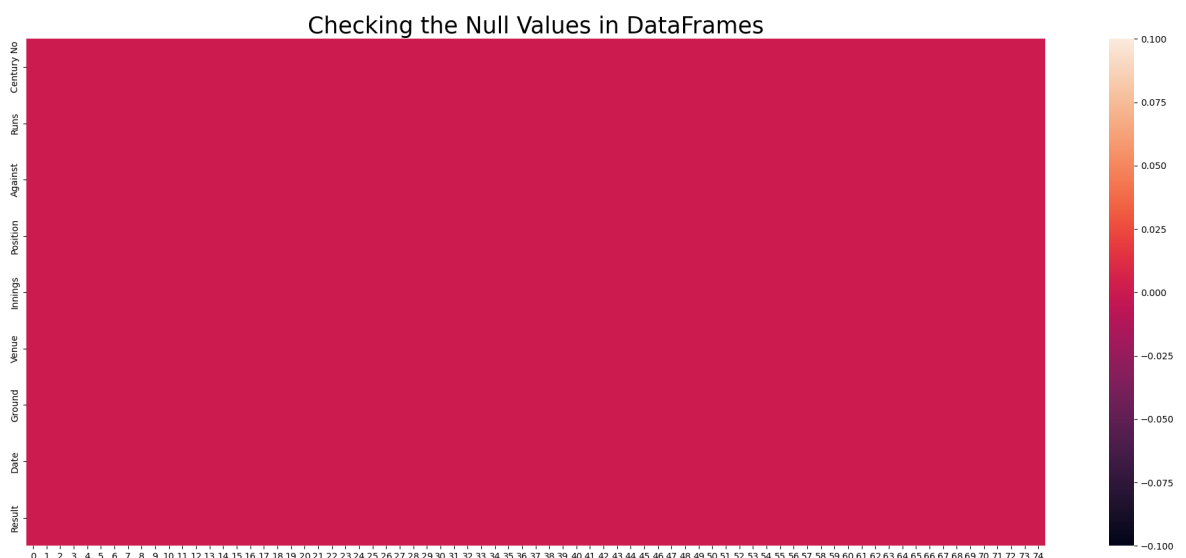
```
In [ ]:  missing_values = df.isnull()
         missing_values
```

	Century No	Runs	Against	Position	Innings	Venue	Ground	Date	Result
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...
70	False	False	False	False	False	False	False	False	False
71	False	False	False	False	False	False	False	False	False
72	False	False	False	False	False	False	False	False	False
73	False	False	False	False	False	False	False	False	False
74	False	False	False	False	False	False	False	False	False

75 rows × 9 columns

```python
# Transpose the boolean DataFrame to match the shape of the input data
missing_values = missing_values.transpose()
```

```python
plt.figure(figsize=(25,10)) # for inscrease the size of graph
plt.title('Checking the Null Values in DataFrames' , fontsize=25)

# Visualize the missing values using a heatmap
sns.heatmap(missing_values)
plt.show()
```


- There is no any Null values in datasets

Check the info with Datatypes

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 75 entries, 0 to 74
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Century No  75 non-null     int64
 1   Runs        75 non-null     object
 2   Against     75 non-null     object
 3   Position    75 non-null     int64
 4   Innings     75 non-null     int64
 5   Venue       75 non-null     object
 6   Ground      75 non-null     object
 7   Date        75 non-null     object
 8   Result      75 non-null     object
dtypes: int64(3), object(6)
memory usage: 5.4+ KB
```

Focusing On:

- Need to change data type of columns 'Runs' and 'Date'

Changing the Data Types of columns 'Runs' and 'Date'

```
In [ ]: df['Date'] = pd.to_datetime(df['Date'])
```

```
C:\Users\rg67266\AppData\Local\Temp\ipykernel_16200\2394721818.py:1: UserWarning:
Could not infer format, so each element will be parsed individually, falling back
to `dateutil`. To ensure parsing is consistent and as-expected, please specify a
format.
  df['Date'] = pd.to_datetime(df['Date'])
```

```
In [ ]: df['Runs'] = df['Runs'].str.replace("*",'')
```

```
In [ ]: df['Runs'] = df['Runs'].astype(np.int64)
```

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 75 entries, 0 to 74
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Century No  75 non-null     int64
 1   Runs        75 non-null     int64
 2   Against     75 non-null     object
 3   Position    75 non-null     int64
 4   Innings     75 non-null     int64
 5   Venue       75 non-null     object
 6   Ground      75 non-null     object
 7   Date        75 non-null     datetime64[ns]
 8   Result      75 non-null     object
dtypes: datetime64[ns](1), int64(4), object(4)
memory usage: 5.4+ KB
```

Data Analysis with EDA

Q1: What are the countries name where Virat Kohali withwhome against made Centuries?

```python
In [ ]: plt.figure(figsize=(25, 10))  # Increase the size of the graph
        plt.title('All Century Counts against Country', fontsize=25)

        ax = sns.countplot(data=df, x='Against')

        plt.xlabel('Country', fontsize=15)  # Set x-axis label
        plt.ylabel('Count', fontsize=15)  # Set y-axis label

        # Add text annotations to the bars
        for p in ax.patches:
            ax.annotate(format(p.get_height(), '.0f'), (p.get_x() + p.get_width() / 2, p

        plt.show()
```
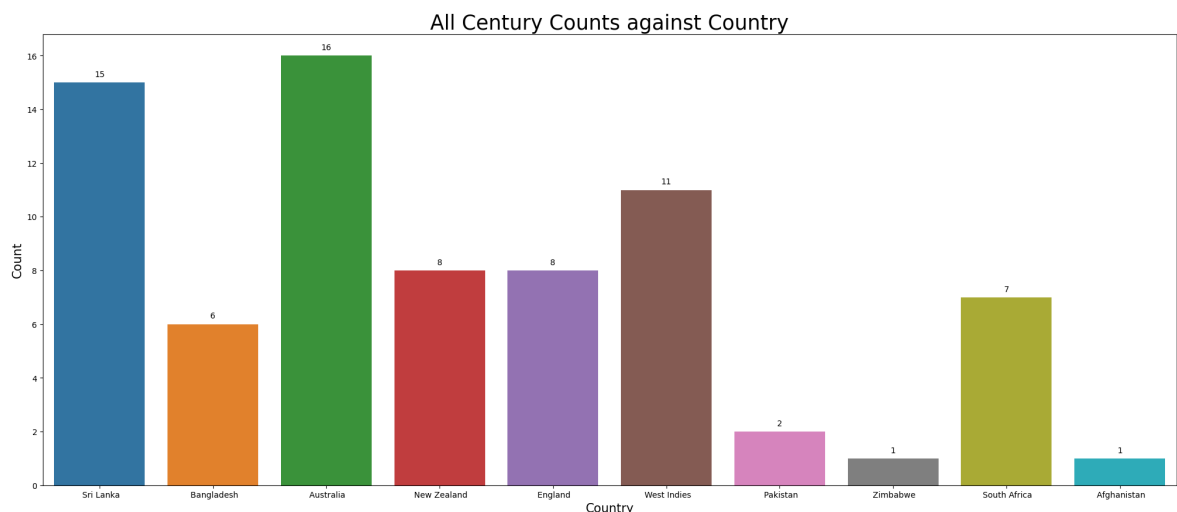
Result:

- Virat has made most 16 century against Austrlia in all formate
- Srilanka is the secend high with 15 century in all formate
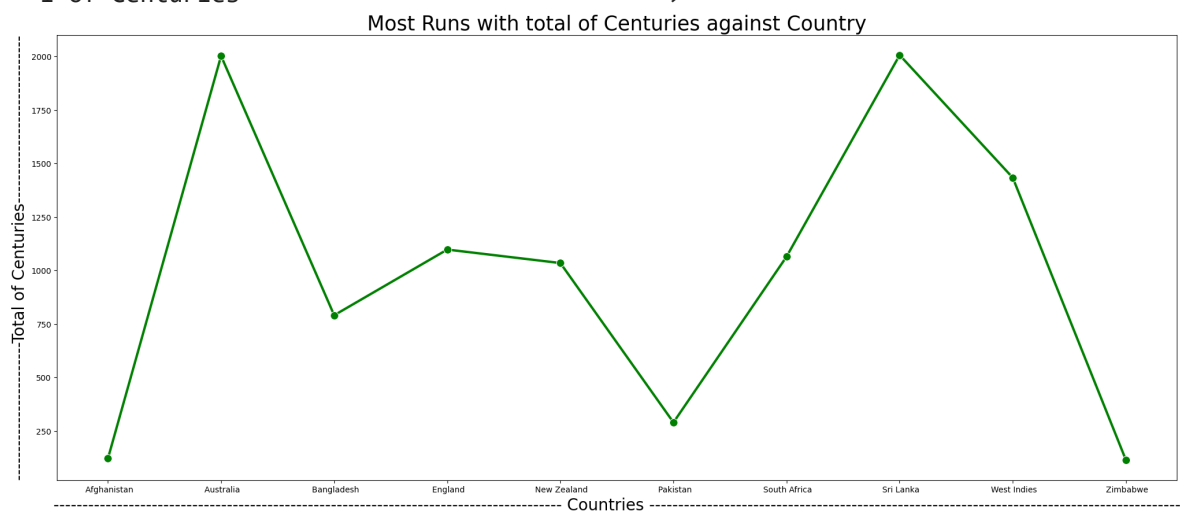
Q2: what are the Most Runs of Only century of Virat Kohali against Countries?

```
In [ ]:  againt_run = df.groupby('Against')['Runs'].sum()
         againt_run
```

```
Out[ ]:  Against
         Afghanistan     122
         Australia      2003
         Bangladesh      791
         England        1098
         New Zealand    1035
         Pakistan        290
         South Africa   1065
         Sri Lanka      2007
         West Indies    1434
         Zimbabwe        115
         Name: Runs, dtype: int64
```

```
In [ ]:  plt.figure(figsize=(25, 10))   # Increase the size of the graph
         plt.title('Most Runs with total of Centuries against Country', fontsize=25)
         plt.xlabel('---------------------------------------------------------------
         plt.ylabel('-----------------------Total of Centuries--------------------


         sns.lineplot(x=againt_run.index, y=againt_run.values, color='g', markers=True, m
```

```
Out[ ]:  <Axes: title={'center': 'Most Runs with total of Centuries against Country'}, x
         label='---------------------------------------------------------------------
         --------------- Countries ---------------------------------------------------
         --------------------------------------', ylabel='-----------------------Tota
         l of Centuries--------------------------'>
```

Result:

- Virat has beat most of runs against Sri Lanka with 2007 runs only in century number
- and Secend one in Austrila with 2003 Runs in century number

Q3: Best inning for Centuries?

```python
def inn(data):
    if data == 1:
        return "First Inning"
    elif data == 2:
        return "Second Inning"
    elif data == 3:
        return "Third Inning"
    elif data == 4:
        return "Forth Inning"
```

```python
df['Inning'] = df['Innings'].apply(inn)
```

```python
df.columns
```

```
Index(['Century No', 'Runs', 'Against', 'Position', 'Innings', 'Venue',
       'Ground', 'Date', 'Result', 'Inning'],
      dtype='object')
```

```python
column = list(df.columns)
new_column = column[0:5] + column[-1:] + column[5:-1]
df[new_column]
```

Out[ ]:

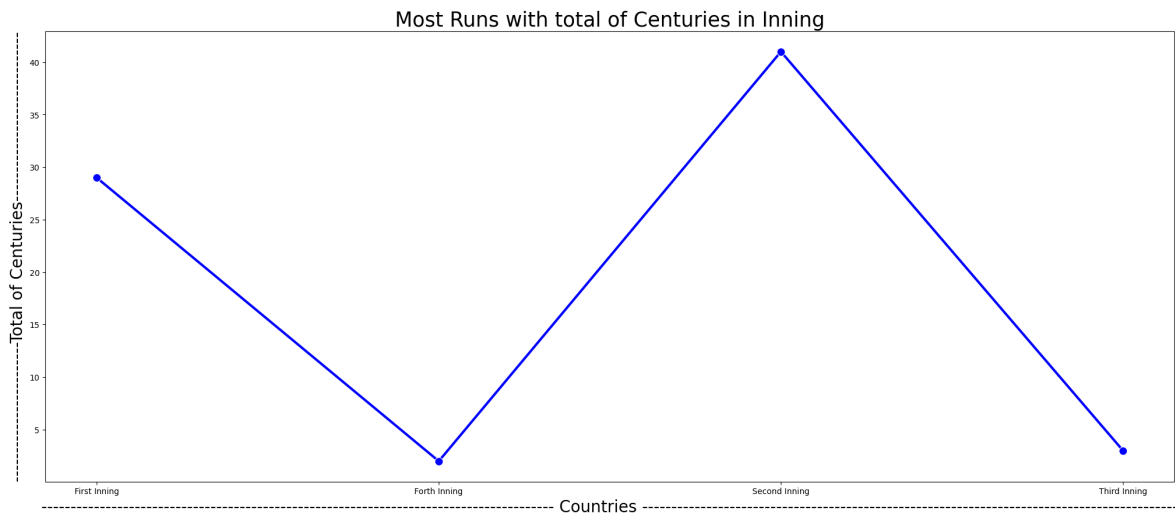|	Century No	Runs	Against	Position	Innings	Inning	Venue	Ground	Da
0	1	107	Sri Lanka	4	2	Second Inning	Eden Gardens, Kolkata	Home	200 12-
1	2	102	Bangladesh	3	2	Second Inning	Sher-e-Bangla Cricket Stadium, Dhaka	Away	201 01-
2	3	118	Australia	3	2	Second Inning	APCA-VDCA Stadium, Visakhapatnam	Home	201 10-
3	4	105	New Zealand	3	1	First Inning	Nehru Stadium, Guwahati	Home	201 11-
4	5	100	Bangladesh	4	1	First Inning	Sher-e-Bangla Cricket Stadium, Dhaka	Away	201 02-
...	...	...	...	...	...	...	...	...	
70	71	122	Afghanistan	2	1	First Inning	Dubai International Cricket Stadium, Dubai	Neutral	202 09-
71	72	113	Bangladesh	3	1	First Inning	Zohur Ahmed Chowdhury Stadium, Chittagong	Away	202 12-
72	73	113	Sri Lanka	3	1	First Inning	ACA Stadium, Guwahati	Home	202 01-
73	74	166	Sri Lanka	3	1	First Inning	Greenfield International Stadium, Thiruvananth...	Home	202 01-
74	75	186	Australia	4	2	Second Inning	Narendra Modi Stadium, Ahmedabad	Home	202 03-

75 rows × 10 columns

```python
inning = df.groupby('Inning')['Runs'].count()
inning
```

Out[ ]:
```
Inning
First Inning     29
Forth Inning      2
Second Inning    41
Third Inning      3
Name: Runs, dtype: int64
```

```
In [ ]:  plt.figure(figsize=(25, 10))   # Increase the size of the graph
         plt.title('Most Runs with total of Centuries in Inning', fontsize=25)
         plt.xlabel('----------------------------------------------------------
         plt.ylabel('-----------------------Total of Centuries---------------

         sns.lineplot(x=inning.index, y=inning.values, color='b', markers=True, marker='c
```

```
Out[ ]:  <Axes: title={'center': 'Most Runs with total of Centuries in Inning'}, xlabel
         ='------------------------------------------------------------------
         ----------- Countries ---------------------------------------------
         --------------------------------', ylabel='-----------------------Total of
         Centuries----------------------------'>
```


Most Runs with total of Centuries in Inning

Result:

- 41 century made in second inning by Virat Kohali

Q4: Which one are the Best ground for Virat Kohali for Century in all over the world.
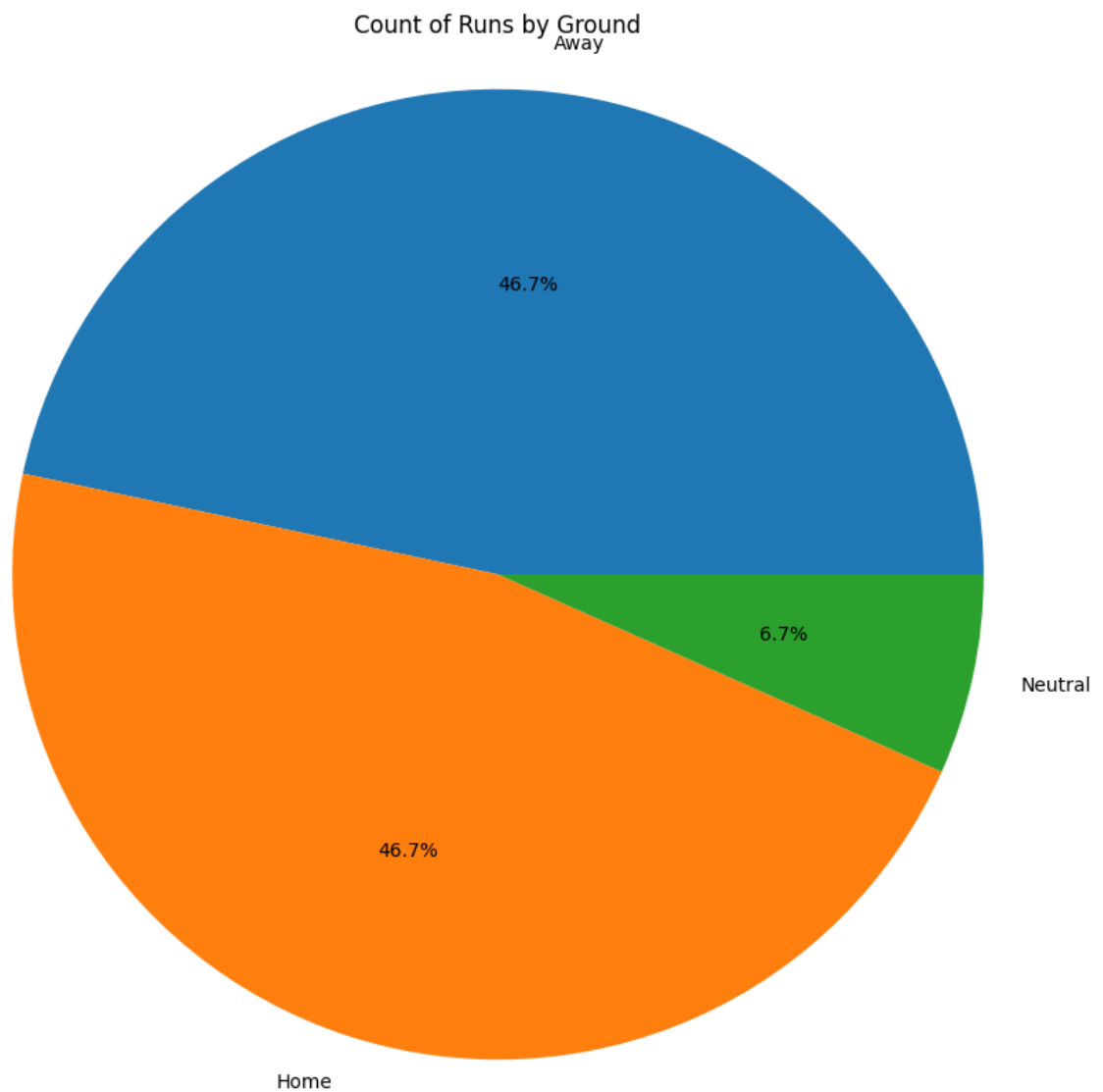
```
In [ ]:  grouped_data = df.groupby('Ground')['Runs'].count()
         grouped_data
```

```
Out[ ]:  Ground
         Away      35
         Home      35
         Neutral    5
         Name: Runs, dtype: int64
```

```
In [ ]:  plt.figure(figsize=(10, 10))   # Set the size of the pie chart
         plt.title('Count of Runs by Ground')   # Set the title of the pie chart

         # Create the pie chart
         plt.pie(grouped_data, labels=grouped_data.index, autopct='%1.1f%%')
```

```
plt.axis('equal')  # Equal aspect ratio ensures a circular pie
plt.show()
```

Count of Runs by Ground

Result

- The number of century made by kohali with equal number in home or away with 35 century.

```
In [ ]: vanue = df.groupby('Venue')['Runs'].count().reset_index()
        vanue
```
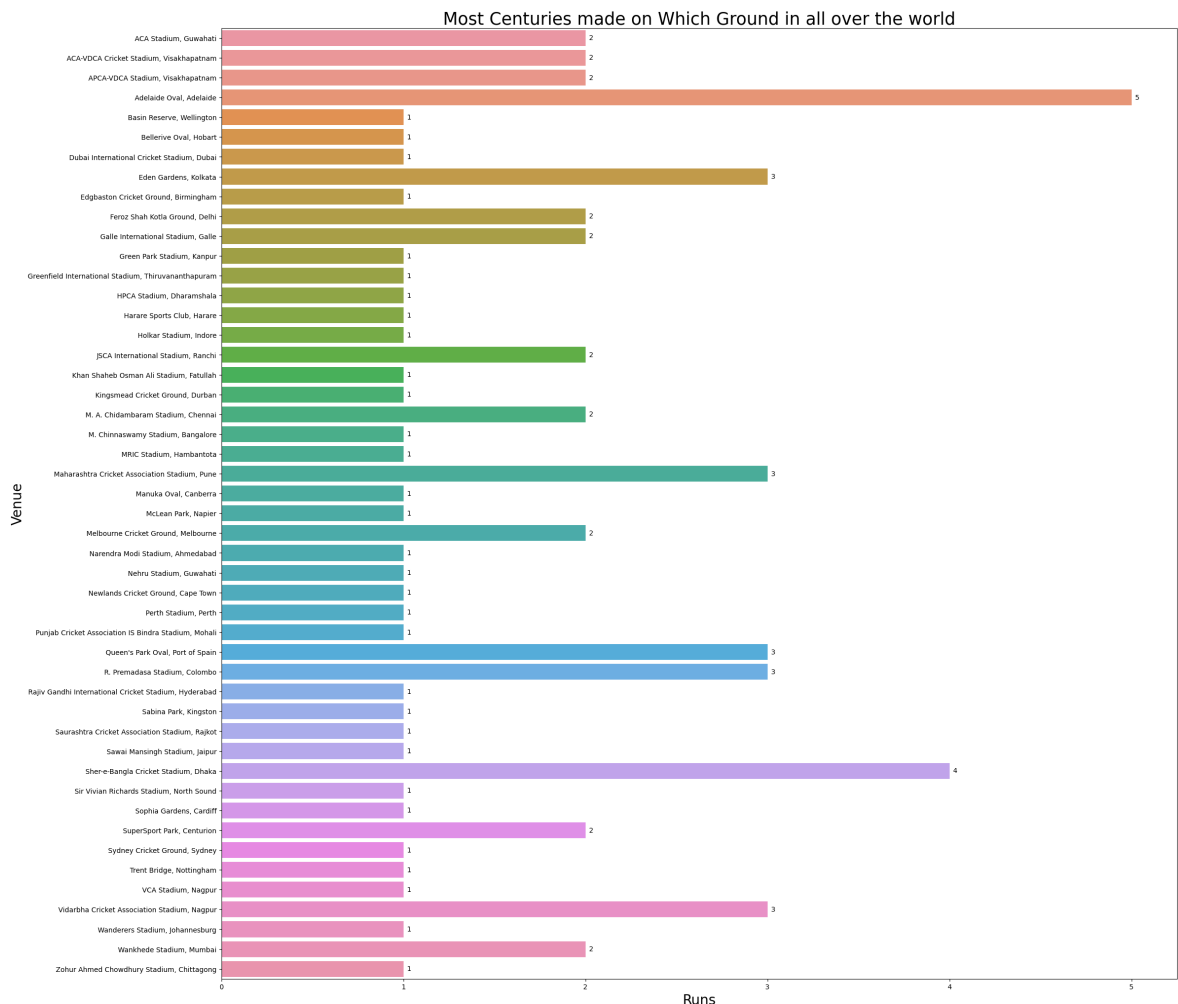
	Venue	Runs
0	ACA Stadium, Guwahati	2
1	ACA-VDCA Cricket Stadium, Visakhapatnam	2
2	APCA-VDCA Stadium, Visakhapatnam	2
3	Adelaide Oval, Adelaide	5
4	Basin Reserve, Wellington	1
5	Bellerive Oval, Hobart	1
6	Dubai International Cricket Stadium, Dubai	1
7	Eden Gardens, Kolkata	3
8	Edgbaston Cricket Ground, Birmingham	1
9	Feroz Shah Kotla Ground, Delhi	2
10	Galle International Stadium, Galle	2
11	Green Park Stadium, Kanpur	1
12	Greenfield International Stadium, Thiruvananth...	1
13	HPCA Stadium, Dharamshala	1
14	Harare Sports Club, Harare	1
15	Holkar Stadium, Indore	1
16	JSCA International Stadium, Ranchi	2
17	Khan Shaheb Osman Ali Stadium, Fatullah	1
18	Kingsmead Cricket Ground, Durban	1
19	M. A. Chidambaram Stadium, Chennai	2
20	M. Chinnaswamy Stadium, Bangalore	1
21	MRIC Stadium, Hambantota	1
22	Maharashtra Cricket Association Stadium, Pune	3
23	Manuka Oval, Canberra	1
24	McLean Park, Napier	1
25	Melbourne Cricket Ground, Melbourne	2
26	Narendra Modi Stadium, Ahmedabad	1
27	Nehru Stadium, Guwahati	1
28	Newlands Cricket Ground, Cape Town	1
29	Perth Stadium, Perth	1
30	Punjab Cricket Association IS Bindra Stadium, ...	1
31	Queen's Park Oval, Port of Spain	3
32	R. Premadasa Stadium, Colombo	3

	Venue	Runs
33	Rajiv Gandhi International Cricket Stadium, Hy…	1
34	Sabina Park, Kingston	1
35	Saurashtra Cricket Association Stadium, Rajkot	1
36	Sawai Mansingh Stadium, Jaipur	1
37	Sher-e-Bangla Cricket Stadium, Dhaka	4
38	Sir Vivian Richards Stadium, North Sound	1
39	Sophia Gardens, Cardiff	1
40	SuperSport Park, Centurion	2
41	Sydney Cricket Ground, Sydney	1
42	Trent Bridge, Nottingham	1
43	VCA Stadium, Nagpur	1
44	Vidarbha Cricket Association Stadium, Nagpur	3
45	Wanderers Stadium, Johannesburg	1
46	Wankhede Stadium, Mumbai	2
47	Zohur Ahmed Chowdhury Stadium, Chittagong	1

In [ ]:
```python
plt.figure(figsize=(25, 25))  # Set the size of the bar plot
plt.title('Most Centuries made on Which Ground in all over the world', fontsize=
plt.xlabel('Runs', fontsize=20)  # Set x-axis label
plt.ylabel('Venue', fontsize=20)  # Set y-axis label

ax = sns.barplot(data=vanue, y='Venue', x='Runs')

# Add text annotations to the bars
for p in ax.patches:
    ax.annotate(format(p.get_width(), '.0f'), (p.get_width(), p.get_y() + p.get_

plt.show()
```

Most Centuries made on Which Ground in all over the world

Result:

- Most highest number of Century in Adelaide Oval, Adelaide in Austrlia with 5 Centuries.
- and Secend one is Sher-e-Bangla Cricket Stadium, Dhaka in Bangladesh with 4 Centuries.

Q5: What is result of centuries number in winning or lossing cose
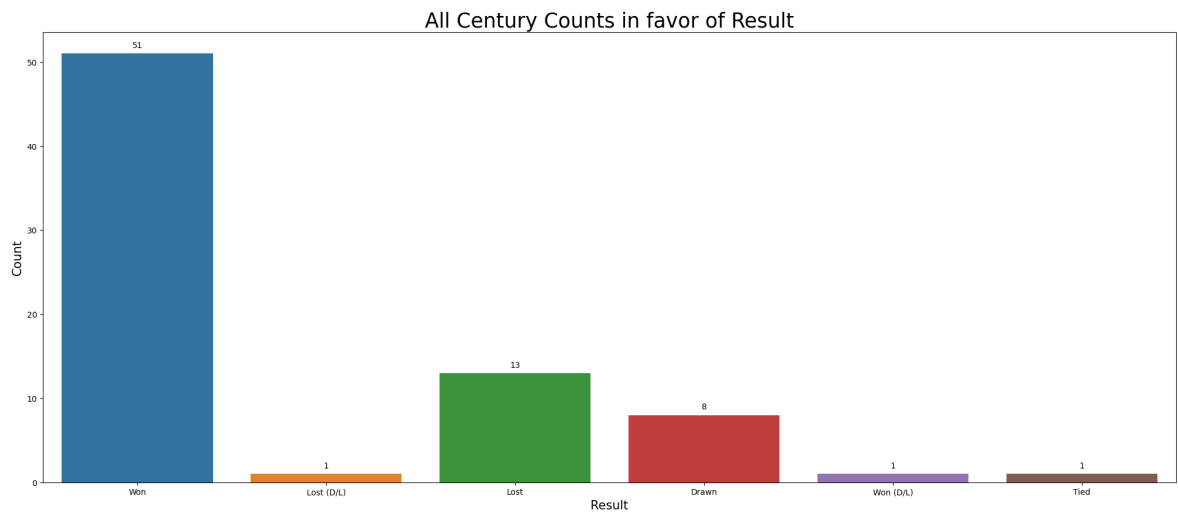
```
In [ ]: plt.figure(figsize=(25, 10))  # Increase the size of the graph
        plt.title('All Century Counts in favor of Result', fontsize=25)

        ax = sns.countplot(data=df, x='Result')

        plt.xlabel('Result', fontsize=15)  # Set x-axis label
        plt.ylabel('Count', fontsize=15)  # Set y-axis label

        # Add text annotations to the bars
        for p in ax.patches:
            ax.annotate(format(p.get_height(), '.0f'), (p.get_x() + p.get_width() / 2, p

        plt.show()
```

All Century Counts in favor of Result

Result:

- Virat has made 50+ centuries in winning favour

Q6: Which year is the best for Virat for Centuries?
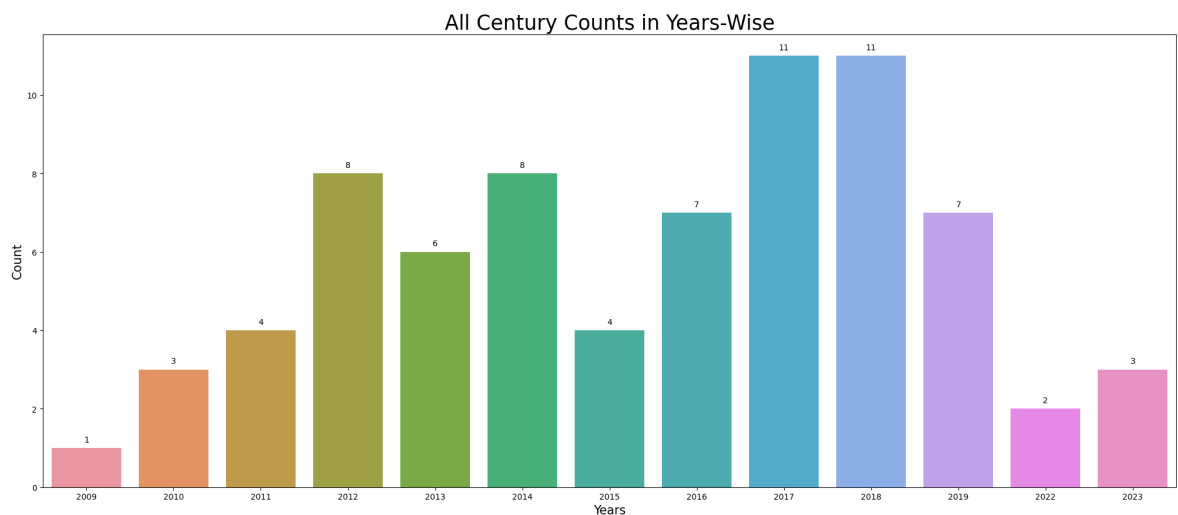
```python
df['Years'] = df['Date'].dt.year
```

```python
plt.figure(figsize=(25, 10))  # Increase the size of the graph
plt.title('All Century Counts in Years-Wise', fontsize=25)

ax = sns.countplot(data=df, x='Years')

plt.xlabel('Years', fontsize=15)  # Set x-axis label
plt.ylabel('Count', fontsize=15)  # Set y-axis label

# Add text annotations to the bars
for p in ax.patches:
    ax.annotate(format(p.get_height(), '.0f'), (p.get_x() + p.get_width() / 2, p

plt.show()
```


All Century Counts in Years-Wise

Result:

- for Virat Kohali's best year for centuries are 2017 and 2018 with equal number of 11.

Q7: Total Centuries of Virat kohali

```
In [ ]: print(f'Total Century of Virat Kohali is: {df.shape[0]}')
```

Total Century of Virat Kohali is: 75

Q8: Last 5 cennturies against which countries?

```
In [ ]: last5 = df.tail()
         last5
```

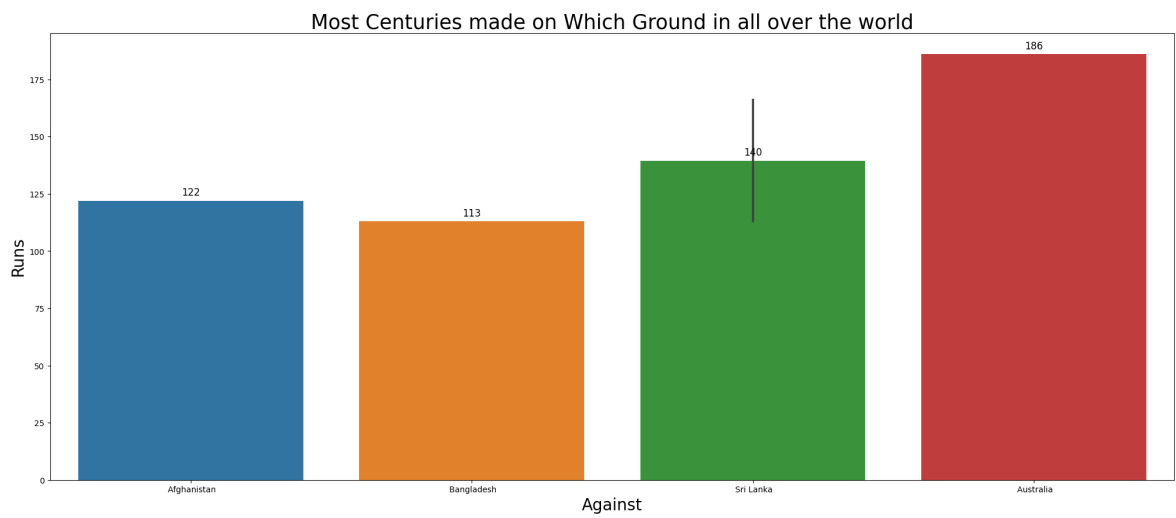Out[ ]:

	Century No	Runs	Against	Position	Innings	Venue	Ground	Date	Resul
70	71	122	Afghanistan	2	1	Dubai International Cricket Stadium, Dubai	Neutral	2022-09-08	Wol
71	72	113	Bangladesh	3	1	Zohur Ahmed Chowdhury Stadium, Chittagong	Away	2022-12-10	Wol
72	73	113	Sri Lanka	3	1	ACA Stadium, Guwahati	Home	2023-01-10	Wol
73	74	166	Sri Lanka	3	1	Greenfield International Stadium, Thiruvananth...	Home	2023-01-15	Wol
74	75	186	Australia	4	2	Narendra Modi Stadium, Ahmedabad	Home	2023-03-09	Drawi

```
In [ ]: plt.figure(figsize=(25, 10))  # Set the size of the bar plot
         plt.title('Most Centuries made on Which Ground in all over the world', fontsize=
         plt.xlabel('Runs', fontsize=20)  # Set x-axis label
         plt.ylabel('Venue', fontsize=20)  # Set y-axis label

         ax = sns.barplot(data=last5, x='Against', y='Runs')

         # Set label numbers on the bars
         for p in ax.patches:
```

```
    ax.annotate(format(p.get_height(), '.0f'), (p.get_x() + p.get_width() / 2.,
            ha = 'center', va = 'center', xytext = (0, 10), textcoords = 'of

plt.show()
```


Most Centuries made on Which Ground in all over the world

Result:

- Last 5 centuries where is 186 against Auctralia in Narendra Modi Statdium is high number.

```
In [ ]:  df
```

Out[ ]:

	Century No	Runs	Against	Position	Innings	Venue	Ground	Date	Resu
0	1	107	Sri Lanka	4	2	Eden Gardens, Kolkata	Home	2009-12-24	Wo
1	2	102	Bangladesh	3	2	Sher-e-Bangla Cricket Stadium, Dhaka	Away	2010-01-11	Wo
2	3	118	Australia	3	2	APCA-VDCA Stadium, Visakhapatnam	Home	2010-10-20	Wo
3	4	105	New Zealand	3	1	Nehru Stadium, Guwahati	Home	2010-11-28	Wo
4	5	100	Bangladesh	4	1	Sher-e-Bangla Cricket Stadium, Dhaka	Away	2011-02-19	Wo
...	...	...	...	...	...	...	...	...	
70	71	122	Afghanistan	2	1	Dubai International Cricket Stadium, Dubai	Neutral	2022-09-08	Wo
71	72	113	Bangladesh	3	1	Zohur Ahmed Chowdhury Stadium, Chittagong	Away	2022-12-10	Wo
72	73	113	Sri Lanka	3	1	ACA Stadium, Guwahati	Home	2023-01-10	Wo
73	74	166	Sri Lanka	3	1	Greenfield International Stadium, Thiruvananth...	Home	2023-01-15	Wo
74	75	186	Australia	4	2	Narendra Modi Stadium, Ahmedabad	Home	2023-03-09	Draw

75 rows × 11 columns