INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

Exploratory Data Analysis on ODI Batting Statistics

Presented by :
Saurabh Shirole

About me

➢ **Name : Saurabh Kailas Shirole**

➢ **Education : Bachelor's in Computer Science**

➢ **http://www.linkedin.com/in/saurabh-shirole-19136a361**

➢ **https://github.com/saurabhks001**

Business Problem

Cricket analysts, and fans struggle to:

➢ **We have thousands of matches and hundreds of players across countries.**
➢ **The data is huge, unorganized, and difficult to compare**.

Objective

To collect and analyze ODI batting data to:

➢ **Find top-performing players and trends.**
➢ **Visualize stats for better understanding.**

Web scrapping process

Import the necessary libraries (BeautifulSoup, requests, regex)

URL = "https://stats.Espncricinfo.Com/ci/engine/stats/index.Html?Class=2;template=results;type=batting"

Request to URL

Get response

Scrape the data

Collect in data frame

Website

First Look of Website

STATISTICS / STATSGURU / ONE-DAY INTERNATIONALS / BATTING RECORDS   2025   ASK CRICINFO   STATSGURU   IPL   WORLD CUP   WPL   MORE ⌄

| Tests | **ODIs** | T20Is | All Test/ODI/T20I | Twenty20 | Women's T20Is | - other - ⌄ |

Batting | Bowling | Fielding | All-round | Partnership | Team | Umpire and referee | Aggregate/overall

View overall figures [change view]
Ordered by runs scored (descending)

Page 1 of 62   Showing 1 - 50 of 3076   ⬛ First ◀ Previous   Next ▶   Last ⬛   ◀◀ Return to query menu
                                                                                    ◀◀ Cleared query menu

Overall figures

Player	Span	Mat	Inns	NO	Runs▼	HS	Ave	BF	SR	100	50	0	
SR Tendulkar (IND)	1989-2012	463	452	41	**18426**	200*	44.83	21368	86.23	49	96	20	▶
KC Sangakkara (Asia/ICC/SL)	2000-2015	404	380	41	**14234**	169	41.98	18048	78.86	25	93	15	▶
V Kohli (IND)	2008-2025	302	290	45	**14181**	183	57.88	15192	93.34	51	74	16	▶
RT Ponting (AUS/ICC)	1995-2012	375	365	39	**13704**	164	42.03	17046	80.39	30	82	20	▶
ST Jayasuriya (Asia/SL)	1989-2011	445	433	18	**13430**	189	32.36	14725	91.20	28	68	34	▶
DPMD Jayawardene (Asia/SL)	1998-2015	448	418	39	**12650**	144	33.37	16020	78.96	19	77	28	▶
Inzamam-ul-Haq (Asia/PAK)	1991-2007	378	350	53	**11739**	137*	39.52	15812	74.24	10	83	20	▶
JH Kallis (Afr/ICC/SA)	1996-2014	328	314	53	**11579**	139	44.36	15885	72.89	17	86	17	▶
SC Ganguly (Asia/IND)	1992-2007	311	300	23	**11363**	183	41.02	15416	73.70	22	72	16	▶
RG Sharma (IND)	2007-2025	273	265	36	**11168**	264	48.76	12034	92.80	32	58	16	▶
R Dravid (Asia/ICC/IND)	1996-2011	344	318	40	**10889**	153	39.16	15285	71.23	12	83	13	▶
MS Dhoni (Asia/IND)	2004-2019	350	297	84	**10773**	183*	50.57	12303	87.56	10	73	10	▶
CH Gayle (ICC/WI)	1999-2019	301	294	17	**10480**	215	37.83	12019	87.19	25	54	25	▶
BC Lara (ICC/WI)	1990-2007	299	289	32	**10405**	169	40.48	13086	79.51	19	63	16	▶
TM Dilshan (SL)	1999-2016	330	303	41	**10290**	161*	39.27	11933	86.23	22	47	11	▶

Extracting the content form table

Name · **Country** · **Start/End = Year** · **Matches** · **Runs** · **Highest Score** · **Average** · **Century**

Overall figures

Player	Span	Mat	Inns	NO	Runs	HS	Ave	BF	SR	100	50
SR Tendulkar (IND)	1989-2012	463	452	41	**18426**	200*	44.83	21368	86.23	49	96
KC Sangakkara (Asia/ICC/SL)	2000-2015	404	380	41	**14234**	169	41.98	18048	78.86	25	93
V Kohli (IND)	2008-2025	302	290	45	**14181**	183	57.88	15192	93.34	51	74
RT Ponting (AUS/ICC)	1995-2012	375	365	39	**13704**	164	42.03	17046	80.39	30	82
ST Jayasuriya (Asia/SL)	1989-2011	445	433	18	**13430**	189	32.36	14725	91.20	28	68
DPMD Jayawardene (Asia/SL)	1998-2015	448	418	39	**12650**	144	33.37	16020	78.96	19	77

Innings · **Not Ball** · **Ball Faced** · **Strike Rate** · **Fifty**

Old Data Frame

	Player	Span	Mat	Inns	NO	Runs	HS	Ave	BF	SR	100	50
0	SR Tendulkar (IND)	1989-2012	463	452	41	18426	200*	44.83	21368	86.23	49	96
1	KC Sangakkara (Asia/ICC/SL)	2000-2015	404	380	41	14234	169	41.98	18048	78.86	25	93
2	V Kohli (IND)	2008-2025	302	290	45	14181	183	57.88	15192	93.34	51	74
3	RT Ponting (AUS/ICC)	1995-2012	375	365	39	13704	164	42.03	17046	80.39	30	82
4	ST Jayasuriya (Asia/SL)	1989-2011	445	433	18	13430	189	32.36	14725	91.20	28	68

New Data Frame

	Player_Names	Country	Start_Year	End_Year	Matches	Innings	Not_Out	Total_Runs	Highest_Score	Batting_Average	Ball_Faced	Strike_Rate	100's	50's
0	SR Tendulkar	India	1989	2012	463	452	41	18426	200	44.83	21368	86.23	49	96
1	KC Sangakkara	Sri Lanka	2000	2015	404	380	41	14234	169	41.98	18048	78.86	25	93
2	V Kohli	India	2008	2025	302	290	45	14181	183	57.88	15192	93.34	51	74
3	RT Ponting	Australia	1995	2012	375	365	39	13704	164	42.03	17046	80.39	30	82
4	ST Jayasuriya	Sri Lanka	1989	2011	445	433	18	13430	189	32.36	14725	91.20	28	68
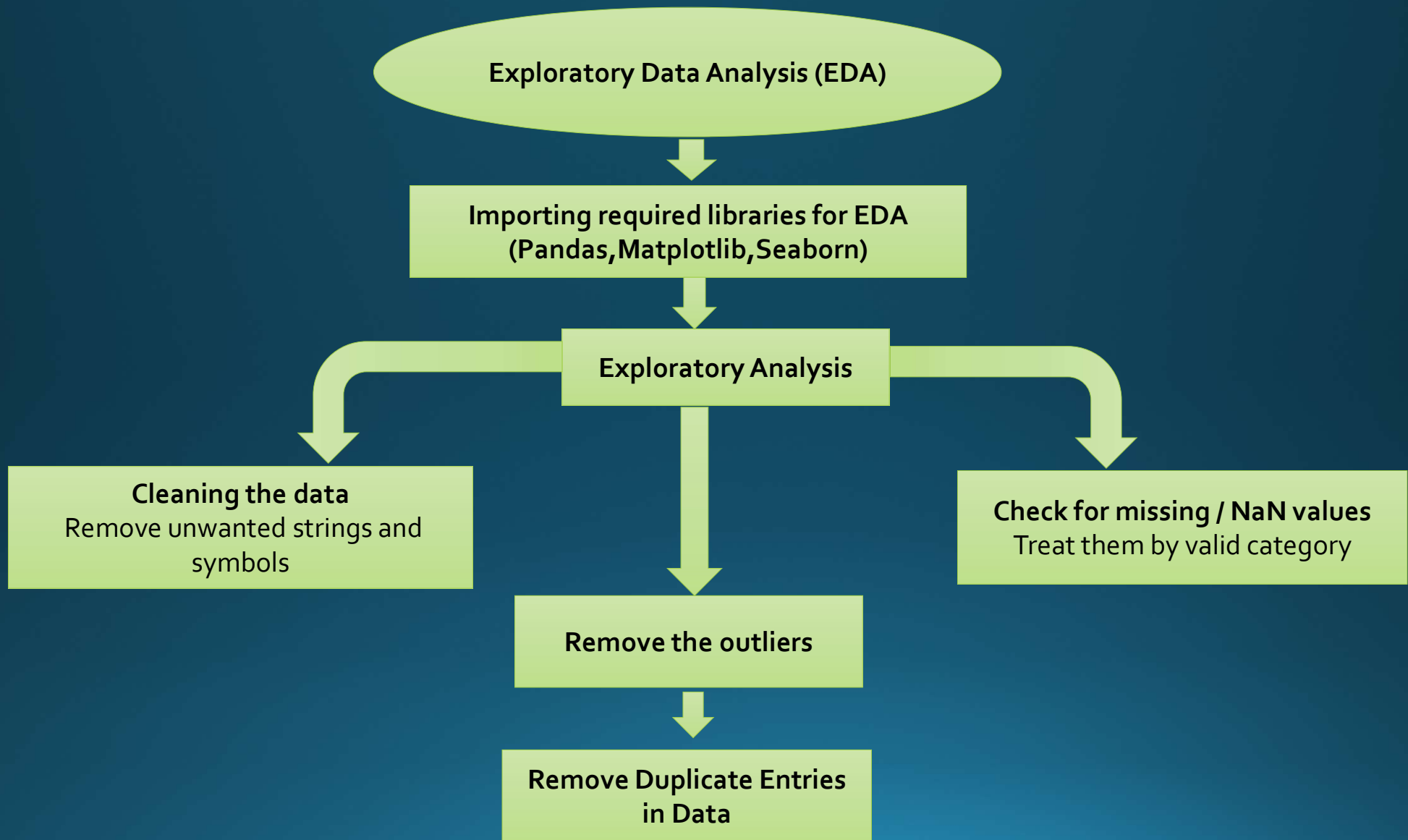
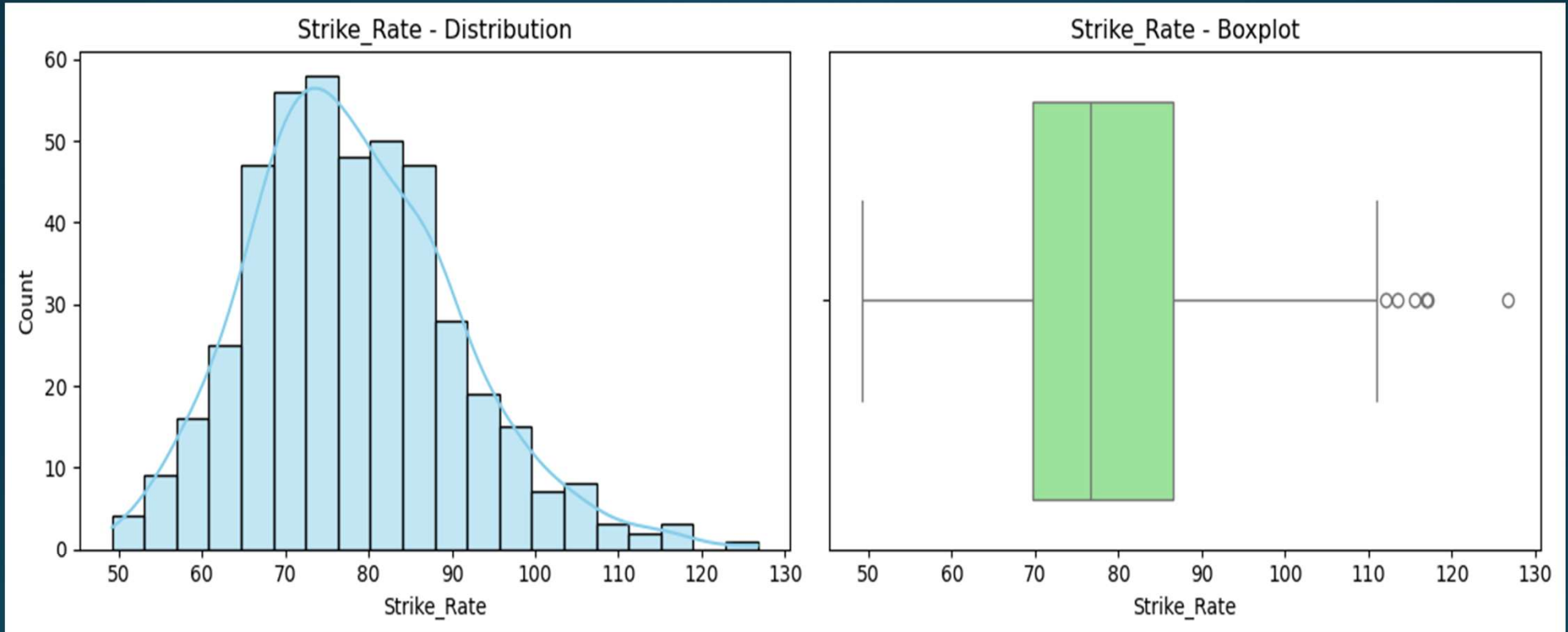# Univariate Numerical Analysis

Matches Distribution of Players.
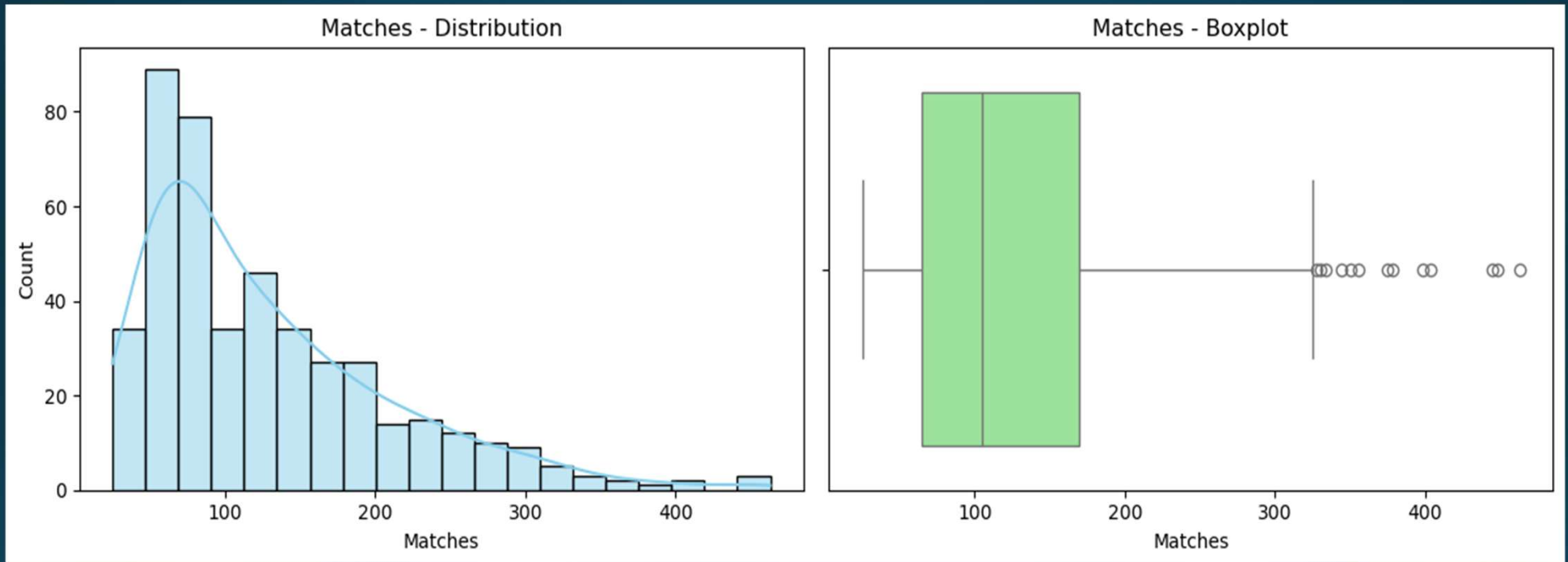
- ➤ Most players play 50-170 matches in their careers.
- ➤ The Data is right-skewed-many play fewer matches, only a few play very high matches.
- ➤ The median career length is about 100 matches.
- ➤ A small number of players are outliers with 300+ matches(Very long careers).
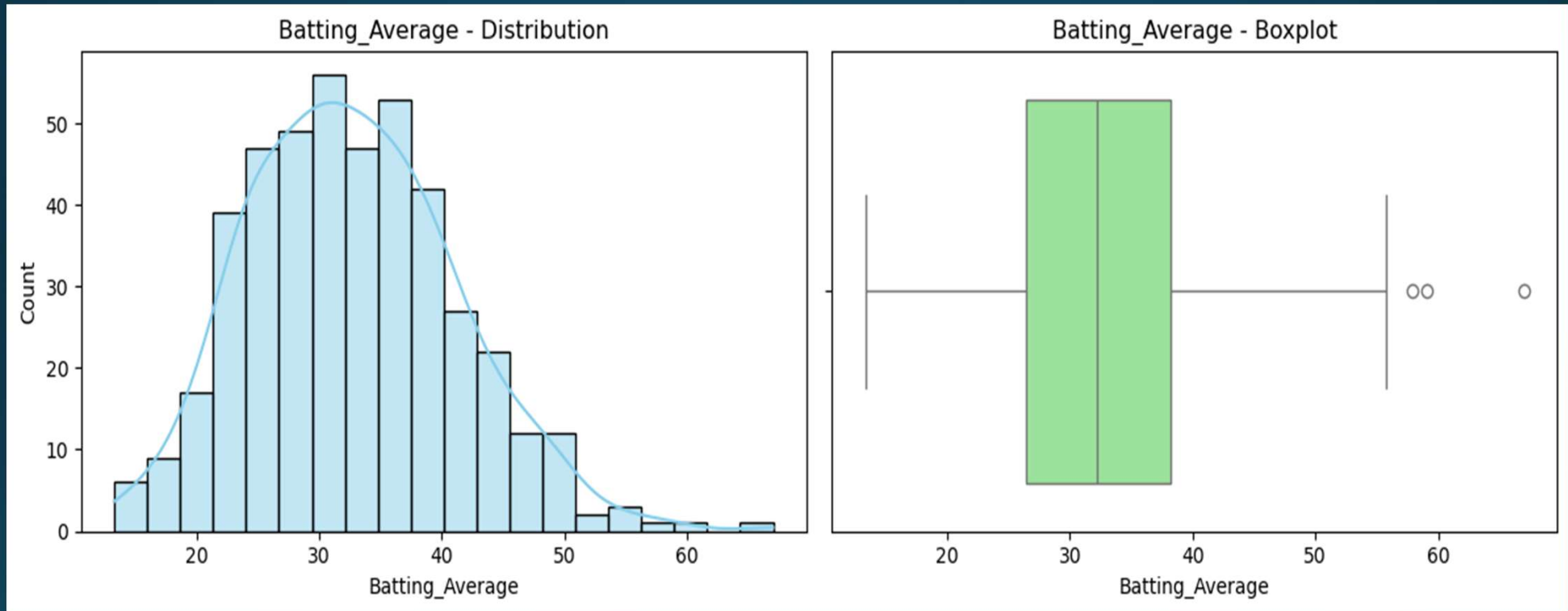
Univariate Numerical Analysis

Strike Rate Distribution of Matches.

> ➤ **Most Players have strike rates around 70-90, while only a few exceptional players have very high strike rates(110+)**

Univariate Numerical Analysis

Batting Average Distribution of Matches.

- ➤ Most players have a batting average between 25-40.
- ➤ The median batting average is around 30-32.
- ➤ A few outliers (60+) indicate exceptionally consistent player with very high averages.

Univariate Categorical Analysis

Top 5 Countries in ODI Batting.
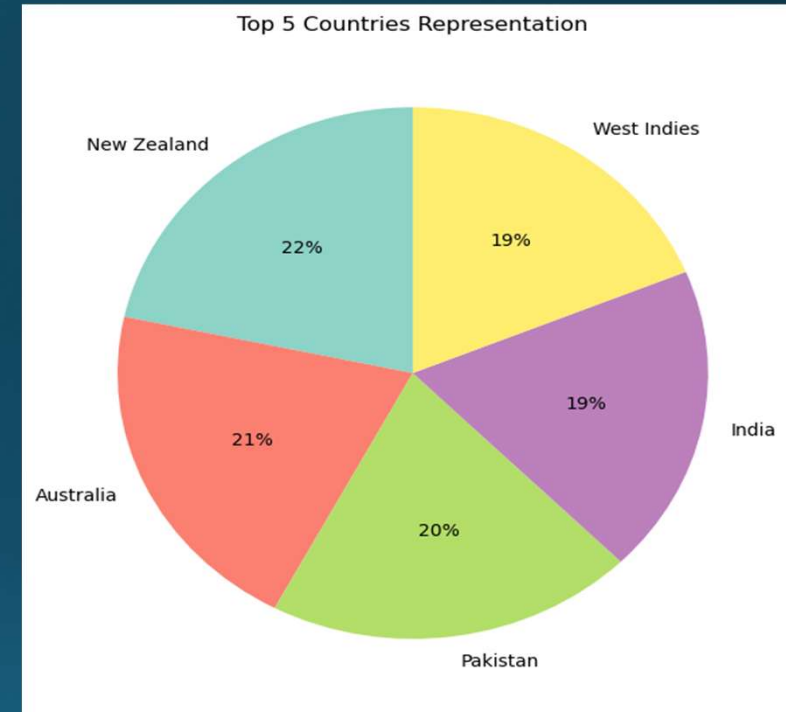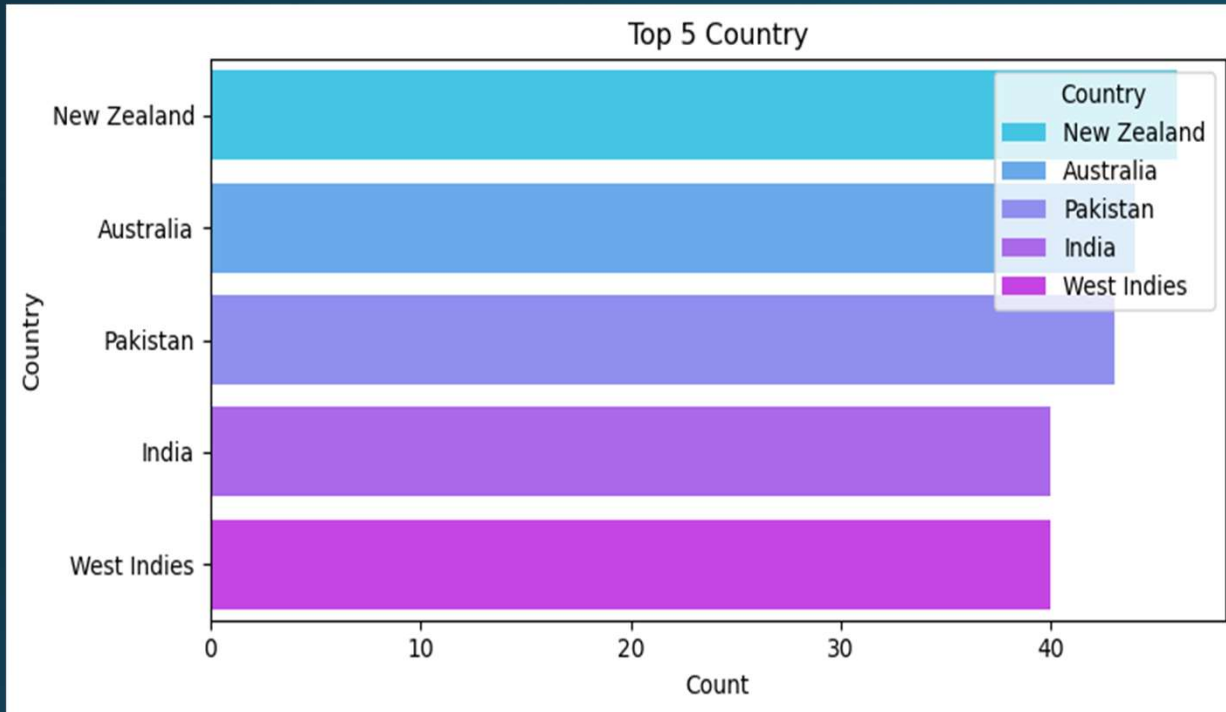
- ➤ New Zealand has the highest representation among the top 5 countries, showing strong player participation or dominance in this dataset.
- ➤ Australia and Pakistan are close contenders, with nearly equal counts, indicating balanced performance or presence.
- ➤ India and West Indies have slightly lower counts, but they are still significant enough to make the top 5, highlighting their consistent contributions
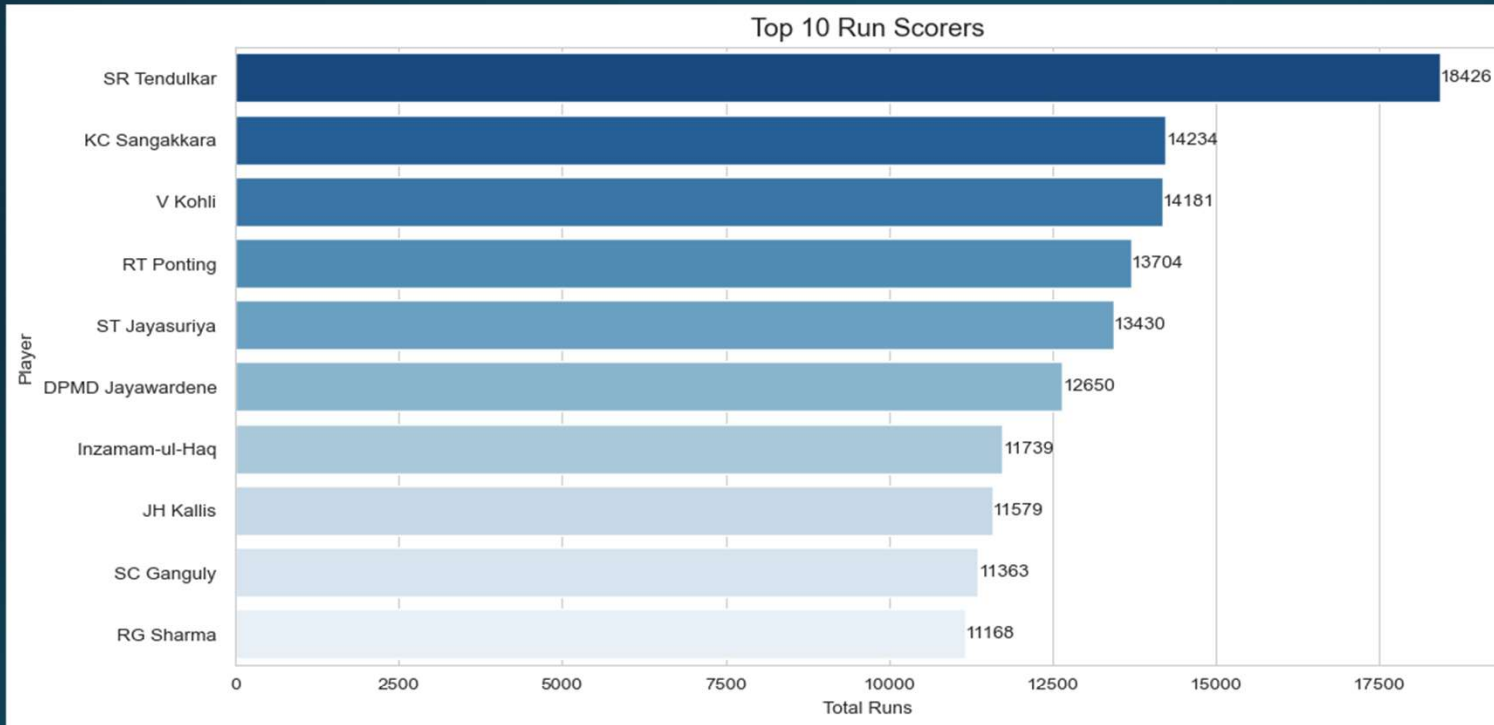
Bivariate Analysis

Number of Players per Country.
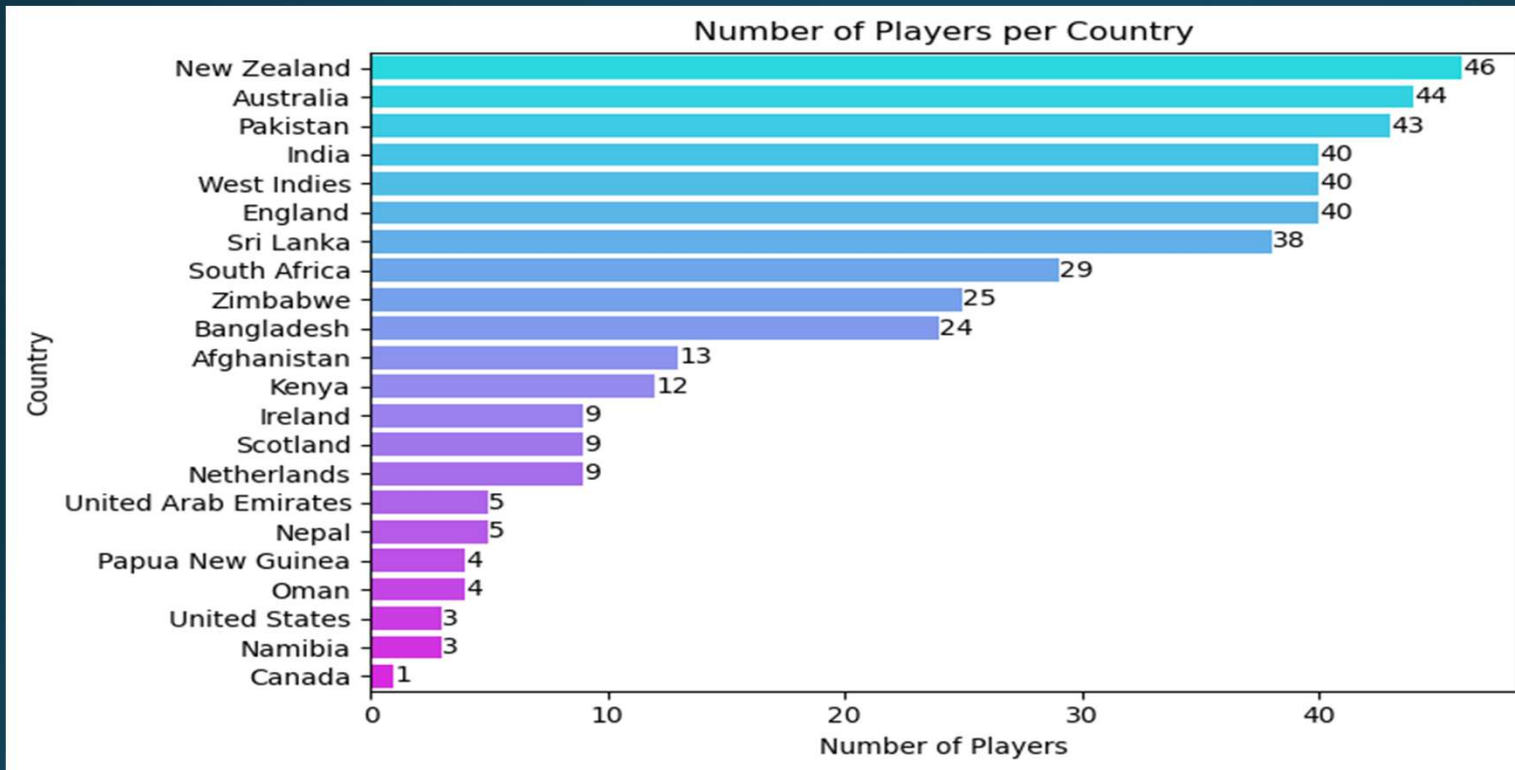
- ➤ Top cricketing nations: New Zealand, Australia, Pakistan, India, West Indies, and England each have 40+ players.
- ➤ Mid-tier nations: Sri Lanka and South Africa have moderate representation.
- ➤ Lower-tier nations: Zimbabwe, Bangladesh, Afghanistan, and others contribute fewer players.

Bivariate Analysis

Top 10 Run Scores.

> ➤ Sachin Tendulkar dominates with over 18,000 runs, far ahead of all other players, showing his unmatched consistency.
> ➤ Kumar Sangakkara and Virat Kohli follow closely behind, both with over 14,000 runs, proving their exceptional performance.
> ➤ The rest of the players (Ponting, Jayasuriya, Jayawardene, etc.) have impressive tallies between 11,000–13,000 runs, but there's a noticeable gap compared to the top scorer.
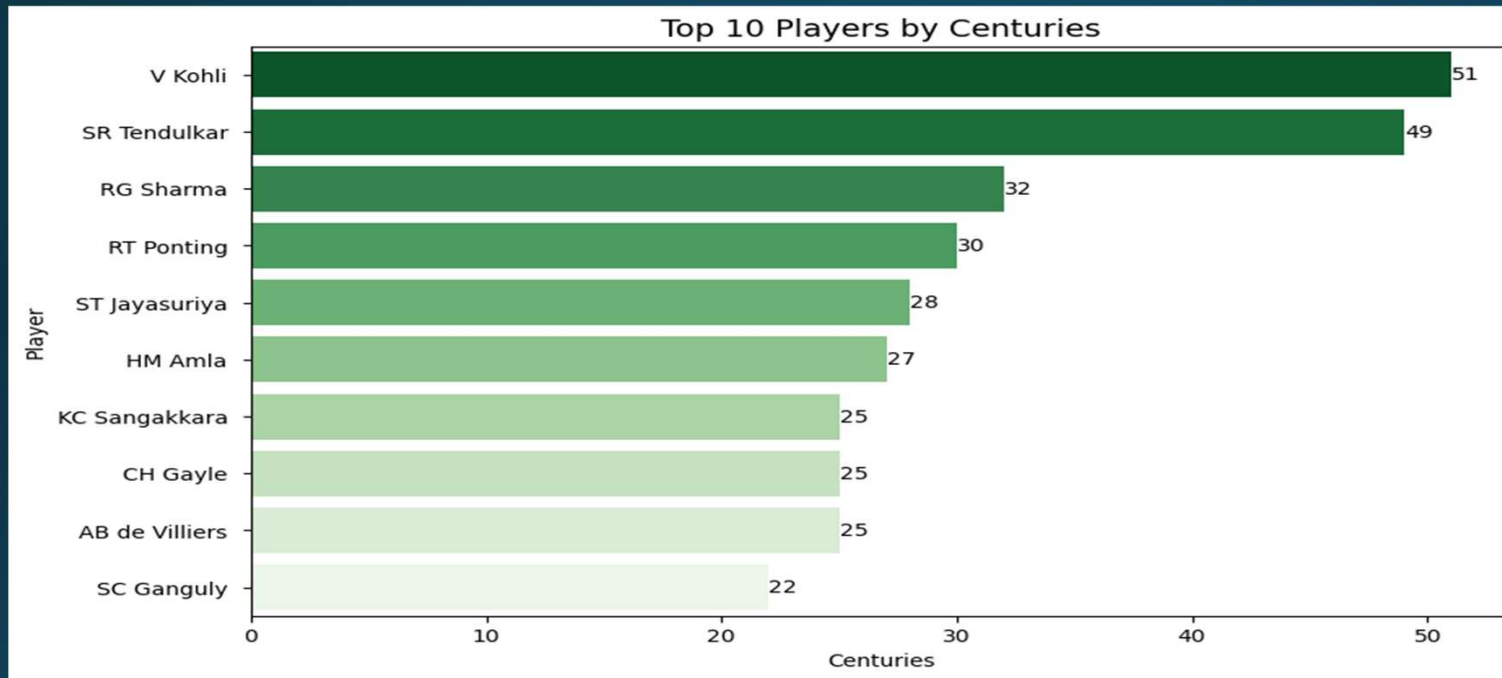
Bivariate Analysis

Total Runs by Countries.

Total Runs by Country (Sum)
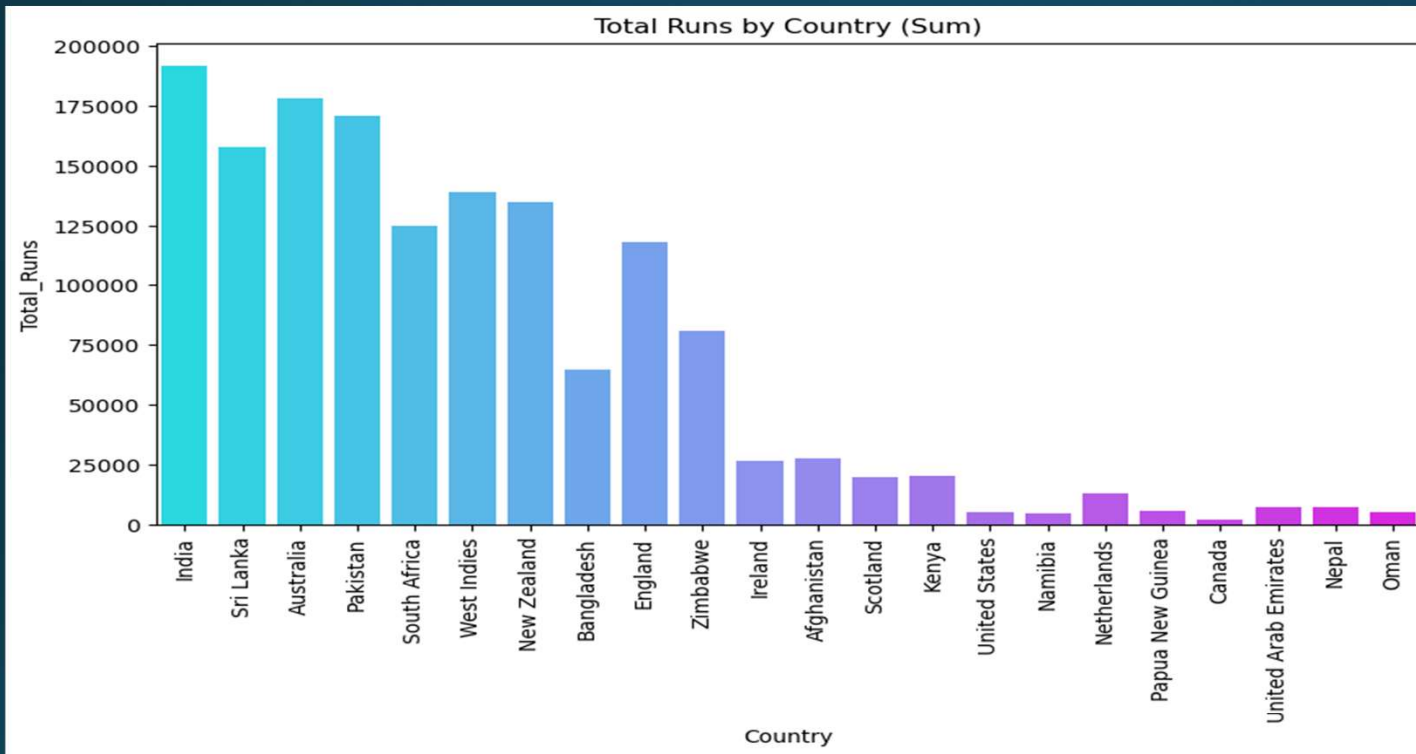
- ➤ India dominates with the highest total runs, showing strong batting depth over time.
- ➤ Sri Lanka, Australia, and Pakistan also have very high totals, forming the next strongest group.
- ➤ Smaller cricket nations like Nepal, Oman, and UAE have much lower totals, highlighting fewer matches or less experience.

Bivariate Analysis

Top 10 Players by Centuries.

Top 10 Players by Centuries

Player	Centuries
V Kohli	51
SR Tendulkar	49
RG Sharma	32
RT Ponting	30
ST Jayasuriya	28
HM Amla	27
KC Sangakkara	25
CH Gayle	25
AB de Villiers	25
SC Ganguly	22

➢ **Top Performers:** Virat Kohli and Sachin Tendulkar are clear leaders with 51 and 49 centuries, far ahead of others.

➢ **Mid-tier Group:** Players like Rohit Sharma, Ponting, and Jayasuriya are in the 30–28 century range, showing consistent performance.

➢ **Close Competition:** Several legends like Sangakkara, Gayle, and AB de Villiers are tied at 25 centuries, highlighting a competitive group.

Multivariate Analysis

Matches vs Runs Distribution :

Bubble Plot: Matches vs Runs (Color=Batting Avg, Size=Strike Rate)
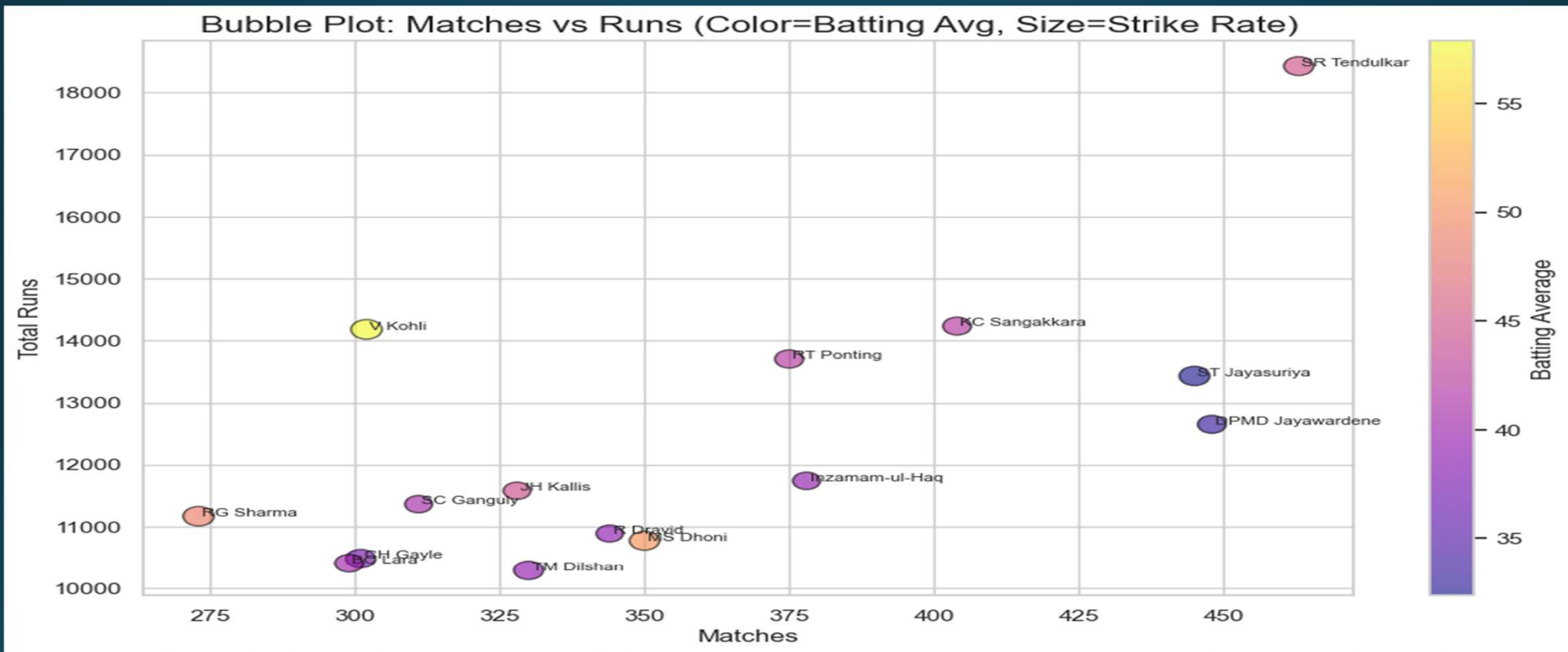
- ➤ **More matches = more runs** – Players who played more matches generally scored more runs.
- ➤ **Kohli has best average** – Even with fewer matches, Virat Kohli has the highest batting average.
- ➤ **Big hitters have high strike rates** – Players like Criss Gayle and Rohit Sharma score faster.

🔑 **Key Insight**

➤ **Top Run-Scorer : Sachine Tendulkar(India) -> 18426, 49 centuries.**

➤ **Highest Strike Rate : Jos Buttler(England) -> SR~118, shows modern aggressive style.**

➤ **Most Centuries in ODIs : Virat Kohli (India) -> 51 centuries.**

➤ **Matches vs Runs : Kumar Sangakkara (Sri Lanka) -> Long career 400+ matches, 14000+ runs.**

➤ **Country Dominance : India, Australia, Sri Lanka consistency produce top ODI batters.**

🔑 Key Insights – Top 5 ODI Batting Legends

Player	Country	Runs	Avg	Strike Rate	100's	50's
Sachin Tendulkar	India	18426	44.8	86.2	49	96
Kumar Sangakkara	Sri Lanka	14234	41.9	78.9	25	93
Ricky Ponting	Australia	13704	42.0	80.4	30	82
Jayasuriya	Sri Lanka	13430	32.4	91.2	28	68
Virat Kohli	India	13848	58.7	93.6	50	70

Conclusion ✅

➢ **I collect and cleaned ODI batting data using web scraping.**

➢ **EDA help me find important patterns in player performance.**

➢ **A few top players score most of the runs and centuries, while many stay in the mid-range.**

➢ **Strike rates are higher in the modern era, showing a more aggressive style.**

➢ **More matches usually mean more runs, but batting averages differ widely.**

➢ **This project shows how data helps compare players, find trends, and support cricket strategies.**

Thank You !