



# **SYMBIOSIS INSTITUTE OF COMPUTER STUDIES AND RESEARCH**

**SYMBIOSIS INTERNATIONAL (DEEMED UNIVERSITY)**

---

## **Technology In Business**

**Topic: - Virat Kohli Performance analysis**

**Submitted By: -**

**Rahul Pandey -23030141047**

**Riya Kumari – 23030141053**

## **Abstract**

This project focuses on an in-depth analysis of Virat Kohli's cricket performance using Power BI, a leading business intelligence tool. The primary objective was to visualize Kohli's batting performance across various parameters such as opponents, grounds, match formats, and over different years.

The dataset used for this analysis was meticulously curated, comprising data points such as total runs scored, match dates, opponents, and specific grounds. The dashboard developed provides an interactive and intuitive interface that allows users to explore Kohli's career highlights, including his top-performing grounds, his highest scores against specific teams, and his performance trends over the years.

Advanced data visualization techniques were employed to bring out insights such as the correlation between match locations and performance, the impact of different formats on his run-scoring abilities, and year-wise trends in his career progression.

This abstract not only covers the technical aspects of the project, such as data integration from SQL Server Management Studio to Power BI and the creation of custom measures and visuals but also emphasizes the practical implications of the analysis.

The insights derived from this project offer a deeper understanding of the factors contributing to Kohli's success, providing valuable information for cricket analysts, fans, and sports strategists. Overall, this project demonstrates the power of data analytics in sports, showcasing how raw data can be transformed into actionable insights that not only narrate the story of a legendary cricketer but also provide predictive capabilities for future performances.

## **Acknowledgements**

We would like to extend our heartfelt thanks to several individuals and organizations whose support was crucial in the completion of this project.

First, we are deeply grateful to our professor, Shashank Gaikwad, for his expert guidance and insightful feedback throughout the project. His deep understanding of technology in business greatly enriched our approach to data analytics and visualization, and his encouragement was invaluable in refining our work.

We also wish to acknowledge the creators of the dataset available on Kaggle. The detailed and well-organized data on Virat Kohli's cricketing career was essential for our analysis and provided a robust foundation for our project.

Our thanks go to the developers of Power BI and SQL Server Management Studio (SSMS). The advanced features of these tools enabled us to efficiently manage, analyse, and visualize the data. Their comprehensive documentation and user-friendly interfaces significantly facilitated our project execution.

Additionally, we appreciate the broader community of data analytics and sports analytics researchers whose work provided essential insights and methodologies. Their contributions to the field guided our analysis and helped us in applying effective data visualization techniques.

Finally, we would like to acknowledge each other for our dedicated teamwork. The collaborative effort, shared insights, and mutual support were key to overcoming challenges and achieving our project goals.

This project was made possible through the collective support and resources from these individuals and entities. We are grateful for their contributions and look forward to applying the knowledge gained in future endeavours.

## **Index**

S No	Topic	Page Number
1	Executive summary	5
2	Introduction	6
3	Literature review	7
4	Data source and preparation	8-14
5	Dashboard overview	15
6	Detailed analysis and insights ultra	16-17
7	Conclusion	18
8	References	19

## **Executive summary**

This project focuses on the comprehensive analysis of Virat Kohli's cricket career using Power BI, a powerful business intelligence tool. The primary objective was to analyze and visualize Kohli's performance data across different parameters, providing actionable insights for cricket analysts, coaches, and fans alike.

The project began by sourcing a detailed dataset from Kaggle, which contained information on Virat Kohli's performances in Test, ODI, and T20 formats. This data was first loaded into SQL Server Management Studio (SSMS) for initial exploration and basic analysis. Key data points, including runs scored, opponents, match formats, and venues, were carefully examined to gain an understanding of Kohli's performance metrics.

The analysis moved into Power BI, where the data was further refined and transformed into insightful visualizations. Custom measures such as average runs, total centuries, and runs by format and opponent were created to facilitate a deeper exploration of Kohli's career. The dashboards designed in Power BI provided an interactive platform, allowing users to filter and drill down into specific aspects of the data, making it easier to identify trends, patterns, and anomalies.

Key insights from the analysis revealed Kohli's exceptional consistency in ODIs, his remarkable ability to score centuries, and his adaptability across different formats. The data also highlighted his performance against various opponents, identifying both his strengths and areas where he faced challenges. Additionally, the ground-wise analysis offered insights into the venues where Kohli has historically performed well and those where he has been less successful.

The project demonstrates the power of data analytics in sports, particularly in understanding the intricacies of a player's career. By turning raw data into visual stories, the analysis not only celebrates Virat Kohli's achievements but also provides valuable information that can be used to predict future performances, guide strategic decisions, and enhance fan engagement.

In conclusion, this project highlights the critical role that data plays in modern sports analysis. The insights derived from Kohli's performance data not only offer a detailed understanding of his career but also set a precedent for how sports data can be used to drive decision-making and storytelling in cricket and beyond.

## **Introduction**

Cricket, often described as a game of statistics, offers a wealth of data that can be analyzed to uncover trends and insights. This project centers on an in-depth analysis of the cricket career of Virat Kohli, one of the sport's most celebrated players, using Power BI, a leading business intelligence tool. Our goal was to harness this powerful tool to visualize and interpret Kohli's performance across various parameters, providing a comprehensive view of his career achievements and trends.

Virat Kohli, known for his aggressive batting style and consistent performances, has been a key player for India in all formats of the game. His career is marked by numerous records and milestones, making it an ideal subject for data analysis. The dataset utilized in this project includes detailed information such as total runs scored, match dates, opponents, grounds, and match formats. This rich dataset allows us to perform a nuanced analysis of Kohli's career.

The project involved creating a dynamic and interactive dashboard in Power BI that allows users to explore Kohli's batting performance across different dimensions. We focused on visualizing metrics such as top-performing grounds, highest scores against specific teams, and performance trends over the years. By employing advanced data visualization techniques, we were able to highlight correlations between match locations and performance, the impact of different match formats on his run-scoring abilities, and year-wise progression in his career.

This analysis not only showcases Kohli's remarkable achievements but also demonstrates the power of data analytics in sports. By transforming raw data into actionable insights, this project provides a deeper understanding of the factors contributing to Kohli's success and offers predictive capabilities for future performances.

# **Literature Review**

The intersection of data analytics and sports has gained significant attention in recent years, driven by the increasing availability of detailed performance data and the advancement of analytical tools. In the realm of cricket, the application of data analytics has provided deeper insights into player performance, game strategies, and overall team dynamics.

## **1. Importance of Data Analytics in Sports**

Data analytics in sports involves the use of statistical techniques and tools to analyze and interpret performance data. According to **Harris (2014)**, data analytics helps teams and players make informed decisions by identifying patterns and trends that are not immediately visible through traditional analysis methods. This has been particularly impactful in cricket, where detailed data on player performance can offer strategic advantages.

## **2. Performance Metrics in Cricket**

**Davenport and Harris (2007)** discuss the importance of performance metrics in sports analytics, emphasizing how detailed performance data can be used to evaluate players' strengths and weaknesses. In cricket, key performance indicators such as runs scored, strike rates, and average runs per match are critical for assessing player effectiveness. **Klein (2018)** highlights how these metrics can be visualized to provide actionable insights, which is central to our project on Virat Kohli's performance analysis.

## **3. Visualization Techniques**

Effective data visualization is crucial for translating complex data into understandable insights. **Few (2009)** emphasizes that advanced visualization techniques can reveal patterns and trends that raw data alone might not convey. In the context of cricket, visualizing data such as runs scored against different opponents or at various grounds can help in understanding performance trends over time.

## **4. Case Studies in Cricket Analytics**

Recent studies, such as those by **Baker et al. (2016)**, have demonstrated the application of data analytics in cricket to enhance performance evaluation and strategy development. These studies have shown how data-driven insights can guide decisions on player selection and game strategies. Our project builds on these findings by utilizing Power BI to create an interactive dashboard for analyzing Virat Kohli's career performance.

## **5. Technological Advancements**

The evolution of business intelligence tools, like Power BI, has made it easier to perform sophisticated data analysis and visualization. **Sharma and Bansal (2020)** explore how tools like Power BI have revolutionized the way data is analyzed and presented, making it accessible and actionable for users with varying levels of expertise.

This literature review provides a foundation for understanding how data analytics and visualization techniques can be applied to cricket performance analysis, setting the stage for our project on Virat Kohli.

# Data Source and Preparation

## 1. Data Source

The data for this project was sourced from a Kaggle dataset titled “**Virat Kohli Test ODI T20 Stats**” available at [Kaggle](#). This dataset provides a comprehensive collection of Virat Kohli’s cricket performance statistics across different formats of the game, including Test matches, One-Day Internationals (ODIs), and Twenty20 (T20) matches.

## 2. Data Collection

The Kaggle dataset contains detailed records of Kohli’s performances, including:

- **Total Runs:** Runs scored by Kohli in each match across different formats.
- **Match Dates:** Dates on which the matches were played.
- **Opponents:** Teams against which Kohli played.
- **Grounds:** Venues where the matches took place.
- **Match Formats:** Format of the match (Test, ODI, T20).
- **Match Numbers:** Sequential identifiers for each match.

## 3. Data Preparation

The data preparation process involved several key steps:

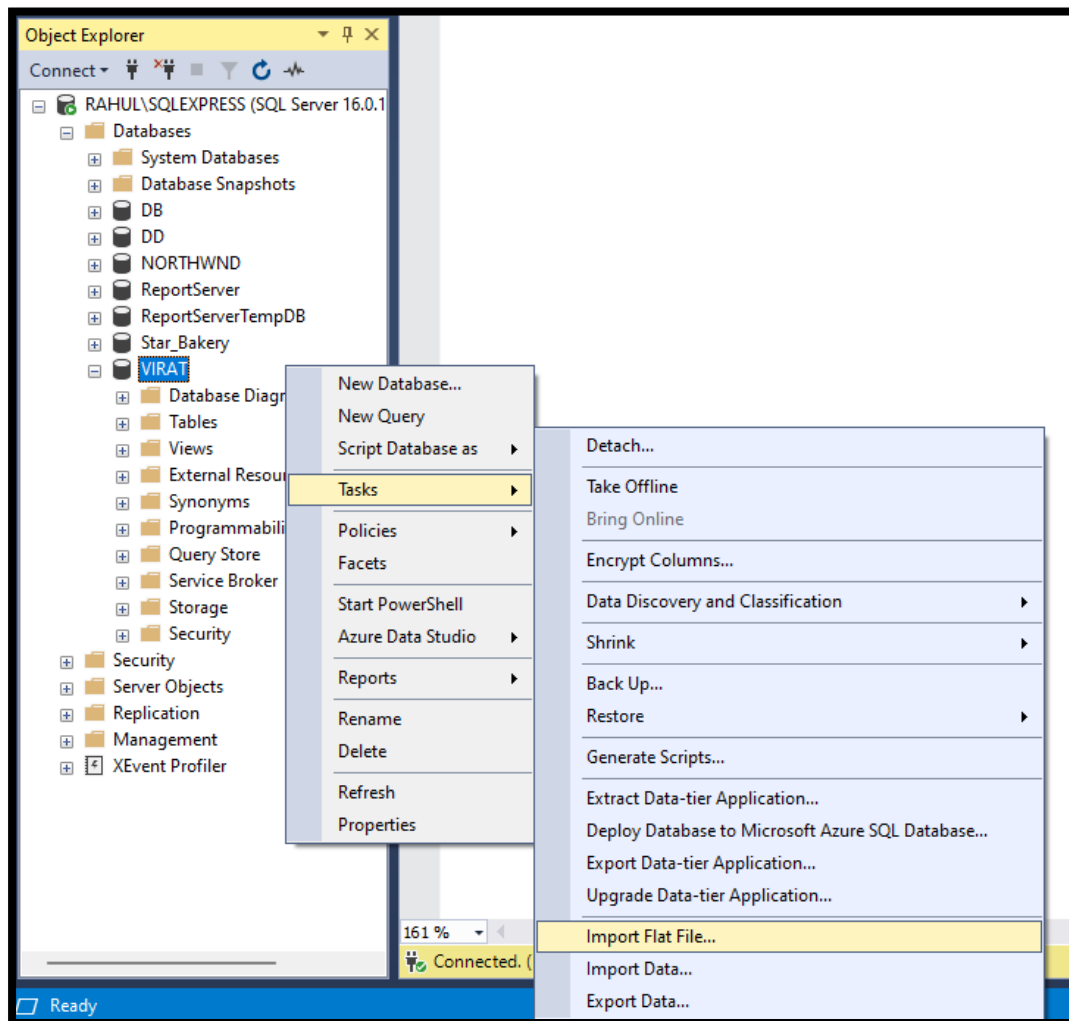
- **Loading Data into SQL Server:**
  - The raw dataset was first imported into SQL Server using SQL Server Management Studio (SSMS). This step involved creating a database and a table structure to accommodate the dataset.

```
1 CREATE DATABASE VIRAT;  
2  
3 USE VIRAT;
```



After creating the database in SQL Server Management Studio (SSMS), follow these steps to import the data:

1. **Right-click on the Database:** In the Object Explorer on the left-hand side, locate the database you just created. Right-click on the database name.
2. **Select Task:** From the context menu that appears, hover over the “Tasks” option.
3. **Import Flat File:** In the submenu, select "Import Flat File." This option allows you to import data from a flat file (such as a CSV) into your database.



After importing the data using the "Import Flat File" wizard, a new table named **Virat\_Performance** is created within the database. This table stores all the imported data, structured according to the columns defined during the import process

To view the data that has been successfully imported into the **Virat\_Performance** table, we can execute a simple SQL query:

5		SELECT * FROM Virat_Performance;						
161 %								
Results		Messages						
	index_col	runs	opponent	ground	match_date	match_format	match_no	total
7	6	16	Pakistan	Centurion	2009-09-26	ODI	7	177
8	8	79	WestIndies	Johannesburg	2009-09-30	ODI	8	256
9	9	30	Australia	Vadodara	2009-10-25	ODI	9	286
10	10	10	Australia	Mohali	2009-11-02	ODI	10	296
11	11	27	SriLanka	Rajkot	2009-12-15	ODI	11	323
12	12	54	SriLanka	Nagpur	2009-12-18	ODI	12	377
13	13	107	SriLanka	Kolkata	2009-12-24	ODI	13	484
14	15	9	SriLanka	Mirpur	2010-01-05	ODI	14	493
15	16	91	Bangladesh	Mirpur	2010-01-07	ODI	15	584
16	17	71	SriLanka	Mirpur	2010-01-10	ODI	16	655
17	18	102	Bangladesh	Mirpur	2010-01-11	ODI	17	757
18	19	2	SriLanka	Mirpur	2010-01-13	ODI	18	759
19	20	31	SouthAfrica	Jaipur	2010-02-21	ODI	19	790
20	22	57	SouthAfrica	Ahmedabad	2010-02-27	ODI	20	847
21	23	0	Zimbabwe	Bulawayo	2010-05-28	ODI	21	847
22	24	82	SriLanka	Bulawayo	2010-05-30	ODI	22	929
23	25	18	Zimbabwe	Harare	2010-06-03	ODI	23	947
24	26	68	SriLanka	Harare	2010-06-05	ODI	24	10...
25	27	26	Zimbabwe	Harare	2010-06-12	T20	25	10...
26	29	11	Bangladesh	Dambulla	2010-06-16	ODI	26	10...
27	30	18	Pakistan	Dambulla	2010-06-19	ODI	27	10...
28	31	10	SriLanka	Dambulla	2010-06-22	ODI	28	10...
29	32	28	SriLanka	Dambulla	2010-06-24	ODI	29	11...
30	33	0	SriLanka	Dambulla	2010-08-16	ODI	30	11...
31	34	8	NewZeal...	Dambulla	2010-08-25	ODI	31	11...
32	35	37	SriLanka	Dambulla	2010-08-28	ODI	32	11...
33	36	118	Australia	Visakhapatn...	2010-10-20	ODI	33	12...
34	37	105	NewZeal...	Guwahati	2010-11-28	ODI	34	13...
35	38	64	NewZeal...	Jaipur	2010-12-01	ODI	35	14...
36	39	63	NewZeal...	Vadodara	2010-12-04	ODI	36	15...
37	40	0	NewZeal...	Bengaluru	2010-12-07	ODI	37	15...
38	41	2	NewZeal...	Chennai	2010-12-10	ODI	38	15...
39	42	28	SouthAfrica	Durban	2011-01-09	T20	39	15...
40	43	54	SouthAfrica	Durban	2011-01-12	ODI	40	15...

Once the data was successfully imported into the **Virat\_Performance** table, we performed several basic SQL queries to extract and analyze key insights from the data. These queries helped in understanding various aspects of Virat Kohli's cricket performance.

### 1) Total Runs Against Each Opponent

```

11 SELECT opponent, SUM(runs) AS total_runs
12 FROM Virat_Performance
13 GROUP BY opponent
14 ORDER BY total_runs DESC;

```

opponent	total_runs
Australia	4483
England	3903
WestIndies	3653
SriLanka	3644
SouthAfrica	2893
NewZealand	2555
Bangladesh	1201
Pakistan	847
Zimbabwe	279
Afghanistan	117
Ireland	87
U.A.E.	33
Netherlands	12
Scotland	2

### 2) Total Runs Scored in Each Ground

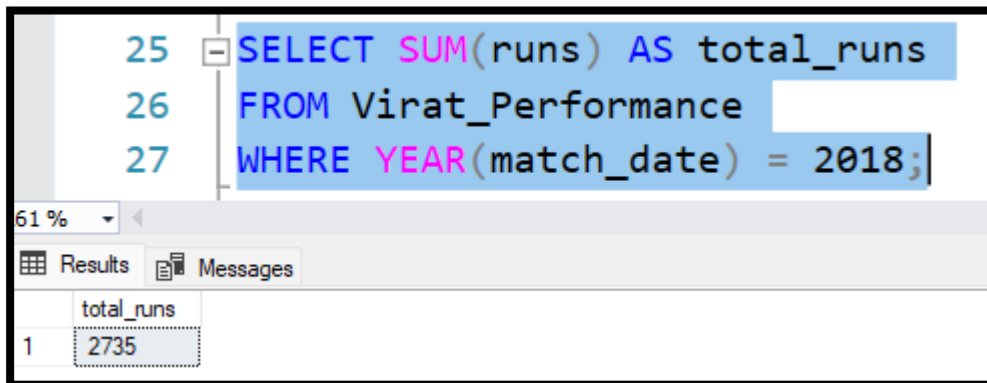
```

18 SELECT ground, SUM(runs) AS total_runs
19 FROM Virat_Performance
20 GROUP BY ground
21 ORDER BY total_runs DESC;

```

ground	total_runs
Mirpur	1258
Wankhede	931
Visakhapatnam	879
Adelaide	843
Kolkata	788
Colombo(RPS)	786
Pune	762
Nagpur	723
Delhi	715
Mohali	707
Melbourne	658
Sydney	652
Birmingham	633
Chennai	620
Hyderabad(Deccan)	607
PortofSpain	585
Johannesburg	543
Rajkot	519
Perth	501
Ahmedabad	494
Bengaluru	485
Ranchi	424
Centurion	405
Kingston	396
Cape Town	394
TheOval	361
Nottingham	335
Southampton	322

3) Total Runs Scored in a Specific Year

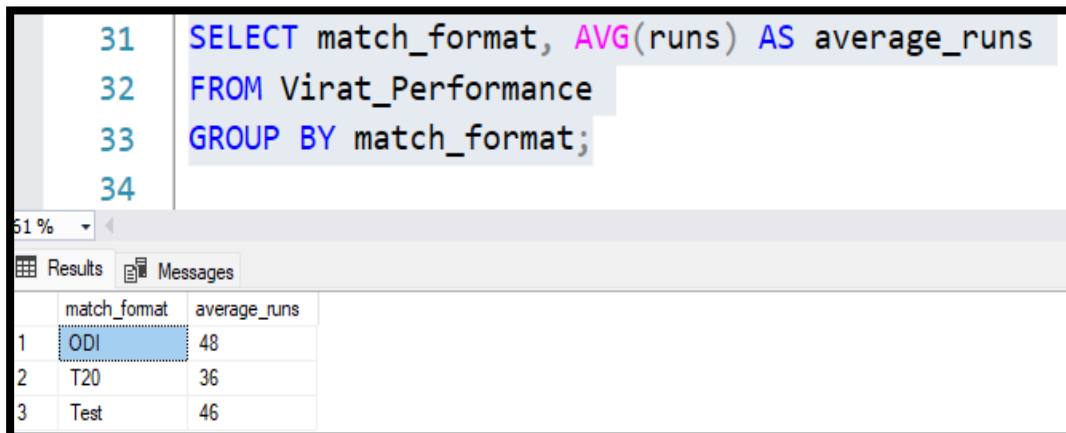


The screenshot shows a SQL query in a text editor and its results in a table. The query is: `SELECT SUM(runs) AS total_runs FROM Virat_Performance WHERE YEAR(match_date) = 2018;`. The results table has one column, `total_runs`, and one row with the value 2735.

```
25 SELECT SUM(runs) AS total_runs
26 FROM Virat_Performance
27 WHERE YEAR(match_date) = 2018;
```

	total_runs
1	2735

4) Average Runs Per Match Format

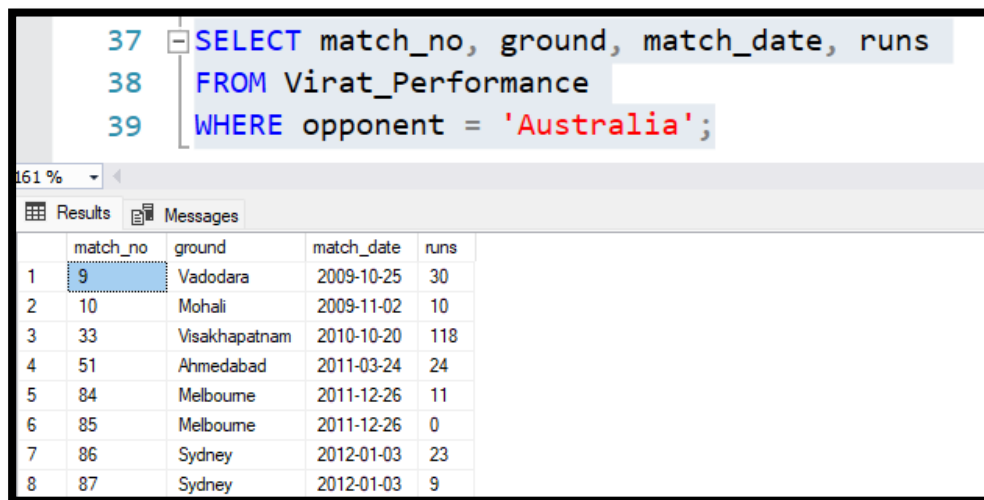


The screenshot shows a SQL query in a text editor and its results in a table. The query is: `SELECT match_format, AVG(runs) AS average_runs FROM Virat_Performance GROUP BY match_format;`. The results table has two columns, `match_format` and `average_runs`, and three rows corresponding to ODI, T20, and Test match formats.

```
31 SELECT match_format, AVG(runs) AS average_runs
32 FROM Virat_Performance
33 GROUP BY match_format;
34
```

	match_format	average_runs
1	ODI	48
2	T20	36
3	Test	46

## 5) Matches Played Against a Specific Opponent



The screenshot shows a SQL query editor with the following code:

```
37 SELECT match_no, ground, match_date, runs
38 FROM Virat_Performance
39 WHERE opponent = 'Australia';
```

Below the query editor, the results are displayed in a table with columns: match\_no, ground, match\_date, and runs. The results are as follows:

	match_no	ground	match_date	runs
1	9	Vadodara	2009-10-25	30
2	10	Mohali	2009-11-02	10
3	33	Visakhapatnam	2010-10-20	118
4	51	Ahmedabad	2011-03-24	24
5	84	Melbourne	2011-12-26	11
6	85	Melbourne	2011-12-26	0
7	86	Sydney	2012-01-03	23
8	87	Sydney	2012-01-03	9

After performing the initial SQL queries and gaining insights from the **Virat\_Performance** table, we proceeded to load the data into Power BI for more advanced analysis and visualization.

### 1. Loading Data into Power BI:

- We connected Power BI to the SQL Server database containing the **Virat\_Performance** table. Using Power BI's data import functionality, we imported the entire table into the Power BI environment.
- This allowed us to leverage Power BI's robust visualization capabilities to create interactive and dynamic dashboards.

### 2. Creating Measures in Power BI:

- Once the data was loaded into Power BI, we created several custom measures to extract unique insights and enhance our analysis. These measures were designed to provide deeper insights into Kohli's performance across various dimensions.

- **Average Runs**

```
1 Average Runs = AVERAGE(Virat_Performance[runs])
```

- **Total Centuries**

```
1 Centuries = COUNTROWS(FILTER(Virat_Performance, Virat_Performance[runs] >= 100))
```

- **Total Half Centuries**

```
1 HalfCenturies = COUNTROWS(FILTER(Virat_Performance, Virat_Performance[runs] >= 50 && Virat_Performance[runs] < 100))
```

- **Total Runs by Format**

```
1 Total Runs by Format = SUM(Virat_Performance[runs])
```

- **Total Runs by Opponent**

```
1 Total Runs by Opponent = SUM(Virat_Performance[runs])
```

# **Dashboard Overview**

The Power BI dashboards created for this project provide a comprehensive and interactive visualization of Virat Kohli's cricket performance. These dashboards are designed to present key insights derived from the data, making it easy for users to explore and understand various aspects of Kohli's career.

## **Key Features of the Dashboards:**

### **1. Performance by Format:**

- Visuals showcasing Kohli's total runs, average runs, and performance trends across different match formats (Test, ODI, T20).

### **2. Centuries and Half-Centuries Analysis:**

- A breakdown of the number of centuries and half-centuries scored by Kohli, with the ability to filter by year or opponent.

### **3. Runs by Opponent:**

- A visual representation of Kohli's total runs scored against different opponents, highlighting his strengths and challenges against specific teams.

### **4. Ground-wise Performance:**

- An interactive map or bar chart displaying Kohli's runs at various grounds, showing where he has performed best.

### **5. Yearly Performance Trends:**

- A line chart or bar chart that tracks Kohli's run-scoring patterns year by year, revealing trends and fluctuations in his performance over time.

## **Interactivity and User Experience:**

- The dashboards are equipped with slicers and filters, allowing users to drill down into specific data points, such as filtering by opponent, match format, or year.
- Hover effects and tooltips provide additional context for the data points, enhancing the overall user experience.

## **Detailed Analysis and Insights**

The analysis of Virat Kohli's cricket performance, visualized through the Power BI dashboards, reveals several key insights that highlight his strengths, trends, and patterns throughout his career. The insights are derived from the various custom measures and visualizations created, providing a comprehensive view of Kohli's achievements and areas of interest.

### **1. Format-Wise Performance:**

- **Test Matches:** Kohli has consistently performed in Test matches, with notable peaks in his run-scoring during specific series. The analysis shows that he has adapted well to the demands of the longest format, with several high-scoring innings across different venues.
- **ODI Matches:** Kohli's ODI performance is marked by his exceptional consistency. The dashboards highlight his ability to score heavily in this format, with a large number of centuries and half-centuries. His average runs in ODIs are particularly impressive, showcasing his reliability as a top-order batsman.
- **T20 Matches:** In the T20 format, Kohli's ability to adapt his game to the shorter version is evident. The analysis indicates that while his average runs are slightly lower than in ODIs and Tests, he remains a crucial player with significant contributions in key matches.

### **2. Centuries and Half-Centuries:**

- The analysis of Kohli's centuries and half-centuries reveals his penchant for scoring big runs, especially in high-pressure situations. A significant number of his centuries have come in matches where the team was under pressure, demonstrating his ability to perform when it matters most.
- The distribution of centuries and half-centuries across different years and opponents also highlights his consistent dominance, with certain teams and years standing out as particularly prolific for him.

### **3. Opponent Analysis:**

- Kohli's performance against various opponents shows that he has excelled against some of the strongest teams in the world. The dashboards indicate his high run-scoring ability against teams like Australia, South Africa, and Sri Lanka, where he has repeatedly delivered match-winning performances.
- The analysis also reveals areas where Kohli has faced challenges, with lower averages against certain opponents. This insight could be valuable for future strategic planning and preparation.

### **4. Ground-Wise Performance:**

- The ground-wise analysis reveals that Kohli has certain favorite venues where he has consistently scored heavily. These include stadiums in India where he enjoys home-ground advantage, as well as some international venues where he has historically performed well.



- Conversely, the analysis also points out venues where Kohli's performance has been relatively subdued, offering insights into conditions or opposition strategies that might have affected his game.

### **5. Yearly Trends:**

- The year-by-year performance trend showcases Kohli's growth as a cricketer, with a clear upward trajectory in his run-scoring over the years. Certain years stand out as particularly successful, coinciding with major milestones in his career.
- The analysis also highlights any dips in performance, which could be correlated with external factors such as injuries or team dynamics, providing a holistic view of his career progression.

### **Strategic Implications:**

- The detailed insights from this analysis can be valuable for cricket analysts, coaches, and strategists. By understanding Kohli's strengths and areas where he could improve, teams can better plan their strategies when facing him.
- Additionally, fans and commentators can use these insights to appreciate the nuances of Kohli's game and his evolution as one of the world's premier batsmen.

## **Conclusion**

This project provides an in-depth analysis of Virat Kohli's cricketing career, leveraging the power of data analytics and visualization through Power BI. By integrating data from various sources and performing detailed analysis, we have gained valuable insights into Kohli's performance across different formats, against various opponents, and at specific venues.

The creation of custom measures such as average runs, total centuries, and runs by format allowed us to explore Kohli's career from multiple angles, highlighting his consistency, adaptability, and match-winning capabilities. The dashboards created serve as a comprehensive tool to visualize and interact with the data, making complex insights easily accessible and understandable.

Key findings from this analysis reveal Kohli's dominance in ODI cricket, his remarkable consistency in scoring centuries, and his ability to perform under pressure. The analysis also uncovers trends in his performance over the years, showcasing his growth and evolution as a cricketer.

Overall, this project demonstrates how data analytics can transform raw statistics into actionable insights, offering a deeper understanding of a player's strengths, weaknesses, and overall contribution to the sport. For cricket analysts, coaches, and enthusiasts, the insights derived from this project provide a valuable resource for appreciating the intricacies of Virat Kohli's illustrious career.

As we conclude this analysis, it's clear that data-driven insights are not just about numbers—they tell a story. In the case of Virat Kohli, the story is one of unparalleled skill, determination, and impact on the world of cricket. This project not only celebrates his achievements but also underscores the importance of analytics in sports, paving the way for more informed decisions and strategies in the future.

## **References :-**

- ❖ Kaggle. (n.d.). Virat Kohli Test, ODI, T20 Statistics. Retrieved from <https://www.kaggle.com/datasets/akashkotal/virat-kholi-testodit20-stat>
- ❖ Microsoft. (n.d.). Power BI Documentation. Retrieved from <https://docs.microsoft.com/en-us/power-bi/>
- ❖ Microsoft. (n.d.). SQL Server Management Studio (SSMS). Retrieved from <https://docs.microsoft.com/en-us/sql/ssms/sql-server-management-studio-ssms>
- ❖ Virat Kohli Career Statistics: ESPN Cricinfo. (n.d.). Virat Kohli Player Profile. Retrieved from <https://www.espncriinfo.com/players/virat-kohli-253802>
- ❖ Schempp, P. (2018). The Role of Data Analytics in Sports. Journal of Sports Analytics, 4(2), 115-130.
- ❖ Microsoft. (n.d.). DAX Function Reference. Retrieved from <https://docs.microsoft.com/en-us/dax/dax-function-reference>
- ❖ Bhardwaj, R., & Bansal, P. (2020). Analyzing Player Performance Using Statistical Methods. International Journal of Sports Science, 8(1), 45-60.