

# SQL-Case-Study-On-Sachin-Tendulkar-s-ODI-Brilliance



## INTRODUCTION

In the annals of cricket history, few names resonate with as much reverence as Sachin Tendulkar. Often hailed as the “**God of Cricket**,” Tendulkar’s monumental impact on the game is immortalized through his extraordinary records and unparalleled achievements. One facet of his illustrious career that stands as a testament to his cricketing genius is his performance in One Day Internationals (**ODIs**). From his debut in **1989** to his retirement in **2012**, Tendulkar’s ODI journey spans over two decades, encapsulating an era of cricket marked by his unwavering dedication, remarkable consistency, and a thirst for records that seemed insatiable.

As we embark on a deep dive into Sachin Tendulkar’s ODI statistics, this **case study** explores the intricacies of his cricketing prowess using the lens of **SQL** (Structured Query Language). SQL, a powerful language for managing and manipulating relational databases, provides us with a unique toolset to unravel the nuances of Tendulkar’s performance. Through SQL queries and analysis, we aim to uncover patterns, milestones, and statistical insights that illuminate the unparalleled career of one of cricket’s greatest icons.

### DATASET LINK:

<https://docs.google.com/spreadsheets/d/1clhicKPjOo8K1uRgwWvsAbDaGn8JaZlqxqtWmBc6hTE/edit?usp=sharing>

The dataset includes only a single table. The name of the table is Scores and it has the following attributes:

**Match\_No: Integer Primary Key**

**Innings: Integer**

**Match\_Date: Date**

**Versus: Text**

**Mode\_Of\_Dismissal: Text**

**Runs: Integer**

**Balls\_Faced: Integer**

**The columns include :-**

**Match\_No**, signifying the match sequence;

**Innings**, indicating the specific inning played by Tendulkar;

**Match\_Date**, providing the date of the match;

**Versus**, specifying the opposing team;

**Mode\_Of\_Dismissal**, detailing how he was dismissed;

**Runs**, enumerating the runs scored by Tendulkar in that particular match; and

**Balls\_faced**, revealing the number of deliveries he faced.

This comprehensive dataset offers a comprehensive overview of Tendulkar's ODI performances, facilitating analysis and insights into his cricketing achievements.

**Before attempting the questions**, feel free to engage in exploratory data analysis on the dataset to enhance your understanding of the data.

**Note:** The table doesn't contain the data points of matches in which **Sachin Tendulkar did not bat**.

## Basic Cricketing Terms:

**Innings:** In cricket, an ‘innings’ denotes the instance in which a team or batsman participates in batting. The distinction between ‘Match\_No’ and ‘Innings’ lies in their sequential progression. While ‘Match\_No’ increases consistently with each match a batsman plays, ‘Innings’ only increments when the batsman actively participates in batting during that match. Therefore, if a batsman is part of a match but doesn’t get the opportunity to bat, the ‘Match\_No’ increases, but the ‘Innings’ value remains constant for that specific match.

**Average:** The batting average in cricket is defined as the total runs scored divided by the total number of dismissals. For eg: A Batsman scores 400 runs in 10 innings and is dismissed 8 times, his average will be  $400/8 = 50$  runs

$$Average = \left( \frac{\text{Total Runs Scored}}{\text{Total No Of Dismissals}} \right)$$

**Strike Rate:** The strike rate in batting is a measure of a batsman’s scoring ability and is expressed as the number of runs scored per 100 balls faced. Its formula is

$$StrikeRate = \left( \frac{\text{Total Balls Faced}}{\text{Total Runs Scored}} \right) \times 100$$

**Hundreds:** A batsman is considered to have achieved a century or hundred when he scores 100 or more runs.

**Fifties:** A batsman is considered to have achieved a half-century or fifty when scoring between 50 and 99 runs. Please note that if a batsman scores a half-century and then converts it into a century, it won’t be counted as a separate half-century.

**Note:** Knowledge of **window functions**, with clause (**CTE**), **subqueries** and **union** is recommended before attempting the below questions. While not mandatory, possessing some basic **cricketing knowledge** can be beneficial for attempting these questions.

1) Write a query to display a summary of Sachin Tendulkar's ODI career, including various batting statistics such as Matches, Innings, Runs, Balls Faced, Not Outs, Highest Score, Average, Strike Rate, Hundreds, and Fifties.

```
SELECT 'Batting Statistic','Value'
UNION
SELECT 'Total Matches',COUNT(Match_No) FROM Scores
UNION
SELECT 'Total Innings',COUNT(Innings) FROM Scores
UNION
SELECT 'Total Runs',SUM(Runs) FROM Scores
UNION
SELECT 'Total Ball Faced',SUM(Balls_faced) FROM Scores
UNION
SELECT "Not Out's",COUNT(mode_of_dismissal) FROM Scores
WHERE mode_of_dismissal='not out'or mode_of_dismissal='retired hurt'
UNION
SELECT 'Highest Score',MAX(Runs) FROM Scores
UNION
SELECT 'Average',(SELECT ROUND((total_runs/(no_of_dismissal)), 2) AS Average
FROM (SELECT(SELECT SUM(runs) FROM scores) AS total_runs,
(SELECT COUNT(*) FROM scores
WHERE mode_of_dismissal Not IN ('not out','retired hurt')) AS no_of_dismissal
) AS subquery) from Scores
UNION
SELECT 'Strike Rate',ROUND((SUM(Runs)/SUM(Balls_faced))*100,2) FROM Scores
UNION
SELECT "Total Hundreded's",COUNT(Runs) FROM Scores
WHERE Runs BETWEEN 100 AND 199
UNION
SELECT "Total Fifty's",COUNT(Runs) FROM Scores
WHERE Runs BETWEEN 50 AND 99 ;
```

Batting Statistic	Value
Batting Statistic	Value
Total Matches	452
Total Innings	452
Total Runs	18426
Total Ball Faced	21367
Not Out's	41
Highest Score	200
Average	44.83
Strike Rate	86.24
Total Hundreded's	48
Total Fifty's	96

2) Find the number of centuries and half-centuries scored by Sachin Tendulkar each year since his debut.

```
SELECT SUBSTRING_INDEX(SUBSTRING_INDEX(match_date, ' ', 1), '-', -1) AS 'YearWise',
SUM(CASE WHEN Runs BETWEEN 100 AND 199 THEN 1 ELSE 0 END ) AS 'Centuries',
SUM(CASE WHEN Runs BETWEEN 50 AND 99 THEN 1 ELSE 0 END) AS 'Half Centuries'
FROM Scores
GROUP BY SUBSTRING_INDEX(SUBSTRING_INDEX(match_date,' ',1),'-', -1);
```

	YearWise	Centuries	Half Centuries
▶	1989	0	0
	1990	0	1
	1991	0	4
	1992	0	6
	1993	0	1
	1994	3	9
	1995	1	1
	1996	6	9
	1997	2	5
	1998	9	7
	1999	3	1
	2000	3	6
	2001	4	3
	2002	2	3
	2003	3	8
	2004	1	5
	2005	1	2
	2006	2	3
	2007	1	13
	2008	1	3
	2009	3	3
	2010	0	0
	2011	2	2
	2012	1	1

3) Find the number of innings, total runs scored and average against each team Sachin Tendulkar played. Return result in ascending order of teams.

```
SELECT Versus,COUNT(Innings) AS 'Total Innings', SUM(Runs) AS 'Runs Scored', AVG(Runs) AS 'Avg Runs'
FROM Scores
GROUP BY Versus
ORDER BY Versus;
```

	Versus	Total Innings	Runs Scored	Avg Runs
▶	Australia	70	3077	43.9571
	Bangladesh	11	496	45.0909
	Bermuda	1	57	57.0000
	England	37	1455	39.3243
	Ireland	2	42	21.0000
	Kenya	9	647	71.8889
	Namibia	1	152	152.0000
	Netherlands	2	79	39.5000
	New Zealand	41	1750	42.6829
	Pakistan	67	2526	37.7015
	South Africa	57	2001	35.1053
	Sri Lanka	80	3113	38.9125
	United Arab...	2	81	40.5000
	West Indies	39	1573	40.3333
	Zimbabwe	33	1377	41.7273

4) Write a query to display Sachin milestone matches ie matches in which Sachin completed 1000, 5000 and 10000 runs. Display the milestone number, milestone runs, the innings and the match\_no in which the milestone was achieved.

```
WITH CTE AS (
SELECT Match_No,Innings,Runs,
SUM(Runs) OVER(ORDER BY Match_No ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW) AS 'Rolling_Sum'
FROM Scores),
CTE2 AS (
SELECT 1 AS 'Milestone_Number',1000 AS 'Milestone_Runs'
UNION ALL
SELECT 2 AS 'Milestone_Number',5000 AS 'Milestone_Runs'
UNION ALL
SELECT 3 AS 'Milestone_Number',10000 AS 'Milestone_Runs'
)
SELECT Milestone_Number,Milestone_Runs,MIN(Match_No) AS 'Matches', MIN(Innings) AS 'Innings' FROM CTE2
INNER JOIN CTE
ON Rolling_Sum>Milestone_Runs
GROUP BY Milestone_Number,Milestone_Runs
ORDER BY Milestone_Number;
```

	Milestone_Number	Milestone_Runs	Matches	Innings
▶	1	1000	36	34
	2	5000	141	138
	3	10000	266	259

5) Sachin Tendulkar holds the record for being in the nineties (90s) the most number of times. Determine the instances when Sachin Tendulkar was dismissed in his 90s, the number of times he remained not out, and the total number of times he scored in the 90s throughout his entire ODI career.

```
SELECT
SUM(CASE WHEN (Runs BETWEEN 90 AND 99) AND (mode_of_dismissal='not out') THEN 1 ELSE 0 END) AS 'Remained Not Out',
SUM(CASE WHEN (Runs BETWEEN 90 AND 99) AND (mode_of_dismissal<>'not out') THEN 1 ELSE 0 END) AS 'Out in 90s',
SUM(CASE WHEN Runs BETWEEN 90 AND 99 THEN 1 ELSE 0 END) AS "90's Scores"
FROM Scores ;
```

Result Grid			
	Remained Not Out	Out in 90s	90's Scores
▶	1	17	18

6) Determine the longest streak of consecutive innings in which Sachin Tendulkar scored 50 or more.

```
WITH CTE AS (SELECT Innings,Runs,
Innings-LAG(Innings,1,NULL) OVER(ORDER BY Innings) AS Diff FROM Scores
WHERE Runs>=50),
CTE2 AS
(SELECT *,
ROW_NUMBER() OVER(ORDER BY Innings),
ROW_NUMBER() OVER(PARTITION BY Diff ORDER BY Innings),
ROW_NUMBER() OVER(ORDER BY Innings)-ROW_NUMBER() OVER(PARTITION BY Diff ORDER BY Innings) AS 'Diff2'
FROM CTE)
SELECT MAX(Cnt)+1 AS 'Longest Streak' FROM
(SELECT Diff2,COUNT(*) Cnt FROM CTE2 GROUP BY Diff2) Subquery ;
```

Result Grid	
	Longest Streak
▶	5

#### Reference –

- Video by Ankit Bansal for Question 4: <https://www.youtube.com/watch?v=7LufPVm01NQ>
- Video for Max Consecutive Values for Question 6: <https://www.youtube.com/watch?v=u1ctlafFferk>
- <https://7criccricket.in/player/sachin-tendulkar/>
- Dataset: <https://docs.google.com/spreadsheets/d/1clhicKPjO08K1uRgwWvsAbDaGn8JaZI9xqtWmBc6hTE/edit?usp=sharing>

Thank You For Reading 😊