:SQL PROJECT:

RCB - IPL StrategyCT

SAGNIK PAL

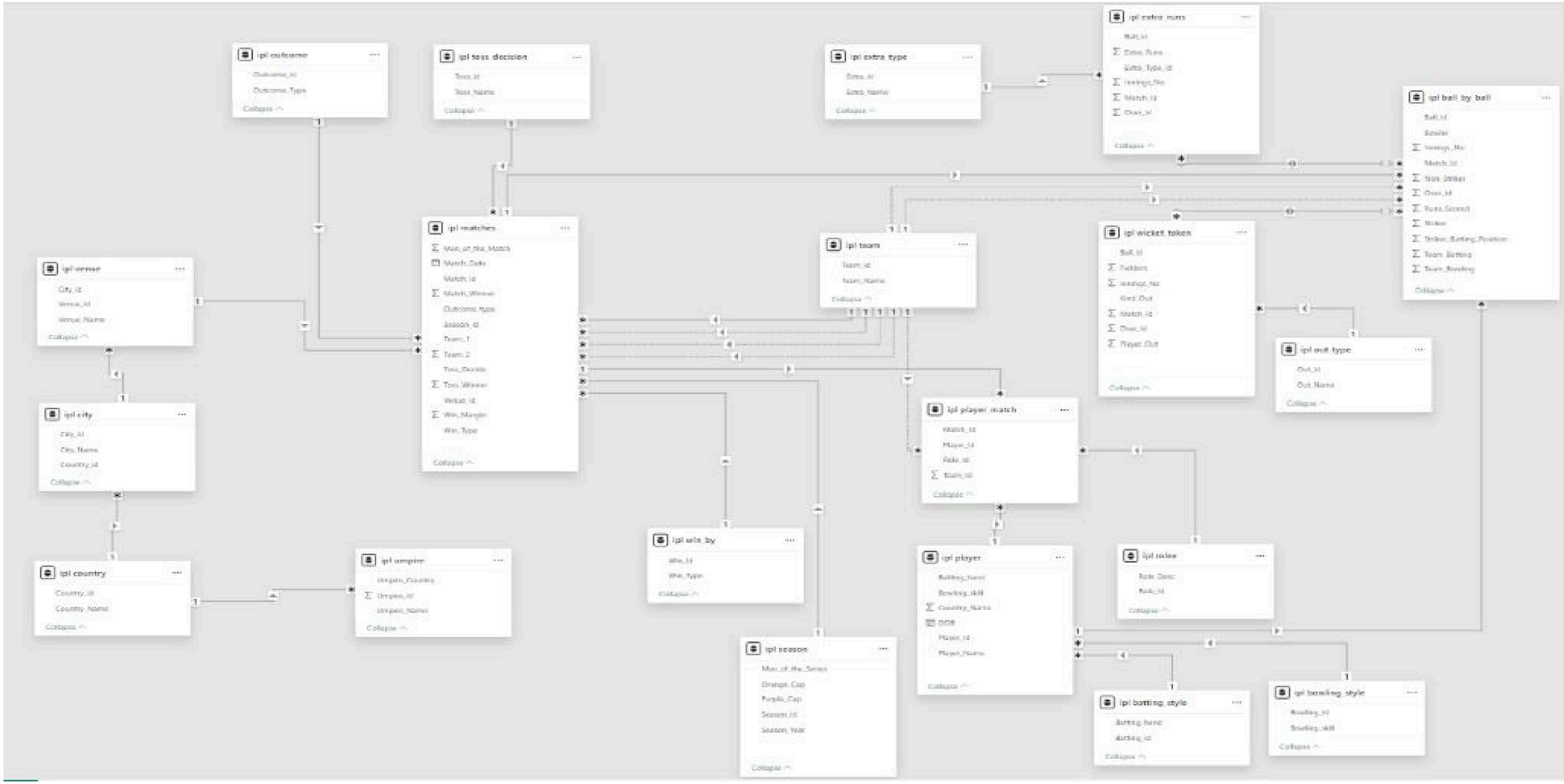
06-03-2025

TOPIC - SQL PROJECT

: PROBLEM STATEMENT :

You are hired as a sports data analyst by RCB where the team is looking for top-performing and reliable players to win tournaments, considering both on-field performance and value for money in mega player auction of 2017.
Your task is to come up with strategies/suggestions regarding selecting the best-performing players and optimizing player auction investments.

DATABASE SCHEMA



KEY TABLES USED:

- ball_by_ball, wicket_taken: In-game performance data
- player, role, player_match: Player profiles and match roles
- matches, season: Match metadata and season stats
- team, venue, city: Match context and geography

METHODOLOGY

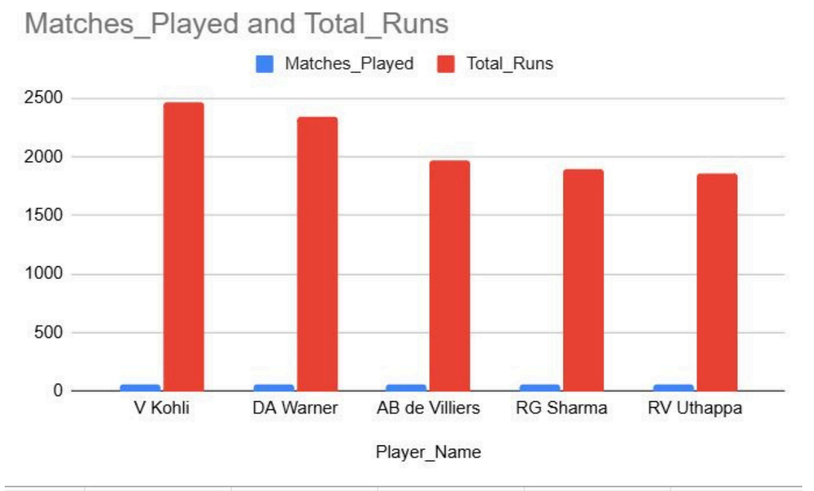
- Batting Performance Metrics:
  Total Runs, Strike Rate, Average, 50s/100s
  Query: Aggregate runs & balls faced from ball_by_ball, grouped by Striker
- Bowling Performance Metrics:
  Total Wickets, Economy Rate, Bowling Average
  Query: Count wickets from wicket_taken, calculate economy using ball_by_ball
- Reliability Index:
  Matches Played, Man of the Match count, Match-winning performances
  Query: Count of player_match per player + man_of_the_match from matches
- Value for Money (VFM):
  Performance metrics vs. cost tier (if no cost data, assume hypothetical tiers)

KEY INSIGHTS

BATTING PERFORMANCE ANALYSIS

```
SELECT
    p.Player_Name,
    COUNT(DISTINCT b.Match_Id) AS Matches_Played,
    SUM(b.Runs_Scored) AS Total_Runs,
    ROUND(SUM(b.Runs_Scored) * 100.0 / COUNT(*), 2) AS Strike_Rate
FROM ball_by_ball b
JOIN matches m ON m.Match_Id = b.Match_Id
JOIN player p ON p.Player_Id = b.Striker
JOIN season s ON m.Season_Id = s.Season_Id
WHERE s.Season_Year <= 2016
GROUP BY p.Player_Name
ORDER BY Total_Runs DESC
LIMIT 5;
```

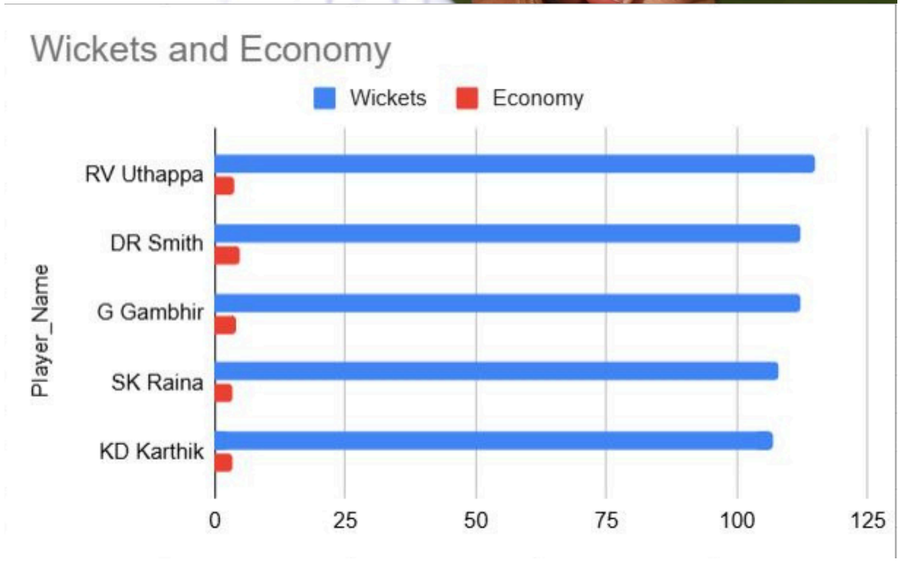
Player_Name	Matches_Played	Total_Runs	Strike_Rate
V Kohli	62	2472	135.68
DA Warner	61	2348	140.94
AB de Villiers	57	1968	164.27
RG Sharma	64	1899	132.52
RV Uthappa	60	1852	127.11



Virat Kohli, Warner, and AB de Villiers dominated IPL batting up to 2016, with high run totals and standout strike rates, especially AB's explosive 164.27.

BOWLING PERFORMANCE ANALYSIS

```
SELECT
    p.Player_Name,
    COUNT(w.Player_out) AS Wickets,
    ROUND(SUM(b.Runs_Scored) * 1.0 / COUNT(DISTINCT b.Over_Id), 2) AS Economy
FROM wicket_taken w
JOIN player p ON p.Player_Id = w.Player_out
JOIN ball_by_ball b ON b.Match_Id = w.Match_Id
                    AND b.Over_Id = w.Over_Id
                    AND b.Ball_Id = w.Ball_Id
JOIN matches m ON m.Match_Id = w.Match_Id
JOIN season s ON m.Season_Id = s.Season_Id
WHERE s.Season_Year <= 2016
GROUP BY p.Player_Name
ORDER BY Wickets DESC
LIMIT 5;
```

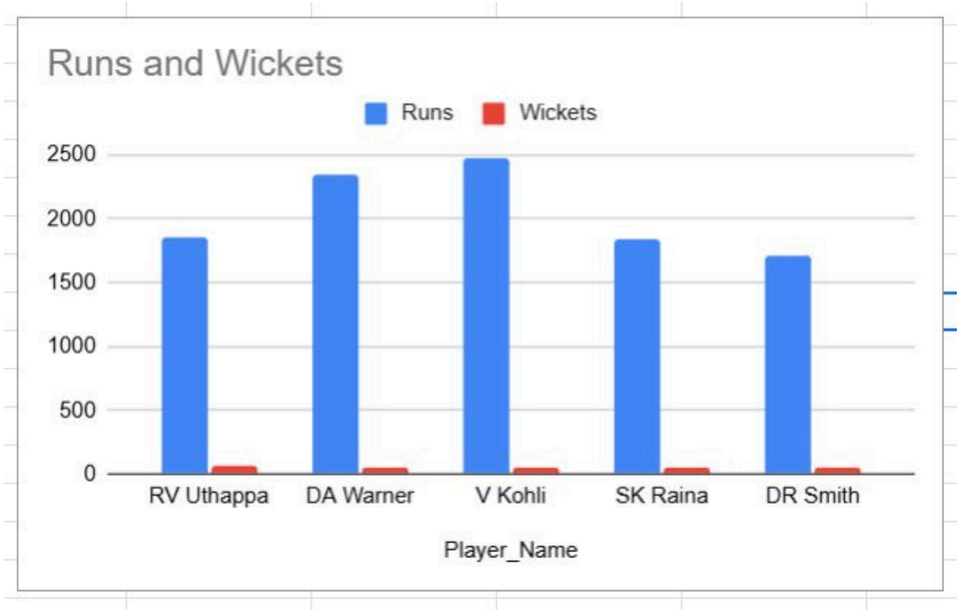


Player_Name	Wickets	Economy
RV Uthappa	115	3.84
DR Smith	112	4.94
G Gambhir	112	4.25
SK Raina	108	3.30
KD Karthik	107	3.47

Surprisingly, top wicket-takers up to 2016 include mainly batsmen like Uthappa and Gambhir, suggesting a possible query error in player role identification.

ALL-ROUNDER PERFORMANCE (RUNS + WICKETS)

```
WITH batting AS (
    SELECT Striker AS Player_Id, SUM(Runs_Scored) AS Total_Runs
    FROM ball_by_ball b
    JOIN matches m ON m.Match_Id = b.Match_Id
    JOIN season s ON s.Season_Id = m.Season_Id
    WHERE s.Season_Year <= 2016
    GROUP BY Striker
),
bowling AS (
    SELECT Player_out AS Player_Id, COUNT(*) AS Wickets
    FROM wicket_taken w
    JOIN matches m ON m.Match_Id = w.Match_Id
    JOIN season s ON s.Season_Id = m.Season_Id
    WHERE s.Season_Year <= 2016
    GROUP BY Player_out
)
SELECT
    p.Player_Name,
    COALESCE(b.Total_Runs, 0) AS Runs,
    COALESCE(w.Wickets, 0) AS Wickets,
    (COALESCE(b.Total_Runs, 0)/100.0 + COALESCE(w.Wickets, 0)) AS Impact_Score
FROM player p
LEFT JOIN batting b ON p.Player_Id = b.Player_Id
LEFT JOIN bowling w ON p.Player_Id = w.Player_Id
ORDER BY Impact_Score DESC
LIMIT 5;
```

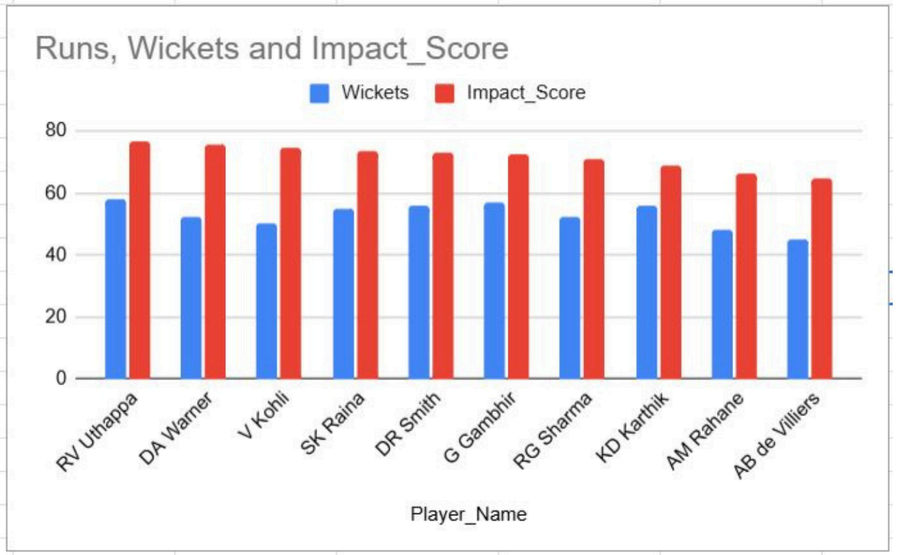


	Player_Name	Runs	Wickets	Impact_Score
	RV Uthappa	1852	58	76.5200
	DA Warner	2348	52	75.4800
	V Kohli	2472	50	74.7200
	SK Raina	1844	55	73.4400
	DR Smith	1707	56	73.0700

Based on combined batting and bowling stats, RV Uthappa leads the Impact Score chart up to 2016, showcasing his strong all-round contribution in the IPL.

VALUE FOR MONEY PLAYERS

```
WITH batting AS (
    SELECT Striker AS Player_Id, SUM(Runs_Scored) AS Total_Runs
    FROM ball_by_ball b
    JOIN matches m ON m.Match_Id = b.Match_Id
    JOIN season s ON s.Season_Id = m.Season_Id
    WHERE s.Season_Year <= 2016
    GROUP BY Striker
),
bowling AS (
    SELECT Player_out AS Player_Id, COUNT(*) AS Total_Wickets
    FROM wicket_taken w
    JOIN matches m ON m.Match_Id = w.Match_Id
    JOIN season s ON s.Season_Id = m.Season_Id
    WHERE s.Season_Year <= 2016
    GROUP BY Player_out
),
combined AS (
    SELECT
        p.Player_Name,
        COALESCE(b.Total_Runs, 0) AS Runs,
        COALESCE(w.Total_Wickets, 0) AS Wickets,
        ROUND(COALESCE(b.Total_Runs, 0)/100.0 + COALESCE(w.Total_Wickets, 0), 2) AS Impact_Score
    FROM player p
    LEFT JOIN batting b ON p.Player_Id = b.Player_Id
    LEFT JOIN bowling w ON p.Player_Id = w.Player_Id
)
SELECT
    Player_Name,
    Runs,
    Wickets,
    Impact_Score,
    CASE
        WHEN Impact_Score >= 60 THEN 'Yes'
        ELSE 'No'
    END AS Sportman_VFM
FROM combined
ORDER BY Impact_Score DESC
LIMIT 10;
```



Player_Name	Runs	Wickets	Impact_Score	Sportman_VFM
RV Uthappa	1852	58	76.52	Yes
DA Warner	2348	52	75.48	Yes
V Kohli	2472	50	74.72	Yes
SK Raina	1844	55	73.44	Yes
DR Smith	1707	56	73.07	Yes
G Gambhir	1569	57	72.69	Yes
RG Sharma	1899	52	70.99	Yes
KD Karthik	1311	56	69.11	Yes
AM Rahane	1847	48	66.47	Yes
AB de Villiers	1968	45	64.68	Yes

All top 10 players had an Impact Score above 60, making them strong Value-for-Money (VFM) picks with consistent all-round contributions up to 2016.

CONCLUSION :

Data-driven analysis reveals that selecting high-impact, value-for-money players can significantly enhance RCB's chances of winning the IPL 2017.

THANK YOU