

|  |
| --- |
| AI CRICKET ANALYST |
|  |
| August 20  Group G: Neeta Chaudhary Naveera Kamal Noor ul ain |



Table of Contents

# INTRODUCTION

# PROJECT SCOPE:

# ER- DAIGRAM:

# APPROACH

# DATA EXPLORATION:

# DASHBOARD VISUALS:

# RECOMMENDATIONS:

## 

INTRODUCTION:

Pakistanis have a passion for cricket, and the Pakistan Super League (PSL) has grown to be an important fixture on the nation's cricket calendar. The "AI Cricket Analyst - PSL 2016-2022 Data Analysis" project seeks to thoroughly analyze the PSL competition data from 2016 to 2022 in order to glean important trends, insights, and patterns. The goal of this project is to provide a thorough study that benefits teams, players, fans, and stakeholders by utilizing data analytics and artificial intelligence.

PROBLEM STATEMENT:

The AI Cricket Analyst for PSL seeks to address the challenge of extracting meaningful insights from the vast array of data generated by the tournament. By utilizing data analytics and AI, the solution aims to provide comprehensive player performance evaluation, strategic decision support for teams, and enhanced fan engagement through interactive visualizations, recognition of historical patterns, predictive modeling, and real-time analysis during matches. Ultimately, the goal is to transform raw data into actionable recommendations that enhance team strategies, player performance, and overall fan experience in the Pakistan Super League, ushering in a new era of data-driven cricket analysis.

OBJECTIVE:

* Comprehensive Data Analysis:

Perform in-depth analysis of PSL data from 2016 to 2022 to extract insights and trends.

* Player Performance Evaluation:

Evaluate batting, bowling, and all-round player performance across seasons.

* Team Dynamics Analysis:

Analyze team performance metrics, including win-loss ratios and net run rates.

* Venue Impact Assessment:

Determine favorable venues for batting and bowling, accounting for seasonal variations.

* Seasonal Trends Identification:

Identify trends and patterns in player and team performance across different seasons.

* Visualizations Creation:

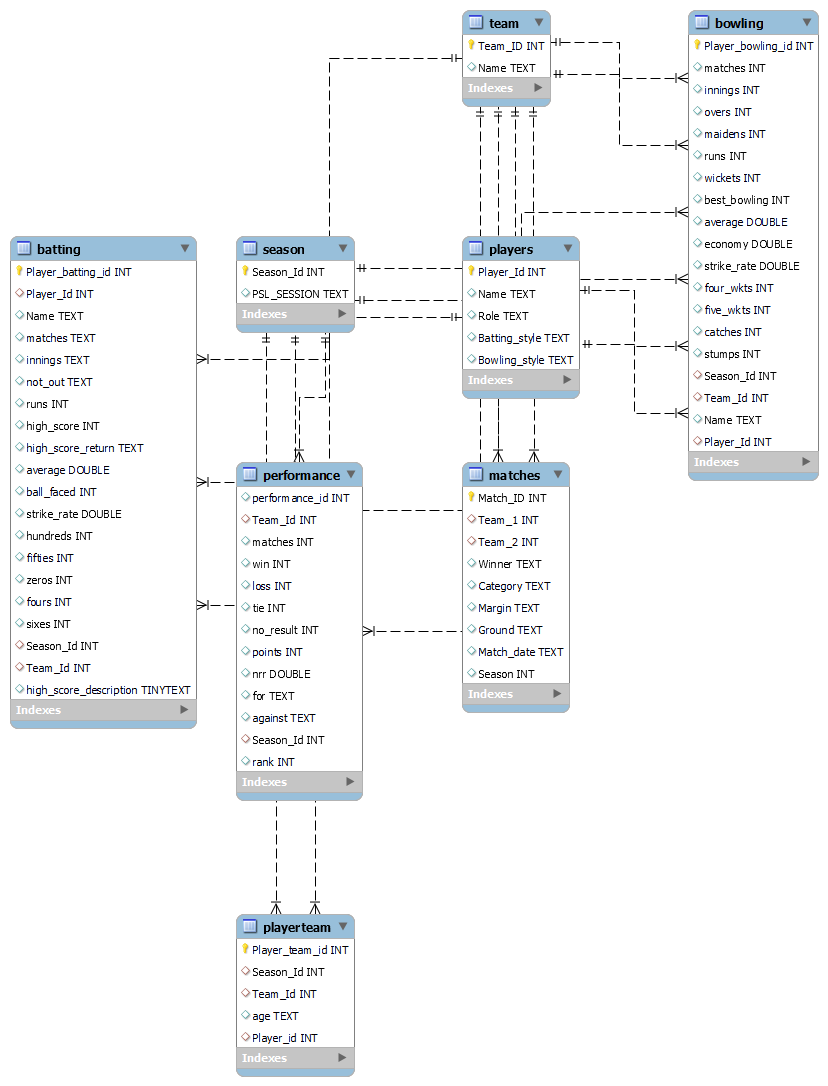
Develop visually appealing charts and graphs to present data insights effectively.

* Strategic Recommendations:

Provide actionable recommendations for teams, players, and organizers to improve strategies.

PROJECT SCOPE:

* Perform extensive data analysis on Pakistan Super League (PSL) data spanning from 2016 to 2022.
* Focus on extracting actionable insights and conducting in-depth analysis across different facets of the tournament.
* Analyze and evaluate player performances, team strategies, and overall tournament dynamics.
* Investigate the impact of different venues on match outcomes, batting, and bowling performances.
* Examine trends and patterns over the years to identify changes in player and team strategies.
* Provide a comprehensive report that encapsulates the findings and insights derived from the analysis.
* Empower decision-making by offering data-driven insights that contribute to strategic planning.
* Offer a deeper understanding of the game by highlighting key statistics and correlations.
* Explore the relationships between player performance, team strategies, and match outcomes.
* The project is limited to PSL data from 2016 to 2022, focusing on its specific timeframe.
* The analysis aims to uncover not only historical trends but also provide actionable insights for future tournaments.

ER- DAIGRAM: 

APPROACH:

Data Collection and Cleaning:

* Gather PSL data (2016-2022) including match scores, player stats, team info, and venue details.
* Clean data to remove inconsistencies, missing values, and errors.
* Rectify incorrect age data types in the Player Team table, affecting 815 rows, through SQL adjustments to ensure accurate age representation.
* Complete missing information within the Player table, specifically filling in fields like role, batting style, and bowling style to enhance player profiles comprehensively.
* Standardize data in the Batting and Bowling tables by introducing hyphens as placeholders for rows lacking information, promoting consistency across the dataset.
* Eliminate duplicate entries of player names from the Players table to ensure accurate and non-redundant player records.

Descriptive Analysis:

* Use basic stats to understand variables like runs, wickets, averages, and strike rates.
* Identify outliers and trends.

Player Performance:

* Calculate batting/bowling averages, strike rates, and other metrics.
* Highlight consistent high performers and notable achievements.

Team Analysis:

* Group data by teams and seasons.
* Analyze team performance trends, win-loss ratios, net run rates, and average scores.

Venue Impact:

* Examine match outcomes, batting/bowling performances at venues.
* Determine favorable conditions for batting/bowling and changes over seasons.

Seasonal Trends:

* Track player and team performance changes across seasons.
* Spot trends, strategies, and patterns linked to successful seasons.

Visualization:

* Create engaging visuals like charts and graphs.
* Present data in a visually appealing manner.

Interpretation:

* Explain findings by connecting them to cricket strategies and player strengths.
* Provide insights into trends and analysis outcomes.

Report Generation:

* Compile all findings, insights, and visuals.
* Offer actionable recommendations for teams, players, and organizers.

Presentation and Sharing:

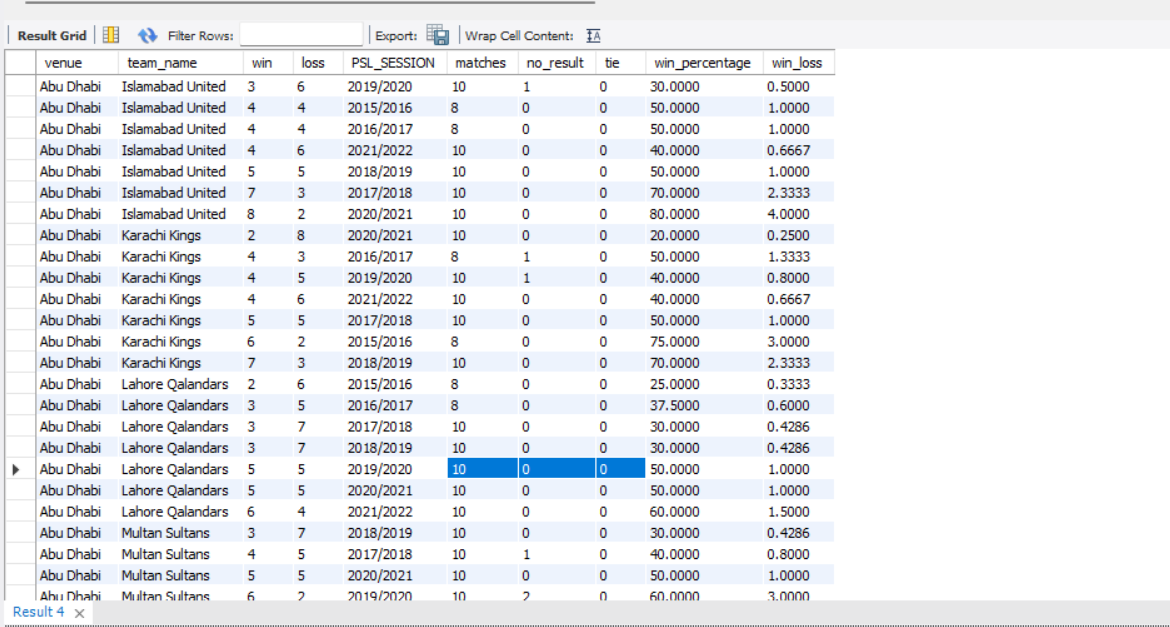
* Share the report with stakeholders including teams, management, and fans.
* Disseminate insights through social media and various platforms.

Continuous Improvement:

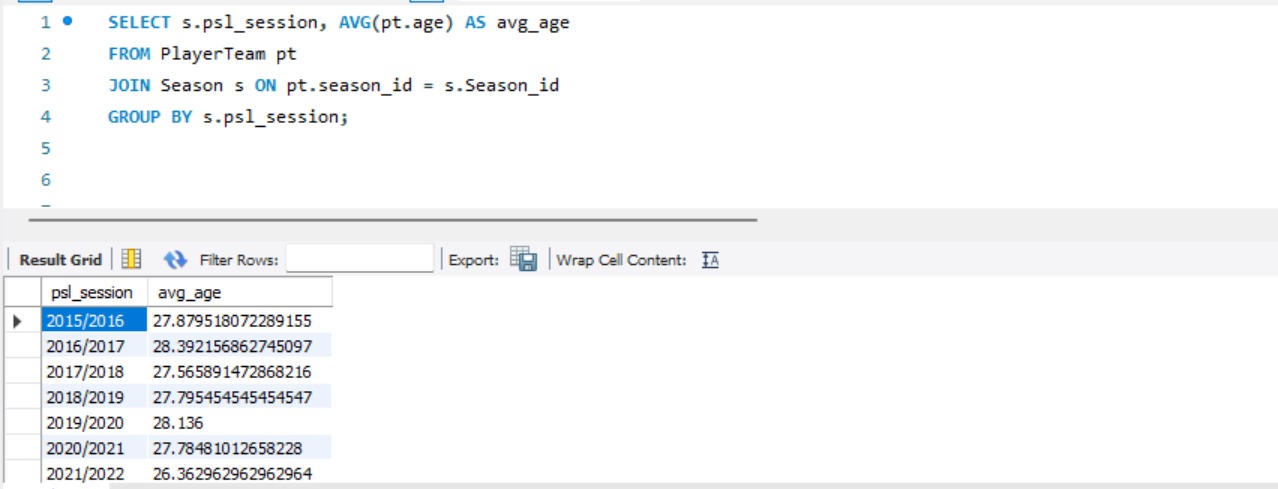
* Collect feedback from stakeholders.
* Incorporate suggestions for refining future analyses.

DATA EXPLORATION:

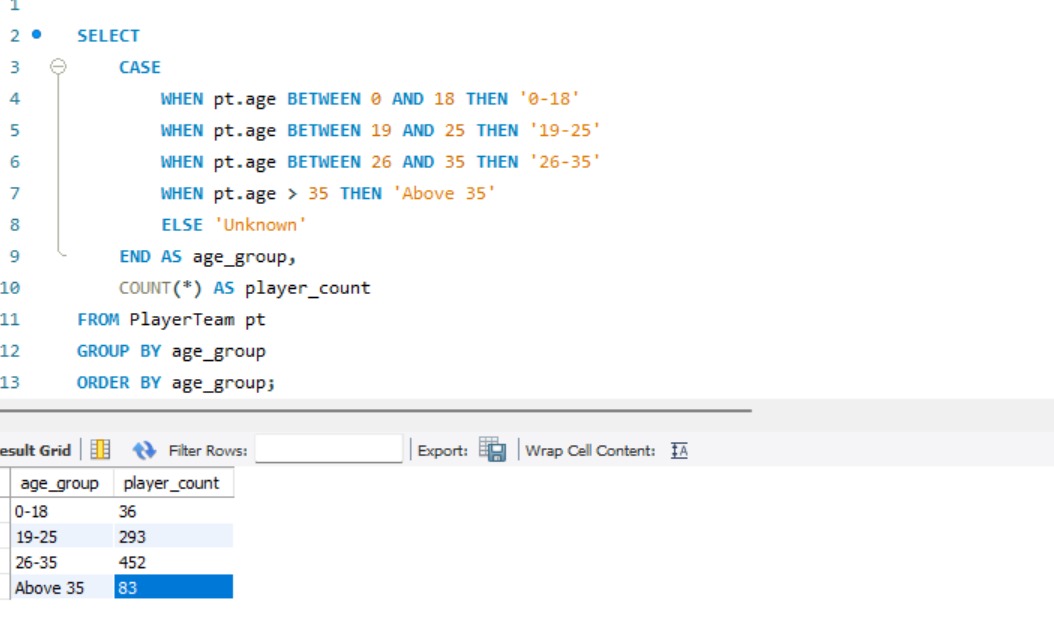
1. Analyze the effects of various venues on team performance and season-long match results.



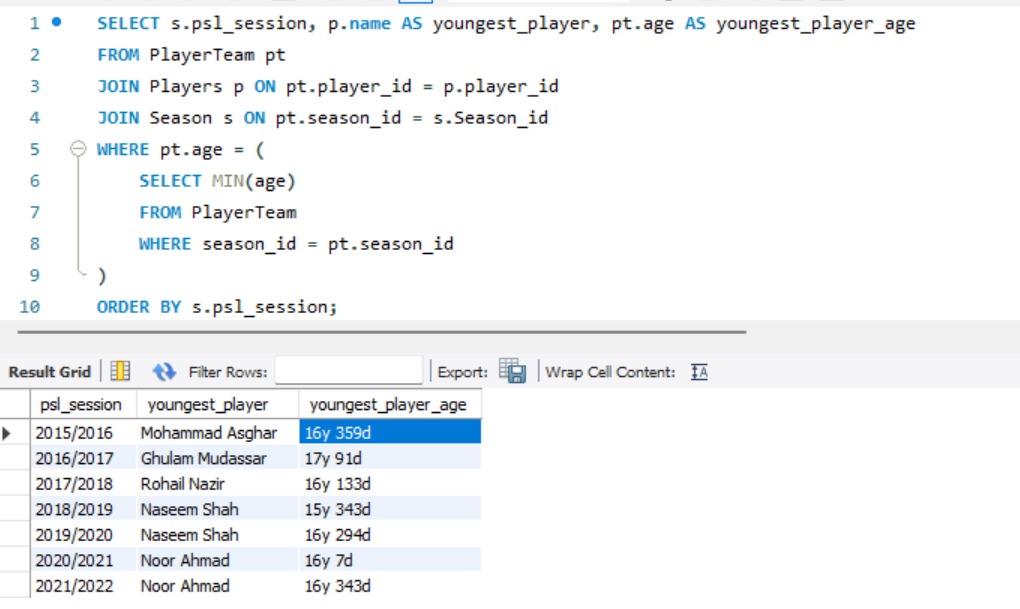
1. What is the average age of players participating in the PSL tournaments over the years?

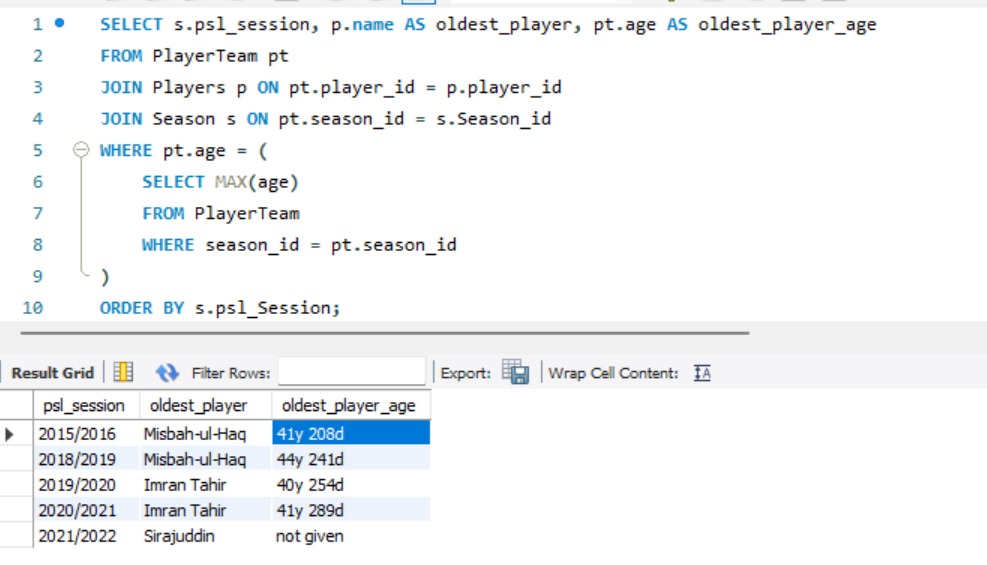


1. What is the overall distribution of player ages in the dataset?



1. Youngest and Experienced Players Age Analysis season wise:

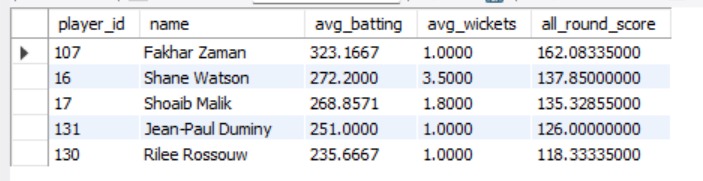




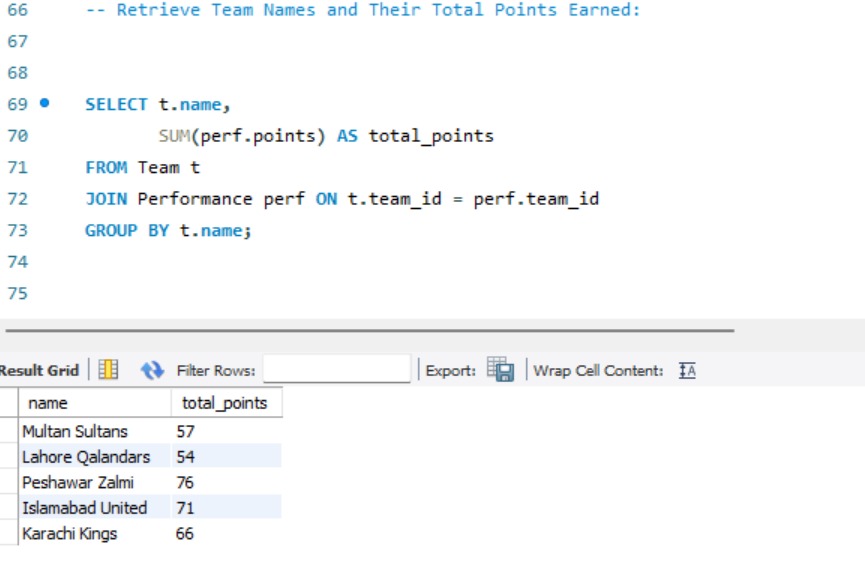
1. Comparing Batting Strike Rate and Bowling Economy Analysis:



1. Analyze whether all-rounders tend to have more balanced performance in both batting and bowling compared to specialized players.

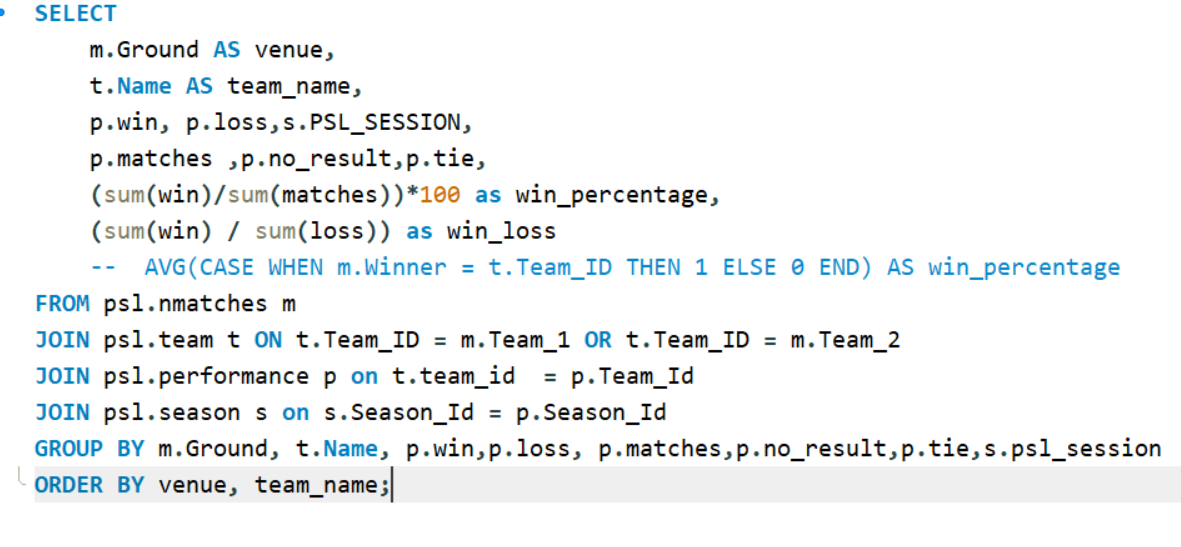


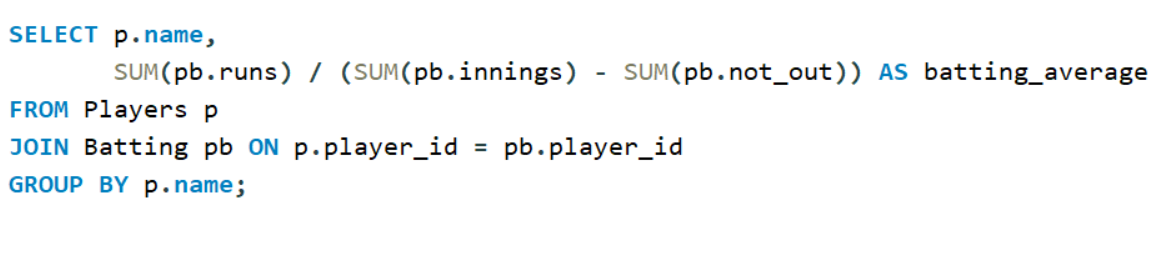
1. Analysis of Team Performances and Points Earned

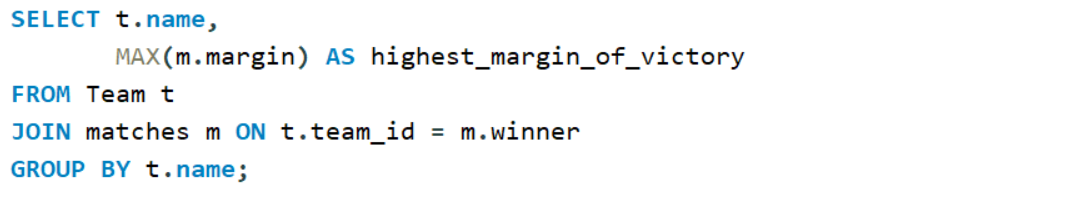


SQL QUERIES:

1. Query #1: Analyze the effects of various venues on team performance and season-long match results.



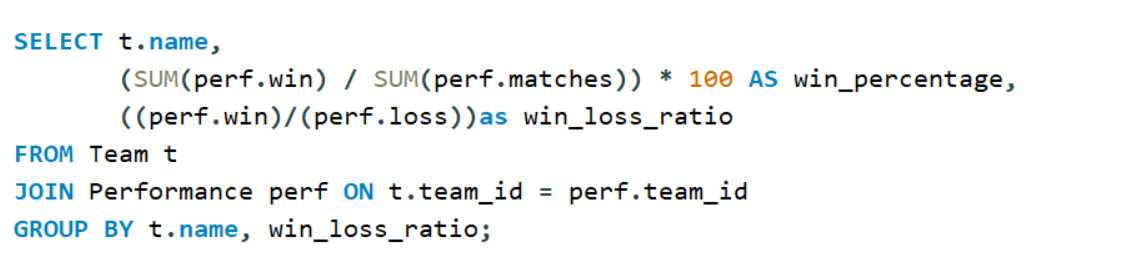
1. Query #2: Retrieve Player Names and Their Batting Averages:
2. Query #3: Retrieve Team Names and Their Highest Margin of Victory:



1. Query #4: Highest wicket taker per season

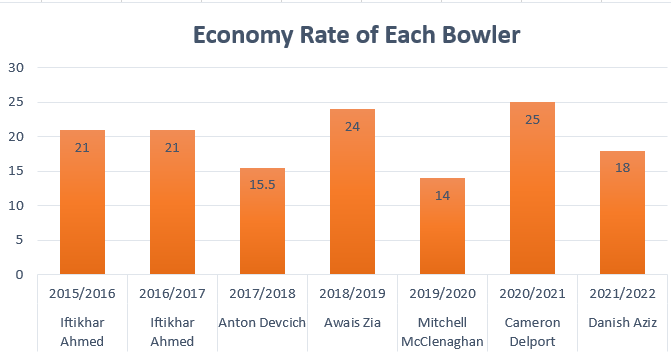


1. Query #5: Retrieve Team Names and Win Percentages and win-loss ratio

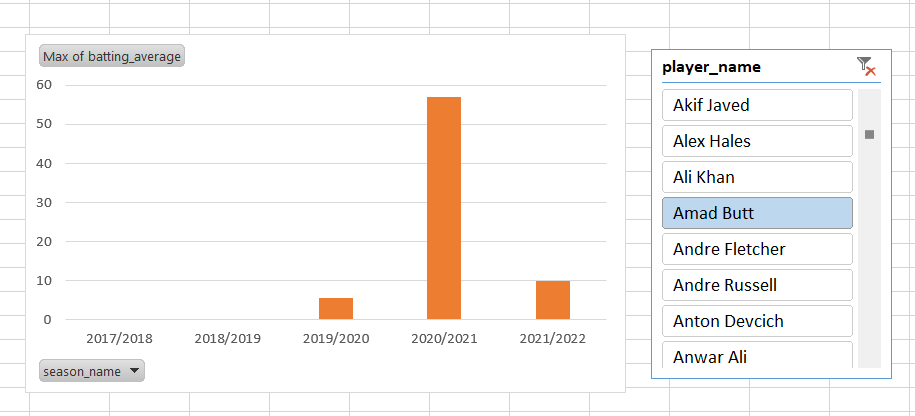


DASHBOARD VISUALS:

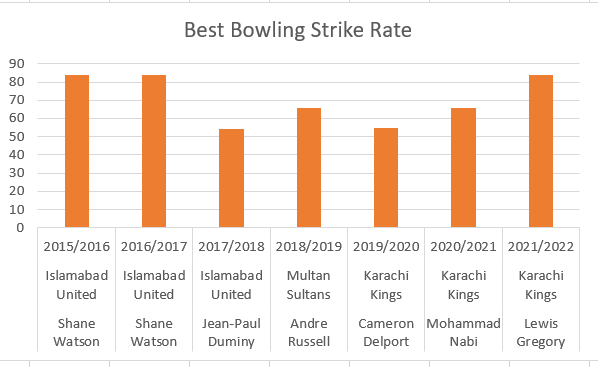
1. Economy rate of each bowler



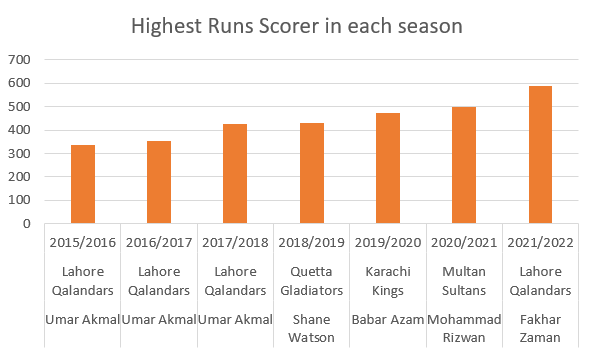
1. Maximum batting average of player



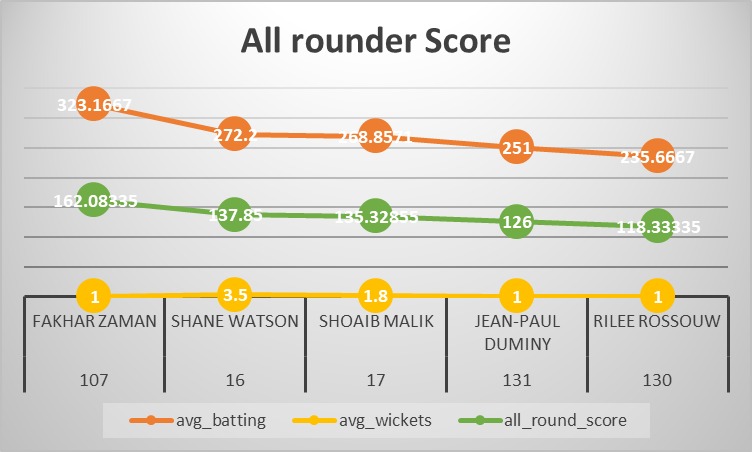
1. Maximum bowling strike rate of player



1. Highest run scorer in each season



1. All-rounder scorer



RECOMMENDATIONS:

* Real-time Analytics:

Develop a real-time analytics dashboard to provide live insights during matches. This could include real-time player statistics, key moments analysis, and predictive analytics for ongoing matches.

* Fan Engagement Insights:

Analyze fan engagement patterns and preferences to provide recommendations for improving fan experiences, both at the stadium and through digital platforms.

* Injury Analysis:

Utilize injury data to identify patterns and factors contributing to player injuries, helping teams design injury prevention strategies.

* Social Media Sentiment Analysis:

Incorporate sentiment analysis of social media discussions to gauge fan reactions, sentiments, and opinions about players, teams, and match outcomes.

* Player Auction Strategy:

Develop a data-driven player auction strategy tool that assists teams in making informed decisions during player drafts, considering player statistics, team composition, and budget constraints.

* Interactive Data Visualization:

Create interactive visualizations that allow users to explore the data themselves, enabling fans, analysts, and stakeholders to uncover insights tailored to their interests.

* Venue-Specific Strategies:

Based on historical venue analysis, provide teams with venue-specific strategies that cater to the strengths and weaknesses of the playing conditions at different stadiums.

* Player Workload Management:

Utilize player workload data to recommend optimal rest and rotation strategies to prevent injuries and maintain player fitness throughout the tournament.

* Opponent Analysis:

Enhance team strategies by offering detailed opponent analysis. Provide insights into opponent weaknesses, key player matchups, and tactical considerations based on historical performance data.