

Problem2- As a data analysis intern, you have to analyze sports data for a client. You are given two datasets

related to IPL (Indian Premier League) cricket matches. One dataset contains ball-by-ball data and the other contains match-wise data. You have to import the datasets into an SQL database and perform the tasks given in this assignment to find important insights from this dataset.

About the Data

The first CSV file is for ball-by-ball data and it has information of all the 193468 balls bowled between the years 2008 and 2020. It has 17 columns and below is the details of those 17

1.write the first query of create a table named 'matches' with appropriate data types for columns.

Ans - CREATE TABLE `matches_table` (
id INT PRIMARY KEY, city
VARCHAR(255), date DATE,
player_of_match VARCHAR(255),
venue VARCHAR(255), neutral_venue
INT, team1 VARCHAR(255), team2
VARCHAR(255), toss_winner
VARCHAR(255), toss_decision
VARCHAR(255), winner
VARCHAR(255), result VARCHAR(255),
result_margin INT, eliminator
VARCHAR(255), method
VARCHAR(255), umpire1

VARCHAR(255), umpire2

VARCHAR(255)

);

2.Create a table named 'deliveries' with appropriate data types for columns

Ans - CREATE TABLE deliveries (

id INT PRIMARY KEY,

inning INT,

`over` INT, ball

INT,

batsman VARCHAR(255),

non_striker VARCHAR(255), bowler

VARCHAR(255), batsman_runs INT,

extra_runs INT, total_runs INT,

is_wicket INT,

dismissal_kind VARCHAR(255), player_dismissed

VARCHAR(255), fielder VARCHAR(255), extras_type

VARCHAR(255), batting_team VARCHAR(255),

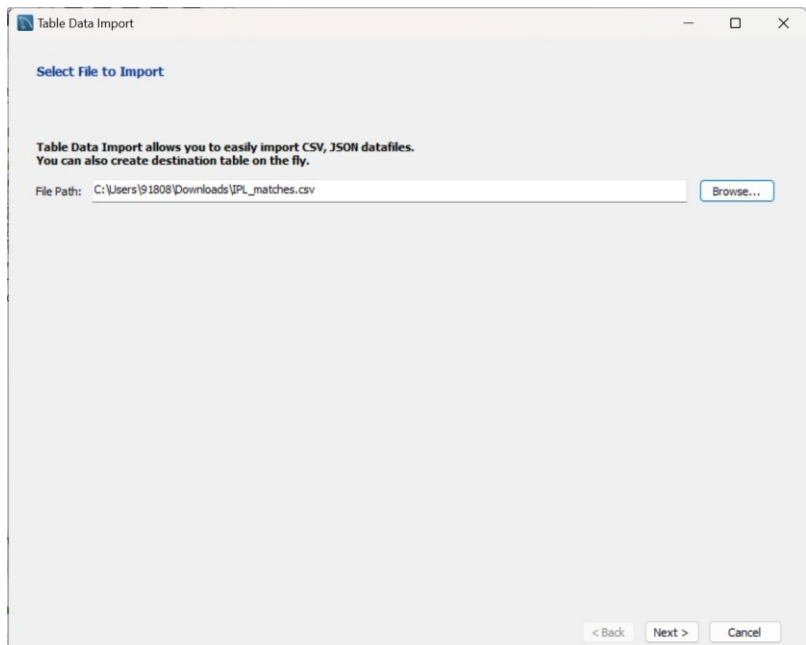
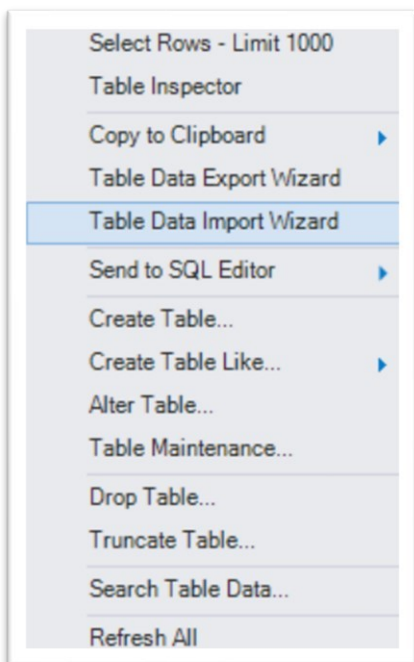
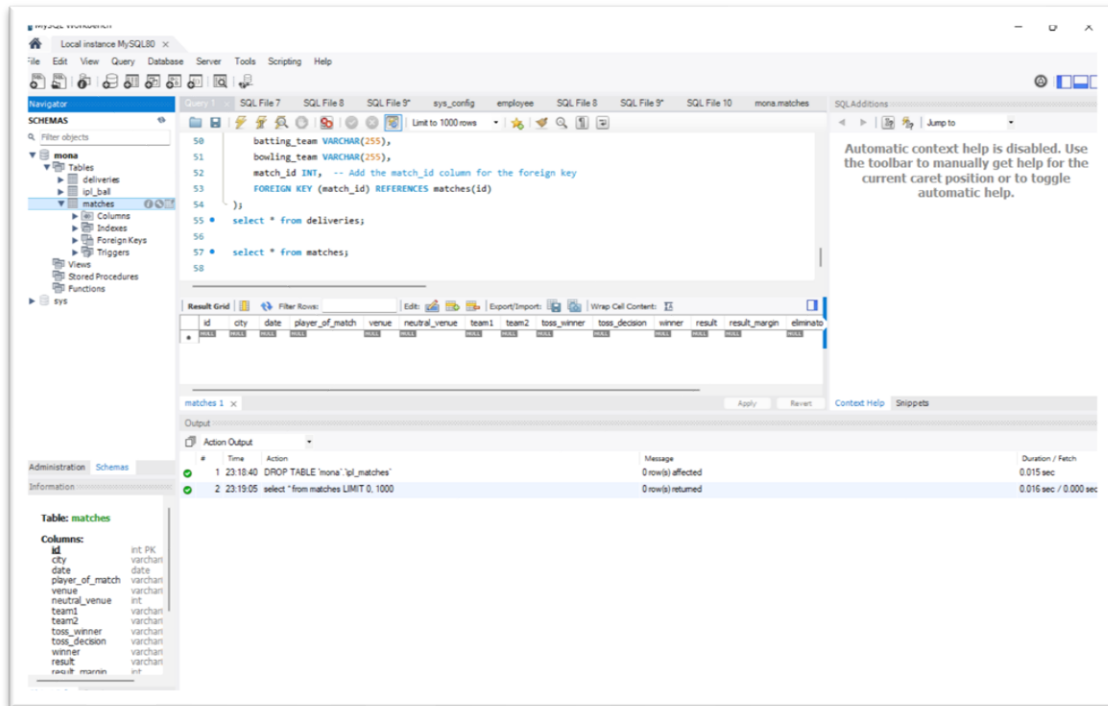
bowling_team VARCHAR(255), match_id INT, -- Add the

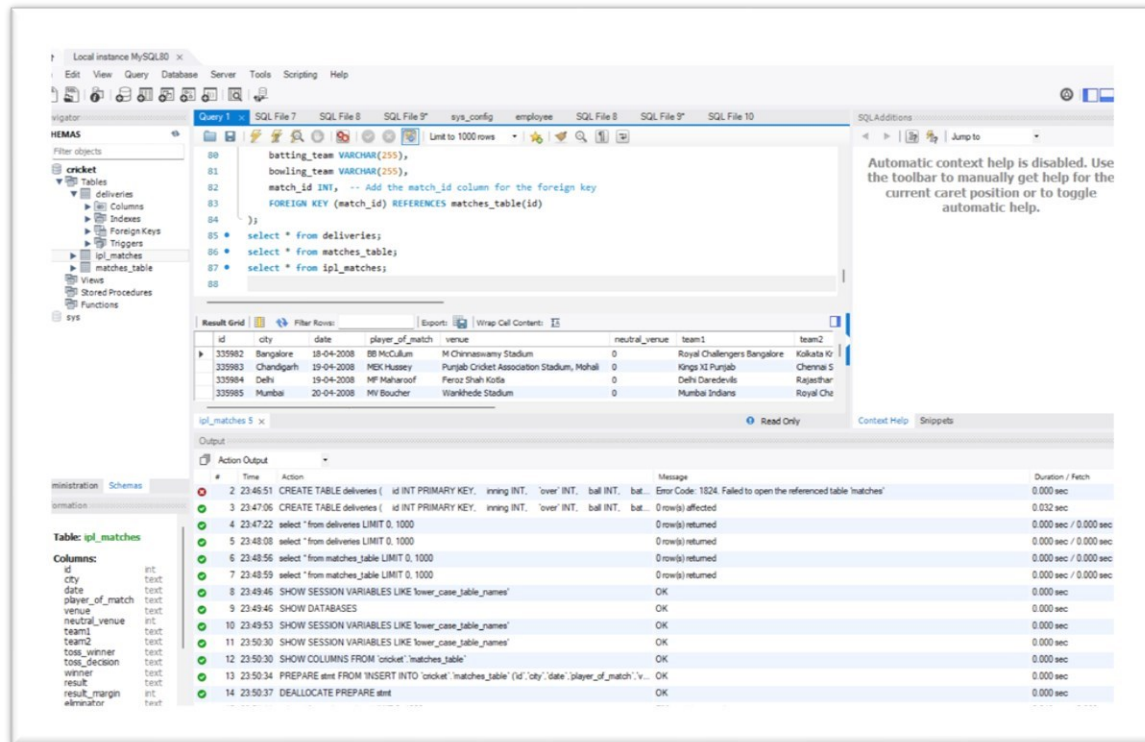
match_id column for the foreign key

FOREIGN KEY (match_id) REFERENCES matches_table(id)

);

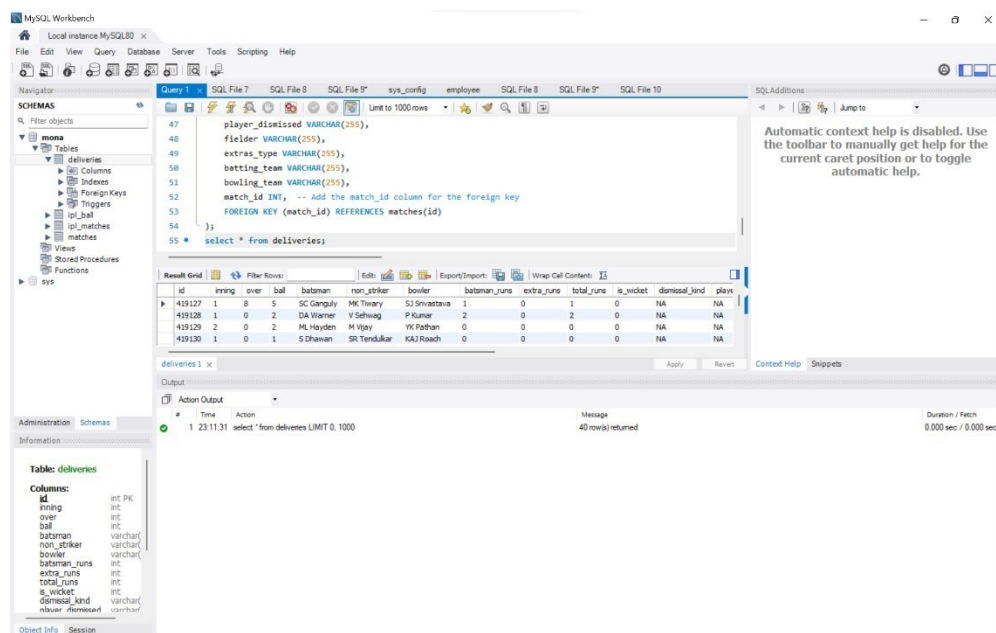
Import data from CSV file 'IPL_matches.csv' attached in resources to 'matches'. Ans

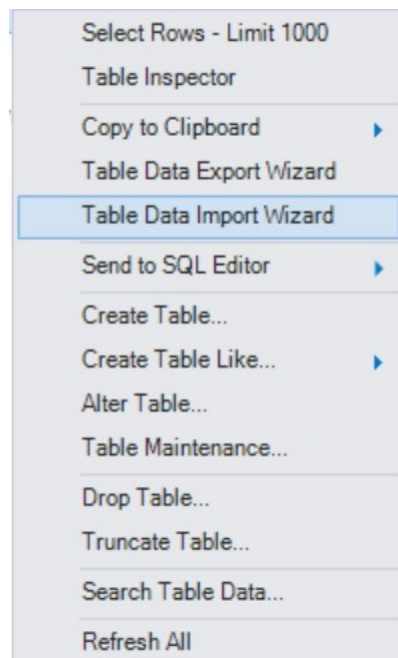


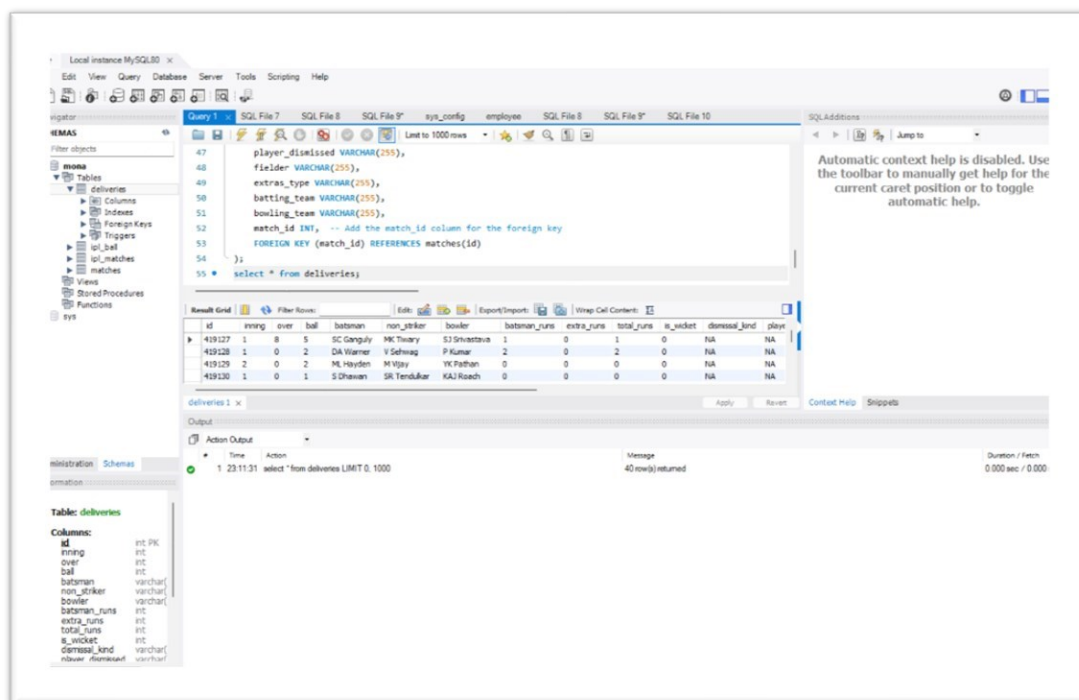
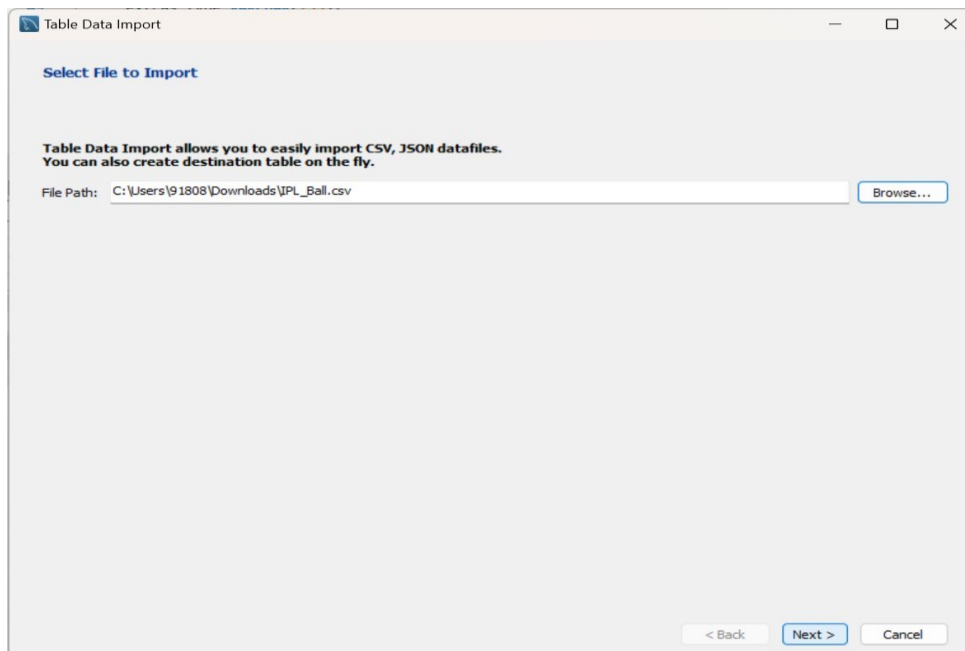


4. Import data from CSV file 'IPL_Ball.csv' attached in resources to 'deliveries'.

Ans-







5. Select the top 20 rows of the deliveries table.

Ans-

```
SELECT *  
FROM deliveries  
LIMIT 20;
```

6. Select the top 20 rows of the matches table.

```
SELECT *  
FROM matches_table  
LIMIT 20;
```

6. Fetch data of all the matches played on 2nd May 2013.

Ans-

```
SELECT *  
FROM matches_table  
WHERE date = '2013-05-02';
```

7. Fetch data of all the matches where the margin of victory is more than 100 runs.

Ans-

```
SELECT *  
FROM matches_table  
WHERE result_margin > 100;
```

8. Fetch data of all the matches where the final scores of both teams tied and order it in descending order of the date.

Ans-

```
SELECT *  
FROM matches_table  
WHERE result = 'tie'  
ORDER BY date DESC;
```

9. Get the count of cities that have hosted an IPL match.

Ans-

```
SELECT COUNT(DISTINCT city) AS total_cities  
FROM matches_table;
```