

Predictive Analytics for Formula 1 Race Outcomes Using Historical Data

Chaitanya Krishna Kommineni

50539270

ckommine@buffalo.edu

Sai Koushik Vimmadisetty

50470993

svimmadi@buffalo.edu

Deeshma Lavanya Gorrela

50470947

deeshmal@buffalo.edu

Abstract— This project aims to develop a predictive model for Formula 1 races, leveraging a comprehensive dataset spanning from 1950 to 2024, including information on drivers, constructors, races, lap times, pit stops, and qualifying sessions. By analyzing key race-day factors such as driver performance, track conditions, tire selection, and pit stop strategies, the model seeks to forecast race outcomes and provide actionable insights for teams. The goal is to help teams adapt strategies in real-time, optimizing performance and enhancing competitiveness, ultimately enabling more precise decision-making and improving race-day results in an ever-evolving sport.

I. PROBLEM STATEMENT

Formula 1 is a sport in which the technologies, strategies, and even the split-second decisions of the runners often hold the key to success or failure. Besides, tire selection and pit stop timing, every bit of race-day decision-making is carefully crafted so as to ensure maximum performance and competitive advantage. The dataset at hand contains detailed information on every aspect of the sport—drivers, constructors, races, lap times, pit stops, qualifying sessions, and more—spanning from the first Formula 1 season in 1950 through to the 2024 season.

The model is to forecast F1 results which will be a major part of the research. It will try to link the key features of the racing day with the winning probability. The data collected about the previous races, which will include driver performance, track conditions, lap times, pit stops, and other race-related factors, as well as the methods the drivers use in the race will be the foundation of our research in finding the connection between them and the correct strategies to be implemented. This model will apply the scientific approach, for example, testing of pit stop times and tire options be one of the scenarios, and furthermore, positively impact the race. Our core objectives are actionable insights for Formula 1 teams and race strategists to come up with alternatives that way around to push issues that arise towards the pits or not.

Through this project, our objective is to develop a system for all F1 teams that will be able to predict the end of the race more precisely. Thus, they will be able to change strategies while a race is going on, which will, in turn, result in improved performance on race day. In this regard, the predictive model

will be a pride achievement for the teams who are interested in elaborating their competitive spirit. Likewise, such teams will be able to come up with more successful strategies for the Formula 1 setting which, as always, will be forever changing.

II. TARGET USERS

Formula 1 teams and engineers are the faces of technological innovation and strategic planning in the world of motor sports. These F1 teams include: McLaren, Red Bull racing, Ferrari, Mercedes, Aston Martin, RB, Haas, etc.

Each of these teams consists of a diverse group of professionals, including race engineers, aerodynamics, data analysts, and mechanics, all working together in order to maximize the performance of their particular cars on race day.

Using this predictive model, F1 strategists and engineers can analyze how parameters such as pit stop timing, tire type, and track-specific conditions really do to a race outcome and take the necessary steps. This approach could also help them to predict the results of possible race winning strategies by enabling them to adapt and make real-time decisions in order to optimize their race performance.

In particular, this model can help:

- Engineers of each team can identify patterns in car performance under different track conditions and can make adjustments that align with the model's recommendations.
- Race strategists can use this model to simulate potential scenarios and select the optimal strategies for race day, including timing of pit stops, tire selections, etc.
- Data analysts and Sports analysts can use these insights from the data to explore trends which can consist of individual driver performances and the factors that are responsible for winning the races.

III. BCNF CHECK AND DECOMPOSITION

To ensure that all the relations of our database are in Boyce-Codd Normal form (BCNF), we have identified whether each table has a primary key that functionally determines all other

attributes in the table, and whether there are no non-trivial FD's on proper subsets of candidate keys. Below are the FD's for each relation:

A. Circuits:

Possible FD's:

1. $\text{circuitId} \rightarrow \text{circuitRef, name, location, country, lat, lng, alt, url}$

This relation is in BCNF because all non-trivial FDs have the primary key (circuitId) as the determinant.

B. Constructors:

Possible FD's:

1. $\text{constructorId} \rightarrow \text{constructorRef, name, nationality, url}$

This relation is in BCNF as the primary key (constructorId) functionally determines all other attributes.

C. Drivers:

Possible FD's:

1. $\text{driverId} \rightarrow \text{driverRef, number, code, forename, surname, dob, nationality, url}$

This relation is in BCNF as driverId is the primary key, and it uniquely determines all other attributes

D. Races:

Possible FD's:

1. $\text{raceId} \rightarrow \text{year, round, circuitId, name, date, time, url}$

The primary key raceId functionally determines all other attributes, thus this relation is in BCNF.

E. Lap_times:

Possible FD's:

1. $\{\text{raceId, driverId, lap}\} \rightarrow \text{position, time, milliseconds}$

The composite key {raceId, driverId, lap} uniquely identifies all the other attributes in the table. Therefore, the table is in **BCNF**.

F. Driver_standings:

Possible FD's:

1. $\text{driverStandingId} \rightarrow \text{raceId, driverId, points, position, positionText, wins}$

This table is in BCNF as the primary key (driverStandingId) functionally determines all other attributes.

G. Constructor_results:

Possible FD's:

1. $\text{constructorResultsId} \rightarrow \text{raceId, constructorId, points}$

This table is in BCNF as the primary key (constructorResultsId) functionally determines all other attributes.

H. Results:

Possible FD's:

1. $\text{resultsId} \rightarrow \text{raceId, driverId, constructorId, number, grid, position, positionText, positionOrder, points, laps, time, milliseconds, fastestLap, ranks, fastestLapTime, fastestSpeed, statusId}$

This table is in BCNF as the primary key (resultsId) functionally determines all other attributes.

I. Qualifying:

Possible FDs:

1. $\text{qualifyId} \rightarrow \text{raceId, driverId, constructorId, number, position, q1, q2, q3}$

This table is in BCNF as the primary key (qualifyId) functionally determines all other attributes.

J. Seasons:

Possible FD's:

1. $\text{year} \rightarrow \text{url}$

The primary key is year, and it functionally determines the url. There are no non-trivial FDs that violate BCNF because year is the only determinant, and it is the primary key. Thus the relation is in BCNF.

K. Sprint_results:

Possible FD's:

1. $\text{resultId} \rightarrow \text{raceId, driverId, constructorId, number, grid, position, positionText, points, laps, time, milliseconds, fastestLap, fastestLapTime, statusId}$

The primary key is resultId, which functionally determines all the other attributes in the table. There are no non-trivial FDs violating BCNF. Therefore, this relation is in BCNF.

L. Status:

Possible FD's:

1. $\text{statusId} \rightarrow \text{status}$

The primary key statusId determines the status description (status). As statusId is the only key, this relation is in **BCNF**.

M. Constructor_standings:

Possible FD's:

1. $\text{constructorStandingsId} \rightarrow \text{raceId, constructorId, points, position, positionText, wins}$

The primary key is constructorStandingsId, and it determines all other attributes. There are no non-trivial FDs violating BCNF.

N. Pit_stops:

Possible FD's:

1. $\{\text{raceId, driverId, stop}\} \rightarrow \text{lap, time, duration, milliseconds}$

The composite primary key {raceId, driverId, stop} determines all other attributes in the table, so this table is in **BCNF**.

Conclusion: All the relations are in BCNF as there are no non-trivial FD's on any proper subset of the primary keys. Therefore, decomposition is not required.

IV. ER DIAGRAM

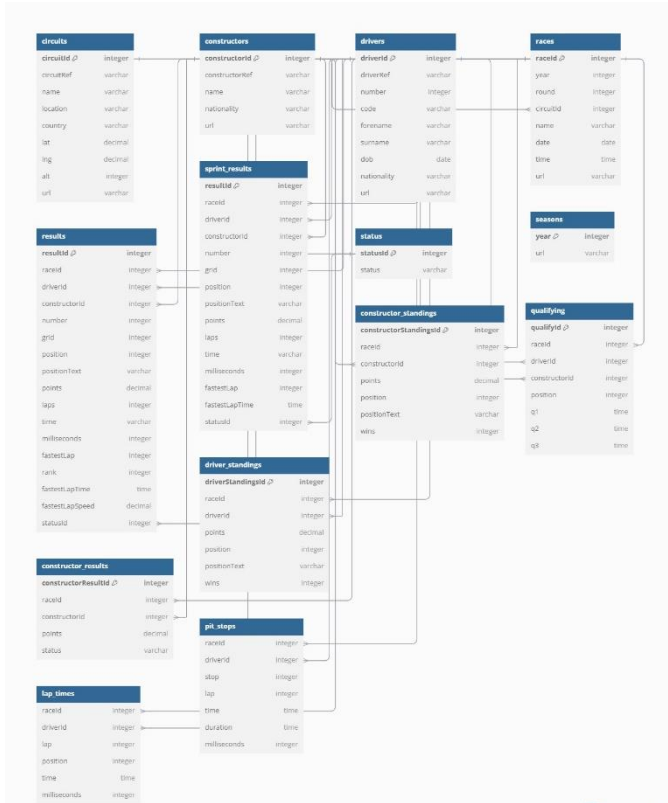


Fig: ER Diagram

V. DATABASE TABLE STRUCTURE

A. Circuits:

- CircuitId (Integer): Unique identifier for each circuit.
- circuitRef (Varchar): Reference code for the circuit.
- Name (varchar): Name of the circuit.
- Location (Varchar): Location of the circuit.
- Country (Varchar): Country where the circuit is located.
- Lat (Decimal): Latitude of the circuit's location.
- Lng (Decimal): Longitude of the circuit's location.
- Alt (Integer): Altitude of the circuit.
- url (Varchar): Web link for more details about the circuit.

B. Constructors:

- constructorId (Integer): Unique identifier for each constructor (team).
- constructorRef (Varchar): Reference code for the constructor.
- Name (Varchar): Name of the constructor.
- Nationality (Varchar): Nationality of the constructor.
- url (Varchar): Web link for more details about the constructor.

C. Drivers:

- driverId (integer): unique identifier for each driver.
- driverref (varchar): reference code for the driver.
- number (integer): driver's racing number.
- code (varchar): short code for the driver.
- forename (varchar): driver's first name.
- surname (varchar): driver's last name.
- dob (date): date of birth of the driver.
- nationality (varcahar): nationality of the driver.
- url (varchar): web link for more details about the driver.

D. Races:

- raceId (Integer): Unique identifier for each race
- year (Integer): Year when the race was held.
- Round (Integer): The round number of the race in the season.
- circuitId (Integer): ID of the circuit where the race was held.
- Name (Varchar): Name of the race.
- Date (date): Date when the race took place.
- Time (time): Time when the race started.
- url (Varchar): Web link for more details about the race.

E. Seasons:

- Year (Integer): Year representing the season.
- url (Varchar): Web link for more details about the season.

F. Sprint_results:

- resultsId (Integer): Unique identifier for each sprint result.
- raceId (Integer): ID of the race associated with the sprint.
- driverId (Integer): ID of the driver in the sprint.
- constructorId (Integer): ID of the constructor associated with the sprint.
- Number (Integer): Racing number of the driver.
- Grid (Integer): Starting position of the driver in the sprint.
- Position (Integer): Finishing position of the driver.
- PositionText (Varchar): Text description of the finishing position.
- Points (Decimal): Points earned by the driver.
- Laps (Integer): Number of laps completed.
- Time (Varchar): Total time taken by the driver.
- Milliseconds (Integer): Total time taken in milliseconds.
- fastestLap(Integer): Fastest lap achieved by the driver.
- fastestLapTime (Time): Time of the fastest lap.
- Statusid (integer): ID representing the driver's status in the sprint.

G. Results

- resultId (Integer): Unique identifier for each race result.
- raceId (Integer): ID of the race associated with the result.
- driverId (Integer): ID of the driver in the race.
- constructorId (Integer): ID of the constructor associated with the result
- Number (Integer): Racing number of the driver.
- Grid (Integer): Starting position of the driver.
- Position (Integer): Finishing position of the driver.
- positionText (Varchar): Text description of the finishing position.
- Points (Decimal): Points earned by the driver.
- Laps (Integer): Number of laps completed.
- Time (Varchar): Total race time.
- Milliseconds (Integer): Total race time in milliseconds.
- fastestLap (Integer): Fastest Lap achieved by the driver.
- Rank (Integer): Rank of the fastest lap.
- fastestLapTime (Time): Time of the fastest lap.
- fastestLapSpeed (Decimal): Speed achieved during the fastest lap.
- statusId (Integer): Unique identifier for each status.

H. Status:

- statusId (Integer): Unique identifier for each status.
- Status (Varchar): Description of the driver's status (e.g., finished, retired).

I. Constructor_standings:

- constructorStandingsId (Integer): Unique identifier for each constructor standing.
- raceId (Integer): ID of the race associated with the constructor standing.
- constructorId (Integer): ID of the constructor.
- Points (Decimal): Points earned by the constructor.
- Position (Integer): Constructor's position in the standings.
- Wins (Integer): Number of wins by the constructor.

J. Driver_Standings:

- driverStandingsId (Integer): Unique identifier for each driver standing.
- raceId (Integer): ID of the race associated with the driver standing.
- driverId (Integer): ID of the driver.
- Points (Decimal): Points earned by the driver.
- Position (Integer): Driver's position in the standings.
- positionText (Varchar): Text description of the position.
- Wins (Integer): Number of wins by the driver.

K. Constructor_results:

- constructorResultId (Integer): Unique identifier for each constructor result.
- raceId (Integer): ID of the constructor.
- constructorId (Integer): ID of the race associated with the constructor result.
- Points (Decimal): Points earned by the constructor.
- Status (Varchar): Status description of the constructor in the race.

L. Lap_times

- raceId (Integer): ID of the race associated with the lap time.
- driverId (Integer): ID of the driver recording the lap time.
- Lap (Integer): The lap number.
- Position (Integer): Driver's position during the lap.
- Time (Time): Total time taken to complete the lap.
- Milliseconds (Integer): Total time taken in milliseconds.

M. Pit_stops:

- raceId (Integer): ID of the race associated with the pit stop.
- driverId (Integer): ID of the driver making the pit stop.
- Stop (Integer): The stop number (i.e., the nth stop during the race).
- Lap (Integer): The lap during which the pit stop occurred.
- Time (Time): Time when the pit stop was made.
- Duration (Time): Total duration of the pit stop.
- Milliseconds (Integer): Duration of the pit stop in milliseconds.

N. Qualifying:

- qualifyId (Integer): Unique identifier for each qualifying session.
- raceId (Integer): ID of the race associated with the qualifying session.
- driverId (Integer): ID of the driver in the qualifying session.
- constructorId (Integer): ID of the constructor associated with the driver.
- Position (Integer): The position achieved by the driver.
- Q1 (Time): Time of the driver's first qualifying session.
- Q2 (Time): Time of the driver's second qualifying session.
- Q3 (Time): Time of the driver's third qualifying session.

VI. SQL QUERIES

We have executed several queries that showcases different clauses such as INSERT, SELECT, DELETE, UPDATE, GROUP BY, HAVING, TRIGGERS, JOINS, etc.

```
Query Query History
1 INSERT INTO results (resultId, raceId, driverId, constructorId, points, position)
2 VALUES (30001, 1020, 45, 12, 18, 1);
3
```

Data Output Messages Explain X Notifications

Successfully run. Total query runtime: 298 msec.
1 rows affected.

Fig1: SQL query that adds a new result to the results table with specified resultId, raceId, driverId, constructorId, points, and position.

```
Query Query History
1 INSERT INTO results (resultId, raceId, driverId, constructorId, points, position)
2 SELECT 30002, 1021, 46, 13, 15, 2
3 WHERE NOT EXISTS (
4 SELECT 1 FROM results WHERE raceId = 1021 AND driverId = 46
5 );
6
```

Data Output Messages Explain X Notifications

Graphical Analysis Statistics

The diagram shows a 'Result' table (represented by a grid icon) with an arrow pointing to an 'Insert' operation (represented by a grid icon with a plus sign). This operation then points to the 'results' table (represented by a grid icon with a downward arrow).

Fig2: SQL query that Inserts a new result only if the combination of raceId and driverId does not already exist in the results table.

```
Query Query History
1 DELETE FROM results WHERE raceId = 1020;
2
```

Data Output Messages Explain X Notifications

Graphical Analysis Statistics

The diagram shows the 'results' table (represented by a grid icon with a downward arrow) with an arrow pointing to a 'Delete' operation (represented by a grid icon with a minus sign).

Fig3: SQL query Deletes all records from the results table for a particular raceId.

```
Query Query History
1 UPDATE results
2 SET points = CASE
3 WHEN position = 1 THEN 25
4 WHEN position = 2 THEN 18
5 ELSE points
6 END
7 WHERE raceId = 1020;
8
```

Data Output Messages Explain X Notifications

Graphical Analysis Statistics

The diagram shows the 'results' table (represented by a grid icon with a downward arrow) with an arrow pointing to an 'Update' operation (represented by a grid icon with a pencil).

Fig4: SQL query that updates the points for results based on the driver's finishing position (e.g., 25 points for 1st place).

Query Query History

```

1 CREATE OR REPLACE FUNCTION update_driver_standings()
2 RETURNS TRIGGER AS $$
3 BEGIN
4     UPDATE driver_standings
5     SET points = points + NEW.points
6     WHERE driverId = NEW.driverId AND raceId = NEW.raceId;
7     RETURN NEW;
8 END;
9 $$ LANGUAGE plpgsql;
10
11 CREATE TRIGGER after_results_insert_
12 AFTER INSERT ON results
13 FOR EACH ROW
14 EXECUTE FUNCTION update_driver_standings();
15

```

Data Output Messages Notifications

CREATE TRIGGER

Query returned successfully in 170 msec.

Total rows: 1 of 1 Query complete 00:00:00.170 Ln 11, Col 37

Fig5: SQL Query that automatically updates the driver_standings table with new points after a new result is inserted into the results table.

Query Query History

```

1 SELECT d.driverId, d.forename, d.surname, SUM(r.points) AS total_points
2 FROM drivers d
3 JOIN results r ON d.driverId = r.driverId
4 JOIN races ra ON r.raceId = ra.raceId
5 WHERE ra.year = 2023
6 GROUP BY d.driverId, d.forename, d.surname
7 ORDER BY total_points DESC;
8

```

Data Output Messages Explain X Notifications

Graphical Analysis Statistics

Total rows: 1 of 1 Query complete 00:00:00.109 Ln 8, Col 1

Fig6: SQL Query that retrieves total points scored by each driver in a specific season, ordered by total points.

Query Query History

```

1 SELECT raceId, constructorId, SUM(points) AS total_points
2 FROM results
3 GROUP BY raceId, constructorId
4 HAVING SUM(points) > 50;
5

```

Data Output Messages Explain X Notifications

Graphical Analysis Statistics

Total rows: 1 of 1 Query complete 00:00:00.092 Ln 4, Col 25

Fig7: SQL query that finds races where constructors have earned more than 50 points by summing points for each constructor in each race.

Query Query History

```

1 SELECT driverId, SUM(points) AS total_points,
2 RANK() OVER (ORDER BY SUM(points) DESC) AS rank
3 FROM results
4 JOIN races ON results.raceId = races.raceId
5 WHERE races.year = 2023
6 GROUP BY driverId;
7

```

Data Output Messages Explain X Notifications

Graphical Analysis Statistics

Total rows: 1 of 1 Query complete 00:00:00.096 Ln 7, Col 1

Fig8: SQL query that ranks drivers by their total points in a season using the rank() window function.

Query Query History

```

1 SELECT r.raceId, d.forename, d.surname, c.name AS constructor, r.points
2 FROM results r
3 JOIN drivers d ON r.driverId = d.driverId
4 JOIN constructors c ON r.constructorId = c.constructorId
5 ORDER BY r.points DESC;
6

```

Data Output Messages Explain X Notifications

Graphical Analysis Statistics

Total rows: 1 of 1 Query complete 00:00:00.099 Ln 3, Col 23

Fig9: SQL query that retrieves race results with driver and constructor names, ordered by points in descending order.

VII. PROBLEMS AND SOLUTIONS

In the raw dataset, the dob (date of birth) column was formatted as dd/mm/yyyy. Since PostgreSQL accepts dates in the yyyy-mm-dd format, the date format was converted to this standard during preprocessing. The following Python code was used to read the CSV file, transform the dob column to the desired format using pandas, and save the updated dataset:

```
4 CODE < TEXT
1 | SELECT pandas_df.dob
2 |
3 | df = df.read_csv('dataset/dob.csv')
4 |
5 | df['dob'] = pd.to_datetime(df['dob'], format='%d/%m/%Y', errors='coerce').dt.strftime('%Y-%m-%d')
6 |
7 | df.to_csv('dataset/dob.csv', index=False)
8 |
9 | print(df)
10 |
11 | driverId  driverName  number  number  nationality  dob
12 | 1  Hamilton  44  44  British  1985-01-07
13 | 2  Vettel  5  5  German  1980-01-08
14 | 3  Alonso  14  14  Spanish  1981-01-27
15 | 4  Sainz  55  55  Spanish  1994-05-05
16 | 5  Verstappen  33  33  Dutch  1997-03-05
17 | 6  Norris  4  4  Irish  1993-02-08
18 | 7  Pirelli  7  7  Italian  1982-02-05
19 | 8  Ricciardo  18  18  Australian  1988-05-24
20 | 9  Leclerc  16  16  Monaco  1998-10-16
21 | 10  Gasly  10  10  French  1996-02-05
22 | 11  Magnussen  20  20  Danish  1985-02-01
23 | 12  Ocon  31  31  French  1997-07-01
24 | 13  Russell  63  63  British  1998-02-03
25 | 14  Stroll  14  14  Canadian  1994-01-09
26 | 15  Alonso  14  14  Spanish  1981-01-27
27 | 16  Sainz  55  55  Spanish  1994-05-05
28 | 17  Verstappen  33  33  Dutch  1997-03-05
29 | 18  Norris  4  4  Irish  1993-02-08
30 | 19  Ricciardo  18  18  Australian  1988-05-24
31 | 20  Leclerc  16  16  Monaco  1998-10-16
32 | 21  Gasly  10  10  French  1996-02-05
33 | 22  Magnussen  20  20  Danish  1985-02-01
34 | 23  Ocon  31  31  French  1997-07-01
35 | 24  Russell  63  63  British  1998-02-03
36 | 25  Stroll  14  14  Canadian  1994-01-09
37 | 26  Alonso  14  14  Spanish  1981-01-27
38 | 27  Sainz  55  55  Spanish  1994-05-05
39 | 28  Verstappen  33  33  Dutch  1997-03-05
40 | 29  Norris  4  4  Irish  1993-02-08
41 | 30  Ricciardo  18  18  Australian  1988-05-24
42 | 31  Leclerc  16  16  Monaco  1998-10-16
43 | 32  Gasly  10  10  French  1996-02-05
44 | 33  Magnussen  20  20  Danish  1985-02-01
45 | 34  Ocon  31  31  French  1997-07-01
46 | 35  Russell  63  63  British  1998-02-03
47 | 36  Stroll  14  14  Canadian  1994-01-09
48 | 37  Alonso  14  14  Spanish  1981-01-27
49 | 38  Sainz  55  55  Spanish  1994-05-05
50 | 39  Verstappen  33  33  Dutch  1997-03-05
51 | 40  Norris  4  4  Irish  1993-02-08
52 | 41  Ricciardo  18  18  Australian  1988-05-24
53 | 42  Leclerc  16  16  Monaco  1998-10-16
54 | 43  Gasly  10  10  French  1996-02-05
55 | 44  Magnussen  20  20  Danish  1985-02-01
56 | 45  Ocon  31  31  French  1997-07-01
57 | 46  Russell  63  63  British  1998-02-03
58 | 47  Stroll  14  14  Canadian  1994-01-09
59 | 48  Alonso  14  14  Spanish  1981-01-27
60 | 49  Sainz  55  55  Spanish  1994-05-05
61 | 50  Verstappen  33  33  Dutch  1997-03-05
62 | 51  Norris  4  4  Irish  1993-02-08
63 | 52  Ricciardo  18  18  Australian  1988-05-24
64 | 53  Leclerc  16  16  Monaco  1998-10-16
65 | 54  Gasly  10  10  French  1996-02-05
66 | 55  Magnussen  20  20  Danish  1985-02-01
67 | 56  Ocon  31  31  French  1997-07-01
68 | 57  Russell  63  63  British  1998-02-03
69 | 58  Stroll  14  14  Canadian  1994-01-09
70 | 59  Alonso  14  14  Spanish  1981-01-27
71 | 60  Sainz  55  55  Spanish  1994-05-05
72 | 61  Verstappen  33  33  Dutch  1997-03-05
73 | 62  Norris  4  4  Irish  1993-02-08
74 | 63  Ricciardo  18  18  Australian  1988-05-24
75 | 64  Leclerc  16  16  Monaco  1998-10-16
76 | 65  Gasly  10  10  French  1996-02-05
77 | 66  Magnussen  20  20  Danish  1985-02-01
78 | 67  Ocon  31  31  French  1997-07-01
79 | 68  Russell  63  63  British  1998-02-03
80 | 69  Stroll  14  14  Canadian  1994-01-09
81 | 70  Alonso  14  14  Spanish  1981-01-27
82 | 71  Sainz  55  55  Spanish  1994-05-05
83 | 72  Verstappen  33  33  Dutch  1997-03-05
84 | 73  Norris  4  4  Irish  1993-02-08
85 | 74  Ricciardo  18  18  Australian  1988-05-24
86 | 75  Leclerc  16  16  Monaco  1998-10-16
87 | 76  Gasly  10  10  French  1996-02-05
88 | 77  Magnussen  20  20  Danish  1985-02-01
89 | 78  Ocon  31  31  French  1997-07-01
90 | 79  Russell  63  63  British  1998-02-03
91 | 80  Stroll  14  14  Canadian  1994-01-09
92 | 81  Alonso  14  14  Spanish  1981-01-27
93 | 82  Sainz  55  55  Spanish  1994-05-05
94 | 83  Verstappen  33  33  Dutch  1997-03-05
95 | 84  Norris  4  4  Irish  1993-02-08
96 | 85  Ricciardo  18  18  Australian  1988-05-24
97 | 86  Leclerc  16  16  Monaco  1998-10-16
98 | 87  Gasly  10  10  French  1996-02-05
99 | 88  Magnussen  20  20  Danish  1985-02-01
100 | 89  Ocon  31  31  French  1997-07-01
101 | 90  Russell  63  63  British  1998-02-03
102 | 91  Stroll  14  14  Canadian  1994-01-09
103 | 92  Alonso  14  14  Spanish  1981-01-27
104 | 93  Sainz  55  55  Spanish  1994-05-05
105 | 94  Verstappen  33  33  Dutch  1997-03-05
106 | 95  Norris  4  4  Irish  1993-02-08
107 | 96  Ricciardo  18  18  Australian  1988-05-24
108 | 97  Leclerc  16  16  Monaco  1998-10-16
109 | 98  Gasly  10  10  French  1996-02-05
110 | 99  Magnussen  20  20  Danish  1985-02-01
111 | 100  Ocon  31  31  French  1997-07-01
112 | 101  Russell  63  63  British  1998-02-03
113 | 102  Stroll  14  14  Canadian  1994-01-09
114 | 103  Alonso  14  14  Spanish  1981-01-27
115 | 104  Sainz  55  55  Spanish  1994-05-05
116 | 105  Verstappen  33  33  Dutch  1997-03-05
117 | 106  Norris  4  4  Irish  1993-02-08
118 | 107  Ricciardo  18  18  Australian  1988-05-24
119 | 108  Leclerc  16  16  Monaco  1998-10-16
120 | 109  Gasly  10  10  French  1996-02-05
121 | 110  Magnussen  20  20  Danish  1985-02-01
122 | 111  Ocon  31  31  French  1997-07-01
123 | 112  Russell  63  63  British  1998-02-03
124 | 113  Stroll  14  14  Canadian  1994-01-09
125 | 114  Alonso  14  14  Spanish  1981-01-27
126 | 115  Sainz  55  55  Spanish  1994-05-05
127 | 116  Verstappen  33  33  Dutch  1997-03-05
128 | 117  Norris  4  4  Irish  1993-02-08
129 | 118  Ricciardo  18  18  Australian  1988-05-24
130 | 119  Leclerc  16  16  Monaco  1998-10-16
131 | 120  Gasly  10  10  French  1996-02-05
132 | 121  Magnussen  20  20  Danish  1985-02-01
133 | 122  Ocon  31  31  French  1997-07-01
134 | 123  Russell  63  63  British  1998-02-03
135 | 124  Stroll  14  14  Canadian  1994-01-09
136 | 125  Alonso  14  14  Spanish  1981-01-27
137 | 126  Sainz  55  55  Spanish  1994-05-05
138 | 127  Verstappen  33  33  Dutch  1997-03-05
139 | 128  Norris  4  4  Irish  1993-02-08
140 | 129  Ricciardo  18  18  Australian  1988-05-24
141 | 130  Leclerc  16  16  Monaco  1998-10-16
142 | 131  Gasly  10  10  French  1996-02-05
143 | 132  Magnussen  20  20  Danish  1985-02-01
144 | 133  Ocon  31  31  French  1997-07-01
145 | 134  Russell  63  63  British  1998-02-03
146 | 135  Stroll  14  14  Canadian  1994-01-09
147 | 136  Alonso  14  14  Spanish  1981-01-27
148 | 137  Sainz  55  55  Spanish  1994-05-05
149 | 138  Verstappen  33  33  Dutch  1997-03-05
150 | 139  Norris  4  4  Irish  1993-02-08
151 | 140  Ricciardo  18  18  Australian  1988-05-24
152 | 141  Leclerc  16  16  Monaco  1998-10-16
153 | 142  Gasly  10  10  French  1996-02-05
154 | 143  Magnussen  20  20  Danish  1985-02-01
155 | 144  Ocon  31  31  French  1997-07-01
156 | 145  Russell  63  63  British  1998-02-03
157 | 146  Stroll  14  14  Canadian  1994-01-09
158 | 147  Alonso  14  14  Spanish  1981-01-27
159 | 148  Sainz  55  55  Spanish  1994-05-05
160 | 149  Verstappen  33  33  Dutch  1997-03-05
161 | 150  Norris  4  4  Irish  1993-02-08
162 | 151  Ricciardo  18  18  Australian  1988-05-24
163 | 152  Leclerc  16  16  Monaco  1998-10-16
164 | 153  Gasly  10  10  French  1996-02-05
165 | 154  Magnussen  20  20  Danish  1985-02-01
166 | 155  Ocon  31  31  French  1997-07-01
167 | 156  Russell  63  63  British  1998-02-03
168 | 157  Stroll  14  14  Canadian  1994-01-09
169 | 158  Alonso  14  14  Spanish  1981-01-27
170 | 159  Sainz  55  55  Spanish  1994-05-05
171 | 160  Verstappen  33  33  Dutch  1997-03-05
172 | 161  Norris  4  4  Irish  1993-02-08
173 | 162  Ricciardo  18  18  Australian  1988-05-24
174 | 163  Leclerc  16  16  Monaco  1998-10-16
175 | 164  Gasly  10  10  French  1996-02-05
176 | 165  Magnussen  20  20  Danish  1985-02-01
177 | 166  Ocon  31  31  French  1997-07-01
178 | 167  Russell  63  63  British  1998-02-03
179 | 168  Stroll  14  14  Canadian  1994-01-09
180 | 169  Alonso  14  14  Spanish  1981-01-27
181 | 170  Sainz  55  55  Spanish  1994-05-05
182 | 171  Verstappen  33  33  Dutch  1997-03-05
183 | 172  Norris  4  4  Irish  1993-02-08
184 | 173  Ricciardo  18  18  Australian  1988-05-24
185 | 174  Leclerc  16  16  Monaco  1998-10-16
186 | 175  Gasly  10  10  French  1996-02-05
187 | 176  Magnussen  20  20  Danish  1985-02-01
188 | 177  Ocon  31  31  French  1997-07-01
189 | 178  Russell  63  63  British  1998-02-03
190 | 179  Stroll  14  14  Canadian  1994-01-09
191 | 180  Alonso  14  14  Spanish  1981-01-27
192 | 181  Sainz  55  55  Spanish  1994-05-05
193 | 182  Verstappen  33  33  Dutch  1997-03-05
194 | 183  Norris  4  4  Irish  1993-02-08
195 | 184  Ricciardo  18  18  Australian  1988-05-24
196 | 185  Leclerc  16  16  Monaco  1998-10-16
197 | 186  Gasly  10  10  French  1996-02-05
198 | 187  Magnussen  20  20  Danish  1985-02-01
199 | 188  Ocon  31  31  French  1997-07-01
200 | 189  Russell  63  63  British  1998-02-03
201 | 190  Stroll  14  14  Canadian  1994-01-09
202 | 191  Alonso  14  14  Spanish  1981-01-27
203 | 192  Sainz  55  55  Spanish  1994-05-05
204 | 193  Verstappen  33  33  Dutch  1997-03-05
205 | 194  Norris  4  4  Irish  1993-02-08
206 | 195  Ricciardo  18  18  Australian  1988-05-24
207 | 196  Leclerc  16  16  Monaco  1998-10-16
208 | 197  Gasly  10  10  French  1996-02-05
209 | 198  Magnussen  20  20  Danish  1985-02-01
210 | 199  Ocon  31  31  French  1997-07-01
211 | 200  Russell  63  63  British  1998-02-03
212 | 201  Stroll  14  14  Canadian  1994-01-09
213 | 202  Alonso  14  14  Spanish  1981-01-27
214 | 203  Sainz  55  55  Spanish  1994-05-05
215 | 204  Verstappen  33  33  Dutch  1997-03-05
216 | 205  Norris  4  4  Irish  1993-02-08
217 | 206  Ricciardo  18  18  Australian  1988-05-24
218 | 207  Leclerc  16  16  Monaco  1998-10-16
219 | 208  Gasly  10  10  French  1996-02-05
220 | 209  Magnussen  20  20  Danish  1985-02-01
221 | 210  Ocon  31  31  French  1997-07-01
222 | 211  Russell  63  63  British  1998-02-03
223 | 212  Stroll  14  14  Canadian  1994-01-09
224 | 213  Alonso  14  14  Spanish  1981-01-27
225 | 214  Sainz  55  55  Spanish  1994-05-05
226 | 215  Verstappen  33  33  Dutch  1997-03-05
227 | 216  Norris  4  4  Irish  1993-02-08
228 | 217  Ricciardo  18  18  Australian  1988-05-24
229 | 218  Leclerc  16  16  Monaco  1998-10-16
230 | 219  Gasly  10  10  French  1996-02-05
231 | 220  Magnussen  20  20  Danish  1985-02-01
232 | 221  Ocon  31  31  French  1997-07-01
233 | 222  Russell  63  63  British  1998-02-03
234 | 223  Stroll  14  14  Canadian  1994-01-09
235 | 224  Alonso  14  14  Spanish  1981-01-27
236 | 225  Sainz  55  55  Spanish  1994-05-05
237 | 226  Verstappen  33  33  Dutch  1997-03-05
238 | 227  Norris  4  4  Irish  1993-02-08
239 | 228  Ricciardo  18  18  Australian  1988-05-24
240 | 229  Leclerc  16  16  Monaco  1998-10-16
241 | 230  Gasly  10  10  French  1996-02-05
242 | 231  Magnussen  20  20  Danish  1985-02-01
243 | 232  Ocon  31  31  French  1997-07-01
244 | 233  Russell  63  63  British  1998-02-03
245 | 234  Stroll  14  14  Canadian  1994-01-09
246 | 235  Alonso  14  14  Spanish  1981-01-27
247 | 236  Sainz  55  55  Spanish  1994-05-05
248 | 237  Verstappen  33  33  Dutch  1997-03-05
249 | 238  Norris  4  4  Irish  1993-02-08
250 | 239  Ricciardo  18  18  Australian  1988-05-24
251 | 240  Leclerc  16  16  Monaco  1998-10-16
252 | 241  Gasly  10  10  French  1996-02-05
253 | 242  Magnussen  20  20  Danish  1985-02-01
254 | 243  Ocon  31  31  French  1997-07-01
255 | 244  Russell  63  63  British  1998-02-03
256 | 245  Stroll  14  14  Canadian  1994-01-09
257 | 246  Alonso  14  14  Spanish  1981-01-27
258 | 247  Sainz  55  55  Spanish  1994-05-05
259 | 248  Verstappen  33  33  Dutch  1997-03-05
260 | 249  Norris  4  4  Irish  1993-02-08
261 | 250  Ricciardo  18  18  Australian  1988-05-24
262 | 251  Leclerc  16  16  Monaco  1998-10-16
263 | 252  Gasly  10  10  French  1996-02-05
264 | 253  Magnussen  20  20  Danish  1985-02-01
265 | 254  Ocon  31  31  French  1997-07-01
266 | 255  Russell  63  63  British  1998-02-03
267 | 256  Stroll  14  14  Canadian  1994-01-09
268 | 257  Alonso  14  14  Spanish  1981-01-27
269 | 258  Sainz  55  55  Spanish  1994-05-05
270 | 259  Verstappen  33  33  Dutch  1997-03-05
271 | 260  Norris  4  4  Irish  1993-02-08
272 | 261  Ricciardo  18  18  Australian  1988-05-24
273 | 262  Leclerc  16  16  Monaco  1998-10-16
274 | 263  Gasly  10  10  French  1996-02-05
275 | 264  Magnussen  20  20  Danish  1985-02-01
276 | 265  Ocon  31  31  French  1997-07-01
277 | 266  Russell  63  63  British  1998-02-03
278 | 267  Stroll  14  14  Canadian  1994-01-09
279 | 268  Alonso  14  14  Spanish  1981-01-27
280 | 269  Sainz  55  55  Spanish  1994-05-05
281 | 270  Verstappen  33  33  Dutch  1997-03-05
282 | 271  Norris  4  4  Irish  1993-02-08
283 | 272  Ricciardo  18  18  Australian  1988-05-24
284 | 273  Leclerc  16  16  Monaco  1998-10-16
285 | 274  Gasly  10  10  French  1996-02-05
286 | 275  Magnussen  20  20  Danish  1985-02-01
287 | 276  Ocon  31  31  French  1997-07-01
288 | 277  Russell  63  63  British  1998-02-03
289 | 278  Stroll  14  14  Canadian  1994-01-09
290 | 279  Alonso  14  14  Spanish  1981-01-27
291 | 280  Sainz  55  55  Spanish  1994-05-05
292 | 281  Verstappen  33  33  Dutch  1997-03-05
293 | 282  Norris  4  4  Irish  1993-02-08
294 | 283  Ricciardo  18  18  Australian  1988-05-24
295 | 284  Leclerc  16  16  Monaco  1998-10-16
296 | 285  Gasly  10  10  French  1996-02-05
297 | 286  Magnussen  20  20  Danish  1985-02-01
298 | 287  Ocon  31  31  French  1997-07-01
299 | 288  Russell  63  63  British  1998-02-03
300 | 289  Stroll  14  14  Canadian  1994-01-09
301 | 290  Alonso  14  14  Spanish  1981-01-27
302 | 291  Sainz  55  55  Spanish  1994-05-05
303 | 292  Verstappen  33  33  Dutch  1997-03-05
304 | 293  Norris  4  4  Irish  1993-02-08
305 | 294  Ricciardo  18  18  Australian  1988-05-24
306 | 295  Leclerc  16  16  Monaco  1998-10-16
307 | 296  Gasly  10  10  French  1996-02-05
308 | 297  Magnussen  20  20  Danish  1985-02-01
309 | 298  Ocon  31  31  French  1997-07-01
310 | 299  Russell  63  63  British  1998-02-03
311 | 300  Stroll  14  14  Canadian  1994-01-09
312 | 301  Alonso  14  14  Spanish  1981-01-27
313 | 302  Sainz  55  55  Spanish  1994-05-05
314 | 303  Verstappen  33  33  Dutch  1997-03-05
315 | 304  Norris  4  4  Irish  1993-02-08
316 | 305  Ricciardo  18  18  Australian  1988-05-24
317 | 306  Leclerc  16  16  Monaco  1998-10-16
318 | 307  Gasly  10  10  French  1996-02-05
319 | 308  Magnussen  20  20  Danish  1985-02-01
320 | 309  Ocon  31  31  French  1997-07-01
321 | 310  Russell  63  63  British  1998-02-03
322 | 311  Stroll  14  14  Canadian  1994-01-09
323 | 312  Alonso  14  14  Spanish  1981-01-27
324 | 313  Sainz  55  55  Spanish  1994-05-05
325 | 314  Verstappen  33  33  Dutch  1997-03-05
326 | 315  Norris  4  4  Irish  1993-02-08
327 | 316  Ricciardo  18  18  Australian  1988-05-24
328 | 317  Leclerc  16  16  Monaco  1998-10-16
329 | 318  Gasly  10  10  French  1996-02-05
330 | 319  Magnussen  20  20  Danish  1985-02-01
331 | 320  Ocon  31  31  French  1997-07-01
332 | 321  Russell  63  63  British  1998-02-03
333 | 322  Stroll  14  14  Canadian  1994-01-09
334 | 323  Alonso  14  14  Spanish  1981-01-27
335 | 324  Sainz  55  55  Spanish  1994-05-05
336 | 325  Verstappen  33  33  Dutch  1997-03-05
337 | 326  Norris  4  4  Irish  1993-02-08
338 | 327  Ricciardo  18  18  Australian  1988-05-24
339 | 328  Leclerc  16  16  Monaco  1998-10-16
340 | 329  Gasly  10  10  French  1996-02-05
341 | 330  Magnussen  20  20  Danish  1985-02-01
342 | 331  Ocon  31  31  French  1997-07-01
343 | 332  Russell  63  63  British  1998-02-03
344 | 333  Stroll  14  14  Canadian  1994-01-09
345 | 334  Alonso  14  14  Spanish  1981-01-27
346 | 335  Sainz  55  55  Spanish  1994-05-05
347 | 336  Verstappen  33  33  Dutch  1997-03-05
348 | 337  Norris  4  4  Irish  1993-02-08
349 | 338  Ricciardo  18  18  Australian  1988-05-24
350 | 339  Leclerc  16  16  Monaco  1998-10-16
351 | 340  Gasly  10  10  French  1996-02-05
352 | 341  Magnussen  20  20  Danish  1985-02-01
353 | 342  Ocon  31  31  French  1997-07-01
354 | 343  Russell  63  63  British  1998-02-03
355 | 344  Stroll  14  14  Canadian  1994-01-09
356 | 345  Alonso  14  14  Spanish  1981-01-27
357 | 346  Sainz  55  55  Spanish  1994-05-05
358 | 347  Verstappen  33  33  Dutch  1997-03-05
359 | 348  Norris  4  4  Irish  1993-02-08
360 | 349  Ricciardo  18  18  Australian  1988-05-24
361 | 350  Leclerc  16  16  Monaco  1998-10-16
362 | 351  Gasly  10  10  French  1996-02-05
363 | 352  Magnussen  20  20  Danish  1985-02-01
364 | 353  Ocon  31  31  French  1997-07-01
365 | 354  Russell  63  63  British  1998-02-03
366 | 355  Stroll  14  14  Canadian  1994-01-09
367 | 356  Alonso  14  14  Spanish  1981-01-27
368 | 357  Sainz  55  55  Spanish  1994-05-05
369 | 358  Verstappen  33  33  Dutch  1997-03-05
370 | 359  Norris  4  4  Irish  1993-02-08
371 | 360  Ricciardo  18  18  Australian  1988-05-24
372 | 361  Leclerc  16  16  Monaco  1998-10-16
373 | 362  Gasly  10  10  French  1996-02-05
374 | 363  Magnussen  20  20  Danish  1985-02-01
375 | 364  Ocon  31  31  French  1997-07-01
376 | 365  Russell  63  
```


2. The below subquery runs for every row in results, causing repetitive execution. This is inefficient, especially with large datasets like ours:

Query	Query History
1	EXPLAIN ANALYZE
2	SELECT * FROM results
3	WHERE constructorId IN (
4	SELECT constructorId FROM constructor_details WHERE nationality = 'Italian'
5);
6	

Data Output	Messages	Notifications
QUERY PLAN		
text		
1	Nested Loop (cost=5.26..257.33 rows=125 width=87) (actual time=0.092..0.092 rows=0 loops=1)	
2	-> Seq Scan on constructor_details (cost=0.00..4.65 rows=1 width=4) (actual time=0.084..0.084 rows=0 loops=1)	
3	Filter: ((nationality)::text = 'Italian')::text	
4	Rows Removed by Filter: 212	
5	-> Bitmap Heap Scan on results (cost=5.26..251.42 rows=126 width=87) (never executed)	
6	Recheck Cond: (constructorid = constructor_details.constructorid)	
7	-> Bitmap Index Scan on idx_results_constructorid (cost=0.00..5.23 rows=126 width=0) (never executed)	
8	Index Cond: (constructorid = constructor_details.constructorid)	
9	Planning Time: 0.885 ms	
10	Execution Time: 5.922 ms	

Fig: SQL query that retrieves all rows from the results table where the constructorId is associated with Italian constructors, using a subquery.

To overcome this, we have created an index on constructorId and nationality in the constructor table and have also replaced the IN clause with a JOIN for better performance. By doing this, we have decreased the execution time from 5.922 ms to 0.076 ms.

Query	Query History
1	CREATE INDEX idx_constructor_details_nationality ON constructor_details(nationality);
2	CREATE INDEX idx_results_constructorId ON results(constructorId);
3	EXPLAIN ANALYZE
4	SELECT r.*
5	FROM results r
6	JOIN constructor_details cd ON r.constructorId = cd.constructorId
7	WHERE cd.nationality = 'Italian';
8	

Data Output	Messages	Notifications
QUERY PLAN		
text		
1	Nested Loop (cost=5.26..257.33 rows=125 width=87) (actual time=0.047..0.047 rows=0 loops=1)	
2	-> Seq Scan on constructor_details cd (cost=0.00..4.65 rows=1 width=4) (actual time=0.046..0.046 rows=0 loops=1)	
3	Filter: ((nationality)::text = 'Italian')::text	
4	Rows Removed by Filter: 212	
5	-> Bitmap Heap Scan on results r (cost=5.26..251.42 rows=126 width=87) (never executed)	
6	Recheck Cond: (constructorid = cd.constructorid)	
7	-> Bitmap Index Scan on idx_results_constructorid (cost=0.00..5.23 rows=126 width=0) (never executed)	
8	Index Cond: (constructorid = cd.constructorid)	
9	Planning Time: 3.380 ms	
10	Execution Time: 0.076 ms	

3. For the query shown below, window functions like rank() are computationally expensive when used over large datasets. Moreover, aggregation (Sum(points)) combined with ranking requires scanning and sorting the dataset.

Query	Query History
1	EXPLAIN ANALYZE
2	SELECT driverId, RANK() OVER (ORDER BY SUM(points) DESC) AS rank
3	FROM results
4	GROUP BY driverId;
5	

Data Output	Messages	Explain X	Notifications
QUERY PLAN			
text			
1	WindowAgg (cost=780.95..795.97 rows=859 width=20) (actual time=8.046..8.421 rows=859 loops=1)		
2	-> Sort (cost=780.94..783.08 rows=859 width=12) (actual time=8.034..8.064 rows=859 loops=1)		
3	Sort Key: (sum(points)) DESC		
4	Sort Method: quicksort Memory: 51kB		
5	-> HashAggregate (cost=730.48..739.07 rows=859 width=12) (actual time=7.772..7.893 rows=859 loops=1)		
6	Group Key: driverid		
7	Batches: 1 Memory Usage: 105kB		
8	-> Seq Scan on results (cost=0.00..597.99 rows=26499 width=12) (actual time=0.023..1.623 rows=26499 loops=1)		
9	Planning Time: 0.144 ms		
10	Execution Time: 8.515 ms		

Total rows: 10 of 10 Query complete 00:00:00.094 Ln 5, Col 1

Fig: SQL query that calculates the rank of each driver based on the total points, using a window function to rank them in descending order of points.

So we have used a materialized view to precompute Sum(points), thereby reducing the execution time from 8.515 ms to 1.314 ms.

Query	Query History
1	CREATE MATERIALIZED VIEW driver_points AS
2	SELECT driverId, SUM(points) AS total_points
3	FROM results
4	GROUP BY driverId;
5	EXPLAIN ANALYZE
6	SELECT driverId, RANK() OVER (ORDER BY total_points DESC) AS rank
7	FROM driver_points;
8	

Data Output	Messages	Notifications
QUERY PLAN		
text		
1	WindowAgg (cost=142.56..178.24 rows=2040 width=20) (actual time=0.496..1.233 rows=859 loops=1)	
2	-> Sort (cost=142.54..147.64 rows=2040 width=12) (actual time=0.488..0.560 rows=859 loops=1)	
3	Sort Key: total_points DESC	
4	Sort Method: quicksort Memory: 51kB	
5	-> Seq Scan on driver_points (cost=0.00..30.40 rows=2040 width=12) (actual time=0.027..0.317 rows=859 loops=1)	
6	Planning Time: 0.672 ms	
7	Execution Time: 1.314 ms	

Total rows: 7 of 7 Query complete 00:00:00.197 Ln 5, Col 16

IX. WEB PORTAL IMPLEMENTATION.

We have created the Formula 1 Race Result Predictor, a machine learning-powered application designed to predict whether a Formula 1 driver will achieve a podium finish (top 3 positions) in a race. The app allows users to input race-related data such as driver experience, constructor experience, grid position, laps completed, and fastest lap speed. Based on these inputs, we predict the likelihood of a podium or non-podium finish. For single predictions, users can interact with sliders to manually input race details, while for batch predictions, the app enables users to upload a CSV file containing multiple race scenarios. It then processes the file, predicts outcomes for all rows, and provides a downloadable CSV with the results.

We have powered the app with a pre-trained Random Forest machine learning model, which has been trained on historical Formula 1 data to identify patterns and factors influencing race outcomes. The app features a user-friendly interface built with Streamlit, making it accessible to both Formula 1 analysts and enthusiasts. Whether used to simulate race strategies, analyze driver performance, or explore the practical application of machine learning, this app offers a seamless way to gain insights into race predictions.

The screenshot shows the 'Input Race Data' section of the web portal. It features five horizontal sliders for inputting race details: Driver Experience (No. of Races) with a value of 50, Constructor Experience (No. of Races) with a value of 50, Grid Position with a value of 10, Number of Laps Completed with a value of 50, and Fastest Lap Speed (km/h) with a value of 259. Below these sliders is a 'Batch Prediction' section with a text input for uploading a CSV file and a 'Browse files' button. A 'Predict' button is located to the right of the input sliders.

Formula 1 Race Result Predictor

This app predicts whether a driver will finish on the podium based on various factors like driver experience, constructor performance, and grid position.

Moreover, our model achieves an accuracy of 89.83%, demonstrating its reliability in predicting race results.

The screenshot displays a Jupyter Notebook with Python code for training a Random Forest classifier and evaluating its performance. The code includes imports for sklearn.metrics, sklearn.model_selection, and sklearn.preprocessing. The evaluation results are shown in a table format:

	precision	recall	f1-score	support
0	0.92	0.96	0.94	1548
1	0.74	0.57	0.64	216

Additional metrics shown include accuracy (0.898779653696773), macro avg (0.83), and weighted avg (0.89).

X. REFERENCES

- [1] [Online]. Available: <https://www.kaggle.com/datasets/rohanrao/formula-1-world-championship-1950-2020>.
- [2] [Online]. Available: <https://www.geeksforgeeks.org/sql-queries-on-clustered-and-non-clustered-indexes/>.