An Analysis of Formula 1 Data Set between the years 1950 & 2017

INFO 8076 – SQL & Data Analytics

Group – 1

Group Members:

Hima Babu

Leelaksh Chadha

Sumedha Galgali

Kanganpreet Kaur

Chukwunonso Okoh

Parth Savani

Jaydeep Sharma

Content:

Introduction:

Formula 1 is a high-speed, globally popular motorsport that involves a series of races (Grands Prix) held on circuits worldwide. Analyzing F1 data can provide valuable insights into various aspects of the sport, including driver performance, team strategies, race outcomes, and historical trends.

Formula 1 racing generates vast amounts of data, covering several aspects of the sport, such as race results, driver statistics, team performances, and details about circuits. Analyzing this data can help teams make informed decisions and to help fans understand the dynamics of the races. From assessing the dominance of legendary drivers to predicting race outcomes based on historical trends, F1 data analysis provides a thrilling journey into the heart of motorsport.

Problem Statement & Objectives:

Formula 1 (F1), being a data-rich sport, presents an opportunity for comprehensive analysis to extract meaningful insights. The problem statement is to conduct a thorough F1 data analysis, exploring various aspects such as driver performance, team strategies, circuit influences, and historical trends. Our main goal is to find the top performing drivers in specific time periods as our dataset consists of F1 information from 1950 – 2017.

Some of the key objectives for our analysis are:

1. Analysis of Driver Performance:
   - Evaluate the performance of drivers over the years.
   - Identify top – performing drivers based on race wins, race participations and points earned.

2. Team Performance & Success:
   - Examine Team Performance in terms of total points earned by constructors.

3. Demographic Insights:
   - Explore demographic insights such as the nationality distribution of drivers and teams.
   - Analyze the correlation between driver nationality and success.

4. Quantification of Dominance:
   - Calculate the extent to which certain drivers or teams dominated a given era.
   - Make use of statistical analyses to identify dominance patterns.

Dataset Description:

The Formula 1 Dataset used in this analysis encompasses an extensive collection of information spanning from 1950 to 2017. This dataset consists of key information and statistics like detailed race information, results of each race, driver information along with their constructors. Below is the detailed explanation of each table in out Formula 1 Database.

1. Drivers:
   - Consists of columns like driver id, driver reference, driver number, driver code, driver forename, driver surname, driver's DOB, and their nationality.
   - Primary Key: Driver ID.
   - Consists detailed information about F1 drivers like Lewis Hamilton, Michael Schumacher and many more.

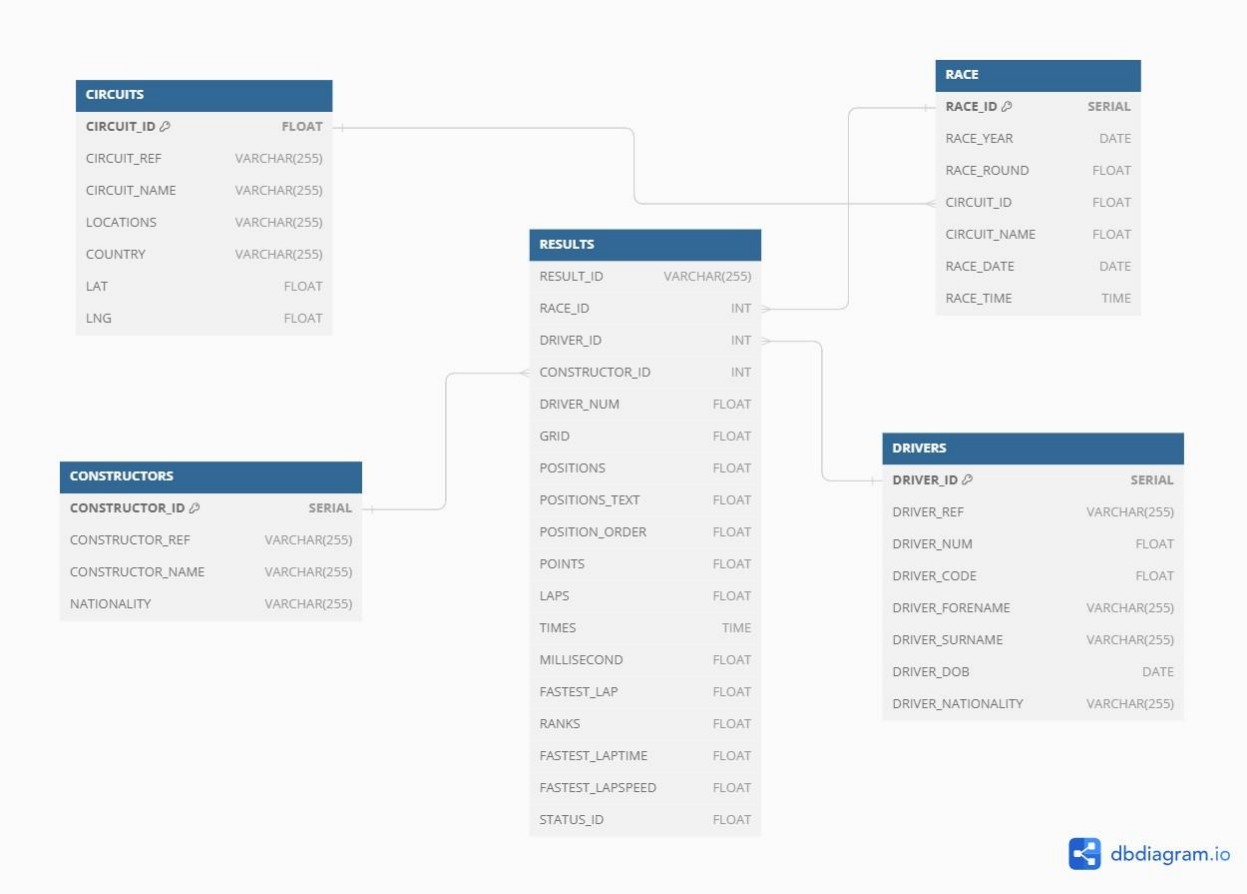
2. Races:
   - Consists of columns like race id, race year, race round, circuit id, circuit name, race date and race time.
   - Primary Key: Race ID
   - Foreign Key: Circuit ID
   - Offers details on F1 races, including the year, round, associated circuit, date, and time.

3. Circuits:
   - Consists of columns like circuit id, circuit name, country, locations, lat, Lng.
   - Primary Key: Circuit ID
   - Contains famous circuits information like Monaco, Silverstone, and Monza.

4. Constructors:
   - Consists of columns like constructor id, constructor ref, constructor name and their nationality.
   - Primary Key: Constructor ID
   - Contains notable constructors' information with the likes of Ferrari, Mercedes, Red Bull, and McLaren.

5. Results:
   - Consists of columns like result id, race id, driver id, constructor id, driver number, grid, positions, positions order, laps, times, millisecond record, fastest lap, rank, fastest lap time, fastest lap speed and status id.
   - Foreign Key: Race ID, Constructor ID, and Driver ID
   - Contains comprehensive results data for F1 races, with information about drivers, constructors, positions, points, laps, and more.

Entity Relationship Diagram (ERD):



Above is the Entity Relationship Diagram (ERD) for our Formula 1 Data Set. It is a visual representation of our tables (entities) in our database and the relationship between them. We have used ERD as they are commonly used in database design to provide a clear and concise overview of the database structure.

- Entities in our ERD are the tables that we have created in SQL i.e., Circuits, Races, Constructors, Drivers, and Results.
- In our database, the Results Table is our fact table as it stores information of various dimension tables like race, circuits, constructors, and drivers. One more reason that result table is considered fact table is that the Result Table is used to store quantitative information and the fact tables are used to store descriptive information.
- Relationships would describe how drivers are associated with races, how constructors are associated with results, and so on. This all is possible with the help of primary and foreign keys in our tables.

Data Analysis with Insights:

Question 1: Find the winners of each race.

SQL Query:

Query | Query History

```sql
SELECT
    R.Race_id,
    R.Driver_id,
    D.driver_forename || ' ' || D.driver_surname AS Driver_Name,
    MAX(R.Points) AS Winning_Points
FROM
    Results R
JOIN
    Drivers D ON R.Driver_id = D.driver_id
GROUP BY
    R.Race_id, R.Driver_id, D.driver_forename, D.driver_surname
ORDER BY
    Winning_Points DESC;
```

SQL Output:

	race_id integer	winner_name text	winning_points double precision
1	918	Lewis Hamilton	50
2	987	Sebastian Vettel	25
3	986	Max Verstappen	25
4	985	Lewis Hamilton	25
5	984	Lewis Hamilton	25
6	983	Max Verstappen	25
7	982	Lewis Hamilton	25
8	981	Lewis Hamilton	25
9	980	Lewis Hamilton	25
10	979	Sebastian Vettel	25
11	978	Lewis Hamilton	25
12	977	Valtteri Bottas	25
13	976	Daniel Ricciardo	25
14	975	Lewis Hamilton	25
15	974	Sebastian Vettel	25
16	973	Lewis Hamilton	25
17	972	Valtteri Bottas	25
18	971	Sebastian Vettel	25
19	337	Fernando Alonso	25
20	338	Jenson Button	25

Insights:

We want to view the winners of each race that has occurred from 1950 – 2017. As the total count of the races is very large, we have only provided a snippet of our output. To get the desired result we first joined the results and the drivers table using join query and to view the player's full name we used string concatenation. It's observable that Lewis Hamilton has won most of the races during his tenure, which is currently ongoing. It is noticeable that Lewis Hamilton won 50 points for Race 918 which is an anomaly in F1. This may be a dataset error, or the error had occurred in the actual race. Apart from Lewis Hamilton other noticeable names are Sebastian Vettel, Michael Schumacher, Max Verstappen, and Daniel Riccardo.

Question 2: Get the total points scored by each driver in a specific year.

SQL Query

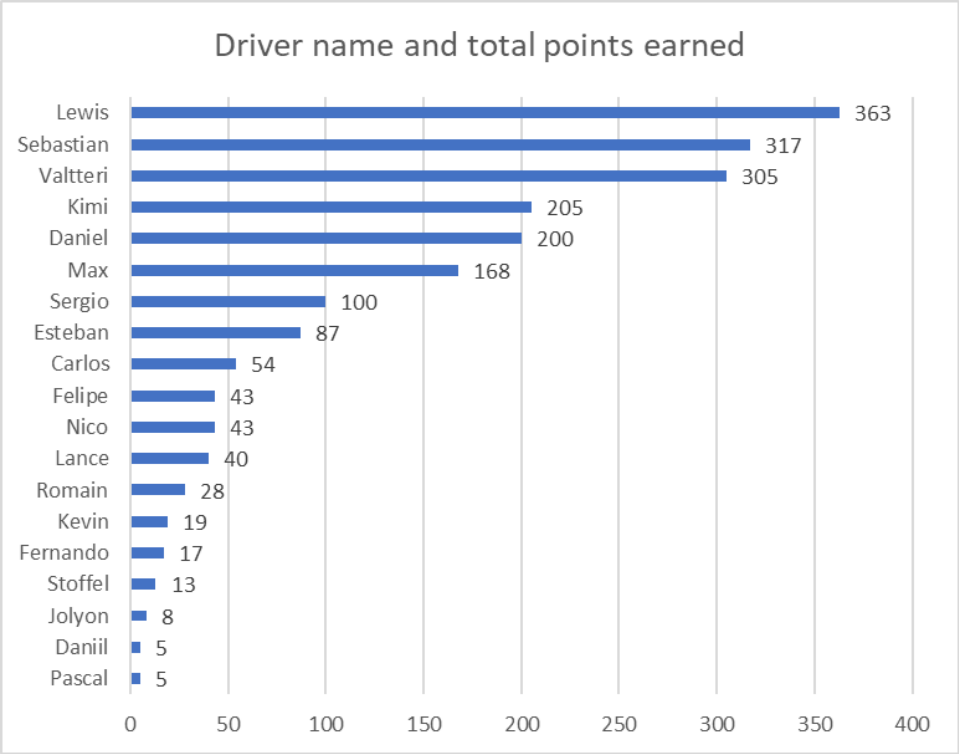
Query | Query History

```
SELECT Drivers.Driver_id, Drivers.Driver_forename, Drivers.Driver_surname, SUM(Results.Points) AS Total_Points
FROM Results
JOIN Drivers ON Results.Driver_id = Drivers.Driver_id
JOIN Race ON Results.Race_id = Race.Race_id
WHERE Race.Race_year = '2017'
GROUP BY Drivers.Driver_id, Drivers.Driver_forename, Drivers.Driver_surname
ORDER BY total_points DESC
```

SQL Output:

	driver_id integer	driver_name text	total_points double precision
1	1	Lewis Hamilton	363
2	20	Sebastian Vettel	317
3	822	Valtteri Bottas	305
4	8	Kimi raikkonen	205
5	817	Daniel Ricciardo	200
6	830	Max Verstappen	168
7	815	Sergio Ṗrez	100
8	839	Esteban Ocon	87
9	832	Carlos Sainz	54
10	13	Felipe Massa	43
11	807	Nico H_lkenberg	43
12	840	Lance Stroll	40
13	154	Romain Grosjean	28
14	825	Kevin Magnussen	19
15	4	Fernando Alonso	17
16	838	Stoffel Vandoorne	13
17	835	Jolyon Palmer	8
18	826	Daniil Kvyat	5
19	836	Pascal Wehrlein	5

Visualization:



Insights:

We wanted to analyze the most recent year in our dataset which was the year 2017. To analyze the F1 Calendar Year 2017 we first joined the drivers and the races table to get the total points accumulated by each player in 2017. Lewis Hamilton leads with an impressive 363 points in this data summary, closely followed by Sebastian Vettel and Valtteri Bottas. This indicates that Lewis Hamilton was the most consistent player in the year 2017 which also resulted in Mercedes winning the Constructors Championship that year. The drivers' observed significant point variance suggests notable differences in their on-track performances, offering essential insights into the Formula 1 competitive environment.

Question 3: Find the TOP 10 driver with the most race wins.

SQL Query

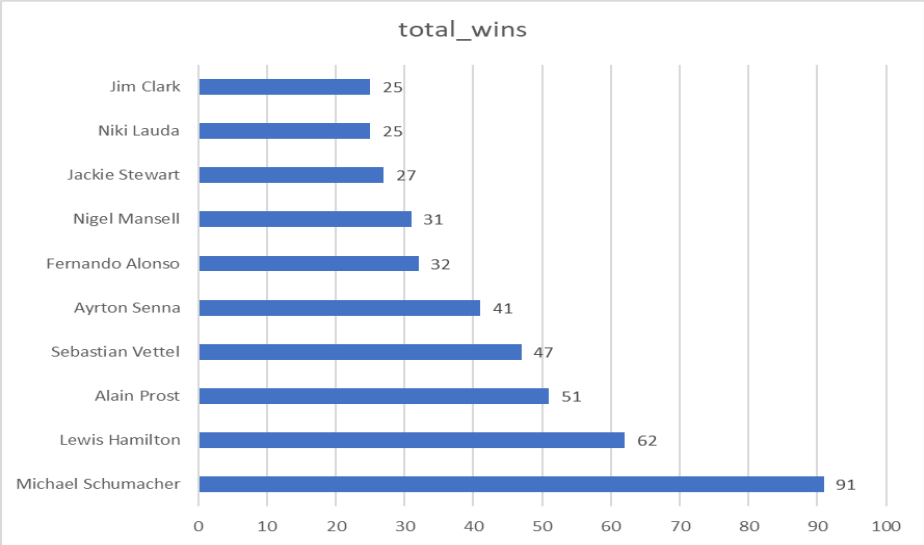
Query | Query History

```sql
SELECT
    D.driver_id,
    D.driver_forename || ' ' || D.driver_surname AS Driver_Name,
    COUNT(*) AS Total_Wins
FROM
    Results R
JOIN
    Drivers D ON R.Driver_id = D.driver_id
WHERE
    R.Points = (SELECT MAX(Points) FROM Results WHERE Race_id = R.Race_id)
GROUP BY
    D.driver_id
ORDER BY
    Total_Wins DESC
LIMIT 10;
```

SQL Output:

	driver_id [PK] integer	driver_name text	total_wins bigint
1	30	Michael Schumacher	91
2	1	Lewis Hamilton	62
3	117	Alain Prost	51
4	20	Sebastian Vettel	47
5	102	Ayrton Senna	41
6	4	Fernando Alonso	32
7	95	Nigel Mansell	31
8	328	Jackie Stewart	27
9	182	Niki Lauda	25
10	373	Jim Clark	25

Visualization:



Insights:

The top 10 drivers in this data summary are ordered by the number of races they have won, and Michael Schumacher leads the pack with an astounding 91 wins, highlighting his legendary status in Formula 1. Other notable drivers like Lewis Hamilton, Alain Prost, and Sebastian Vettel are also mentioned on the list, which broadens the pool of accomplished individuals in the motorsports industry. It is important to note that we only have data from 1950 to 2017 and in the current year of 2023, Lewis Hamilton has overtaken Michael Schumacher in the greatest number of wins in F1 History. In summary, this query aims to identify the top drivers who have the highest total wins based on the points they scored in individual races, providing insights into the driver's success in terms of race victories.

Question 4: Drivers with participation in most races

SQL Query

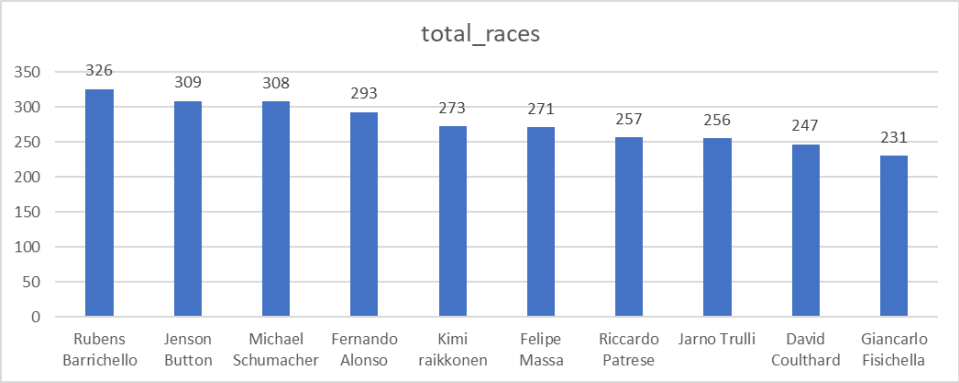
Query | Query History

```
SELECT
    D.driver_id,
    D.driver_forename || ' ' || D.driver_surname AS Driver_Name,
    COUNT(*) AS Total_Races
FROM
    Results R
JOIN
    Drivers D ON R.Driver_id = D.driver_id
GROUP BY
    D.driver_id
ORDER BY
    Total_Races DESC
LIMIT 10;
```

SQL Output

	driver_id [PK] integer	driver_name text	total_race_participations bigint
1	22	Rubens Barrichello	326
2	18	Jenson Button	309
3	30	Michael Schumacher	308
4	4	Fernando Alonso	293
5	8	Kimi raikkonen	273
6	13	Felipe Massa	271
7	119	Riccardo Patrese	257
8	15	Jarno Trulli	256
9	14	David Coulthard	247
10	21	Giancarlo Fisichella	231

Visualization:



Insights:

This data, which gives you a glimpse into drivers' racing careers, is comparable to getting a backstage pass to assess how long they have been involved in the sport and how much experience they have amassed. We have used the join query to join the results and the drivers table to get the desired output. Here we can see that players like Rubens Barrichello and Jenson Button have the most experience in terms of participation in F1 races. It's also noticeable that having the most experience doesn't mean getting the most wins as Rubens and Jenson were not in the top 10 list of players with the most wins whereas Hamilton is not the one with the most experience but still, he is second in terms of most race wins. The one name which is common in both the analysis is Michael Schumacher who has loads of experience, but he also has won the greatest number of races with his experience.

Question 5: Number of players by nationality

SQL Query:

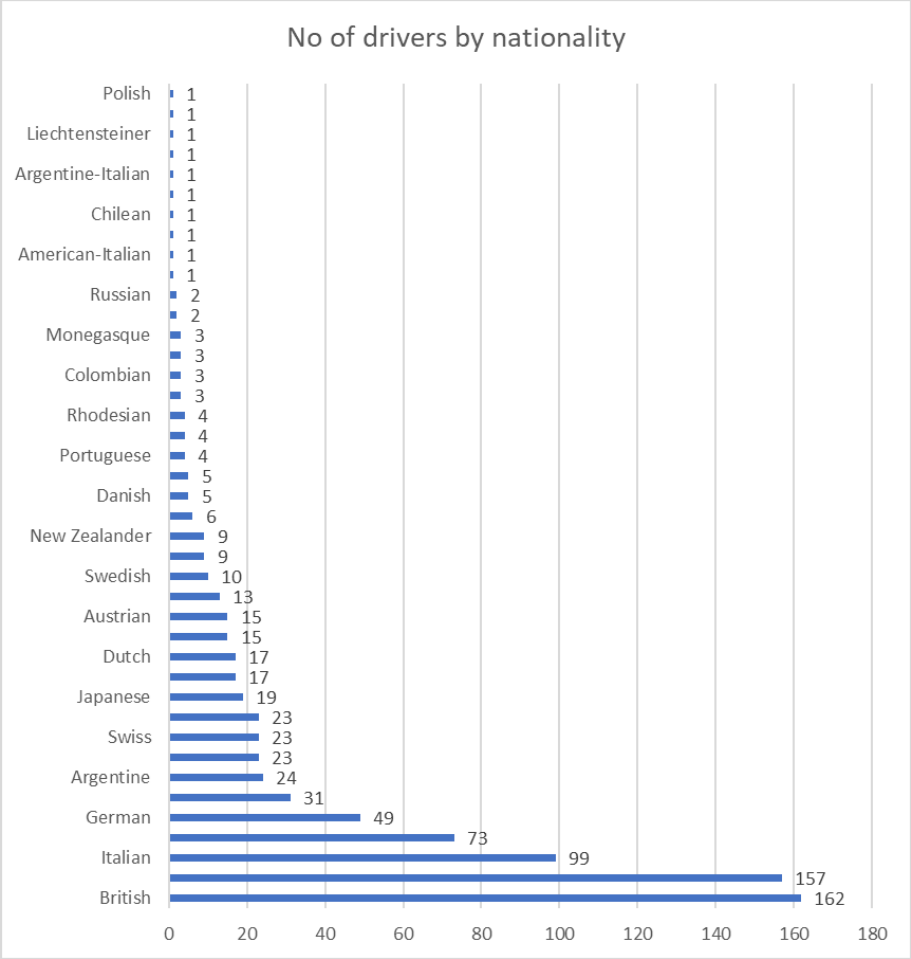
Query | Query History

```sql
SELECT
    driver_nationality,
    COUNT(driver_id) AS number_of_drivers
FROM
    drivers
GROUP BY
    driver_nationality
ORDER BY
    number_of_drivers DESC;
```

SQL Output:

	driver_nationality character varying (255)	total_players bigint
1	British	162
2	American	157
3	Italian	99
4	French	73
5	German	49
6	Brazilian	31
7	Argentine	24
8	Belgian	23
9	Swiss	23
10	South African	23
11	Japanese	19
12	Australian	17
13	Dutch	17
14	Spanish	15
15	Austrian	15

Visualization:



Insights:

Next, we move to demographic analysis where we aim to identify which country has got the most F1 racers which may directly correlate to more success in F1 Sport. Here we can see that British has the most F1 players of their nationality followed by American and Italian. Some noticeable names are Lewis Hamilton from Britain, Alberto Ascari from Italy, and Michael Schumacher from Germany.

Question 6: List all time points leaders, with birthday and nationality.

SQL Query

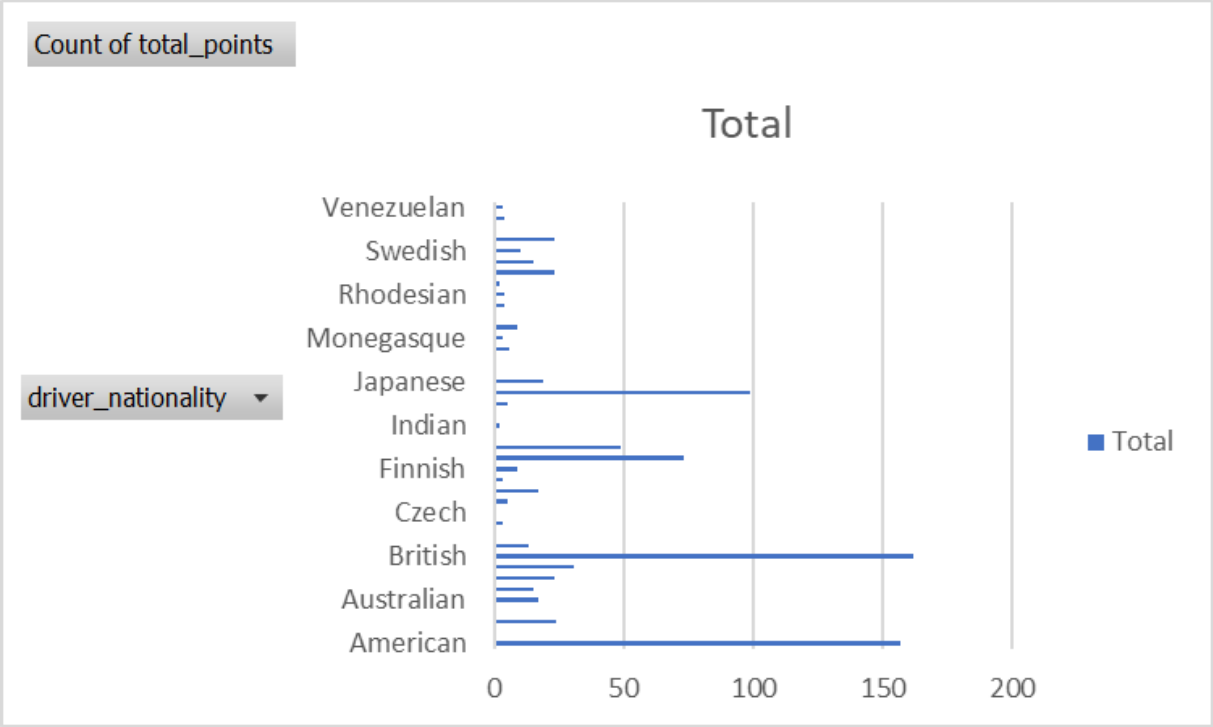
Query | Query History

```sql
SELECT
    d.driver_id,
    d.driver_forename,
    d.driver_surname,
    d.driver_dob,
    d.driver_nationality,
    SUM(r.points) AS total_points
FROM
    drivers d
JOIN
    results r ON d.driver_id = r.driver_id
GROUP BY
    d.driver_id, d.driver_forename, d.driver_surname, d.driver_dob, d.driver_nationality
ORDER BY
    total_points DESC;
```

SQL Output

	driver_id [PK] integer	driver_name text	birthday date	nationality character varying (255)	total_points double precision
1	1	Lewis Hamilton	1985-01-07	British	2610
2	20	Sebastian Vettel	1987-07-03	German	2425
3	4	Fernando Alonso	1981-07-29	Spanish	1849
4	3	Nico Rosberg	1985-06-27	German	1594.5
5	30	Michael Schumacher	1969-01-03	German	1566
6	8	Kimi raikkonen	1979-10-17	Finnish	1565
7	18	Jenson Button	1980-01-19	British	1235
8	13	Felipe Massa	1981-04-25	Brazilian	1167
9	17	Mark Webber	1976-08-27	Australian	1047.5
10	817	Daniel Ricciardo	1989-07-01	Australian	816
11	117	Alain Prost	1955-02-24	French	798.5
12	822	Valtteri Bottas	1989-08-28	Finnish	716
13	22	Rubens Barrichello	1972-05-23	Brazilian	658
14	102	Ayrton Senna	1960-03-21	Brazilian	614
15	14	David Coulthard	1971-03-27	British	535
16	137	Nelson Piquet	1952-08-17	Brazilian	485.5
17	95	Nigel Mansell	1953-08-08	British	482
18	815	Sergio Pˌrez	1990-01-26	Mexican	467
19	830	Max Verstappen	1997-09-30	Dutch	421
20	182	Niki Lauda	1949-02-22	Austrian	420.5

Visualization:



Insights:

In the next demographic analysis, we aim to find out which countries have won the most points in F1 races. This argument will support our previous analysis where in we found out which countries have the most F1 players. The drivers from America, Japanese, and the United Kingdom have the highest total point. This might be a result of these countries' historical dominance in Formula One racing as well as the quantity of drivers they have produced. The drivers from Venezuela, Sweden, Rhodesia, Japan, India, and Sweden have the lowest total point totals. This could mean that these countries are underrepresented or have not had as much success in Formula One racing as the other countries, or because they are newer to the sport. This provides key insights on which country has more interest regarding F1 and which country does not.

Question 7: Identifying Top Performing Teams (Constructors) in F1.

SQL Query

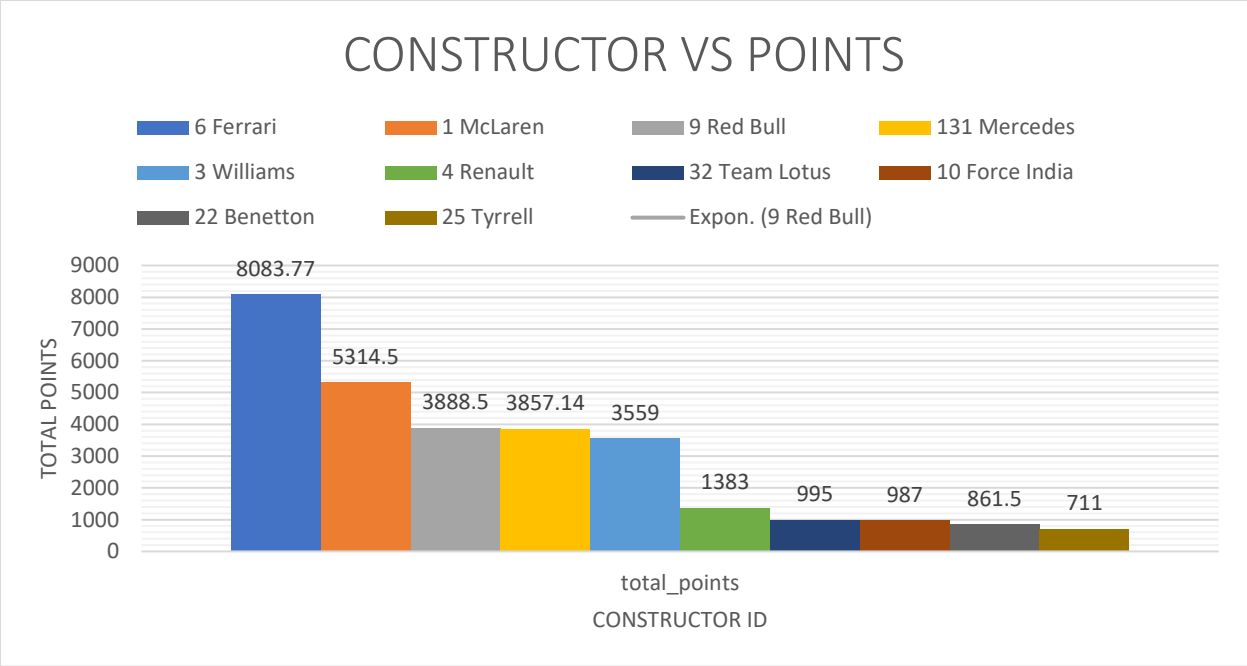
Query | Query History

```sql
SELECT
    C.constructor_id,
    C. constructor_name,
    SUM(points) AS total_points
FROM
    results R
JOIN
    constructors C ON R.constructor_id = C.constructor_id
GROUP BY
    C.constructor_id, C.constructor_name
ORDER BY
    total_points DESC
LIMIT 10;
```

SQL Output:

	constructor_id [PK] integer	constructor_name character varying (255)	total_points double precision
1	6	Ferrari	8083.77
2	1	McLaren	5314.5
3	9	Red Bull	3888.5
4	131	Mercedes	3857.14
5	3	Williams	3559
6	4	Renault	1383
7	32	Team Lotus	995
8	10	Force India	987
9	22	Benetton	861.5
10	25	Tyrrell	711

Visualization:



Insights:

Next, we move to Team Performance and how successful they are in F1. In F1, there are a lot of teams between the years 1950 – 2017 but to identify the top 10 most successful teams in terms of points won, we analyzed that by joining the results and constructor's table. We can see that; Ferrari leads the pack with the most points scored and one of the reasons for that is the presence of Michael Schumacher in the Ferrari Team during his tenure. The next teams are McLaren, Red Bull, and Mercedes. It is to be noted that post 2017, Mercedes has picked up 5 championships and hence they have overtaken the likes of Red Bull and McLaren but still behind Ferrari.

Question 8: Players with dob after 1985 who have the most wins.

SQL Query

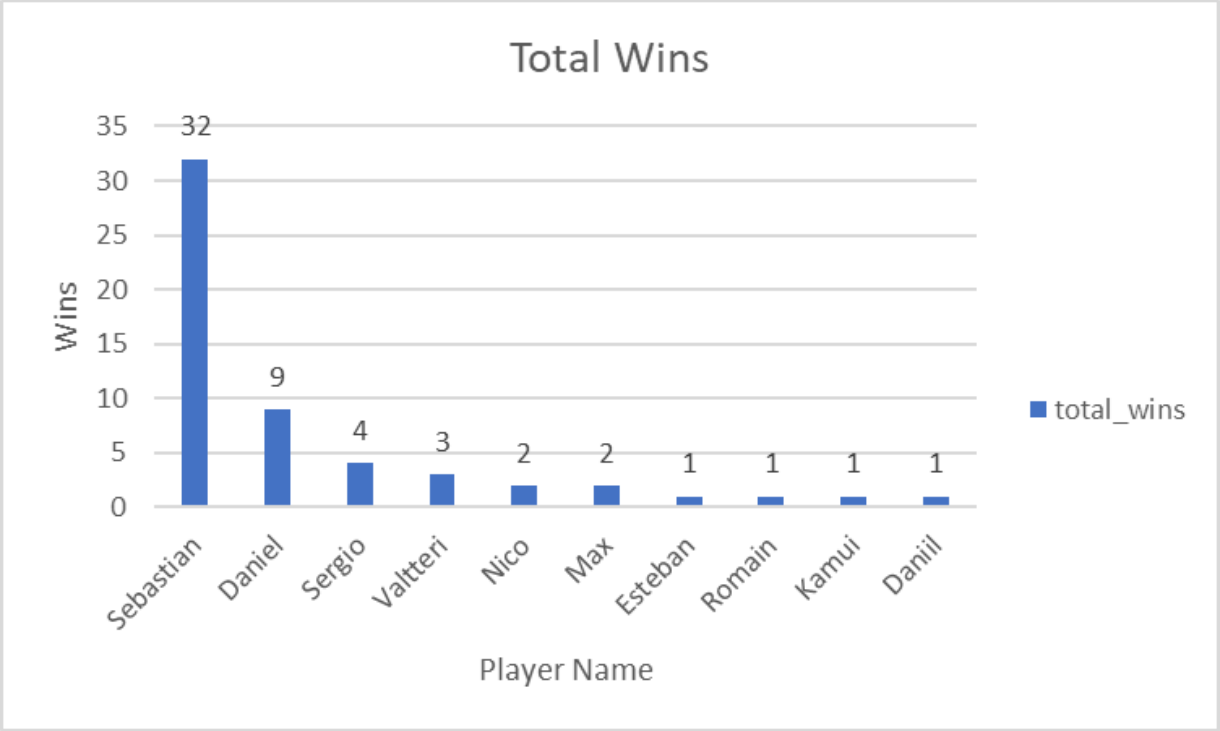
Query | Query History

```sql
SELECT
    d.driver_id,
    d.driver_forename,
    d.driver_surname,
    d.driver_dob,
    d.driver_nationality,
    COUNT(r.result_id) AS total_wins
FROM
    drivers d
JOIN
    results r ON d.driver_id = r.driver_id
JOIN
    race ra ON r.race_id = ra.race_id
WHERE
    EXTRACT(YEAR FROM d.driver_dob) > 1985
    AND r.ranks = 1
GROUP BY
    d.driver_id, d.driver_forename, d.driver_surname, d.driver_dob, d.driver_nationality
ORDER BY
    total_wins DESC;

```

SQL Output

	driver_id [PK] integer	driver_name text	birthday date	nationality character varying (255)	total_wins bigint
1	20	Sebastian Vettel	1987-07-03	German	32
2	817	Daniel Ricciardo	1989-07-01	Australian	9
3	815	Sergio Ṗrez	1990-01-26	Mexican	4
4	822	Valtteri Bottas	1989-08-28	Finnish	3
5	807	Nico H_lkenb…	1987-08-19	German	2
6	830	Max Verstappen	1997-09-30	Dutch	2
7	821	Esteban Gutịrrez	1991-08-05	Mexican	1
8	154	Romain Grosjean	1986-04-17	French	1
9	155	Kamui Kobayashi	1986-09-13	Japanese	1
10	826	Daniil Kvyat	1994-04-26	Russian	1

Visualization:



Insights:

Next, to analyze players born after 1985, we want to see which players have accomplished more in a very little time frame. This is a very interesting graph in which we can see that Sebastian Vettel as outperformed his peers of similar age drastically by winning 32 races. The next player is Daniel Riccardo with a mere 9 wins compared to that of Sebastian. Max Verstappen has only 2 wins, but it is to be noted that after 2017, he picked up and he has now almost the same wins as Sebastian Vettel.

Question 9: The top drivers with the most race win in three different time periods: 1950-1970, 1970-2000, and 2000-2017

SQL Query

Query Query History

```sql
WITH TopDriverWins AS (
    SELECT
        D.driver_id,
        D.driver_forename || ' ' || D.driver_surname AS Driver_Name,
        COUNT(*) AS Total_Wins,
        MIN(CAST(Race.race_year AS INTEGER)) AS Start_Year,
        MAX(CAST(Race.race_year AS INTEGER)) AS End_Year
    FROM
        Results R
    JOIN
        Drivers D ON R.Driver_id = D.driver_id
    JOIN
        Race Race ON R.Race_id = Race.Race_id
    WHERE
        CAST(Race.race_year AS INTEGER) BETWEEN 1950 AND 1970
        AND R.Points = (SELECT MAX(Points) FROM Results WHERE Race_id = R.Race_id)
    GROUP BY
        D.driver_id

    UNION ALL

    SELECT
        D.driver_id,
        D.driver_forename || ' ' || D.driver_surname AS Driver_Name,
        COUNT(*) AS Total_Wins,
        MIN(CAST(Race.race_year AS INTEGER)) AS Start_Year,
        MAX(CAST(Race.race_year AS INTEGER)) AS End_Year
    FROM
        Results R
    JOIN
        Drivers D ON R.Driver_id = D.driver_id
    JOIN
```

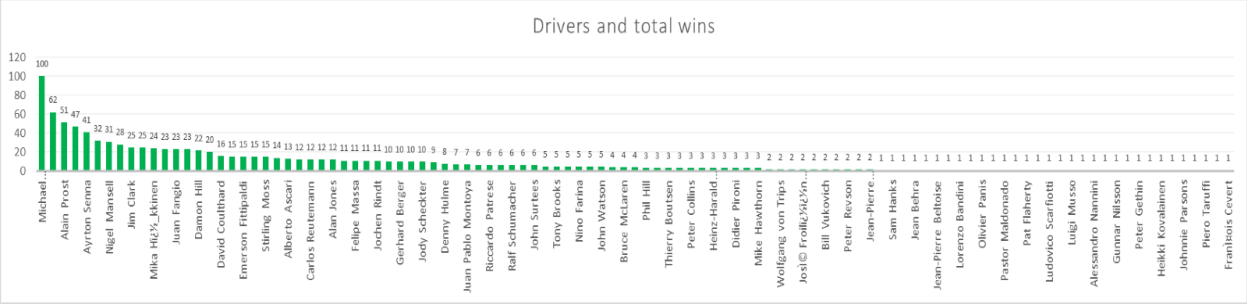
Query Query History

```sql
WITH TopDriverWins AS (
    SELECT
        D.driver_id,
        D.driver_forename || ' ' || D.driver_surname AS Driver_Name,
        COUNT(*) AS Total_Wins,
        MIN(CAST(Race.race_year AS INTEGER)) AS Start_Year,
        MAX(CAST(Race.race_year AS INTEGER)) AS End_Year
    FROM
        Results R
    JOIN
        Drivers D ON R.Driver_id = D.driver_id
    JOIN
        Race Race ON R.Race_id = Race.Race_id
    WHERE
        CAST(Race.race_year AS INTEGER) BETWEEN 1950 AND 1970
        AND R.Points = (SELECT MAX(Points) FROM Results WHERE Race_id = R.Race_id)
    GROUP BY
        D.driver_id

    UNION ALL

    SELECT
        D.driver_id,
        D.driver_forename || ' ' || D.driver_surname AS Driver_Name,
        COUNT(*) AS Total_Wins,
        MIN(CAST(Race.race_year AS INTEGER)) AS Start_Year,
        MAX(CAST(Race.race_year AS INTEGER)) AS End_Year
    FROM
        Results R
    JOIN
        Drivers D ON R.Driver_id = D.driver_id
    JOIN
    Start_Year,
    End_Year
FROM
    TopDriverWins
ORDER BY
    Total_Wins DESC;
```

SQL Output:

	driver_id integer	driver_name text	total_wins bigint	start_year integer	end_year integer
1	1	Lewis Hamilton	62	2007	2017
2	30	Michael Schumacher	56	2000	2006
3	117	Alain Prost	51	1981	1993
4	20	Sebastian Vettel	47	2008	2017
5	30	Michael Schumacher	44	1992	2000
6	102	Ayrton Senna	41	1985	1993
7	4	Fernando Alonso	32	2003	2013
8	95	Nigel Mansell	31	1985	1994
9	373	Jim Clark	25	1962	1968
10	182	Niki Lauda	25	1974	1985
11	579	Juan Fangio	23	1950	1957
12	137	Nelson Piquet	23	1980	1991
13	3	Nico Rosberg	23	2012	2016
14	71	Damon Hill	22	1993	1998
15	8	Kimi raikkonen	20	2003	2013
16	57	Mika H_kkinen	18	1997	2000
17	328	Jackie Stewart	16	1970	1973
18	475	Stirling Moss	15	1955	1961
19	18	Jenson Button	15	2006	2012
20	289	Graham Hill	14	1962	1969

Visualization:


Drivers and total wins
120
100
80
60
40
20
0
100
62
51
47
41
32
31
28
25
25
24
23
23
23
22
20
16
15
15
15
14
13
12
12
12
11
11
11
11
10
10
10
10
9
8
7
7
6
6
6
6
6
5
5
5
5
5
4
4
4
3
3
3
3
3
3
3
3
3
3
2
2
2
2
2
2
2
2
1
1
1

Insights:

The analysis identifies drivers who excelled in different decades, offering insights into the evolving landscape of Formula 1 racing. Drivers with a higher count of total wins demonstrate sustained success across multiple seasons, solidifying their status as legendary figures in the sport. By associating drivers with specific years, the analysis provides historical context, enabling a deeper understanding of Formula 1's rich tapestry. he "Total Wins" column facilitates the easy identification of the most successful drivers, offering a snapshot of their achievements and contributions to Formula 1 history. The reason there are multiple players in our output is that there were several players in that time who have performed well but we need to focus on the top 3 rows where we can see the top 3 players namely Lewis Hamilton, Michael Schumacher, and Alain Prost. They were the most successful players in the given time mentioned in the output.

Excel Pivot Tables Used for Analysis:

1. **Question 9 Pivot Table (Pivot Table was large so we have attached a snippet):**

Row Labels	Sum of total_wins
Michael Schumacher	100
Lewis Hamilton	62
Alain Prost	51
Sebastian Vettel	47
Ayrton Senna	41
Fernando Alonso	32
Nigel Mansell	31
Jackie Stewart	28
Jim Clark	25
Niki Lauda	25
Mika Hï¿½_kkinen	24
Nelson Piquet	23
Juan Fangio	23
Nico Rosberg	23
Damon Hill	22
Kimi raikkonen	20
David Coulthard	16
Jack Brabham	15
Emerson Fittipaldi	15
Jenson Button	15
Stirling Moss	15
Graham Hill	14
Alberto Ascari	13
Mario Andretti	12

2. **Question 2 Pivot Table:**

Row Labels	Sum of total_points
Pascal	5
Daniil	5
Jolyon	8
Stoffel	13
Fernando	17
Kevin	19
Romain	28
Lance	40
Nico	43
Felipe	43
Carlos	54
Esteban	87
Sergio	100
Max	168
Daniel	200
Kimi	205
Valtteri	305
Sebastian	317
Lewis	363
Grand Total	**2020**

3. **Question 6 Pivot Table (Pivot Table was large so we have attached a snippet):**

Row Labels	Count of total_points
American	157
American-Italian	1
Argentine	24
Argentine-Italian	1
Australian	17
Austrian	15
Belgian	23
Brazilian	31
British	162
Canadian	13
Chilean	1
Colombian	3
Czech	1
Danish	5
Dutch	17
East German	3
Finnish	9
French	73
German	49
Hungarian	1
Indian	2
Indonesian	1
Irish	5
Italian	99
Japanese	19
Liechtensteiner	1
Malaysian	1
Mexican	6
Monegasque	3
New Zealander	9
Polish	1
Portuguese	4

Conclusion:

To conclude this Analysis that we have done on Formula 1 Data set which contains data from Year 1950- 2017. We have the five different Datasets like Circuits, Races, Result, Constructor and Drivers which provide valuable insights and Recommendations into various aspect of Formula 1. Here are the key takeaways from the analysis:

1. Total Points Scored by driver in 2017: Lewis Hamilton is on the first place with total score of 363 points, Runner up is Sebastian Vettel with the total score of 317

2. Top 10 driver with most race win: We also want to find out top 10 racer has the most wins, and the number 1 is Michael Schumacher with 91 wins, followed by Lewis Hamilton, Alain Prost, and others

3. We also find out which nationality has the most points in the Formula 1. British take the lead in it with total of 162 points, Followed by American 157 points.

4. We also did the Analysis on which younger player (born after 1990) has the most wins: Vettel Sebastian has the most wins in the younger generation with the total of 32 wins, where other younger player is falling way behind him

In conclusion, this F1 data analysis unveiled intriguing insights into the sport's history from 1950 to 2017. By identifying top-performing drivers within distinct time periods and showcasing their consistent success, we gained a nuanced understanding of the evolving landscape of Formula 1 racing. The analysis not only celebrated legendary drivers but also highlighted the dynamic nature of competition across different eras. This comprehensive exploration provided valuable perspectives on race wins, emphasizing the enduring legacy of exceptional talents in the thrilling world of Formula 1.

Data Set Snippets:

Below are the dataset snippets for each of our 5 tables in our PostgreSQL database. We cannot display the entire dataset as our dataset consists of more than 24000 rows of data from the years 1950 to 2017.

1. Drivers Table:

	driver_id [PK] integer	driver_ref character varying (255)	driver_num double precision	driver_code character varying (255)	driver_forename character varying (255)	driver_surname character varying (255)	driver_dob date	driver_nationality character varying (255)
1	1	hamilton	44	HAM	Lewis	Hamilton	1985-01-07	British
2	2	heidfeld	[null]	HEI	Nick	Heidfeld	1977-10-05	German
3	3	rosberg	6	ROS	Nico	Rosberg	1985-06-27	German
4	4	alonso	14	ALO	Fernando	Alonso	1981-07-29	Spanish
5	5	kovalainen	[null]	KOV	Heikki	Kovalainen	1981-10-19	Finnish
6	6	nakajima	[null]	NAK	Kazuki	Nakajima	1985-01-11	Japanese
7	7	bourdais	[null]	BOU	Ṣbastien	Bourdais	1979-02-28	French
8	8	raikkonen	7	RAI	Kimi	raikkonen	1979-10-17	Finnish
9	9	kubica	[null]	KUB	Robert	Kubica	1984-12-07	Polish
10	10	glock	[null]	GLO	Timo	Glock	1982-03-18	German
11	11	sato	[null]	SAT	Takuma	Sato	1977-01-28	Japanese
12	12	piquet_jr	[null]	PIQ	Nelson	Piquet Jr.	1985-07-25	Brazilian
13	13	massa	19	MAS	Felipe	Massa	1981-04-25	Brazilian
14	14	coulthard	[null]	COU	David	Coulthard	1971-03-27	British
15	15	trulli	[null]	TRU	Jarno	Trulli	1974-07-13	Italian
16	16	sutil	99	SUT	Adrian	Sutil	1983-01-11	German
17	17	webber	[null]	WEB	Mark	Webber	1976-08-27	Australian
18	18	button	22	BUT	Jenson	Button	1980-01-19	British
19	19	davidson	[null]	DAV	Anthony	Davidson	1979-04-18	British
20	20	vettel	5	VET	Sebastian	Vettel	1987-07-03	German

2. Races Table:

	race_id [PK] integer	race_year character varying (255)	race_round double precision	circuit_id double precision	circuit_name character varying (255)	race_date date	race_time time without time zone
1	1	2009	1	1	Australian Grand Prix	[null]	06:00:00
2	2	2009	2	2	Malaysian Grand Prix	[null]	09:00:00
3	3	2009	3	17	Chinese Grand Prix	[null]	07:00:00
4	4	2009	4	3	Bahrain Grand Prix	[null]	12:00:00
5	5	2009	5	4	Spanish Grand Prix	[null]	12:00:00
6	6	2009	6	6	Monaco Grand Prix	[null]	12:00:00
7	7	2009	7	5	Turkish Grand Prix	[null]	12:00:00
8	8	2009	8	9	British Grand Prix	[null]	12:00:00
9	9	2009	9	20	German Grand Prix	[null]	12:00:00
10	10	2009	10	11	Hungarian Grand Prix	[null]	12:00:00
11	11	2009	11	12	European Grand Prix	[null]	12:00:00
12	12	2009	12	13	Belgian Grand Prix	[null]	12:00:00
13	13	2009	13	14	Italian Grand Prix	[null]	12:00:00
14	14	2009	14	15	Singapore Grand Prix	[null]	12:00:00
15	15	2009	15	22	Japanese Grand Prix	[null]	05:00:00
16	16	2009	16	18	Brazilian Grand Prix	[null]	16:00:00
17	17	2009	17	24	Abu Dhabi Grand Prix	[null]	11:00:00
18	18	2008	1	1	Australian Grand Prix	[null]	04:30:00
19	19	2008	2	2	Malaysian Grand Prix	[null]	07:00:00
20	20	2008	3	3	Bahrain Grand Prix	[null]	11:30:00

3. Circuits Table:

	circuit_id [PK] double precision	circuit_ref character varying (255)	circuit_name character varying (255)	locations character varying (255)	country character varying (255)	lat double precision	lng double precision
1	1	albert_park	Albert Park Grand Prix Circuit	Melbourne	Australia	-37.8497	144.968
2	2	sepang	Sepang International Circuit	Kuala Lumpur	Malaysia	2.76083	101.738
3	3	bahrain	Bahrain International Circuit	Sakhir	Bahrain	26.0325	50.5106
4	4	catalunya	Circuit de Barcelona-Catalunya	MontmelÃ_	Spain	41.57	2.26111
5	5	istanbul	Istanbul Park	Istanbul	Turkey	40.9517	29.405
6	6	monaco	Circuit de Monaco	Monte-Carlo	Monaco	43.7347	7.42056
7	7	villeneuve	Circuit Gilles Villeneuve	Montreal	Canada	45.5	-73.5228
8	8	magny_cours	Circuit de Nevers Magny-Cours	Magny Cours	France	46.8642	3.16361
9	9	silverstone	Silverstone Circuit	Silverstone	UK	52.0786	-1.01694
10	10	hockenheimring	Hockenheimring	Hockenheim	Germany	49.3278	8.56583
11	11	hungaroring	Hungaroring	Budapest	Hungary	47.5789	19.2486
12	12	valencia	Valencia Street Circuit	Valencia	Spain	39.4589	-0.331667
13	13	spa	Circuit de Spa-Francorchamps	Spa	Belgium	50.4372	5.97139
14	14	monza	Autodromo Nazionale di Monza	Monza	Italy	45.6156	9.28111
15	15	marina_bay	Marina Bay Street Circuit	Marina Bay	Singapore	1.2914	103.864
16	16	fuji	Fuji Speedway	Oyama	Japan	35.3717	138.927
17	17	shanghai	Shanghai International Circuit	Shanghai	China	31.3389	121.22
18	18	interlagos	AutÃ_dromo JosÃ© Carlos Pace	SÃ£o Paulo	Brazil	-23.7036	-46.6997
19	19	indianapolis	Indianapolis Motor Speedway	Indianapolis	USA	39.795	-86.2347
20	20	nurburgring	NÃ_rburgring	NÃ_rburg	Germany	50.3356	6.9475

4. Constructors Table:

	constructor_id [PK] integer	constructor_ref character varying (255)	constructor_name character varying (255)	nationality character varying (255)
1	1	mclaren	McLaren	British
2	2	bmw_sauber	BMW Sauber	German
3	3	williams	Williams	British
4	4	renault	Renault	French
5	5	toro_rosso	Toro Rosso	Italian
6	6	ferrari	Ferrari	Italian
7	7	toyota	Toyota	Japanese
8	8	super_aguri	Super Aguri	Japanese
9	9	red_bull	Red Bull	Austrian
10	10	force_india	Force India	Indian
11	11	honda	Honda	Japanese
12	12	spyker	Spyker	Dutch
13	13	mf1	MF1	Russian
14	14	spyker_mf1	Spyker MF1	Dutch
15	15	sauber	Sauber	Swiss
16	16	bar	BAR	British
17	17	jordan	Jordan	Irish
18	18	minardi	Minardi	Italian
19	19	jaguar	Jaguar	British
20	20	prost	Prost	French

5. Results Table:

	result_id character var	race_id integer	driver_id integer	constructor_i integer	driver_num double precis	grid double precis	positions double precis	positions_tex character var	position_orde integer	points double precis	laps double precis	times text	millisecond double precis	fastest_lap double precis	ranks double precis
1	1	18	1	1	22	1	1	1	1	10	58	34:50.6	5690616	39	2
2	2	18	2	2	3	5	2	2	2	8	58	5.478	5696094	41	3
3	3	18	3	3	7	7	3	3	3	6	58	8.163	5698779	41	5
4	4	18	4	4	5	11	4	4	4	5	58	17.181	5707797	58	7
5	5	18	5	1	23	3	5	5	5	4	58	18.014	5708630	43	1
6	6	18	6	3	8	13	6	6	6	3	57	[null]	[null]	50	14
7	7	18	7	5	14	17	7	7	7	2	55	[null]	[null]	22	12
8	8	18	8	6	1	15	8	8	8	1	53	[null]	[null]	20	4
9	9	18	9	2	4	2	[null]	R	9	0	47	[null]	[null]	15	9
10	10	18	10	7	12	18	[null]	R	10	0	43	[null]	[null]	23	13
11	11	18	11	8	18	19	[null]	R	11	0	32	[null]	[null]	24	15
12	12	18	12	4	6	20	[null]	R	12	0	30	[null]	[null]	20	16
13	13	18	13	6	2	4	[null]	R	13	0	29	[null]	[null]	23	6
14	14	18	14	9	9	8	[null]	R	14	0	25	[null]	[null]	21	11
15	15	18	15	7	11	6	[null]	R	15	0	19	[null]	[null]	18	10
16	16	18	16	10	20	22	[null]	R	16	0	8	[null]	[null]	8	17
17	17	18	17	9	10	14	[null]	R	17	0	0	[null]	[null]	[null]	[null]
18	18	18	18	11	16	12	[null]	R	18	0	0	[null]	[null]	[null]	[null]

Dataset Preparation:

1. CIRCUIT TABLE:

Query | Query History

```
1  CREATE TABLE CIRCUITS(
2      CIRCUIT_ID SERIAL PRIMARY KEY ,
3      CIRCUIT_REF VARCHAR(255),
4      CIRCUIT_NAME VARCHAR(255),
5      LOCATIONS VARCHAR(255),
6      COUNTRY VARCHAR(255),
7      LAT FLOAT,
8      LNG FLOAT
9  )
10
```

2. CONSTRUCTORS TABLE:

Query | Query History

```
1  CREATE TABLE CONSTRUCTORS(
2      CONSTRUCTOR_ID SERIAL PRIMARY KEY,
3      CONSTRUCTOR_REF VARCHAR(255),
4      CONSTRUCTOR_NAME VARCHAR(255),
5      NATIONALITY VARCHAR(255)
6  )
7
```

3. DRIVERS TABLE:

Query | Query History

```
1  CREATE TABLE DRIVERS(
2  DRIVER_ID SERIAL PRIMARY KEY,
3  DRIVER_REF VARCHAR(255),
4  DRIVER_NUM INT,
5  DRIVER_CODE INT,
6  DRIVER_FORENAME VARCHAR(255),
7  DRIVER_SURNAME VARCHAR(255),
8  DRIVER_DOB DATE,
9  DRIVER_NATIONALITY VARCHAR(255)
10 )
11
```

4. RACE TABLE:

Query | Query History

```
CREATE TABLE RACE(
RACE_ID SERIAL PRIMARY KEY,
RACE_YEAR DATE,
RACE_ROUND INT,
CIRCUIT_ID INT REFERENCES CIRCUITS(CIRCUIT_ID),
CIRCUIT_NAME VARCHAR(255),
RACE_DATE DATE,
RACE_TIME TIME
)
```

5. RESULTS TABLE:

Query | Query History

```
CREATE TABLE RESULTS(
    RESULT_ID INT REFERENCES RESULTS (RESULT_ID),
    RACE_ID INT REFERENCES RACE (RACE_ID),
    DRIVER_ID INT REFERENCES DRIVERS (DRIVER_ID),
    CONSTRUCTOR_ID INT REFERENCES CONSTRUCTORS (CONSTRUCTORS_ID)
    DRIVER_NUM INT,
    GRID INT,
    POSITIONS FLOAT,
    POSITIONS_TEXT VARCHAR(255),
    POSITION_ORDER INT,
    POINTS FLOAT,
    LAPS FLOAT,
    TIMES TIME,
    MILLISECOND FLOAT,
    FASTEST_LAP FLOAT,
    RANKS FLOAT,
    FASTEST_LAPTIME FLOAT,
    FASTEST_LAPSPEED FLOAT,
    STATUS_ID INT
)
```