

# Formula One Data Analysis using SQL

Dikshita Jain



# Objectives

Evaluate Driver Performance across Seasons and Constructors.

- Analyze the performance of drivers across different seasons and constructors.
- Identify the top-performing drivers for each season and constructor combination.
- Understand the factors that contribute to a driver's success within a specific constructor or season.

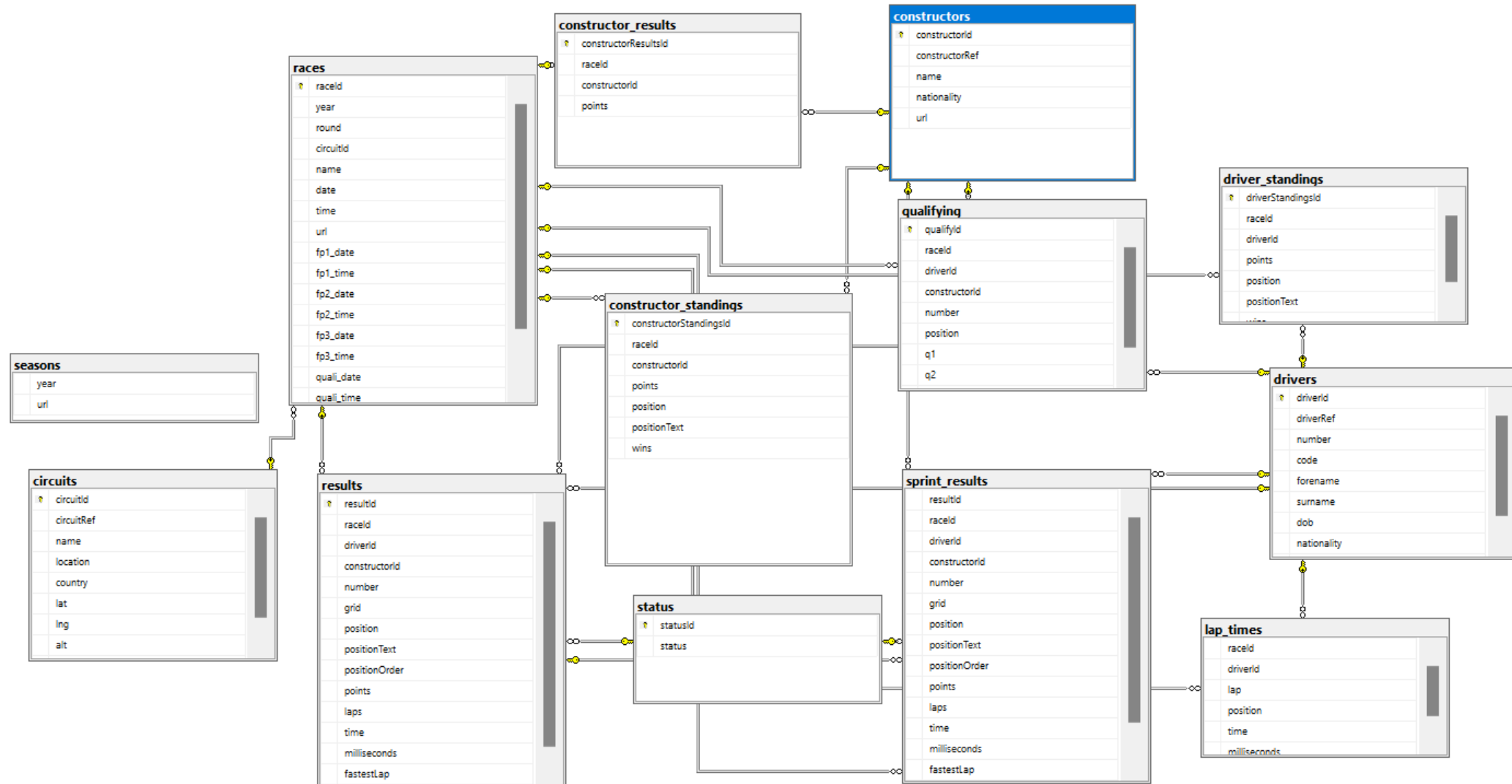
Assess Pit Stop Strategies and Their Impact on Race Results.

Evaluate Constructor Dominance and Championship Performance.

Evaluate Driver Dominance and Championship Performance.



# ERD Diagram



# Top Drivers by Season and Constructor

Problem Statement: Find the top 3 drivers by points scored for each season and constructor combination.

Output

```
with cte1 as
(Select d.driverId,year,sum(points) as total,
  rank() over (partition by year order by sum(points) desc) as rank,
  c.name as Constructor
 from drivers d
   JOIN results res on res.driverId=d.driverId
   JOIN races r on r.raceId=res.raceId
   JOIN constructors c on c.constructorId=res.constructorId
  group by d.driverId,year,c.name)
SELECT Year, CONCAT(d.forename, ' ', d.surname) AS Fullname, Rank, Constructor
FROM   cte1 AS c INNER JOIN
       drivers AS d ON d.driverId = c.driverId
WHERE  (Rank IN (1, 2, 3))
ORDER BY Year desc, Rank;
```

	Year	Fullname	Rank	Constructor
1	2023	Max Verstappen	1	Red Bull
2	2023	Sergio Pérez	2	Red Bull
3	2023	Lewis Hamilton	3	Mercedes
4	2022	Max Verstappen	1	Red Bull
5	2022	Charles Leclerc	2	Ferrari
6	2022	Sergio Pérez	2	Red Bull
7	2021	Max Verstappen	1	Red Bull
8	2021	Lewis Hamilton	2	Mercedes
9	2021	Valtteri Bottas	3	Mercedes
10	2020	Lewis Hamilton	1	Mercedes
11	2020	Valtteri Bottas	2	Mercedes
12	2020	Max Verstappen	3	Red Bull
13	2019	Lewis Hamilton	1	Mercedes
14	2019	Valtteri Bottas	2	Mercedes

## Insights:

For each season, we can view the top 3 drivers and their teams.



# Fastest Lap Times on Each Circuit

Find the fastest lap time recorded on each circuit, along with the driver's name, constructor, and race information.

```
with cte2 as
(Select c.name,driverId,constructorId,fastestLapTime,year,
rank() over (partition by c.name order by res.fastestLapTime desc) as Rank
from circuits c
JOIN races r on r.circuitId=c.circuitId
JOIN results res on res.raceId=r.raceId
where fastestLapTime is not null)
SELECT cte.name AS CircuitName, CONCAT(d.forename, ' ', d.surname) AS DriverName,
con.name AS ConstructorName, fastestLapTime AS FastestLapTime
FROM cte2 AS cte INNER JOIN
drivers AS d ON d.driverId = cte.driverId INNER JOIN
constructors AS con ON con.constructorId = cte.constructorId
WHERE (Rank = 1)
ORDER BY FastestLapTime DESC;
```

## Insights:

For each circuit, fastest lap time can be noted along with their driver name and constructor.

## Output

	CircuitName	DriverName	ConstructorName	FastestLapTime
1	Nürburgring	Vitantonio Liuzzi	Toro Rosso	03:22.3
2	Circuit de Spa-Francorchamps	Pastor Maldonado	Williams	02:56.2
3	Silverstone Circuit	Pascal Wehrlein	Manor Marussia	02:48.8
4	Suzuka Circuit	Robert Kubica	Renault	02:41.4
5	Shanghai International Circuit	Jarno Trulli	Toyota	02:39.6
6	Circuit of the Americas	Valtteri Bottas	Williams	02:31.9
7	Albert Park Grand Prix Circuit	Bruno Senna	HRT	02:27.3
8	Hockenheimring	Jarno Trulli	Lotus	02:14.5
9	Circuit de Monaco	Jenson Button	McLaren	02:14.3
10	Autodromo Internazionale del Mugello	Carlos Sainz	McLaren	02:14.1
11	Marina Bay Street Circuit	Fernando Alonso	McLaren	02:13.6
12	Bahrain International Circuit	Lucas di Grassi	Virgin	02:09.4
13	Buddh International Circuit	Timo Glock	Virgin	02:09.0
14	Sepang International Circuit	Romain Grosjean	Lotus F1	02:08.5
15	Circuit Gilles Villeneuve	Franck Montagny	Super Aguri	02:07.7
16	Fuji Speedway	Alexander Wurz	Williams	02:05.6
17	Korean International Circuit	Jarno Trulli	Lotus	02:05.2



# Pit Stop Strategy Analysis

Calculate the average time spent in the pits, the number of pit stops, and the time delta between pit stops for each driver and race.

```
With cte3 as
(Select *,
datediff(minute, lag(time) over (partition by raceId,driverID order by time), time) as delta
from pit_stops ps
where duration is not null)
SELECT r.name as CircuitName, CONCAT(d.forename, ' ', d.surname) AS FullName,
COUNT(stop) AS NoOfStops, AVG(duration) AS AvgTimeSpent, AVG(delta) AS Delta
FROM cte3 AS c INNER JOIN
drivers AS d ON d.driverId = c.driverID
INNER JOIN races r on r.raceId=c.raceId
GROUP BY CONCAT(d.forename, ' ', d.surname), r.name
order by r.name;
```

Output

	CircuitName	FullName	NoOfStops	AvgTimeSpent	Delta
1	70th Anniversary Grand Prix	Valtteri Bottas	2	28.2980003356934	29
2	70th Anniversary Grand Prix	Carlos Sainz	2	32.3179988861084	19
3	70th Anniversary Grand Prix	Daniel Ricciardo	3	28.2703329722087	17
4	70th Anniversary Grand Prix	Lando Norris	2	28.4114999771118	27
5	70th Anniversary Grand Prix	Pierre Gasly	2	28.331000328064	25
6	70th Anniversary Grand Prix	Charles Leclerc	1	29.4190006256104	NULL
7	70th Anniversary Grand Prix	Esteban Ocon	1	28.0930004119873	NULL
8	70th Anniversary Grand Prix	Daniil Kvyat	2	28.2004995346069	24
9	70th Anniversary Grand Prix	Lance Stroll	2	28.1045007705688	20
10	70th Anniversary Grand Prix	Nico Hulkenberg	3	28.701000213623	22
11	70th Anniversary Grand Prix	Romain Grosjean	2	29.5900001525879	32
12	70th Anniversary Grand Prix	Nicholas Latifi	3	28.6103331247966	29
13	70th Anniversary Grand Prix	Max Verstappen	2	29.0504999160767	9
14	70th Anniversary Grand Prix	Kevin Magnussen	2	29.1614999771118	18
15	70th Anniversary Grand Prix	Kimi Räikkönen	1	29.2959995269775	NULL

## Insights:

For each circuit and driver, number of pit stops taken by each driver and the difference between two pit stops time.

This analysis can be used to improve the Pit Stop strategy for each teams.



# Constructor Championship Dominance

Analyze the dominance of constructors in the championship over multiple seasons. Calculate the percentage of points scored compared to the total points scored by all constructors in each season.

```
with cte4 as
(Select year,c.name as Constructor,
    SUM(res.points) AS totalSeasonPts,
    ROUND(SUM(res.points) * 100.0 /
        (SELECT SUM(res.points) AS Expr1
         FROM   results AS res INNER JOIN
              races AS ra ON ra.raceId = res.raceId
              WHERE (ra.year = r.year))), 2) AS PercComparedToOthers,
    rank() over (partition by year order by sum(points) desc) as rank
from constructors c
    JOIN results res on res.constructorId=c.constructorId
    JOIN races r on r.raceId=res.raceId
group by c.name,year)
SELECT *
FROM   cte4
WHERE  (rank IN (1, 2, 3))
ORDER BY year, rank;
```

## Output

	year	Constructor	totalSeasonPts	PercComparedToOthers	rank
1	1950	Alfa Romeo	89	52.98	1
2	1950	Ferrari	21	12.5	2
3	1950	Talbot-Lago	20	11.9	3
4	1951	Ferrari	86	44.79	1
5	1951	Alfa Romeo	75	39.06	2
6	1951	Kurtis Kraft	19	9.9	3
7	1952	Ferrari	120.5	62.76	1
8	1952	Gordini	15	7.81	2
9	1952	Kurtis Kraft	13	6.77	3
10	1953	Ferrari	122.5	56.71	1
11	1953	Maserati	65.5	30.32	2
12	1953	Kurtis Kraft	24	11.11	3



# Grand Prix dominance by Drivers

```
Select r.name as GrandPrix, concat(forename, ' ', surname) as DriverName, count(*) as #OfWins from races r
JOIN circuits c on c.circuitId=r.circuitId
JOIN results res on res.raceId=r.raceId
JOIN drivers d on d.driverId=res.driverId
where position = 1
group by r.name, concat(forename, ' ', surname)
having count(*)>1
order by r.name, count(*) desc;
```

## Output

	GrandPrix	DriverName	#OfWins
1	Abu Dhabi Grand Prix	Lewis Hamilton	5
2	Abu Dhabi Grand Prix	Max Verstappen	3
3	Abu Dhabi Grand Prix	Sebastian Vettel	3
4	Argentine Grand Prix	Juan Fangio	4
5	Argentine Grand Prix	Damon Hill	2
6	Argentine Grand Prix	Emerson Fittipaldi	2
7	Australian Grand Prix	Michael Schumacher	4
8	Australian Grand Prix	Jenson Button	3
9	Australian Grand Prix	Sebastian Vettel	3
10	Australian Grand Prix	Nico Rosberg	2
11	Australian Grand Prix	Lewis Hamilton	2
12	Australian Grand Prix	Kimi Räikkönen	2
13	Australian Grand Prix	Gerhard Berger	2
14	Australian Grand Prix	Damon Hill	2
15	Australian Grand Prix	David Coulthard	2
16	Australian Grand Prix	Alain Prost	2

## Insights:

We can analyze for each circuit, how a driver performs. Some drivers tend to perform better in some circuits due to familiarity and other factors.





# Grand Prix dominance by Constructors

```
Select r.name as GrandPrix, con.name, count(*) as #OfWins from races r
JOIN circuits c on c.circuitId=r.circuitId
JOIN results res on res.raceId=r.raceId
JOIN constructors con on con.constructorId=res.constructorId
where position = 1
group by r.name, con.name
having count(*)>1
order by r.name, count(*) desc;
```

## Output

GrandPrix	name	#OfWins
Abu Dhabi Grand Prix	Mercedes	6
Abu Dhabi Grand Prix	Red Bull	6
Argentine Grand Prix	Williams	4
Argentine Grand Prix	Ferrari	4
Argentine Grand Prix	McLaren	2
Argentine Grand Prix	Team Lotus	2
Argentine Grand Prix	Maserati	2
Australian Grand Prix	McLaren	11
Australian Grand Prix	Ferrari	10
Australian Grand Prix	Williams	5
Australian Grand Prix	Mercedes	4
Australian Grand Prix	Red Bull	2
Australian Grand Prix	Renault	2

## Insights:

We can analyze for each circuit, how a constructor car performs. On the based of this analysis, teams can improve their cars to adapt to different circuit tracks.



# Constructor Dominance and their Performance

## Output

```
with SeasonConstructorWinners as
(Select year,c.name as Constructor, sum(points) as total,
  rank() over (partition by year order by sum(points) desc) as rank
from constructors c
  JOIN results res on res.constructorId=c.constructorId
  JOIN races r on r.raceId=res.raceId
group by c.name,year)
SELECT s1.Constructor, count(*) as #OfWins,
STUFF((Select ', ' + cast(year as VARCHAR)
  from SeasonConstructorWinners s2
  where s1.Constructor=s2.Constructor
  and rank=1 order by year desc FOR XML PATH('')), 1, 1, '') as YearsWon
FROM SeasonConstructorWinners s1
WHERE (rank IN (1))
group by Constructor
order by #OfWins desc,Constructor;
```

	Constructor	#OfWins	YearsWon
1	Ferrari	22	2008, 2004, 2003, 2002, 2001, 2000, 1999, 1983,...
2	McLaren	9	2007, 1998, 1991, 1990, 1989, 1988, 1985, 1984,...
3	Mercedes	9	2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014,...
4	Williams	9	1997, 1996, 1994, 1993, 1992, 1987, 1986, 1981,...
5	Red Bull	6	2023, 2022, 2013, 2012, 2011, 2010
6	BRM	3	1965, 1964, 1962
7	Brabham-...	2	1967, 1966
8	Cooper-Ci...	2	1960, 1959
9	Renault	2	2006, 2005
10	Tyrrell	2	1973, 1971
11	Alfa Romeo	1	1950
12	Benetton	1	1995
13	Brawn	1	2009
14	Lotus-Cli...	1	1963
15	Lotus-Ford	1	1968
16	Maserati	1	1957
17	Matra-Ford	1	1969
18	Team Lotus	1	1978

## Insights:

Ferrari has one the most amount of Championships till date with 22 wins.

McLaren, Mercedes and Williams coming a close second with 9 wins.

Red Bull holds third position with 6 wins in years 2023, 2022, 2013, 2012, 2011, 2010.



# Driver Dominance and their Performance

```
with SeasonDriverWinners as
(Select year,CONCAT(forename,' ',surname) as DriverName, sum(points) as total,
  rank() over (partition by year order by sum(points) desc) as rank
from Drivers d
  JOIN results res on res.driverId=d.driverId
  JOIN races r on r.raceId=res.raceId
group by CONCAT(forename,' ',surname), year)
SELECT d1.DriverName, count(*) as #OfWins,
STUFF((Select ', ' + cast(year as VARCHAR)
  from SeasonDriverWinners d2
  where d1.DriverName=d2.DriverName
  and rank=1 order by year desc FOR XML PATH('')), 1, 1, '') as YearsWon
FROM SeasonDriverWinners d1
WHERE (rank IN (1))
group by DriverName
order by #OfWins desc,DriverName;
```

## Output

	DriverName	#OfWins	YearsWon
1	Lewis Hamilton	7	2020, 2019, 2018, 2017, 2015, 2014, 2008
2	Michael Schumacher	7	2004, 2003, 2002, 2001, 2000, 1995, 1994
3	Alain Prost	5	1993, 1989, 1988, 1986, 1985
4	Juan Fangio	5	1957, 1956, 1955, 1954, 1951
5	Sebastian Vettel	4	2013, 2012, 2011, 2010
6	Graham Hill	3	1968, 1964, 1962
7	Jack Brabham	3	1966, 1960, 1959
8	Jackie Stewart	3	1973, 1971, 1969
9	Max Verstappen	3	2023, 2022, 2021
10	Nelson Piquet	3	1987, 1983, 1981
11	Niki Lauda	3	1984, 1977, 1975
12	Alberto Ascari	2	1953, 1952
13	Ayrton Senna	2	1991, 1990
14	Emerson Fittipaldi	2	1974, 1972
15	Fernando Alonso	2	2006, 2005
16	Jim Clark	2	1965, 1963
17	Mika Häkkinen	2	1999, 1998
18	Alan Jones	1	1980

## Insights:

Lewis Hamilton and Michael Schumacher holds the first position in championship with 7 wins.

Alain Prost and Juan Fangio hold second position with 5 wins.

Sebastian Vettel holds the third position with 4 wins in consecutive years from 2010 to 2013.

Max Verstappen has been winning for the past three years from 2021 to 2023.

