

A high-angle, slightly rear-quarter view of a red Ferrari Formula 1 car on a dark asphalt racetrack. The car is angled towards the right. It features prominent black tires with green accents on the sidewalls. The rear wing is large and white with red and black checkered patterns. The driver's helmet is visible in the cockpit, featuring a colorful design. Various sponsor logos are visible on the car, including the Ferrari prancing horse, Shell, and Fiat. A semi-transparent blue banner with white text is overlaid at the bottom of the image.

Assignment: Formula One

1 Formula One

Formula 1 (F1) is the top level of international car racing, where drivers compete in some of the fastest and most advanced cars in the world. These races take place on special race tracks (circuits) and city streets across different countries, making it a truly global sport. Drivers compete to finish races as quickly as possible, earning points for themselves and their teams. Over the course of a season, these points decide who becomes the **World Champion Driver** and which team wins the **Constructors' Championship**.

What makes F1 unique is the combination of driver skill, team strategy, and cutting-edge technology. Teams such as Ferrari, Mercedes, Red Bull, and McLaren spend millions designing cars that push the limits of speed, safety, and performance. Success doesn't depend only on the driver's talent but also on teamwork; engineers, mechanics, and strategists all play a huge role in race outcomes.

The name "Formula 1" comes from the idea that all participants must follow a strict set of rules, or a "formula", that governs car design and racing. The "1" shows that it is the highest class of single-seater racing, above lower categories like Formula 2 or Formula 3. This mix of strict rules, fierce competition, and constant innovation makes F1 not just a sport, but also a showcase of human skill and engineering excellence.

1.1 F1 Season

A Formula 1 **season** is the full calendar of races, called **Grands Prix**, that take place across different countries in a year. A season usually has around **20-24 races**, starting in early spring and ending in late autumn. Each race gives drivers and teams the chance to score points based on where they finish. The season is designed to test consistency, skill, and teamwork across many different tracks, weather conditions, and challenges.

The winners are determined by **total points**. The driver with the most points at the end of the season becomes the **Drivers' World Champion**, while the team with the most combined points from both of its drivers wins the **Constructors' Championship**. This system means that winning many races helps, but consistency (i.e., finishing in high positions often) is just as important to becoming champion.

1.2 Individual Races

A single Formula 1 race, called a **Grand Prix**, takes place on Sunday and usually lasts about 1.5 to 2 hours, covering just over **300 kilometers** (except for Monaco, which is shorter). All drivers line up on the **starting grid** based on their qualifying results, with the fastest driver starting from **pole position** at the front. When the lights go out, the race begins, and drivers compete to complete the set number of laps as quickly as possible.

The winner is the driver who crosses the finish line first after completing the full race distance. Along the way, drivers must manage tire changes with pit stops, follow fuel and safety rules, and sometimes adapt to weather or accidents that can bring out the safety car. Points are then awarded based on finishing positions, with the top 10 drivers scoring championship points that count toward both the **Drivers' and Constructors' Championships**.

1.3 Qualifyings

Qualifying is the session that decides the starting order of the cars for the race. It usually takes place the day before the race and is all about setting the fastest single lap time. The driver with the quickest lap earns **pole position**, which means starting the race at the very front, giving them the best chance to win.

Qualifying is split into three parts: **Q1**, **Q2**, and **Q3**. In Q1, all drivers try to set fast times, and the slowest ones are eliminated. The same happens in Q2, leaving only the fastest 10 drivers to battle in Q3 for the top positions on the grid. This format adds excitement and strategy, as teams must choose the best moment to send their cars out and pick the right tires for the fastest lap.

1.4 Pit Stops

A **pit stop** is when a driver quickly pulls into the team's garage area, called the pit lane, during a race. The pit crew (usually around 20 mechanics) works together at incredible speed to change the car's tires, fix small issues, or adjust settings. A tire change can take just 2-3 seconds, making pit stops one of the most intense and precise parts of a race.

Pit stops are also a big part of race strategy. Since tires wear out and lose grip, drivers need fresh ones to stay fast and safe. Teams must carefully decide when to pit, balancing speed with time lost in the stop. A well-timed pit stop can help a driver overtake opponents, while a slow or badly timed one can cost valuable positions.

1.5 Sprint Races

In Formula 1, a **sprint** is a shorter race that takes place on certain weekends, usually on the day before the main race. Unlike the full race, which often lasts over 300 km, a sprint is about **100 km** with no mandatory pit stops, so it's flat-out racing from start to finish. The goal is simple: finish as high as possible in a quick, intense battle.

Sprints add extra excitement because they give drivers a chance to score extra championship points and sometimes help decide the starting order for the main race. Since they are shorter and more aggressive, sprints often feature more overtaking and risk-taking, making them a fan favorite.

1.6 Constructors

In Formula 1, a **constructor** is the team that designs and builds the car, including its chassis (the body and structure). Famous constructors include Ferrari, Mercedes, Red Bull, and McLaren. Each constructor enters two cars in every race, driven by two different drivers. While drivers compete for the **Drivers' Championship**, constructors compete for the **Constructors' Championship**, which is awarded to the team that scores the most points across both of its cars during the season.

The Constructors' Championship is very important because it not only shows which team built the best and most consistent car, but it also affects how much prize money the team earns. Success depends on both the drivers' performances and the teamwork of engineers, mechanics, and strategists behind the scenes. This makes F1 as much a **team sport** as it is an individual competition.

2 Database

The dataset is publicly available on **Kaggle**' [1]. The database is available on Canvas [Files](#) ► [Cases](#) ► [f1db](#). The data has been adopted for the database; see more details below

2.1 Data Preparation

The available dataset comes in [.csv](#) files that do not adhere to good database design (similar to the motivation of the project). Thus the following data preparation steps have been performed:

- The database contains only the F1 Sessions from 1994 to 2024.
- All artificial keys have been replaced with natural keys.
- For drivers, the attribute “nationality” has been replaced with “country” (of birth) for consistency.
- Inconsistencies in the spelling of some countries have been cleaned.
- All timings have been converted to milliseconds for consistency.

2.2 Correctness

With respect to the data, we cannot guarantee the correctness of the data with respect to the real-world event. The data may be incomplete or inaccurate when compared to the actual races. However, we will assume that the data is **our source of correctness**. In other words, you should be concerned only with the data that are present in the database. For example, if Lewis Hamilton actually finished in second place but recorded as finished in first place for a race on 2015-03-15, then we will assume that he indeed finished first place.

With respect to your answer, it should be correct with respect to the question and not the data.

- You should not “hardcode” values unless they are mentioned in the question. For instance, you should not use `'Sepang International Circuit'` when the question specifies circuits in `'Malaysia'`. There may be a change of name.
- You cannot assume the **order of the data**. We will test with data that are randomly inserted. For instance,

- you cannot assume that the driver that finished first will always be at the top of the table (i.e., cannot assume position is sorted).
- you cannot assume that the races are going to be in order.
- etc.
- You should ensure that your code can run and follow the format required for submission.
- We will test with other dataset (e.g., different years).

If you truly think that your solution is correct but does not match the expected result, you are welcomed to contact us.

2.3 Grading

Grading will be automatic.

- **You must use the template for assignment 1.** The template can be found in:
Canvas [Files](#) ► [Assignments](#) ► [A01](#) ► [Assignment01.sql](#).
- Your output should have the same number of columns.
- Your output should have the column in the same order.
- Your output should have the rows in the same order **if sorting is required by the question**. Otherwise, you may output the rows in any order.

No partial marks for outputs that are incorrect (i.e., wrong number of columns, wrong column order, wrong rows, incorrect sorting, etc) or cannot run. Please ensure that your code can be executed with only query permission (i.e., no table creations, etc).

You are encouraged to follow the column name, but you may choose a different column name.

2.4 Checker

We provide an online checker to help you quickly check your solution. You are expected to check that your code can run locally. Passing the online checker does not guarantee full mark as we will also check with other dataset. This includes (but not limited to) data that was inserted in random order, data from different dataset, etc. Your query should finish within 10s.

The checker is located at <https://thisisadi.yoga/Database//Tools/A01/>.

You can find a file matching the template given in the tab named “CODES”. You may download the file from here to be submitted on Canvas.

3 Submission

Submit the file [Assignment01.sql](#) on Canvas [Assignments](#) ► [Assignment 1](#). Ensure that the **file names are correct before submission** (i.e., the file names in your computer). Canvas may perform renaming; that is normal. Only the latest submission will be graded.

4 Questions

- (1 point) Print the seasons with the least number of withdrawals. A withdrawal is recorded with a 'W' in column `position_text` of the table `results`. If there are multiple seasons with the same least number of withdrawals, print all of the seasons.

Important: You should not hardcode the number 0.

Output only the season and number of withdrawals. Sort the result in ascending order of season. See the sample output below.

season	count
2018	0
2019	0

- (1 point) Print the different constructors that never had a car that finished in the top-3. Exclude any constructor that has not participated in any race. A constructor has participated in a race if the constructor name is recorded in `constructor` column of the table `results`.

Output only the constructor name as well as the first and the last seasons the constructor participated in a race. Sort the result in ascending order of constructor name. See the sample output below.

constructor	first	last
Alfa Romeo	2019	2023
Haas F1 Team	2016	2024
Manor Marussia	2015	2016
RB F1 Team	2024	2024
Sauber	2015	2024

- (1 point) For each driver, print the number of times they had the pole position. A driver had a pole position if the column `position` of the table `qualifyings` has a value of 1. Include drivers that never had pole position (i.e., 0 times).

Output only the driver forename, surname, and the number of times they had pole position. Sort the result in descending order of the count, followed by ascending order of surname, followed by ascending order of forename. The sample output below is truncated. There should be 52 rows. We show the first and last three rows only.

forename	surname	count
Lewis	Hamilton	69
Max	Verstappen	43
Charles	Leclerc	23
:	:	:
Stoffel	Vandoorne	0
Pascal	Wehrlein	0
Guanyu	Zhou	0

4. (1 point) Print the different races that the winner of the race did not start from pole position. A driver wins the race if the column `position` of the table `results` has a value of `1`. A driver had a pole position if the column `position` of the table `qualifyings` has a value of `1`.

Output only the race date, driver forename, driver surname, and the qualifying position of the driver. Sort the result in ascending order of the race date. The sample output below is truncated. There should be 96 rows. We show the first and last three rows only.

race	forename	surname	pole
2015-03-29	Sebastian	Vettel	2
2015-05-24	Nico	Rosberg	2
2015-06-21	Nico	Rosberg	2
:	:	:	:
2024-09-15	Oscar	Piastri	2
2024-10-20	Charles	Leclerc	4
2024-11-03	Max	Verstappen	12

5. (1 point) For each season, print the driver champion and the number of race won by the champion. A driver is a driver champion of a season if the driver has the most cumulative sum of the column `points` in the table `results` for the given season. A driver wins the race if the column `position` of the table `results` has a value of `1`.

In the current dataset, the champion of 2024 season is Max Verstappen with 399 points and 9 race wins. Note that for 2016 season, the champion is Nico Rosberg with 385 points and 9 race wins. However, Lewis Hamilton has 10 race wins but with only 380 points.

Output only the season, driver forename, driver surname, and the number of race wins. Sort the result in ascending order of the season. See the sample output below.

season	forename	surname	wins
2015	Lewis	Hamilton	10
2016	Nico	Rosberg	9
2017	Lewis	Hamilton	9
2018	Lewis	Hamilton	11
2019	Lewis	Hamilton	11
2020	Lewis	Hamilton	11
2021	Max	Verstappen	10
2022	Max	Verstappen	15
2023	Max	Verstappen	19
2024	Max	Verstappen	9

References

- [1] *Kaggle: Your Machine Learning and Data Science Community*. <https://www.kaggle.com/>. [Online; last accessed 2025].