



FORMULA 1

SQL PROJECT →

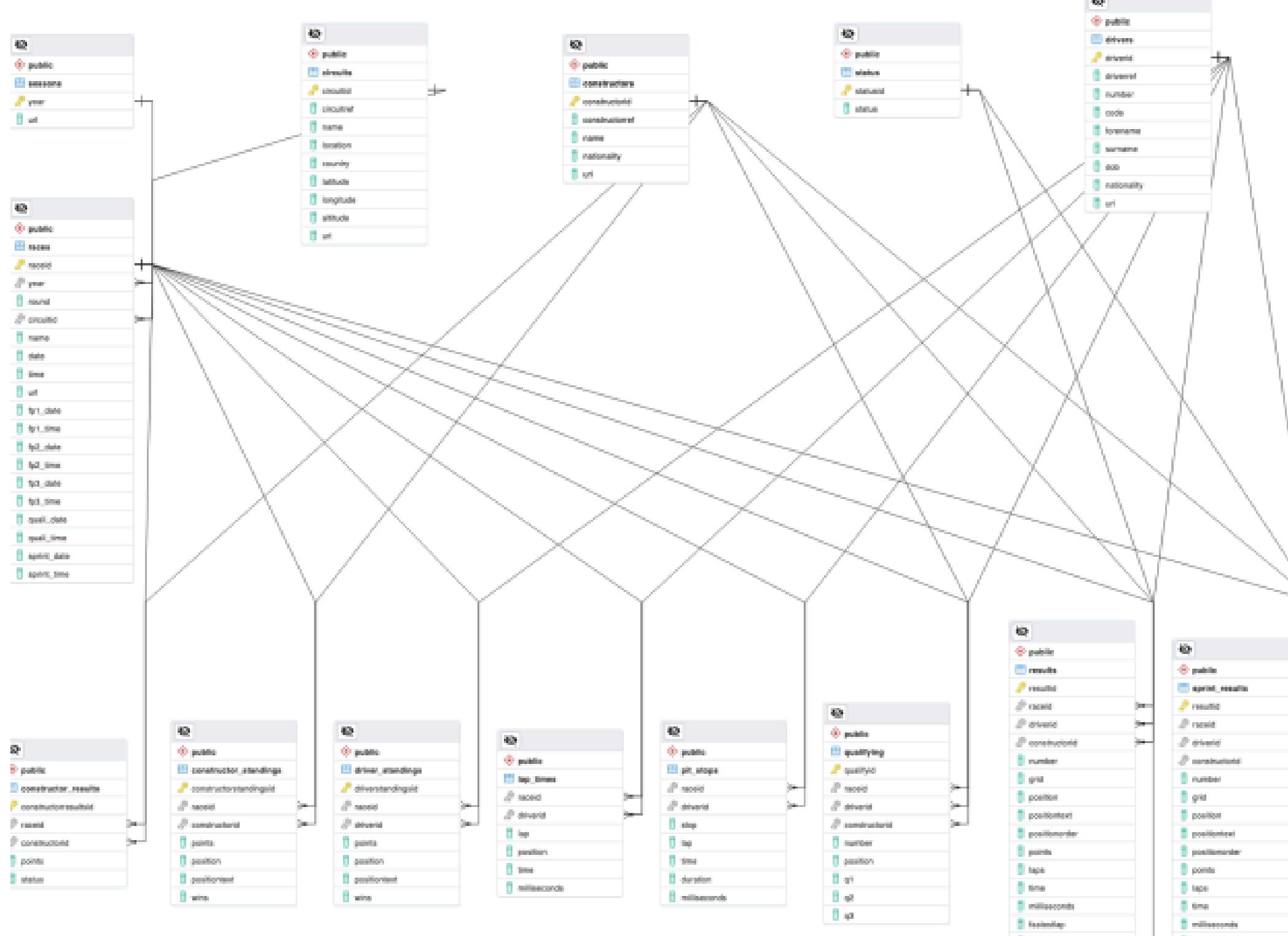
DISCRIPATION



Formula One (more commonly known as Formula 1 or F1) is the highest class of international racing for open-wheel single-seater formula racing cars sanctioned by the Federation Internationale l'Automobile (FIA).



For this case study, we have taken the dataset from kaggle. The dataset consists of all information on the Formula 1 races, drivers, constructors, qualifying, circuits, lap times, pit stops, championships from 1950 till the latest 2023 season.



IDENTIFY THE COUNTRY WHICH HAS PRODUCED THE MOST F1 DRIVERS.

Code:

```
select nationality,count(1) from drivers group by nationality  
order by 2 desc limit 1;
```



WHICH COUNTRY HAS PRODUCED THE MOST NO OF F1 CIRCUITS

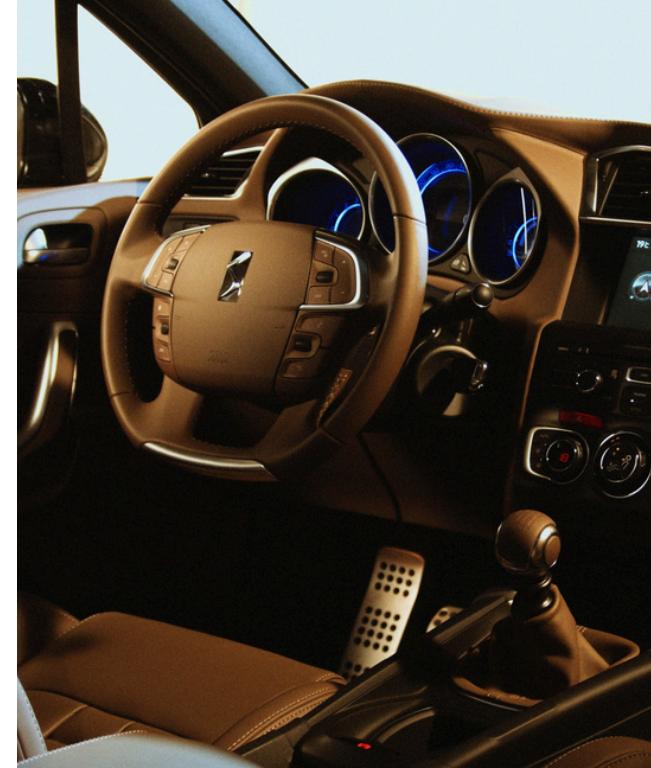
Code:

```
select country,count(1) from circuits group by country order by 2 desc limit 1;
```

WHICH COUNTRIES HAVE PRODUCED EXACTLY 5 CONSTRUCTORS?

Code:

```
select nationality,count(1) from constructors group by  
nationality having count(1) = 5;
```



Code:

```
select year, count(1) from races group by year order by year desc;
```

LIST DOWN THE NO OF RACES THAT HAVE TAKEN PLACE EACH YEAR AND MENTIONED WHICH WAS THE FIRST AND THE LAST RACE OF EACH SEASON.



```
select distinct year, first_value(name) over(partition by year order by date) as  
first_race, last_value(name) over(partition by year order by date  
range between unbounded preceding and unbounded following) as  
last_race  
, count(*) over(partition by year) as no_of_races  
from races  
order by year desc
```

WHICH CIRCUIT HAS HOSTED THE MOST NO OF RACES. DISPLAY THE CIRCUIT NAME, NO OF RACES, CITY AND COUNTRY.



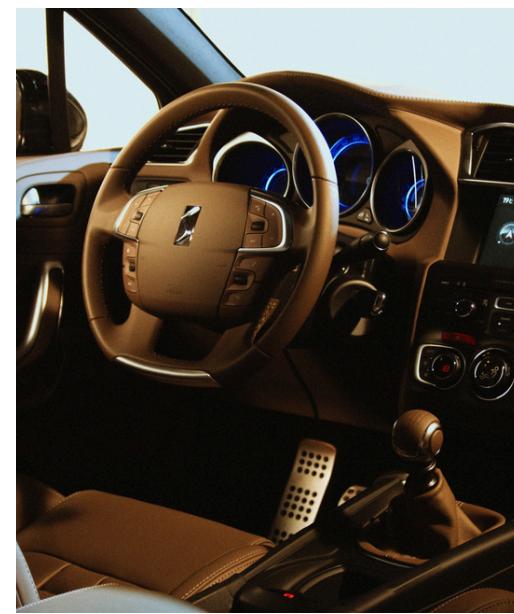
```
with cte as  
(select c.name as circuit_name, count(1) no_of_races  
 , rank() over(order by count(1) desc) as rnk  
      from races r  
     join circuits c on c.circuitid=r.circuitid  
      group by c.name)  
select circuit_name, no_of_races, c.location as city, c.country  
      from circuits c  
     join cte on cte.circuit_name=c.name  
      where rnk=1;
```

DISPLAY THE FOLLOWING FOR 2022 SEASON:

YEAR, RACE_NO, CIRCUIT NAME, DRIVER NAME, DRIVER RACE POSITION, DRIVER RACE POINTS, FLAG TO INDICATE IF WINNER
, CONSTRUCTOR NAME, CONSTRUCTOR POSITION, CONSTRUCTOR POINTS, , FLAG TO INDICATE IF CONSTRUCTOR IS WINNER
, RACE STATUS OF EACH DRIVER, FLAG TO INDICATE FASTEST LAP FOR WHICH DRIVER, TOTAL NO OF PIT STOPS BY EACH DRIVER

```
select r.raceid, r.year, r.round as race_no, r.name as circuit_name, concat(d.forename, ' ',d.surname) as driver_name
, ds.position as driver_position, ds.points as driver_points, case when ds.position=1 then 'WINNER' end as winner_flag
, c.name as constructor_name, cs.position as constructor_position, cs.points as constructor_points
, case when cs.position=1 then 'WINNER' end as cons_winner_flag, sts.status
, case when lp.driverid is not null then 'Faster Lap' end as fastest_lap_indi, pt.no_of_stops
from races r
join driver_standings ds on ds.raceid=r.raceid
join drivers d on d.driverid = ds.driverid
join constructor_standings cs on cs.raceid=r.raceid
join constructors c on c.constructorid=cs.constructorid
join results res on res.raceid=r.raceid and res.driverid=ds.driverid and res.constructorid=cs.constructorid
join status sts on sts.statusid=res.statusid
left join (
    select lp.raceid, lp.driverid
    from lap_times lp
    join ( select raceid, min(time) as fastest_lap
            from lap_times
            group by raceid) x on x.raceid=lp.raceid and x.fastest_lap=lp.time
    ) lp on lp.driverid = ds.driverid and lp.raceid=r.raceid
left join (
    select raceid,driverid, count(1) as no_of_stops
    from pit_stops
    group by raceid,driverid) pt on pt.driverid = ds.driverid and pt.raceid=r.raceid
where year=2022 --and r.raceid=1074
order by year, race_no, driver_position;
```

LIST DOWN THE NAMES OF ALL F1 CHAMPIONS AND THE NO OF TIMES THEY HAVE WON IT.



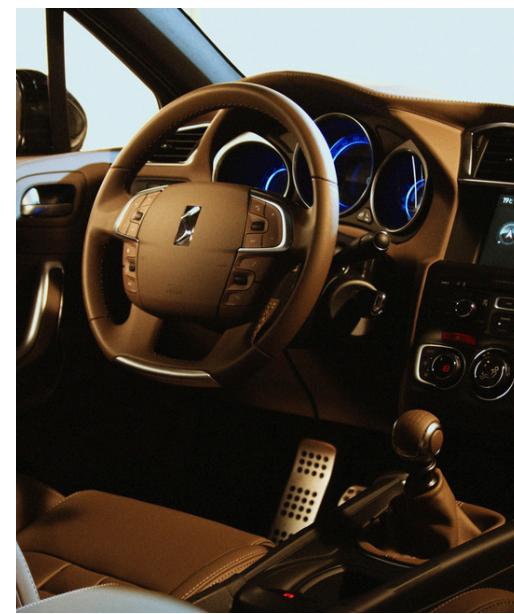
with cte as

```
(select r.year, concat(d.forename,' ',d.surname) as driver_name  
, sum(res.points) as tot_points  
, rank() over(partition by r.year order by sum(res.points) desc) as rnk  
from races r  
join driver_standings ds on ds.raceid=r.raceid  
join drivers d on d.driverid=ds.driverid  
join results res on res.raceid=r.raceid and res.driverid=ds.driverid  
--and res.constructorid=cs.constructorid  
--where r.year>=2000  
group by r.year, res.driverid, concat(d.forename,' ',d.surname) ),  
cte_rnk as  
(select * from cte where rnk=1)  
select driver_name, count(1) as no_of_championships  
from cte_rnk  
group by driver_name  
order by 2 desc;
```

WHO HAS WON THE MOST CONSTRUCTOR CHAMPIONSHIPS

with cte as

```
(select r.year, c.name as constructor_name  
, sum(res.points) as tot_points  
, rank() over(partition by r.year order by sum(res.points) desc) as rnk  
from races r  
join constructor_standings cs on cs.raceid=r.raceid  
join constructors c on c.constructorid = cs.constructorid  
join constructor_results res on res.raceid=r.raceid and res.constructorid=cs.constructorid  
--and res.constructorid=cs.constructorid  
--where r.year>=2022  
group by r.year, res.constructorid, c.name),  
cte_rnk as  
(select * from cte where rnk=1)  
select constructor_name, count(1) as no_of_championships  
from cte_rnk  
group by constructor_name  
order by 2 desc;
```



HOW MANY RACES HAS INDIA HOSTED?



```
select c.name as circuit_name,c.country, count(1) no_of_races  
      from races r  
join circuits c on c.circuitid=r.circuitid  
      where c.country='India'  
group by c.name,c.country;
```

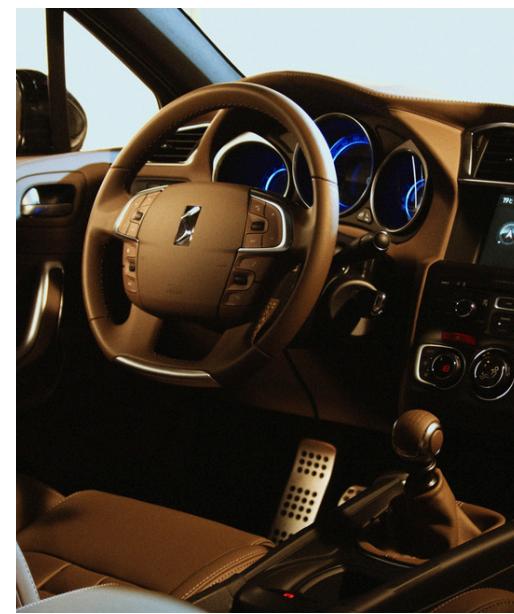


DISPLAY THE TOP 10 DRIVERS WITH MOST WINS.

```
select driver_name, race_wins
from (
    select ds.driverid, concat(d.forename, ' ', d.surname) as driver_name
    , count(1) as race_wins
    , rank() over(order by count(1) desc) as rnk
    from driver_standings ds
    join drivers d on ds.driverid=d.driverid
    where position=1
    group by ds.driverid, concat(d.forename, ' ', d.surname)
    order by race_wins desc, driver_name) x
where rnk <= 10;
```

DISPLAY THE TOP 3 CONSTRUCTORS OF ALL TIME.

```
select constructor_name, race_wins  
from (  
    select cs.constructorid, c.name as constructor_name  
    , count(1) as race_wins  
    , rank() over(order by count(1) desc) as rnk  
    from constructor_standings cs  
    join constructors c on c.constructorid=cs.constructorid  
    where position = 1  
    group by cs.constructorid, c.name  
    order by race_wins desc) x  
where rnk <= 3;
```



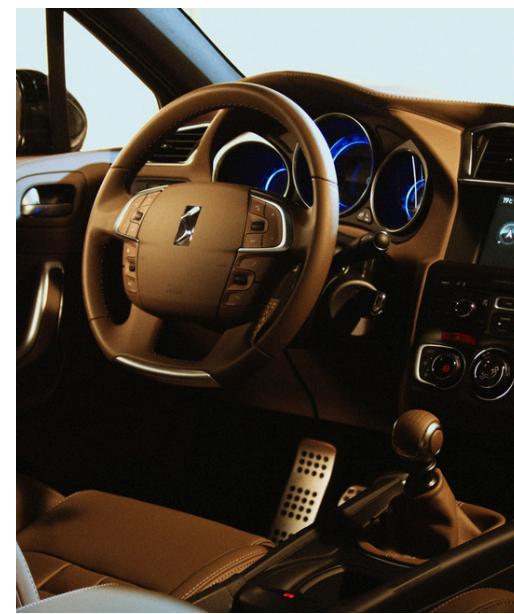
IDENTIFY THE DRIVERS WHO HAVE WON RACES WITH MULTIPLE TEAMS.

```
select driverid, driver_name, string_agg(constructor_name, ',')  
from (  
    select distinct r.driverid  
    , concat(d.forename, ' ', d.surname) as driver_name  
    , c.name as constructor_name  
    from results r  
    join drivers d on d.driverid=r.driverid  
    join constructors c on c.constructorid=r.constructorid  
    where r.position=1) x  
group by driverid, driver_name  
having count(1) > 1  
order by driverid, driver_name;
```



HOW MANY DRIVERS HAVE NEVER WON ANY RACE.

```
select d.driverid  
, concat(d.forename,' ',d.surname) as driver_name  
, nationality  
from drivers d  
where driverid not in (select distinct driverid  
    from driver_standings ds  
    where position=1)  
order by driver_name;
```



HOW MANY RACES HAS EACH OF THE TOP 5 CONSTRUCTORS WON IN THE LAST 10 YEARS.

with top_5_teams as

```
(select constructorid, constructor_name
from (
  select cs.constructorid, c.name as constructor_name
  , count(1) as race_wins
  , rank() over(order by count(1) desc) as rnk
  from constructor_standings cs
  join constructors c on c.constructorid=cs.constructorid
  where position = 1
  group by cs.constructorid, c.name
  order by race_wins desc) x
where rnk <= 5)
select cte.constructorid, cte.constructor_name, coalesce(cs.wins,0) as wins
from top_5_teams cte
left join ( select cs.constructorid, count(1) as wins
  from constructor_standings cs
  join races r on r.raceid=cs.raceid
  where cs.position = 1
  and r.year >= (extract(year from current_date) - 10)
  group by cs.constructorid
) cs
on cte.constructorid = cs.constructorid
order by wins desc;
```

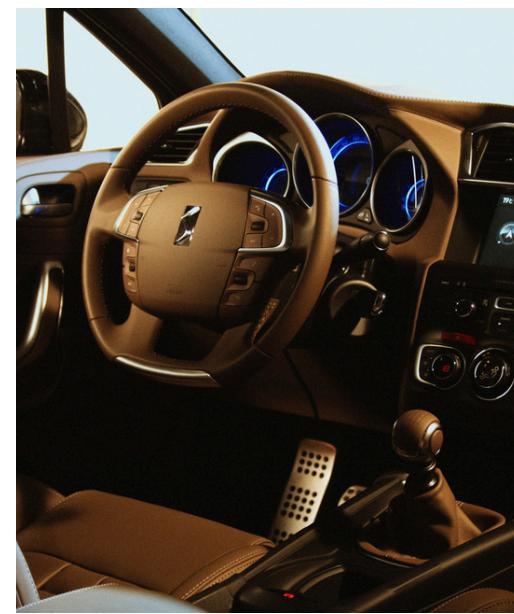


DISPLAY THE WINNERS OF EVERY SPRINT SO FAR IN F1

```
select r.year, r.name, concat(d.forename, ' ', d.surname) as driver_name  
from sprint_results sr  
join drivers d on d.driverid=sr.driverid  
join races r on r.raceid=sr.raceid  
where sr.position=1  
order by 1,2;
```

FIND THE DRIVER WHO HAS THE MOST NO OF DID NOT QUALIFY DURING THE RACE.

```
select driver_name, cnt
from (
    select r.driverid
    , concat(d.forename,' ',d.surname) as driver_name
    , count(1) as cnt
    , rank() over(order by count(1) desc) as rnk
    from status s
    join results r on r.statusid=s.statusid
    join drivers d on d.driverid=r.driverid
    where s.status='Did not qualify'
    group by r.driverid, concat(d.forename,' ',d.surname)
    order by cnt desc) x
where rnk=1;
```



DURING THE LAST RACE OF 2022 SEASON, IDENTIFY THE DRIVERS WHO DID NOT FINISH THE RACE AND THE REASON FOR IT.

```
select concat(d.forename, ' ', d.surname) as
driver_name
, s.status
from results r
join status s on s.statusid=r.statusid
join drivers d on d.driverid=r.driverid
where r.raceid = (select max(raceid) from races
where year=2022)
and r.statusid<>1;
```



WHO WON THE DRIVERS CHAMPIONSHIP WHEN INDIA HOSTED F1 FOR THE FIRST TIME?

```
with driver_champ_points as
    (select r.year, concat(d.forename, ' ', d.surname) as driver_name
     , sum(res.points) as tot_points
     , rank() over(partition by r.year order by sum(res.points) desc) as rnk
      from races r
      join driver_standings ds on ds.raceid=r.raceid
      join drivers d on d.driverid=ds.driverid
      join results res on res.raceid=r.raceid and res.driverid=ds.driverid --and res.constructorid=cs.constructorid
      where r.year in (2011,2012,2013)
      group by r.year, res.driverid, concat(d.forename, ' ', d.surname) ),
driver_champ as
    (select * from driver_champ_points where rnk=1),
india_first_year as
    (select min(year) as first_yr from races
     where circuitid in (select circuitid from circuits
                           where country='India'))
select year, driver_name
from driver_champ
where year = (select first_yr from india_first_year);
```

WHICH DRIVER HAS DONE THE MOST LAP TIME IN F1 HISTORY?

```
select driver_name, total_lap_time
from (
    select lt.driverid
    , concat(d.forename,' ',d.surname) as driver_name
    , sum(time) as total_lap_time
    , rank() over(order by sum(time) desc) as rnk
    from lap_times lt
    join drivers d on d.driverid=lt.driverid
    group by lt.driverid, concat(d.forename,' ',d.surname)) x
where rnk=1;
```



NAME THE TOP 3 DRIVERS WHO HAVE GOT THE MOST PODIUM FINISHES IN F1 (TOP 3 RACE FINISHES)

```
select driver_name, no_of_podiums
from (select ds.driverid, concat(d.forename,' ',d.surname) as
driver_name
, count(1) as no_of_podiums
, rank() over(order by count(1) desc) as rnk
from driver_standings ds
join drivers d on d.driverid=ds.driverid
where ds.position <= 3
group by ds.driverid, concat(d.forename,' ',d.surname)) x
where rnk<=3;
```

WHICH DRIVER HAS THE MOST POLE POSITION (NO 1 IN QUALIFYING)

*** DATA IS MISSING FOR SOME RACE QUALIFICATIONS

```
select driver_name, pole_positions
  from (select q.driverid, concat(d.forename, ' ', d.surname) as
driver_name
       , count(1) as pole_positions
       , rank() over(order by count(1) desc) as rnk
     from qualifying q
    join drivers d on d.driverid=q.driverid
   where position=1
  group by q.driverid, concat(d.forename, ' ', d.surname)
 order by pole_positions desc) x
 where rnk=1;
```