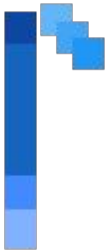
F1 Data Analytics Project



WATANABE

By Nate Vu

Introduction

Purpose of the project

- Exposure to database management
- Enhance data modeling skills
- Get used with ETL using Python, SQL, etc.

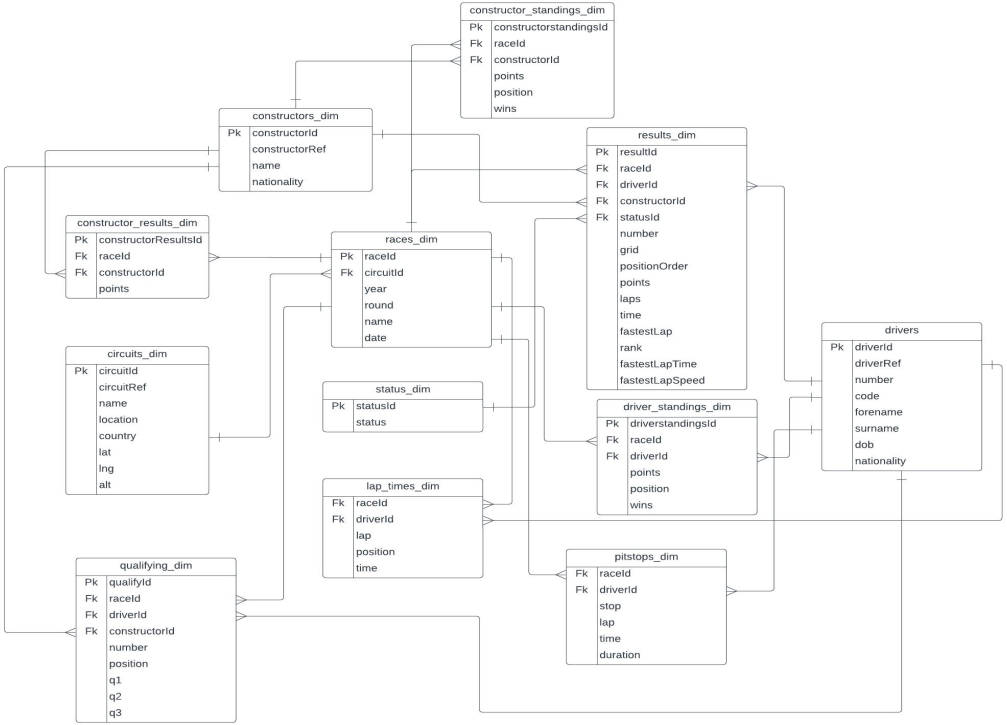
Questions to consider for this project

- Who has accumulated the most points during their career?
- What is the most dangerous circuit?

Project Implementation Steps

1. Model data using LucidChart
2. Assess, clean and transform data before exporting to Cloud Storage (data source)
3. Perform SQL queries to extract data on BigQuery

Data Model


constructor_standings_dim
Pk constructorstandingsId
Fk raceId
Fk constructorId
points
position
wins
constructors_dim
Pk constructorId
constructorRef
name
nationality
results_dim
Pk resultId
Fk raceId
Fk driverId
Fk constructorId
Fk statusId
number
grid
positionOrder
points
laps
time
fastestLap
rank
fastestLapTime
fastestLapSpeed
constructor_results_dim
Pk constructorResultsId
Fk raceId
Fk constructorId
points
races_dim
Pk raceId
Fk circuitId
year
round
name
date
drivers
Pk driverId
driverRef
number
code
forename
surname
dob
nationality
circuits_dim
Pk circuitId
circuitRef
name
location
country
lat
lng
alt
status_dim
Pk statusId
status
driver_standings_dim
Pk driverstandingsId
Fk raceId
Fk driverId
points
position
wins
lap_times_dim
Fk raceId
Fk driverId
lap
position
time
pitstops_dim
Fk raceId
Fk driverId
stop
lap
time
duration
qualifying_dim
Pk qualifyId
Fk raceId
Fk driverId
Fk constructorId
number
position
q1
q2
q3

Sample SQL queries

Find drivers' career points

F1-bestdrivers-2023-07-19 1...:25 RUN SAVE SHARE SCHEDULE MORE Query ...

```sql
--Create new table to find career points of each driver based on driverId
SELECT d.driverId, SUM(r.points) AS AccumulatePts
FROM `hardy-answer-335705.F1_project.results_dim` r
JOIN `hardy-answer-335705.F1_project.drivers_dim` d on d.driverId = r.driverId
GROUP BY d.driverId
--save this query as new table called 'bestdriver_dim'

SELECT d.forename, d.surname, d.dob, d.nationality, b.AccumulatePts
FROM `hardy-answer-335705.F1_project.drivers_dim` d
JOIN `hardy-answer-335705.F1_project.bestdriver_dim` b on d.driverId = b.driverId;
```

Press Alt+F1 for Accessibility Options.

Query results SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	forename	surname	dob	nationality	AccumulatePts
1	Lewis	Hamilton	1985-01-07	British	4396.5
2	Sebastian	Vettel	1987-07-03	German	3098.0
3	Fernando	Alonso	1981-07-29	Spanish	2061.0
4	Max	Verstappen	1997-09-30	Dutch	1983.5
5	Kimi	Räikkönen	1979-10-17	Finnish	1873.0
6	Valtteri	Bottas	1989-08-28	Finnish	1778.0
7	Nico	Rosberg	1985-06-27	German	1594.5
8	Michael	Schumacher	1969-01-03	German	1566.0
9	Daniel	Ricciardo	1989-07-01	Australian	1307.0
10	Jenson	Button	1980-01-19	British	1235.0

Results per page: 50 1 – 50 of 855

Find circuits with the most collisions

F1-stats-2023-07-19 16:08:33 RUN SAVE SHARE SCHEDULE MORE Query com...

```sql
--Finding circuits with most collisions
SELECT c.name, COUNT(
  status) AS Collisions
FROM `hardy-answer-335705.F1_project.status_dim` s
JOIN `hardy-answer-335705.F1_project.results_dim` r on r.statusId = s.statusId
JOIN `hardy-answer-335705.F1_project.races_dim` ra on ra.raceId = r.raceId
JOIN `hardy-answer-335705.F1_project.circuits_dim` c on c.circuitId = ra.circuitId
--setting condition to collision only (4 = Collision)
WHERE r.statusId = 4
GROUP BY c.name
```

Press Alt+F1 for Accessibility Options.

Query results SAVE RESULTS EXPLORE DATA

JOB INFORMATION | RESULTS | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	name	Collisions
1	Circuit de Monaco	72
2	Circuit de Spa-Francorchamps	47
3	Albert Park Grand Prix Circuit	46
4	Circuit Gilles Villeneuve	43
5	Autódromo José Carlos Pace	42
6	Silverstone Circuit	42
7	Autodromo Nazionale di Monza	40
8	Circuit de Barcelona-Catalunya	39
9	Hungaroring	34
10	Hockenheimring	34

Results per page: 50 1 – 50 of 53

Codes

Python transformation codes:
https://github.com/nateVu0681147/watanabe-projects/blob/main/F1-ETLcodes.ipynb

SQL codes:
https://github.com/nateVu0681147/watanabe-projects/blob/main/F1-SQL%20queries.ipynb

Thank you so much for having viewed this!