File Edit View Run Help   Last edit was 6 minutes ago   Provide feedback

▶ Run all   ● dbkcourse_cluster ∨   🗓 Schedule   Share

Cmd 1

```sql
USE f1_processed;
```

OK

Command took 0.41 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 8:00:16 PM on dbkcourse_cluster

Cmd 2

```sql
CREATE TABLE f1_presentation.calculated_race_results_sql
USING parquet
AS
SELECT races.race_year,
       constructors.name AS team_name,
       drivers.name AS driver_name,
       results.position,
       results.points,
       11 - results.position As calculated_points
FROM results
JOIN f1_processed.drivers ON (results.driver_id = drivers.driver_id)
JOIN f1_processed.constructors ON (results.constructor_id = constructors.constructor_id)
JOIN f1_processed.races ON (results.race_id = races.race_id)
WHERE results.position <= 10
```

▶ (8) Spark Jobs

OK

Command took 21.96 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 8:08:18 PM on dbkcourse_cluster

▶ Run all   ● dbkcourse_cluster ∨   📅 Schedule   Share

▶ (8) Spark Jobs

OK

Command took 21.96 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 8:08:18 PM on dbkcourse_cluster

Cmd 3

SQL ▶∨ ▥ ∨ — ✕

```sql
SELECT * FROM f1_presentation.calculated_race_results_sql;
```

▶ (2) Spark Jobs

Table ∨   +

	race_year	team_name	driver_name	position	points	calculated_points
1	1987	Ligier	René Arnoux	10	0	1
2	1987	Brabham	Riccardo Patrese	9	0	2
3	1987	Zakspeed	Christian Danner	8	0	3
4	1987	McLaren	Stefan Johansson	7	0	4
5	1987	Arrows	Eddie Cheever	6	1	5
6	1987	Williams	Nigel Mansell	5	2	6
7	1987	Ferrari	Gerhard Berger	4	3	7

⬇ ∨   10,000 rows | Truncated data | 0.56 seconds runtime   Refreshed 5 minutes ago

Command took 0.56 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 8:08:59 PM on dbkcourse_cluster

Shift+Enter to run
Shift+Ctrl+Enter to run selected text

New

Workspace

Repos

Recents

Data

Compute

Workflows

1.find_dominant_drivers SQL ✓

File   Edit   View   Run   Help   Last edit was 2 hours ago   Provide feedback

▶ Run all   ● dbkcourse_cluster ✓   📅 Schedule   Share

Cmd 1

SQL ▶▾ ⏸ ✓ ─ ✕

```sql
SELECT driver_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
GROUP BY driver_name
HAVING COUNT(1) >= 50
ORDER BY avg_points DESC
```

▶ (2) Spark Jobs

Table ✓   +

	driver_name	total_races	total_points	avg_points
1	Alain Prost	138	1156	8.376811594202898
2	Ayrton Senna	104	869	8.35576923076923
3	Jackie Stewart	61	507	8.311475409836065
4	Lewis Hamilton	232	1903	8.202586206896552
5	Michael Schumacher	226	1796	7.946902654867257
6	Nigel Mansell	90	713	7.9222222222222225
7	Niki Lauda	81	639	7.888888888888889

⬇ 64 rows | 0.97 seconds runtime

Refreshed 2 hours ago

Command took 0.97 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 8:43:18 PM on dbkcourse_cluster

Marketplace   NEW

Partner Connect

1.find_dominant_drivers SQL ∨

File   Edit   View   Run   Help   Last edit was 2 hours ago   Provide feedback

▶ Run all   ● dbkcourse_cluster ∨   📅 Schedule   Share

New
Workspace
Repos
Recents
Data
Compute
Workflows

Marketplace NEW
Partner Connect

Cmd 2

```sql
SELECT driver_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
 WHERE race_year BETWEEN 2011 AND 2020
GROUP BY driver_name
HAVING COUNT(1) >= 50
ORDER BY avg_points DESC
```

▶ (2) Spark Jobs

Table ∨   +

	driver_name	total_races	total_points	avg_points
1	Lewis Hamilton	177	1478	8.350282485875706
2	Sebastian Vettel	164	1282	7.817073170731708
3	Nico Rosberg	93	673	7.236559139784946
4	Max Verstappen	88	605	6.875
5	Valtteri Bottas	117	793	6.777777777777778
6	Fernando Alonso	95	584	6.147368421052631
7	Kimi Räikkönen	121	721	5.958677685950414

↓ 14 rows | 0.51 seconds runtime   Refreshed 2 hours ago

Command took 0.51 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 8:43:18 PM on dbkcourse_cluster

Cmd 3

► Run all   ● dbkcourse_cluster ⌄   📅 Schedule   Share

Cmd 1

```sql
SELECT team_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
GROUP BY team_name
HAVING COUNT(1) >= 100
ORDER BY avg_points DESC
```

► (2) Spark Jobs

Table ⌄   +

	team_name	total_races	total_points	avg_points
1	Mercedes	385	2931	7.612987012987013
2	Ferrari	1478	10500	7.104194857916103
3	McLaren	1032	6865	6.6521317829457365
4	Red Bull	433	2851	6.584295612009238
5	Williams	745	4759	6.387919463087249
6	Benetton	296	1806	6.101351351351352
7	Team Lotus	352	2100	5.965909090909091

⤓   23 rows | 7.58 seconds runtime

Refreshed 1 minute ago

Command took 7.58 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 16:42:02 PM on dbkcourse_cluster

Cmd 2

Data Science & Engi...▾

New

Workspace

Repos

Recents

Data

Compute

Workflows

Marketplace NEW

Partner Connect

1/4 Tasks Completed

2.find_dominant_teams SQL ▾

File Edit View Run Help   Last edit was 2 minutes ago   Provide feedback

▶ Run all   ● dbkcourse_cluster ▾   📅 Schedule   Share

Cmd 2

```sql
SELECT team_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
 WHERE race_year BETWEEN 2011 AND 2020
GROUP BY team_name
HAVING COUNT(1) >= 100
ORDER BY avg_points DESC
```

▶ (2) Spark Jobs

Table ▾   +

	team_name	total_races	total_points	avg_points
1	Mercedes	331	2580	7.794561933534744
2	Red Bull	313	2201	7.031948881789138
3	Ferrari	314	2095	6.671974522292993
4	McLaren	204	984	4.823529411764706
5	Williams	125	531	4.248
6	Force India	189	671	3.5502645502645502
7	Toro Rosso	119	316	2.6554621848739495

↓ 7 rows | 1.11 seconds runtime   Refreshed 2 minutes ago

Command took 1.11 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 10:42:02 PM on dbkcourse_cluster

Cmd 3

New

Workspace

Repos

Recents

Data

Compute

Workflows

```python
%python
html = """<h1 style="color:Black;text-align:center;font-family:Ariel">Report on Dominant Formula 1 Drivers </h1>"""
displayHTML(html)
```

Report on Dominant Formula 1 Drivers

Command took 0.17 seconds — by avinashchinta@my.unt.edu at 7/13/2023, 10:53:32 PM on dbkcourse_cluster

Cmd 2

```sql
CREATE OR REPLACE TEMP VIEW v_dominant_drivers
AS
SELECT driver_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points,
       RANK() OVER(ORDER BY AVG(calculated_points) DESC) driver_rank
  FROM f1_presentation.calculated_race_results_sql
GROUP BY driver_name
HAVING COUNT(1) >= 50
ORDER BY avg_points DESC
```

OK

Command took 0.55 seconds — by avinashchinta@my.unt.edu at 7/13/2023, 10:53:32 PM on dbkcourse_cluster

Cmd 3

SQL ▶▾ 📊 ▾ — ✕

```sql
SELECT * FROM v_dominant_drivers;
```

▶ (3) Spark Jobs

Table ▾ +

	driver_name	total_races	total_points	avg_points	driver_rank
1	Alain Prost	138	1156	8.376811594202898	1
2	Ayrton Senna	104	869	8.35576923076923	2
3	Jackie Stewart	61	507	8.311475409836065	3
4	Lewis Hamilton	232	1903	8.202586206896552	4
5	Michael Schumacher	226	1796	7.946902654867257	5
6	Nigel Mansell	90	713	7.922222222222225	6
7	Niki Lauda	81	639	7.888888888888889	7

Marketplace

Partner Connect

1/4 Tasks Completed

Enable new UI

Menu options

3.viz_dominant_drivers SQL ∨

File   Edit   View   Run   Help   Last edit was 2 minutes ago   Provide feedback

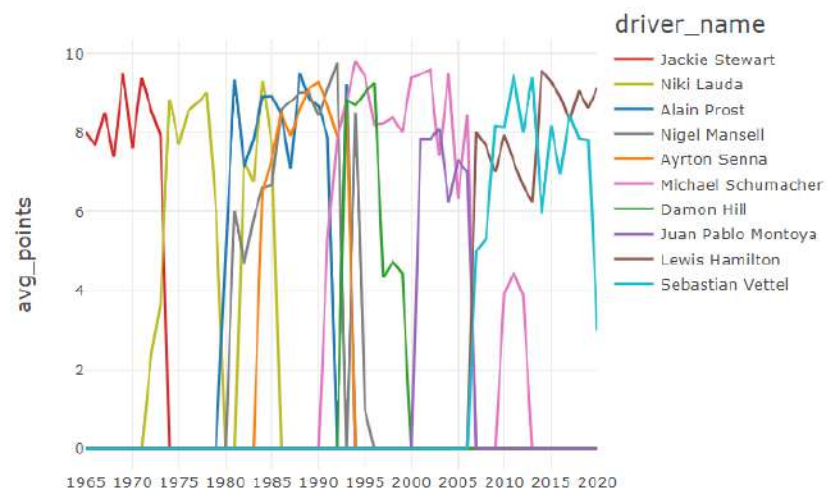▶ Run all   ● dbkcourse_cluster ∨   🗓 Schedule   Share   ∧

```sql
SELECT race_year,
       driver_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
 WHERE driver_name IN (SELECT driver_name FROM v_dominant_drivers WHERE driver_rank <= 10)
 GROUP BY race_year, driver_name
 ORDER BY race_year, avg_points DESC
```
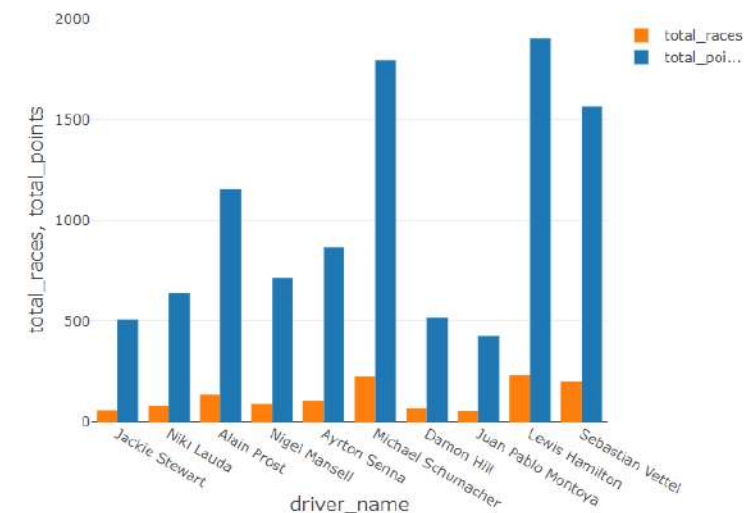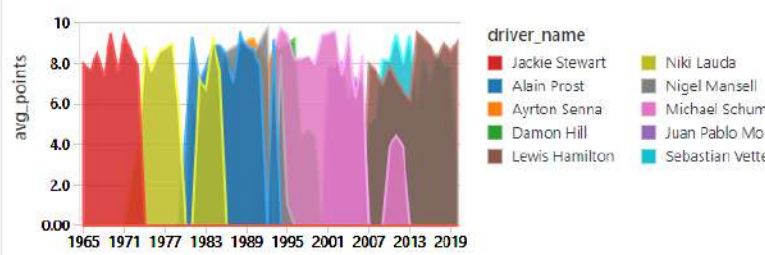
▶ (5) Spark Jobs

Visualization ∨   +

▶ (5) Spark Jobs

New

Workspace

Repos

Recents

Data

Compute

Workflows

Marketplace   NEW

Partner Connect

1/4 Tasks Completed

Enable new UI   NEW

Menu options

3.viz_dominant_drivers SQL ∨

File   Edit   View   Run   Help   Last edit was 3 minutes ago   Provide feedback

▶ Run all   ● dbkcourse_cluster ∨   🗓 Schedule   **Share**   ∧

New

Workspace

Repos

Recents

Data

Compute

Workflows

```sql
SELECT race_year,
       driver_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
 WHERE driver_name IN (SELECT driver_name FROM v_dominant_drivers WHERE driver_rank <= 10)
GROUP BY race_year, driver_name
ORDER BY race_year, avg_points DESC
```

▶ (5) Spark Jobs

Visualization ∨   +

▶ (5) Spark Jobs

Marketplace   **NEW**

Partner Connect

1/4 Tasks Completed

Enable new UI   **NEW**

3.viz_dominant_drivers SQL ∨

▸ Run all   ● dbkcourse_cluster ∨   🗓 Schedule   **Share**   ⌃

Cmd 6

```sql
SELECT race_year,
       driver_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
 WHERE driver_name IN (SELECT driver_name FROM v_dominant_drivers WHERE driver_rank <= 10)
GROUP BY race_year, driver_name
ORDER BY race_year, avg_points DESC
```

▸ (5) Spark Jobs

Visualization ∨   +

▸ (5) Spark Jobs

⬇ 118 rows | 1.65 seconds runtime   Refreshed 7 minutes ago

Command took 1.65 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 10:53:32 PM on dbkcourse_cluster

Cmd 7

```python
%python
html = """<h1 style="color:Black;text-align:center;font-family:Ariel">Report on Dominant Formula 1 Teams </h1>"""
displayHTML(html)
```

Report on Dominant Formula 1 Teams

Command took 0.07 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 11:03:18 PM on dbkcourse_cluster

Cmd 2

```sql
CREATE OR REPLACE TEMP VIEW v_dominant_teams
AS
SELECT team_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points,
       RANK() OVER(ORDER BY AVG(calculated_points) DESC) team_rank
  FROM f1_presentation.calculated_race_results_sql
GROUP BY team_name
HAVING COUNT(1) >= 100
ORDER BY avg_points DESC
```

OK

Command took 0.38 seconds -- by avinashchinta@my.unt.edu at 7/13/2023, 11:03:18 PM on dbkcourse_cluster

Cmd 3

```sql
SELECT * FROM v_dominant_teams;
```

▶ (3) Spark Jobs

Table ∨   +

	team_name	total_races	total_points	avg_points	team_rank
1	Mercedes	385	2931	7.612987012987013	1
2	Ferrari	1478	10500	7.104194857916103	2
3	McLaren	1032	6865	6.6521317829457365	3
4	Red Bull	433	2851	6.584295612009238	4

4.viz_dominant_teams SQL ▾

File   Edit   View   Run   Help   Last edit was 4 minutes ago   Provide feedback
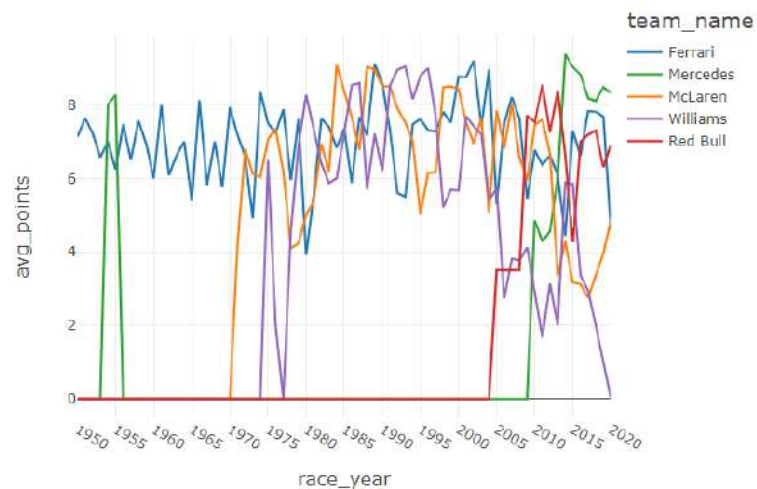
▶ Run all   ● dbkcourse_cluster ▾   📅 Schedule   Share

New
Workspace
Repos
Recents
Data
Compute
Workflows

Marketplace **NEW**
Partner Connect
1/4 Tasks Completed
Enable new UI **NEW**
Menu options

Cmd 4

```sql
SELECT race_year,
       team_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
 WHERE team_name IN (SELECT team_name FROM v_dominant_teams WHERE team_rank <= 5)
GROUP BY race_year, team_name
ORDER BY race_year, avg_points DESC
```

▶ (5) Spark Jobs

Visualization ▾   +

▶ (5) Spark Jobs

team_name
— Ferrari
— Mercedes
— McLaren
— Williams
— Red Bull

4.viz_dominant_teams SQL ∨

File Edit View Run Help   Last edit was 4 minutes ago   Provide feedback

▶ Run all   ● dbkcourse_cluster ∨   📅 Schedule   Share

New

Workspace

Repos

Recents

Data

Compute

Workflows

Command took 1.37 seconds — by avinashchinta@my.unt.edu at 7/13/2023, 11:03:18 PM on dbkcourse_cluster

Cmd 5

SQL ▶▼ ⬛ ∨ − ✕

```sql
SELECT race_year,
       team_name,
       COUNT(1) AS total_races,
       SUM(calculated_points) AS total_points,
       AVG(calculated_points) AS avg_points
  FROM f1_presentation.calculated_race_results_sql
 WHERE team_name IN (SELECT team_name FROM v_dominant_teams WHERE team_rank <= 5)
GROUP BY race_year, team_name
ORDER BY race_year, avg_points DESC
```

▶ (5) Spark Jobs

Visualization ∨   +

▶ (5) Spark Jobs
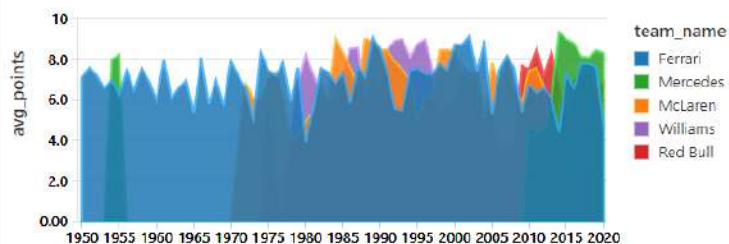
194 rows | 1.10 seconds runtime

Refreshed 4 minutes ago

Command took 1.10 seconds — by avinashchinta@my.unt.edu at 7/13/2023, 11:03:18 PM on dbkcourse_cluster

Marketplace

Partner Connect

1/4 Tasks Completed

New

Workspace

Repos

Recents

Data

Compute

Workflows

Marketplace NEW

Partner Connect

1/4 Tasks Completed

Enable new UI NEW

3.viz_dominant_drivers SQL ⌄
File   Edit   View   Run   Help   Provide feedback

▶ Run all   ● dbkcourse_cluster ⌄   📅 Schedule   **Share**   ⌃

Report on Dominant Formula 1 Drivers

Dominant Drivers

driver_name
- Jackie Stewart
- Niki Lauda
- Alain Prost
- Nigel Mansell
- Ayrton Senna
- Michael Schumac...
- Damon Hill
- Juan Pablo Montoya
- Lewis Hamilton
- Sebastian Vettel

avg_points

race_year

Dominant Drivers

driver_name
- Jackie Stewart
- Niki Lauda
- Alain Prost
- Nigel Mansell
- Ayrton Senna
- Michael Schumacher
- Damon Hill
- Juan Pablo Montoya
- Lewis Hamilton
- Sebastian Vettel

avg_points

F1 Dominant Drivers ✎

View of notebook:
3.viz_dominant_drivers

▶ Present Dashboard

Layout option:

Stack   Float

Dashboard width:

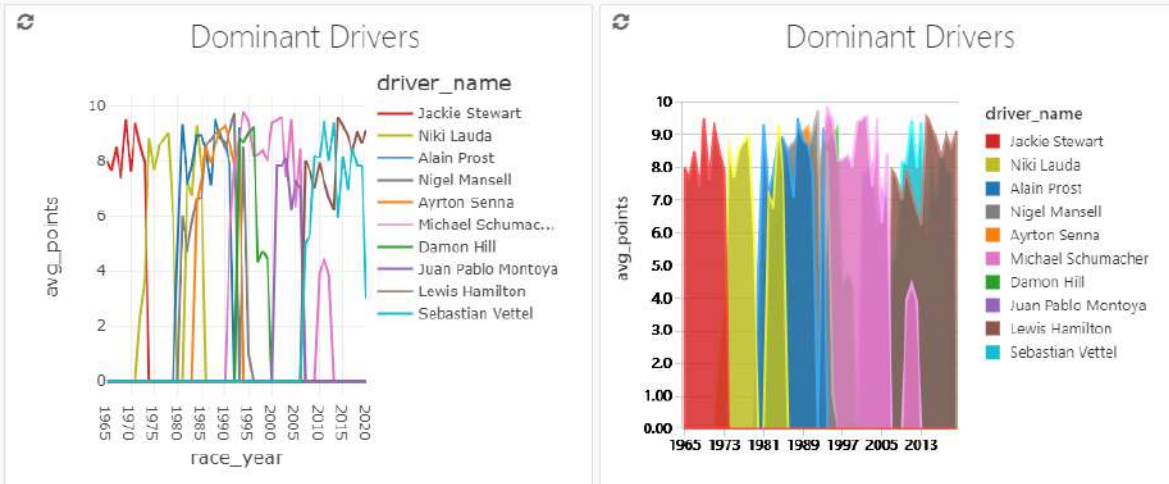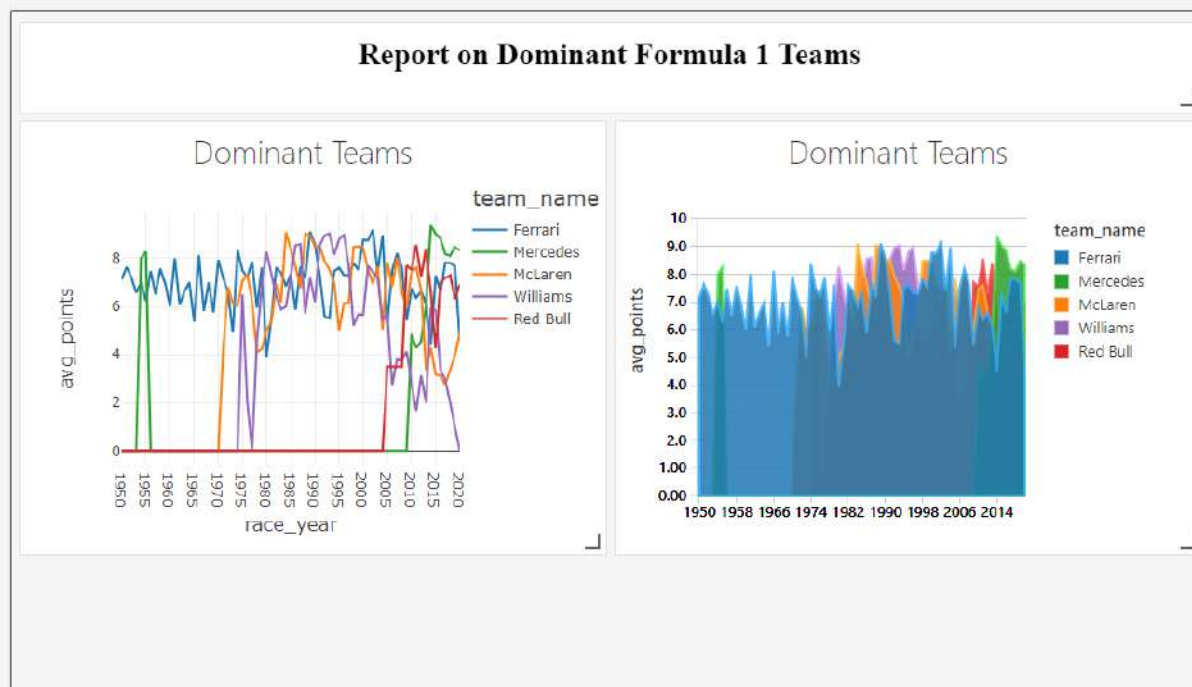1024px   ⌄

✏ Remove all graphs

🗑 Delete this dashboard

F1 Dominant Drivers

Report on Dominant Formula 1 Drivers