

Data Analytics Portfolio

Power BI : DAX, Power Query

- Project Overview
- Dataset
- Data Transformation (Power Query)
- Data Model & Relationships
- DAX Measures
- Final Dashboard & Key Insights

Katarzyna Szczęsna

Project Overview

Key Achievements:

Purpose:

The purpose of this project was to analyze the 2023 and 2024 Formula 1 seasons and extract actionable insights across drivers, teams, and races. The dashboard highlights key performance metrics, identifies trends, and supports strategic planning, performance optimization, and competitive benchmarking.

- Identified top-performing drivers and teams and highlighted patterns of consistency and peak performance
- Revealed gaps between qualifying and race performance, uncovering operational and strategic opportunities
- Provided interactive insights enabling scenario analysis and comparison across drivers and teams
- Demonstrated ability to translate complex datasets into actionable insights applicable to business performance tracking



Dataset

The dataset was sourced from Kaggle and consists of 16 individual CSV files (separate data for each season) and two Excel files created manually (images)

- Season Calendar: Grand Prix names, dates, and circuit details
- Race & Sprint Results: Driver positions, lap times, and race outcomes
- Sprint Shootout & Qualifying: Session times and grid positions
- Driver: Driver profiles including nationality, team, career highlights
- Team: Team details and historical performance
- Driver of the Day: Fan voting results per race
- Images: Manually created tables with driver and circuit images

A screenshot of Microsoft Excel showing a table titled "Formula1_2023season_raceR...". The table contains data for 442 races, with columns for Track, Position, No, Driver, Team, Starting Grid, Laps, Time/Retired, Points, Set Fastest Lap, and Fastest Lap Time. The data includes entries for various drivers like Abu Dhabi, Lewis Hamilton, and Valtteri Bottas across different teams like McLaren, Mercedes, and Haas.

A	B	C	D	E	F	G	H	I
426	Abu Dhabi	5,4	Lando Norris	McLaren Mercedes	5,58,+24.284,10	No,1:28.164		
427	Abu Dhabi	6,61	Oscar Piastri	McLaren Mercedes	3,58,+31.487,8	No,1:28.138		
428	Abu Dhabi	7,14	Fernando Alonso	Aston Martin Aramco Mercedes	7,58,+39.512,6	No,1:28.256		
429	Abu Dhabi	8,22	Yuki Tsunoda	AlphaTauri Honda RBPT	6,58,+43.088,4	No,1:29.256		
430	Abu Dhabi	9,44	Lewis Hamilton	Mercedes	11,58,+44.424,2	No,1:28.372		
431	Abu Dhabi	10,18	Lance Stroll	Aston Martin Aramco Mercedes	13,58,+55.632,1	No,1:28.050		
432	Abu Dhabi	11,3	Daniel Ricciardo	AlphaTauri Honda RBPT	15,58,+56.229,0	No,1:28.571		
433	Abu Dhabi	12,31	Esteban Ocon	Alpine Renault	12,58,+66.373,0	No,1:30.033		
434	Abu Dhabi	13,10	Pierre Gasly	Alpine Renault	10,58,+70.360,0	No,1:29.016		
435	Abu Dhabi	14,23	Alexander Albon	Williams Mercedes	14,58,+73.184,0	No,1:27.845		
436	Abu Dhabi	15,27	Nico Hulkenberg	Haas Ferrari	8,58,+83.696,0	No,1:29.217		
437	Abu Dhabi	16,21	Logan Sargeant	Williams Mercedes	20,58,+87.791,0	No,1:28.580		
438	Abu Dhabi	17,24	Guanyu Zhou	Alfa Romeo Ferrari	19,58,+89.422,0	No,1:28.746		
439	Abu Dhabi	18,55	Carlos Sainz	Ferrari	16,57,DFN,0	No,1:29.452		
440	Abu Dhabi	19,77	Valterri Bottas	Alfa Romeo Ferrari	18,57,+1 lap,0	No,1:29.863		
441	Abu Dhabi	20,20	Kevin Magnussen	Haas Ferrari	17,57,+1 lap,0	No,1:29.934		
442								

Data Transformation (Power Query)

- Merged separate source tables for the 2023 and 2024 seasons into a unified dataset and created a custom RaceID to distinguish seasons and races
- Standardized and corrected data types
- Performed pivot and unpivot transformations to optimize the data model
- Cleaned the dataset by removing, renaming, and reordering columns

A screenshot of the Power Query Editor showing the "Query Settings" pane and the "Applied Steps" pane. The "Properties" section shows the name "DriverOfTheDay". The "Applied Steps" pane lists the following steps:

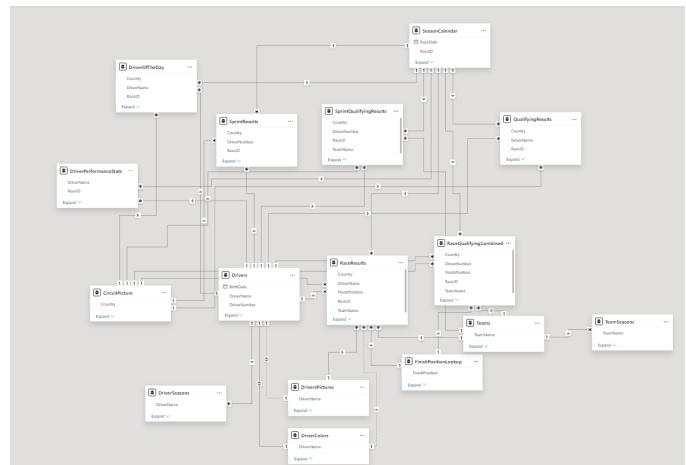
- Source
- Renamed Columns
- Replaced Value
- Added Custom
- Unpivoted Columns
- Split Column by Delimiter
- Changed Type
- Renamed Columns1
- Pivoted Column
- Renamed Columns2
- Replaced Value1
- Renamed Columns3
- Replaced Value2
- Extracted First Characters
- Changed Type1

```
= Table.NestedJoin#"Renamed Columns", {"Country", "Year"},  
SeasonCalendar, {"Country", "Year"}, "SeasonCalendar",  
JoinKind.LeftOuter)  
  
= Table.AddColumn#"Removed Columns", "RaceID", each [Country] & "_" &  
Text.From([Year]))
```

Data Model & Relationships

The Power BI semantic model was built on cleansed and transformed source data, with a focus on scalability, analytical flexibility, and performance.

- Designed a multi-table relational data model integrating race results, qualifying, sprint, and calendar data
- Used RaceID and DriverNumber as primary identifiers to ensure consistent relationships across seasons
- Structured the model to support cross-filtering, time-based analysis, and advanced DAX calculations



DAX Measures

- Developed dynamic performance KPIs for drivers and teams, including total points and ranking metrics
- Implemented statistical performance measures (average, min, max finishing positions) for comparative team analysis
- Built dynamic text measures to generate automated analytical insights within the dashboard
- Designed cumulative and ranking calculations to enable longitudinal season analysis

```
1 TeamsWithConsistentPointsText =
2 VAR TeamsWithPoints =
3   FILTER(
4     VALUES('RaceResults'[TeamName]),
5     CALCULATE([WorstTeamPoints]) > 0
6   )
7 VAR TeamList =
8   CONCATENATEX(TeamsWithPoints, 'RaceResults'[TeamName], " & ")
9 RETURN
10 TeamList & " are the only teams who scored points in each race"
```

```
1 TotalPoints =
2 CALCULATE(
3   SUM(RaceResults[RacePoints]),
4   FILTER(
5     RaceResults,
6     RaceResults[DriverName] = DriverSeasons[DriverName]
7       && RaceResults[Year] = DriverSeasons[Year]
8   )
9 )
```

```
1 Wins =
2 CALCULATE(
3   COUNTROWS(RaceResults),
4   RaceResults[FinishPositionNum] = 1
5 )
```

DAX Measures - Calculated Tables

Created FinishPositionLookup table to enable proper sorting and ranking of finish positions containing non-numeric values (DNF, NC).

```

1 FinishPositionLookup =
2 DATATABLE(
3     "FinishPosition", STRING, "FinishSort", INTEGER,
4     {
5         {"1", 1},
6         {"2", 2},
7         {"3", 3},
8         {"4", 4},
9         {"5", 5},
10        {"6", 6},
11        {"7", 7},
12        {"8", 8},
13        {"9", 9},
14        {"10", 10},
15        {"11", 11},
16        {"12", 12},
17        {"13", 13},
18        {"14", 14},
19        {"15", 15},
20        {"16", 16},
21        {"17", 17},
22        {"18", 18},
23        {"19", 19},
24        {"20", 20},
25        {"DQ", 21},
26        {"NC", 22}
27    }
)

```

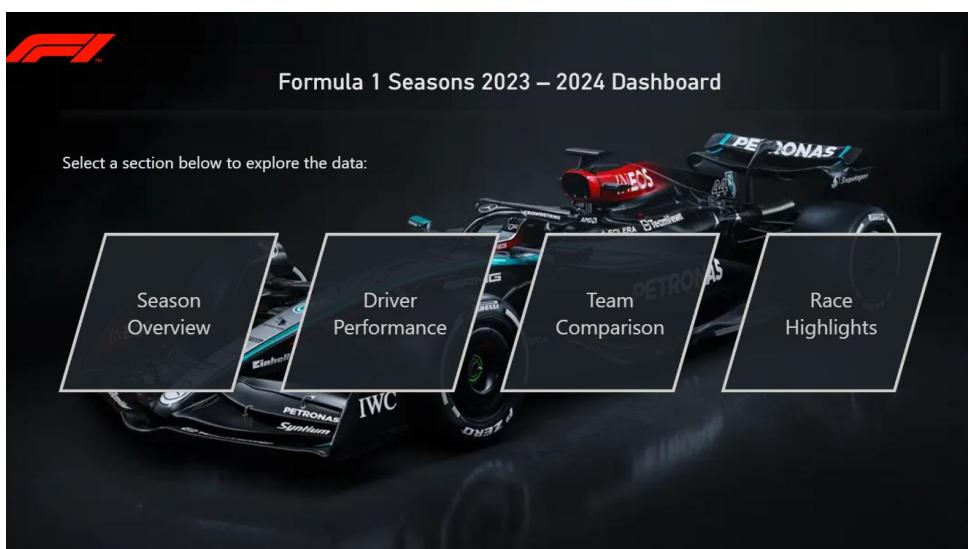
Created DriverPerformanceStats to aggregate driver achievements such as wins, podiums, and fastest laps for statistical analysis.

```

1 DriverPerformanceStats =
2 UNION(
3     SELECTCOLUMNS(
4         FILTER(RaceResults, RaceResults[FinishPositionNum] = 1),
5         "DriverName", RaceResults[DriverName],
6         "RaceID", RaceResults[RaceID],
7         "StatType", "win",
8         "Year", LOOKUPVALUE(SeasonCalendar[Year], SeasonCalendar[RaceID], RaceResults[RaceID]),
9         "GrandPrix", LOOKUPVALUE(SeasonCalendar[GrandPrix], SeasonCalendar[RaceID], RaceResults[RaceID]),
10        "Country", LOOKUPVALUE(SeasonCalendar[Country], SeasonCalendar[RaceID], RaceResults[RaceID])
11    ),
12    SELECTCOLUMNS(
13        FILTER(RaceResults, RaceResults[FinishPositionNum] IN {1, 2, 3}),
14        "DriverName", RaceResults[DriverName],
15        "RaceID", RaceResults[RaceID],
16        "StatType", "Podium",
17        "Year", LOOKUPVALUE(SeasonCalendar[Year], SeasonCalendar[RaceID], RaceResults[RaceID]),
18        "GrandPrix", LOOKUPVALUE(SeasonCalendar[GrandPrix], SeasonCalendar[RaceID], RaceResults[RaceID]),
19        "Country", LOOKUPVALUE(SeasonCalendar[Country], SeasonCalendar[RaceID], RaceResults[RaceID])
20    ),
21    SELECTCOLUMNS(
22        FILTER(RaceResults, RaceResults[FastestLapFlag] = "Yes"),
23        "DriverName", RaceResults[DriverName],
24        "RaceID", RaceResults[RaceID],
25        "StatType", "Fastest Lap",
26        "Year", LOOKUPVALUE(SeasonCalendar[Year], SeasonCalendar[RaceID], RaceResults[RaceID]),
27        "GrandPrix", LOOKUPVALUE(SeasonCalendar[GrandPrix], SeasonCalendar[RaceID], RaceResults[RaceID]),
28        "Country", LOOKUPVALUE(SeasonCalendar[Country], SeasonCalendar[RaceID], RaceResults[RaceID])
)
)

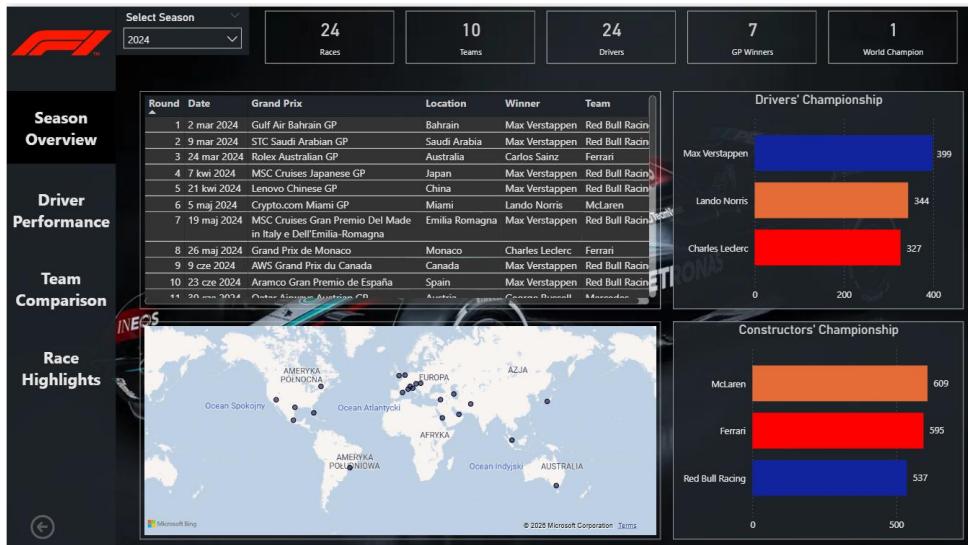
```

Final Dashboard & Key Insights



- Presents the dashboard structure
- Enables navigation to one of the four analytical sections

Season Overview



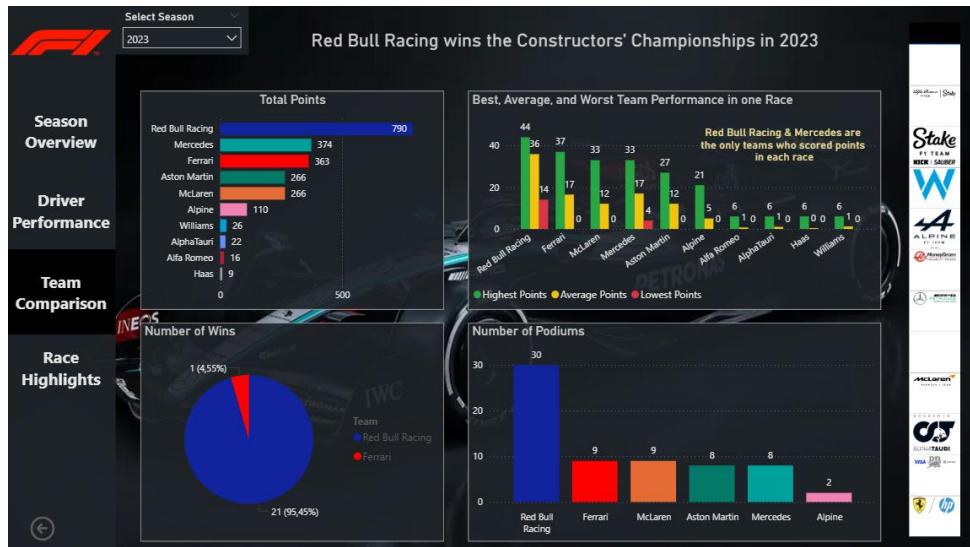
- Identified competitive dominance patterns across drivers and teams
- Analyzed global race distribution to assess market footprint and expansion opportunities
- Provided executive KPIs summarizing season outcomes and structure

Driver Performance



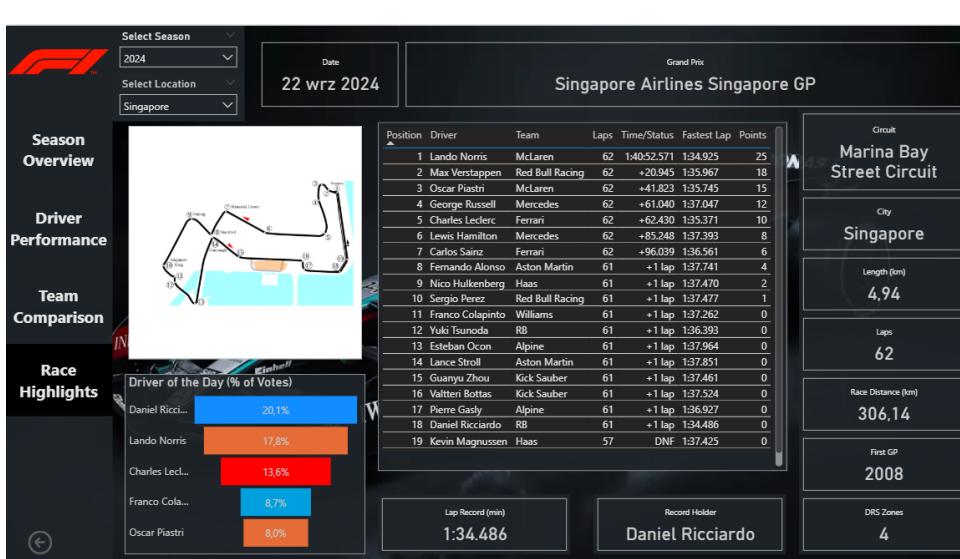
- Benchmarked drivers using points, wins, and podium metrics
- Analyzed consistency and cumulative performance trends across the season
- Identified gaps between qualifying potential and race execution
- Identified high-impact races and outlier results

Team Comparison



- Highlighted championship dominance driven by consistent team performance
- Showed trends in results to support decision-making and optimization
- Compared team performance to uncover gaps in strategy or execution

Race Highlights



- Extracted key track factors to guide technical and operational decisions
- Used fan-engagement data to surface commercial value beyond performance
- Surfaced race-specific performance patterns useful for forecasting future results