
# 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

**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.

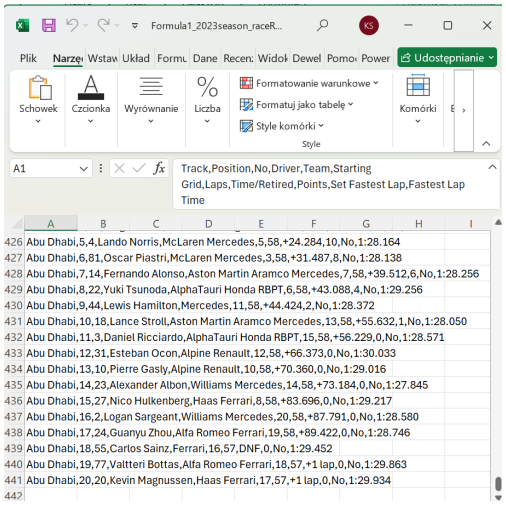
**Key Achievements:**

- 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

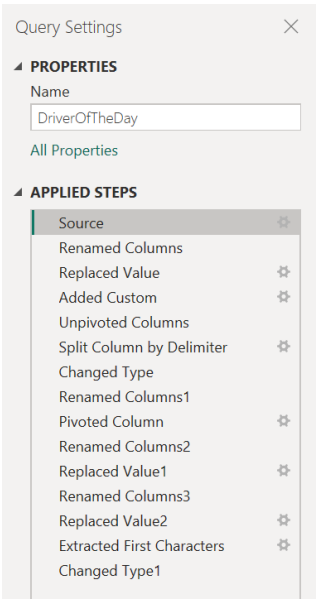



## Data Transformation (Power Query)

- 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

```
= Table.NestedJoin(#"Renamed Columns", {"Country", "Year"},
    SeasonCalendar, {"Country", "Year"}, "SeasonCalendar",
    JoinKind.LeftOuter)
```

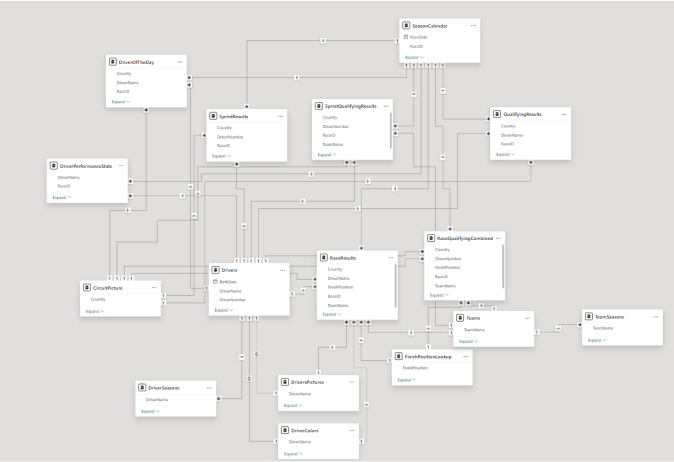
```
= Table.AddColumn(#"Removed Columns", "RaceID", each [Country] & "_" &
    Text.From([Year]))
```



## 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

```
TeamsWithConsistentPointsText =
VAR TeamsWithPoints =
    FILTER(
        VALUES('RaceResults'[TeamName]),
        CALCULATE([WorstTeamPoints]) > 0
    )
VAR TeamList =
    CONCATENATEX(TeamsWithPoints, 'RaceResults'[TeamName], " & ")
RETURN
    TeamList & " are the only teams who scored points in each race"
```

```
TotalPoints =
CALCULATE(
    SUM(RaceResults[RacePoints]),
    FILTER(
        RaceResults,
        RaceResults[DriverName] = DriverSeasons[DriverName]
            && RaceResults[Year] = DriverSeasons[Year]
    )
)
```

```
Wins =
CALCULATE(
    COUNTROWS(RaceResults),
    RaceResults[FinishPositionNum] = 1
)
```

## DAX Measures - Calculated Tables

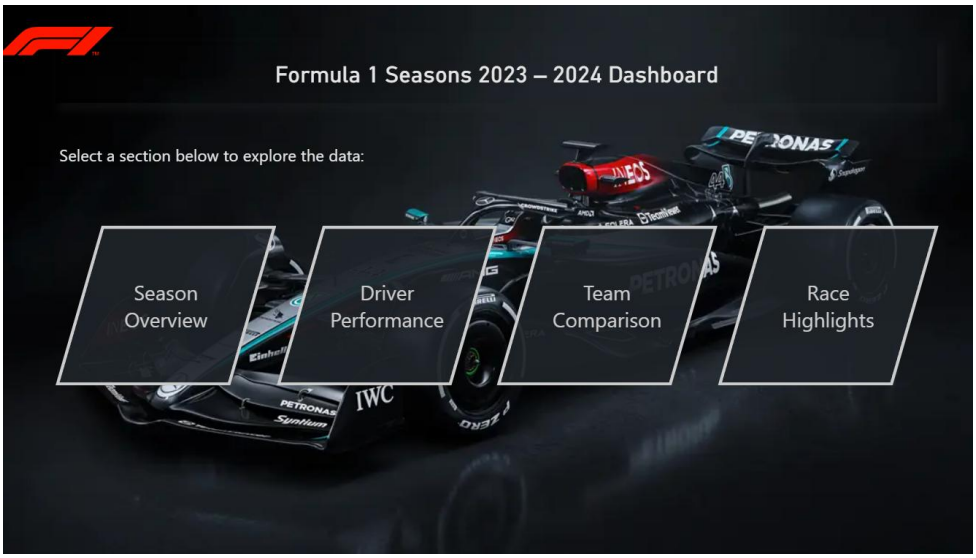
**Created FinishPositionLookup table to enable proper sorting and ranking of finish positions containing non-numeric values (DNF, NC).**

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

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

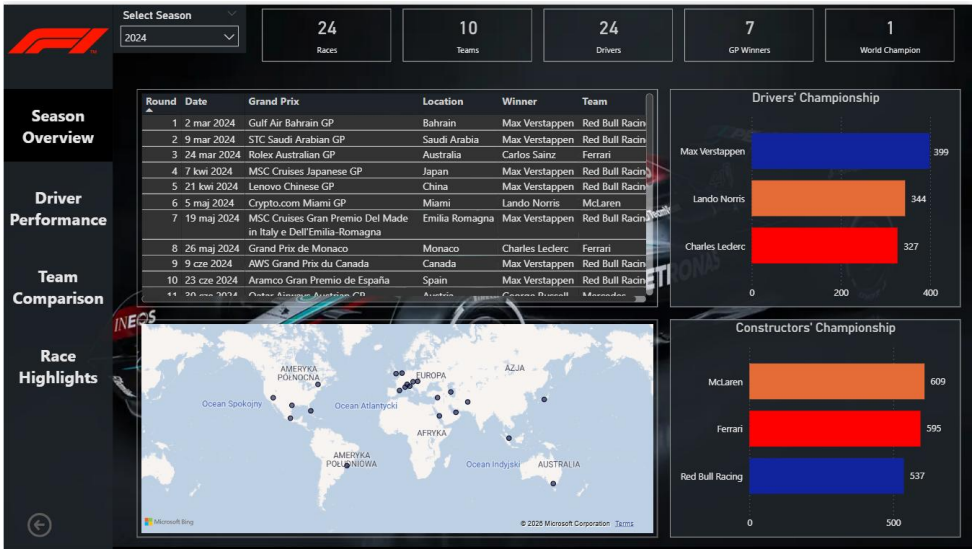
```
DriverPerformanceStats =
UNION(
    SELECTCOLUMNS(
        FILTER(RaceResults, RaceResults[FinishPositionNum] = 1),
        "DriverName", RaceResults[DriverName],
        "RaceID", RaceResults[RaceID],
        "StatType", "Win",
        "Year", LOOKUPVALUE(SeasonCalendar[Year], SeasonCalendar[RaceID], RaceResults[RaceID]),
        "GrandPrix", LOOKUPVALUE(SeasonCalendar[GrandPrix], SeasonCalendar[RaceID], RaceResults[RaceID]),
        "Country", LOOKUPVALUE(SeasonCalendar[Country], SeasonCalendar[RaceID], RaceResults[RaceID])
    ),
    SELECTCOLUMNS(
        FILTER(RaceResults, RaceResults[FinishPositionNum] IN {1, 2, 3}),
        "DriverName", RaceResults[DriverName],
        "RaceID", RaceResults[RaceID],
        "StatType", "Podium",
        "Year", LOOKUPVALUE(SeasonCalendar[Year], SeasonCalendar[RaceID], RaceResults[RaceID]),
        "GrandPrix", LOOKUPVALUE(SeasonCalendar[GrandPrix], SeasonCalendar[RaceID], RaceResults[RaceID]),
        "Country", LOOKUPVALUE(SeasonCalendar[Country], SeasonCalendar[RaceID], RaceResults[RaceID])
    ),
    SELECTCOLUMNS(
        FILTER(RaceResults, RaceResults[FastestLapFlag] = "Yes"),
        "DriverName", RaceResults[DriverName],
        "RaceID", RaceResults[RaceID],
        "StatType", "Fastest Lap",
        "Year", LOOKUPVALUE(SeasonCalendar[Year], SeasonCalendar[RaceID], RaceResults[RaceID]),
```

## Final Dashboard & Key Insights




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

## Season Overview

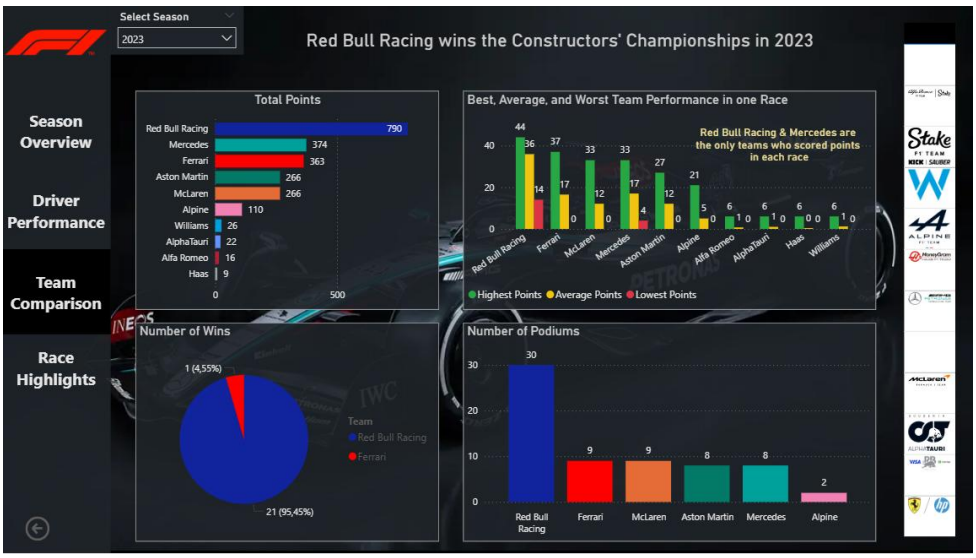



- Identified competitive dominance patterns across drivers and teams
- Analyzed global race distribution to assess market footprint and expansion opportunities
- Provided executive KPIs summarizing season outcomes and structure

## Driver Performance

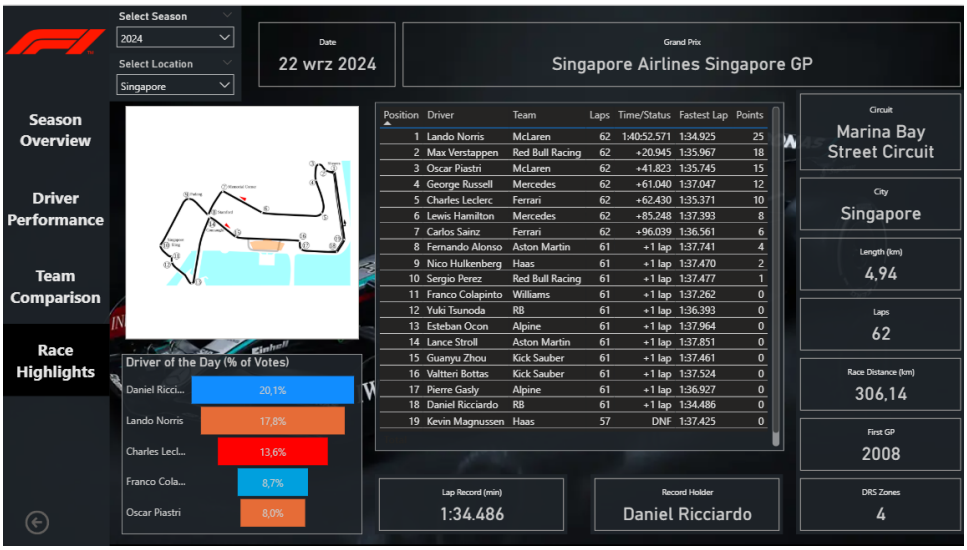

- Benchmarked drivers using points, wins, and podium metrics
- Analyzed consistency and cumulative performance trends across the season
- Identified gaps between qualifying potential and race execution
- Identified high-impact races and outlier results

## 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