Andrew Morcos

IST659\_Project\_Deliverable

Formula\_1

Summary:

Formula 1 driving is the epitome of speed, skill, politics and financial leverage married with ingenuity. In addition to the race, a plethora of data is generated such as preseason testing, practice rounds and qualifiers. All of this data such as lap times, top speed, fastest lap, seasonal point system, drivers, constructors must be organized into a database. The organization of this data will eventually lead to a concise deduction of the winner (pole position) and the rest of the top 10 which are the ones who are awarded points. With 20 drivers and 10 constructors, the stakes to keep a seat and to succeed are so high.

In this database journey, a mass amount of data has been collected since the 1950’s. For the scope of this project, we will focus on years of 2011-2018. So much data is collected in the world of Formula 1, so to dial it all down is a colossal feat. The statistics of each race by each driver and their constructors have been organized into a database system. Tracking of the date of the race, location, lap times, position, points awarded for each and every race has been done. The circuits themselves have been pinpointed by their exact location so that everything is accounted for. There are always two drivers for each constructor and 20 constructors per season. Drivers can switch constructors so you will notice some drivers are not on the same constructor for the year.

Conceptual Model:

Diagram

Description automatically generated

Normalized Model

|  |  |  |  |
| --- | --- | --- | --- |
| **Position** | **Driver** | **Constructor** | **Power Supply** |
| 1 | Lewis Hamilton | Mercedes | Mercedes AMG F1 M11 |
| 2 | Valtteri Bottas | Mercedes | Mercedes AMG F1 M11 |
| 3 | Max Verstappen | Red Bull | Honda RA620H |
| 7 | Alex Albon | Red Bull | Honda RA620H |
| 13 | Sebastian Vettel | Scuderia Ferrari | Ferrari 065 |
| 8 | Charles Leclerc | Scuderia Ferrari | Ferrari 065 |
| 6 | Carlos Sainz | McLaren | Renault E-Tech 20 |
| 9 | Lando Norris | McLaren | Renault E-Tech 20 |
| 5 | Daniel Ricciardo | Renault | Renault E-Tech 20 |
| 12 | Esteban Ocon | Renault | Renault E-Tech 20 |
| 10 | Pierre Gasly | AlphaTauri | Honda RA620H |
| 14 | Daniil Kvyat | AlphaTauri | Honda RA620H |
| 4 | Sergio Perez | Racing Point | BWT Mercedes |
| 11 | Lance Stroll | Racing Point | BWT Mercedes |
| 15 | Niko Hulkenberg | Racing Point | BWT Mercedes |
| 16 | Kimi Raikkonen | Alpha Romeo | Ferrari 065 |
| 17 | Antonio Giovinazzi | Alpha Romeo | Ferrari 065 |
| 19 | Romain Grosjean | Haas Racing | Ferrari 065 |
| 20 | Kevin Magnussen | Haas Racing | Ferrari 065 |
| 23 | Pietro Fittipaldi | Haas Racing | Ferrari 065 |
| 18 | George Russell | Williams | Mercedes AMG F1 M11 |
| 21 | Nicholas Latifi | Williams | Mercedes AMG F1 M11 |
| 22 | Jack Aitken | Williams | Mercedes AMG F1 M11 |

Database Creation Code (GitHub):

<https://gist.github.com/amorcos892/f0fd80f6d44285ab74ef371036f017e1>

Data Question Answers Code (GitHub):

<https://gist.github.com/amorcos892/3e87804021f15ade7456b9680aa90cad>

Data Questions:

Which circuit hosted the most Formula 1 races from 2011-2018?

Graphical user interface, text

Description automatically generated

How many races does each country and each circuit host from 2011-2018?

Text

Description automatically generated

Physical Database Design:

DROP TABLE IF EXISTS seasonsf1

DROP TABLE IF EXISTS resultsf1

DROP TABLE IF EXISTS racesf1

DROP TABLE IF EXISTS lapTimesf1

DROP TABLE IF EXISTS driverStandingsf1

DROP TABLE IF EXISTS driversf1

DROP TABLE IF EXISTS constructorStandingsf1

DROP TABLE IF EXISTS constructorsf1

DROP TABLE IF EXISTS constructorResultsf1

DROP TABLE IF EXISTS circuitsf1

USE [FORMULA1db]

GO

/\*\*\*\*\*\* Object: Table [dbo].[circuitsf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[circuitsf1](

[circuitId] [nvarchar](50) NOT NULL,

[circuitRef] [nvarchar](50) NOT NULL,

[name] [nvarchar](50) NOT NULL,

[location] [nvarchar](50) NOT NULL,

[country] [nvarchar](50) NOT NULL,

[lat] [float] NOT NULL,

[lng] [float] NOT NULL,

[alt] [int] NULL,

[url] [nvarchar](100) NOT NULL,

CONSTRAINT [PK\_circuitsf1] PRIMARY KEY CLUSTERED

(

[circuitId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[constructorResultsf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[constructorResultsf1](

[constructorResultsId] [int] NOT NULL,

[raceId] [int] NOT NULL,

[constructorId] [nvarchar](50) NOT NULL,

[points] [nvarchar](50) NOT NULL,

[status] [nvarchar](50) NOT NULL,

CONSTRAINT [PK\_constructorResultsf1] PRIMARY KEY CLUSTERED

(

[constructorResultsId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[constructorsf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[constructorsf1](

[constructorId] [nvarchar](50) NOT NULL,

[constructorRef] [nvarchar](50) NOT NULL,

[name] [nvarchar](50) NOT NULL,

[nationality] [nvarchar](50) NOT NULL,

[url] [nvarchar](100) NOT NULL,

CONSTRAINT [PK\_constructorsf1] PRIMARY KEY CLUSTERED

(

[constructorId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[constructorStandingsf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[constructorStandingsf1](

[constructorStandingsId] [int] NOT NULL,

[raceId] [int] NOT NULL,

[constructorId] [nvarchar](50) NOT NULL,

[points] [nvarchar](50) NOT NULL,

[position] [nvarchar](50) NOT NULL,

[positionText] [nvarchar](50) NOT NULL,

[wins] [nvarchar](50) NOT NULL,

CONSTRAINT [PK\_constructorStandingsf1] PRIMARY KEY CLUSTERED

(

[constructorStandingsId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[driversf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[driversf1](

[driverId] [nvarchar](50) NOT NULL,

[driverRef] [nvarchar](50) NOT NULL,

[number] [int] NULL,

[code] [nvarchar](50) NULL,

[forename] [nvarchar](50) NOT NULL,

[surname] [nvarchar](50) NOT NULL,

[dob] [nvarchar](50) NULL,

[nationality] [nvarchar](50) NOT NULL,

[url] [nvarchar](100) NULL,

CONSTRAINT [PK\_driversf1] PRIMARY KEY CLUSTERED

(

[driverId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[driverStandingsf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[driverStandingsf1](

[driverStandingsId] [int] NOT NULL,

[raceId] [int] NOT NULL,

[driverId] [nvarchar](50) NOT NULL,

[points] [nvarchar](50) NOT NULL,

[position] [nvarchar](50) NOT NULL,

[positionText] [nvarchar](50) NOT NULL,

[wins] [nvarchar](50) NOT NULL,

CONSTRAINT [PK\_driverStandingsf1] PRIMARY KEY CLUSTERED

(

[driverStandingsId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[lapTimesf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[lapTimesf1](

[raceId] [int] NOT NULL,

[driverId] [nvarchar](50) NOT NULL,

[lap] [nvarchar](50) NOT NULL,

[position] [nvarchar](50) NULL,

[time] [nvarchar](50) NOT NULL,

[milliseconds] [int] NULL

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[racesf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[racesf1](

[raceId] [int] NOT NULL,

[year] [int] NOT NULL,

[round] [nvarchar](50) NOT NULL,

[circuitId] [nvarchar](50) NOT NULL,

[name] [nvarchar](50) NOT NULL,

[date] [datetime2](7) NOT NULL,

[time] [datetime2](7) NOT NULL,

[url] [nvarchar](100) NULL,

CONSTRAINT [PK\_racesf1] PRIMARY KEY CLUSTERED

(

[raceId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[resultsf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

CREATE TABLE [dbo].[resultsf1](

[resultId] [int] NOT NULL,

[raceId] [int] NOT NULL,

[driverId] [nvarchar](50) NOT NULL,

[constructorId] [nvarchar](50) NOT NULL,

[number] [nvarchar](50) NOT NULL,

[grid] [nvarchar](50) NOT NULL,

[position] [nvarchar](50) NULL,

[positionText] [nvarchar](50) NOT NULL,

[positionOrder] [nvarchar](50) NOT NULL,

[points] [nvarchar](50) NOT NULL,

[laps] [nvarchar](50) NOT NULL,

[time] [nvarchar](50) NULL,

[milliseconds] [int] NULL,

[fastestLap] [int] NULL,

[rank] [nvarchar](50) NOT NULL,

[fastestLapTime] [nvarchar](50) NULL,

[statusId] [nvarchar](50) NOT NULL,

CONSTRAINT [PK\_resultsf1] PRIMARY KEY CLUSTERED

(

[resultId] ASC

)WITH (PAD\_INDEX = OFF, STATISTICS\_NORECOMPUTE = OFF, IGNORE\_DUP\_KEY = OFF, ALLOW\_ROW\_LOCKS = ON, ALLOW\_PAGE\_LOCKS = ON, OPTIMIZE\_FOR\_SEQUENTIAL\_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO

/\*\*\*\*\*\* Object: Table [dbo].[seasonsf1] Script Date: 6/27/2021 8:47:33 PM \*\*\*\*\*\*/

SET ANSI\_NULLS ON

GO

SET QUOTED\_IDENTIFIER ON

GO

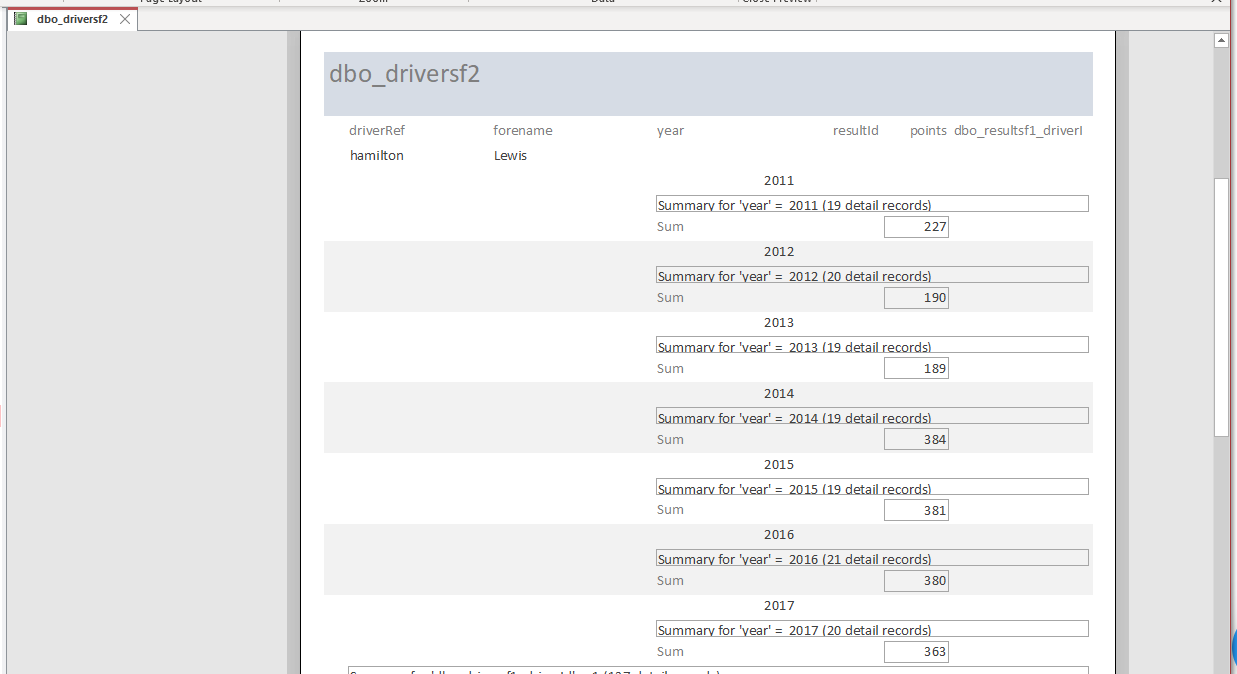
CREATE TABLE [dbo].[seasonsf1](

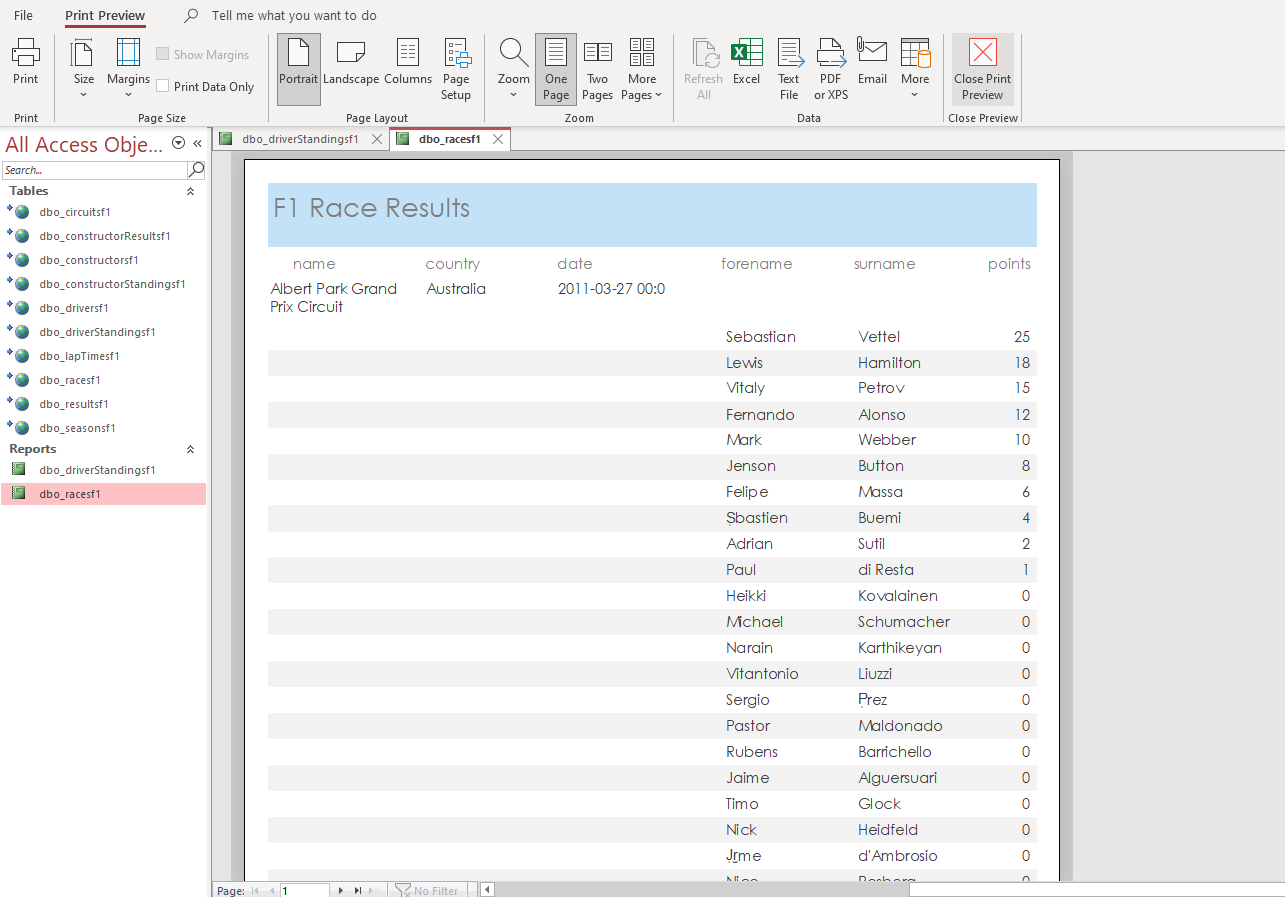
[year] [int] NOT NULL,

[url] [nvarchar](100) NOT NULL

) ON [PRIMARY]

GO





Reflection:

My data changed drastically from the start because of the amount of available data. My original scope changed due to some data that is not revealed by constructors such as details of the engine or specs of the car for that season. Being that I had to rely solely on statistics of the driver to that specific race and was able to narrow it down to that season as you see on the graphic when watching on tv. Knowing now how much detail goes into and the privacy made by Formula 1 teams and their engineers, I would have shifted my focus on the stats early on possibly doing some statistical analysis and making a predictive measure for future races.