

Text  Description automatically generated with low confidence

**Formula 1 Drivers DA Project**

**Submitted by- Uday Kumar**

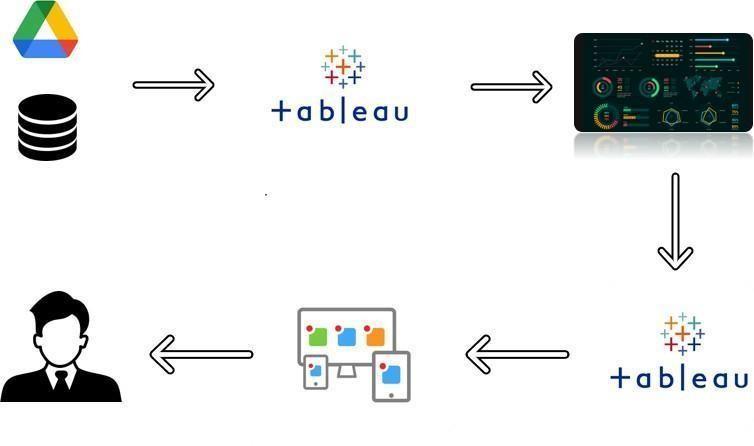
**F1 Drivers Analysis**

**Project Description:**

The “Formula 1 Drivers Analysis” project delves into the performance and accomplishments of Formula 1 drivers, uncovering the road to their triumphs. Through an examination of race wins, championships, podium finishes, and other metrics, this project aims to reveal patterns and insights that highlight the path to success in the highly esteemed motor sport competition. By analyzing the distribution of drivers' nationalities, tracking performance trends over seasons, exploring the correlations between performance metrics, and showcasing notable drivers and their contributions, we aim to paint a comprehensive picture of the Formula 1 drivers' journey. Through interactive visualizations and data-driven storytelling, this project offers a captivating exploration of the fascinating world of Formula 1 racing.

By presenting a comprehensive data story, the Formula 1 Drivers Analysis project not only highlights the individual achievements of drivers but also provides valuable insights into the sport as a whole. Through visualizations that capture the dynamics of nationality distribution, performance trends over seasons, and comparisons between drivers, we bring to light the factors that contribute to success in Formula 1 racing. The project offers a unique opportunity to appreciate the dedication, skill, and determination required to excel in this highly competitive and exhilarating sport, creating a captivating narrative that engages both motor sport enthusiasts and data enthusiasts alike.

**Technical Architecture:**



**Pre-Requisites:**

* Microsoft SQL server management studio - <https://youtu.be/VKWIiJUl70A>
* Create a Tableau Desktop Account with a student’s ID
* Create a Tableau Public account with a personal ID

**Prior Knowledge:**

Basic Concept of Tableau

<https://help.tableau.com/current/pro/desktop/en-us/gettingstarted_overview.htm>

**Project Objectives:**

By the end of this project, you will

* Know to connect any database with Tableau.
* Know fundamental concepts and can work on Tableau.
* Gain a broad understanding of plotting different graphs/charts.
* Able to create meaningful dashboards
* Able to integrate with web, using Bootstrap templates.

**Project Flow:**

To accomplish this, we have to complete all the activities listed below,

* + Define Problem / Problem Understanding
    - Specify the business problem
    - Business requirements
    - Literature Survey
  + Data Collection & Extraction from Database
    - Collect the dataset
    - Storing Data in DB
    - Perform SQL Operations
    - Connect DB with Tableau
  + Data Preparation
    - Prepare the Data for Visualization
  + Data Visualizations
    - No of Unique Visualizations
  + Dashboard
    - Responsive and Design of Dashboard
  + Story
    - No of Scenes of Story
  + Performance Testing
    - Amount of Data Rendered to tableau
    - Utilization of Data Filters
    - No of Calculation Fields
    - No of Visualizations/ Graphs
  + Web Integration
    - Dashboard and Story Integrating in Webpage and Flask Integration
  + Project Demonstration & Documentation
    - Record explanation Video for project end to end solution
    - Project Documentation-Step by step project development procedure

# Milestone 1: Define Problem / Problem Understanding

## Activity 1: Specify the business problem

Refer Project Description

## Activity 2: Business requirements

**Comprehensive Data Analysis:** The project should provide a thorough analysis of Formula 1 drivers' performance metrics, including race wins, championships, podium finishes, and other relevant factors. The analysis should cover multiple seasons and highlight trends, patterns, and notable achievements.

**Interactive Visualization:** The project should leverage interactive visualizations to allow users to explore and interact with the data. Users should be able to filter and drill down into specific drivers, seasons, or performance metrics to gain deeper insights and understand the factors contributing to success.

**Comparative Analysis:** The project should enable users to compare the performance of multiple drivers side by side. Users should be able to select specific drivers and view their race wins, podium finishes, points earned, or other metrics in a comparative format.

**Data Accuracy and Reliability:** The project should ensure the accuracy and reliability of the data used for analysis. Regular data updates and verification processes should be in place to maintain the integrity of the information presented.

**Data Security:** The project should adhere to data security best practices, ensuring the protection of sensitive or confidential information. Access controls and authentication mechanisms should be in place to restrict unauthorized access to the data and analysis.

## Activity 3: Literature Survey (Student Will Write)

A literature survey conducted by students exploring YouTube channels would typically involve researching existing studies, academic papers, and publications related to the topic.

# Milestone 2: Data Collection & Extraction from Database

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes and generate insights from the data.

## Activity 1: Collect the dataset

Please use the link to download the dataset:

<https://www.kaggle.com/datasets/dubradave/formula-1-drivers-dataset>

## Activity 1.1: Understand the data

Data contains all the meta information regarding the columns described in the CSV files. We have provided a csv file.

The data source of this project contains a list of Simpsons characters with roles and other features.

Fields include:

Driver: This column represents the name of the driver.

Nationality: It indicates the nationality of the driver.

Seasons: This column specifies the seasons in which the driver participated in Formula 1. It lists the years or seasons during which the driver competed.

Championships: It denotes the number of championships won by the driver.

Race\_Entries: This column represents the total number of races the driver entered.

Race\_Starts: It indicates the total number of races the driver actually started.

Pole\_Positions: This column indicates the number of pole positions earned by the driver. A pole position means starting a race from the front of the grid.

Race\_Wins: It denotes the number of races won by the driver.

Podiums: This column represents the total number of podium finishes achieved by the driver. A podium finish means finishing in the top three positions in a race.

Fastest\_Laps: It indicates the number of fastest laps achieved by the driver. The fastest lap is the lap with the quickest time during a race.

Points: This column represents the total number of points earned by the driver.

Active: It indicates whether the driver is currently active in Formula 1 (TRUE) or not (FALSE).

Championship Years: This column indicates the specific years in which the driver won a championship.

Decade: It represents the decade during which the driver was active in Formula 1.

Pole\_Rate: This column calculates the rate at which the driver achieved pole positions, expressed as a decimal.

Start\_Rate: It calculates the rate at which the driver started races, expressed as a decimal.

Win\_Rate: This column calculates the rate at which the driver won races, expressed as a decimal.

Podium\_Rate: It calculates the rate at which the driver achieved podium finishes, expressed as a decimal.

FastLap\_Rate: This column calculates the rate at which the driver achieved the fastest lap, expressed as a decimal.

Points\_Per\_Entry: It calculates the average number of points earned per race entry.

Years\_Active: This column represents the total number of years the driver was active in Formula 1.

Champion: It indicates whether the driver has won a championship (TRUE) or not (FALSE).

### Activity 2: Storing Data in DB & Connect DB with Tableau

**Explanation video link:**

<https://drive.google.com/file/d/1n2ba7mH9B5lHXDJQAtefa4YkM6pibX_R/view?usp=sharing>

# Milestone 3: Data Preparation

### Activity 1: Prepare the Data for Visualization

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into our analysis.

**Milestone 4: Data Visualization charts**

Data visualization is the process of creating graphical representations of data in order to help people understand and explore information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

The number of unique visualizations that can be created with a given dataset. Some common types of visualizations that can be used to analyze the performance and efficiency of project include bar charts, line charts, heat maps, scatter plots, pie charts, Maps etc. These visualizations can be used to compare performance, track changes over time, show distribution, and relationships between variables.

**Activity 1:** **Number of race wins**

**Explanation video link:**

<https://drive.google.com/file/d/1-sfB8VWiscVJIclC9GQ1H0qUsVpUE0Dc/view?usp=drive_link>

**Activity 2: Points per seasons**

**Explanation video link:**

<https://drive.google.com/file/d/1IeM3_Uee1dGSJORB-aMqg4btaHrqVCQG/view?usp=sharing>

**Activity 3: Pole positions vs Race wins**

**Explanation video link:**

<https://drive.google.com/file/d/12JT-ihOgj-HdoN9mNMzkxT2IhEatnHu6/view?usp=drive_link>

**Activity 4: Highest Championships**

**Explanation video link:**

<https://drive.google.com/file/d/1Ue310webotBxKlGFWlwFZge_GA7-Yorf/view?usp=drive_link>

**Activity 5: Podiums by seasons**

**Explanation video link:**

<https://drive.google.com/file/d/1gGxcddFAX6Y9yF77tMkdEd8l4ZmUT4ER/view?usp=drive_link>

**Activity 6: Seasonal Race Entries**

**Explanation video link:**

<https://drive.google.com/file/d/1yRT6kKc3M6lzIqYxhmaNE1-MjULiHc-x/view?usp=drive_link>

**Activity 7: Top 5 Race Winners**

**Explanation video link:**

<https://drive.google.com/file/d/1H3vwnQDc2xDLmpJgiYiwCBZbqLH9V3J7/view?usp=drive_link>

**Activity 8: Championships by Nation**

**Explanation video link:**

<https://drive.google.com/file/d/1fIG7dQA2BofoCZ765dXhxwIk9SD_Ns_R/view?usp=drive_link>

**Activity 9: Win Rate by Country**

**Explanation video link:**

<https://drive.google.com/file/d/12OwNuxZLs1fZ-lWy5opnYYRElEHCPes5/view?usp=drive_link>

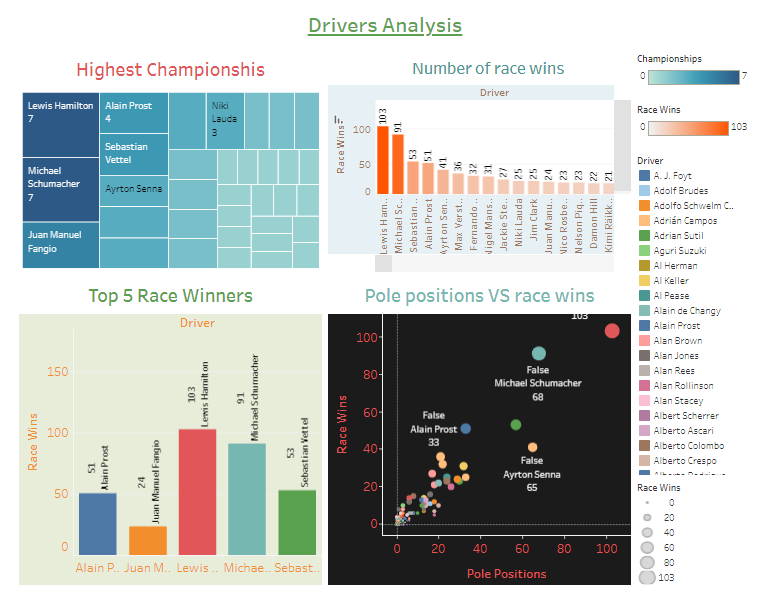
**Milestone 5: Create Tableau Dashboard**

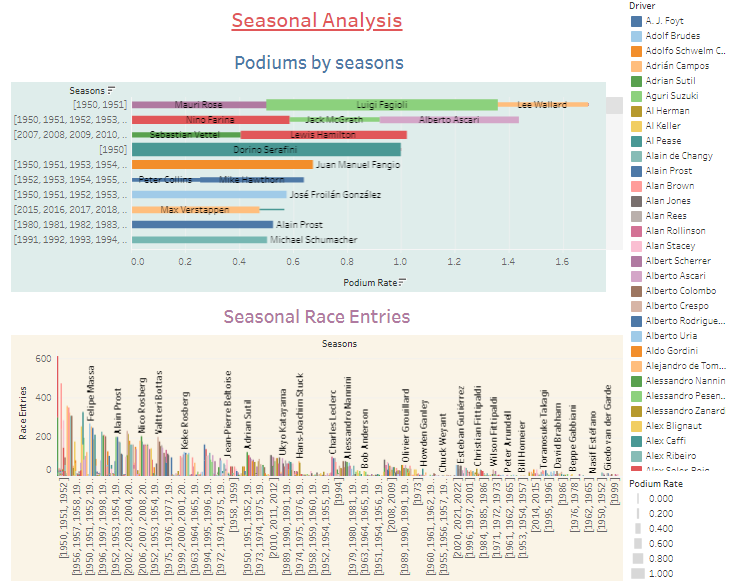
Once you have created views on different sheets in Tableau, you can pull them into a dashboard.

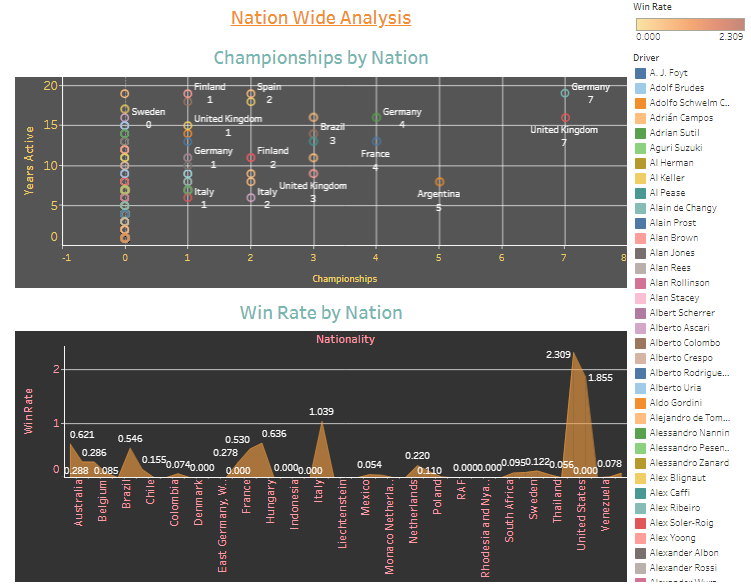
A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data, and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

## Activity :1- Responsive and Design of Dashboard

The responsiveness and design of a dashboard for Data-Driven insights on YouTube channels Analysis is crucial to ensure that the information is easily understandable and actionable. Key considerations for designing a responsive and effective dashboard include user-centered design, clear and concise information, interactivity, data-driven approach, accessibility, customization, and security. The goal is to create a dashboard that is user-friendly, interactive, and data-driven, providing actionable insights.







**Drivers Analysis Dashboard 1 Explanation video link:**

<https://drive.google.com/file/d/1CkGASvNS_HjiAIgk5U_ZWJDgpi-v_eHz/view?usp=drive_link>

**Seasonal Analysis Dashboard Explanation video link:**

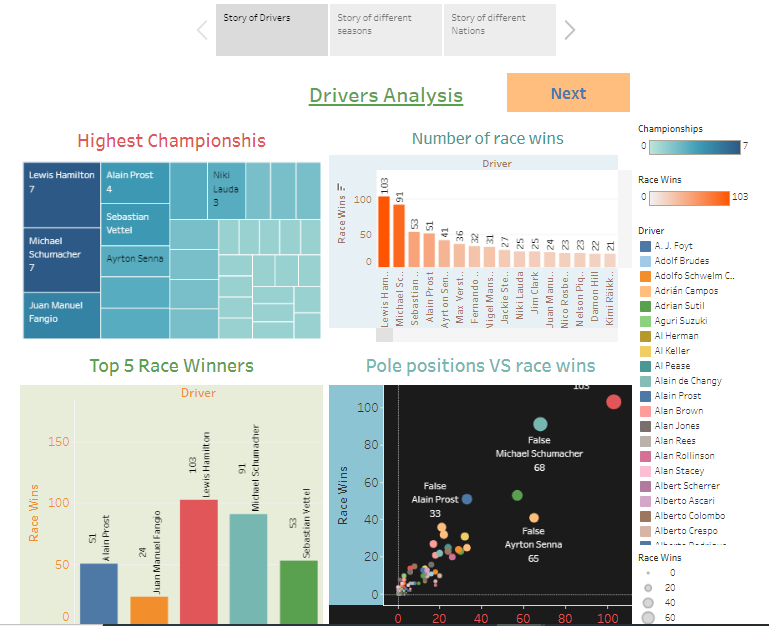
<https://drive.google.com/file/d/1-H6v4O75gmIPTaEAoae3koPaQL5api8O/view?usp=drive_link>

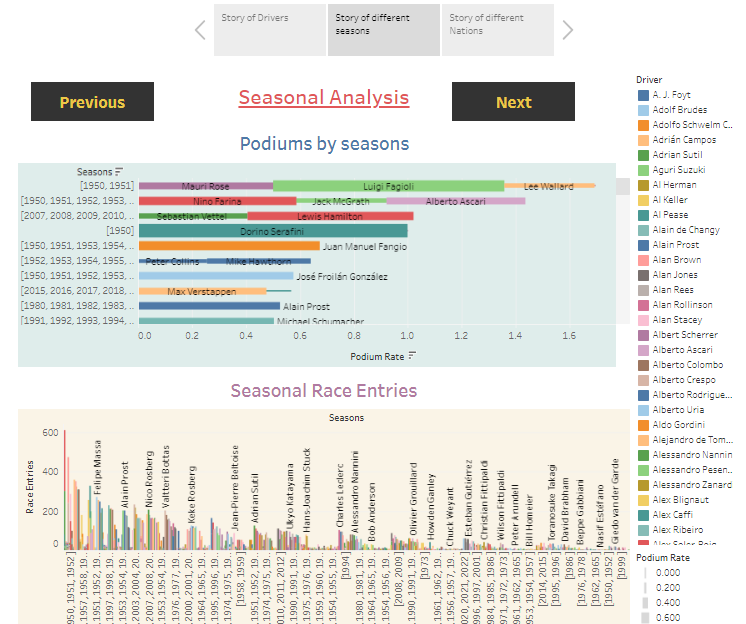
**Nation Wide Analysis Dashboard Explanation video link:**

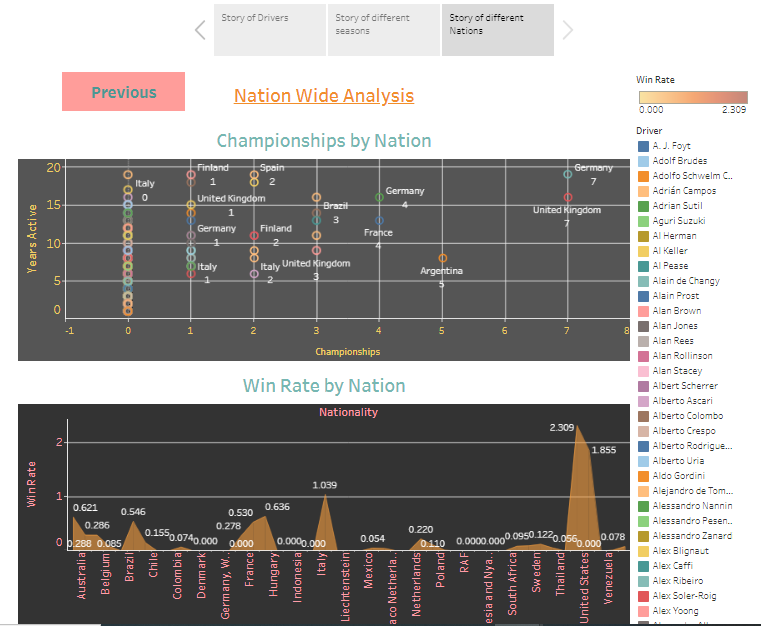
<https://drive.google.com/file/d/1mW1QJWdlgbxP4davt9PmoDf4V5AtJjQk/view?usp=drive_link>

**Milestone 6: Create Tableau Story**

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.







**Story for Drivers explanation video link 1:**

<https://drive.google.com/file/d/14yqYVGf8qSmuQB6sqyDijcR4ua1SUQuZ/view?usp=drive_link>

**Story of different Seasons explanation video link 1:**

<https://drive.google.com/file/d/1w5ZpImZ7Or7kOApNrMn1ffJZFtD1PWjD/view?usp=drive_link>

**Story for different Nations Explanation video link 1:**

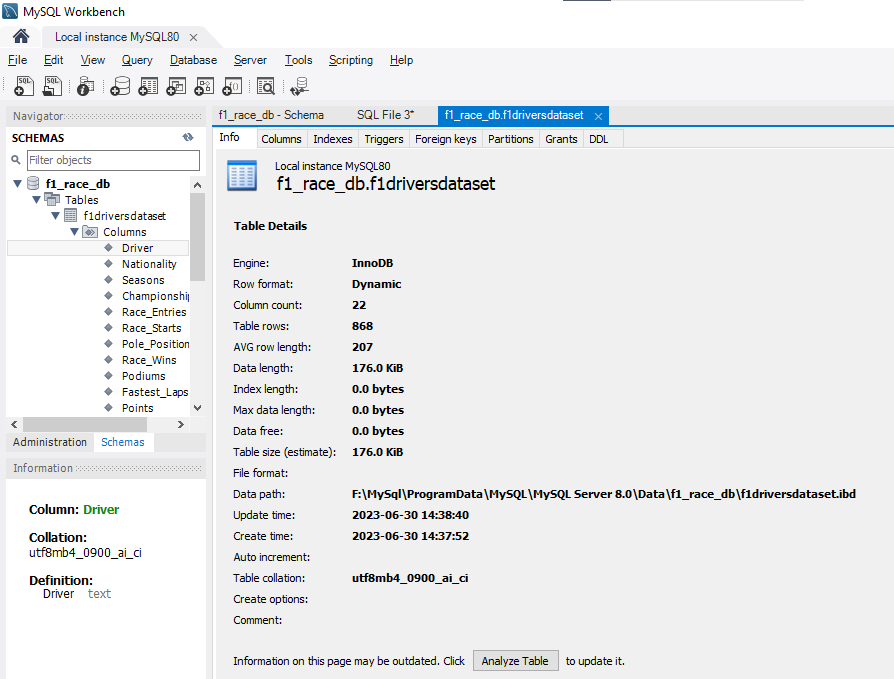
<https://drive.google.com/file/d/1sH4qVTfiYhsEHe6zGQw1zcOR7xPvtycA/view?usp=drive_link>

# Milestone 7: Performance Testing

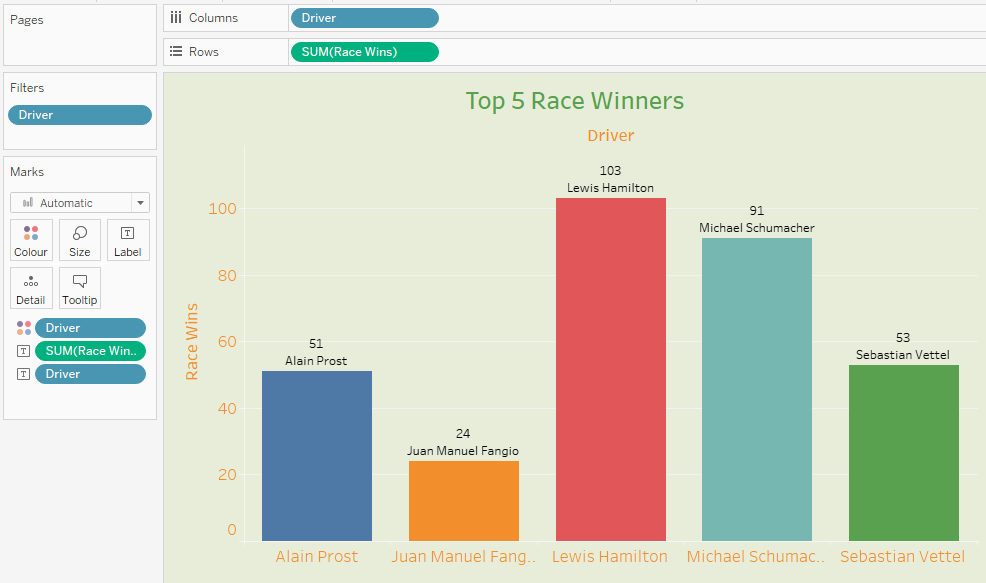
## Activity 1: Amount of Data Rendered to tableau

The amount of data that is rendered to a database depends on the size of the dataset and the capacity of the database to store and retrieve data.

Open the MySQL Workbench, go to the database then click to expand the tables, select the table and click on (i) button to get the information related to the table such as column count, table rows etc.



## Activity 2: Utilization of Filters



**Activity 3: No of Visualizations/ Graphs**

1. Bar graph depicts the number of race wins
2. Line chart is showing points per season
3. Scatter plot tells the correlation between pole positions and race wins
4. Tree map shows drivers highest number of championship
5. Vertical bar graph displays the Podiums by seasons
6. Stacked bars reveals the seasonal race entries
7. Bar chart for top 5 race winners
8. Circle shape chart for championships from different countries
9. Area chart tells the win rate by nation

# Milestone 8: Web integration

Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others.

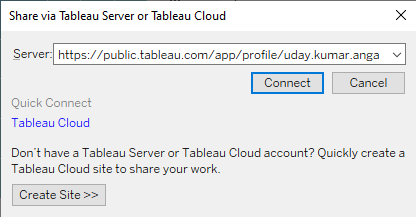
**Activity 1 : Create Tableau Public Account with personal ID**

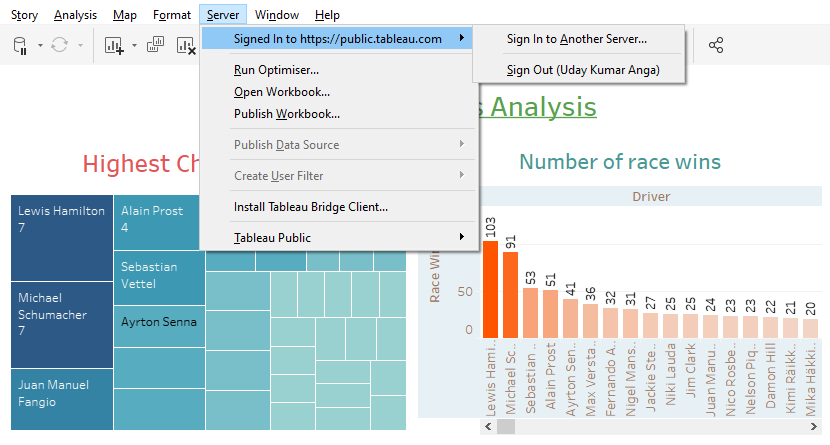
**Reference link:** <https://public.tableau.com/app/discover>

**Activity 2 : Publish into Tableau Public**

Step 1: Go to Dashboard/story, click on share button on the top ribbon.

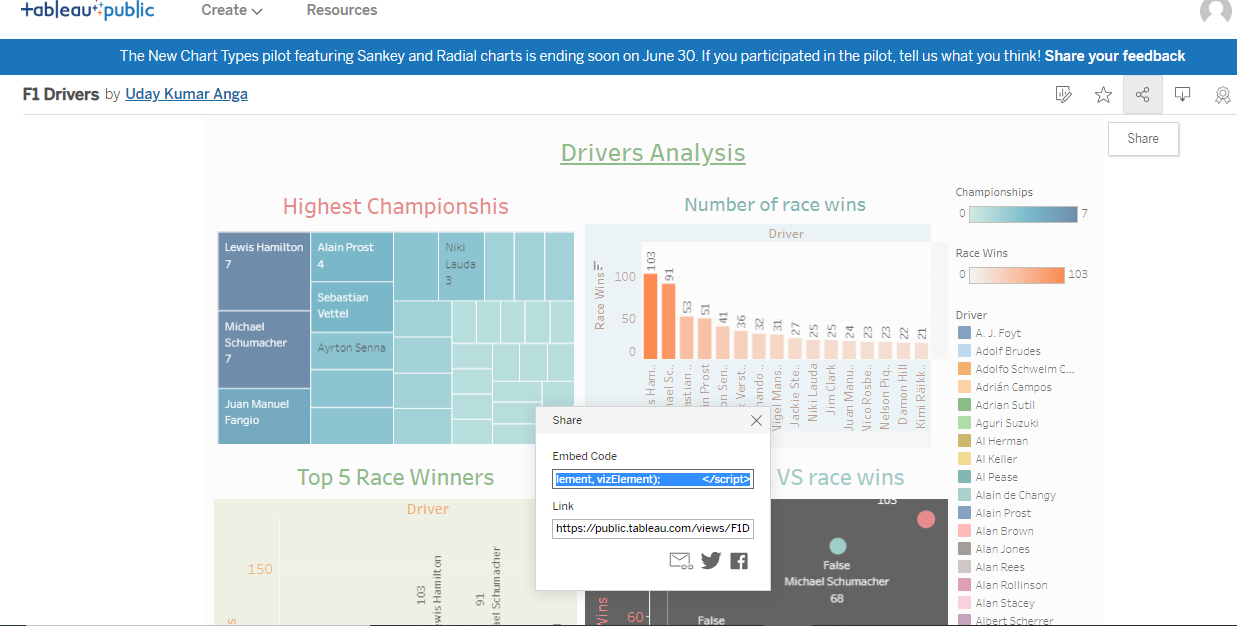
Step 2: Give the server address of your tableau public account and click on connect.





**Activity 3: Integrating Web with Embed code**

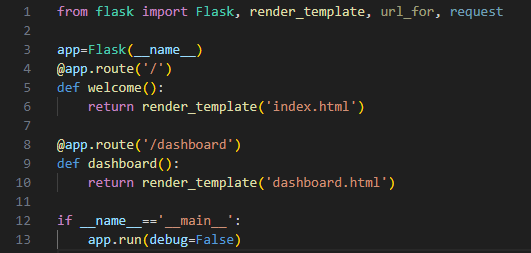
Step 1: Click on Share button and copy the Embed code

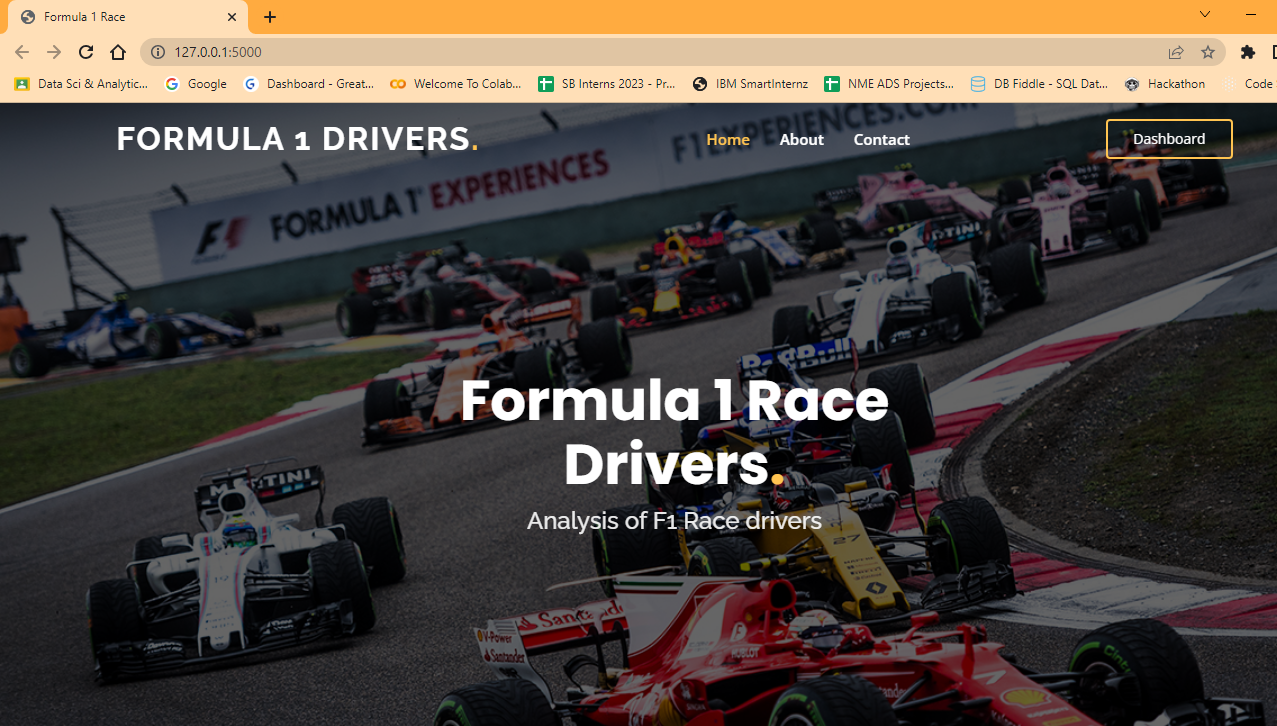


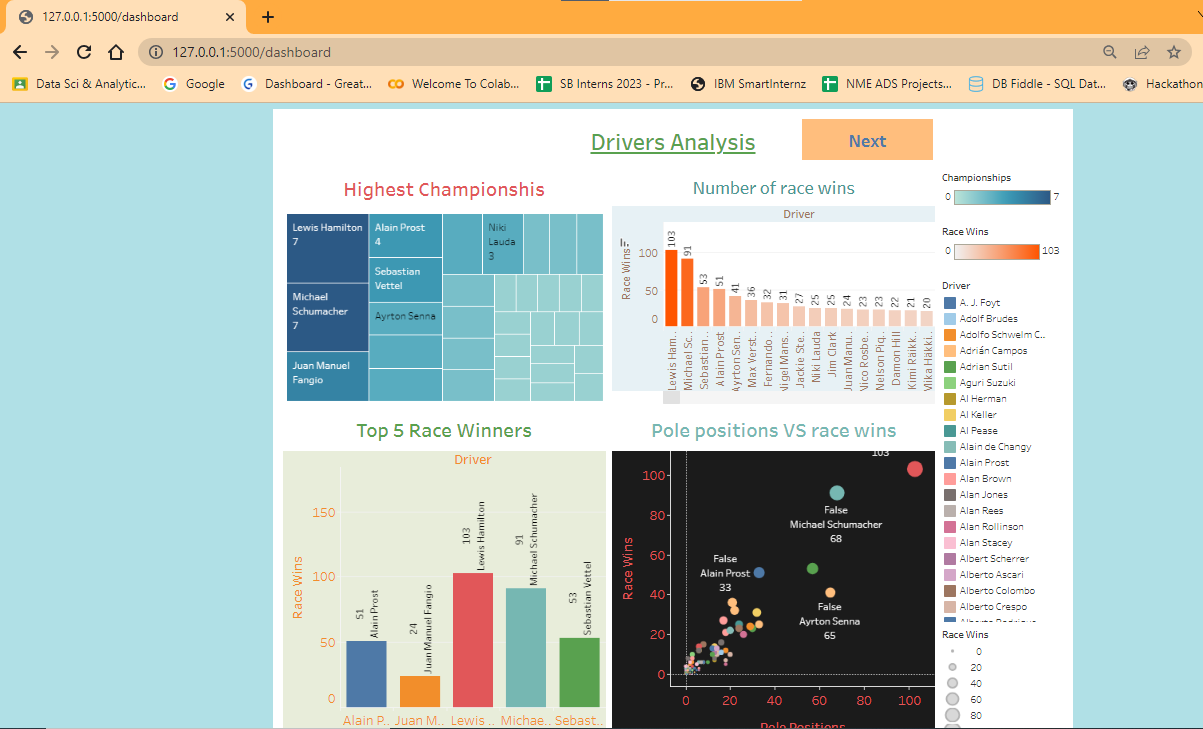
**Explanation video link:**

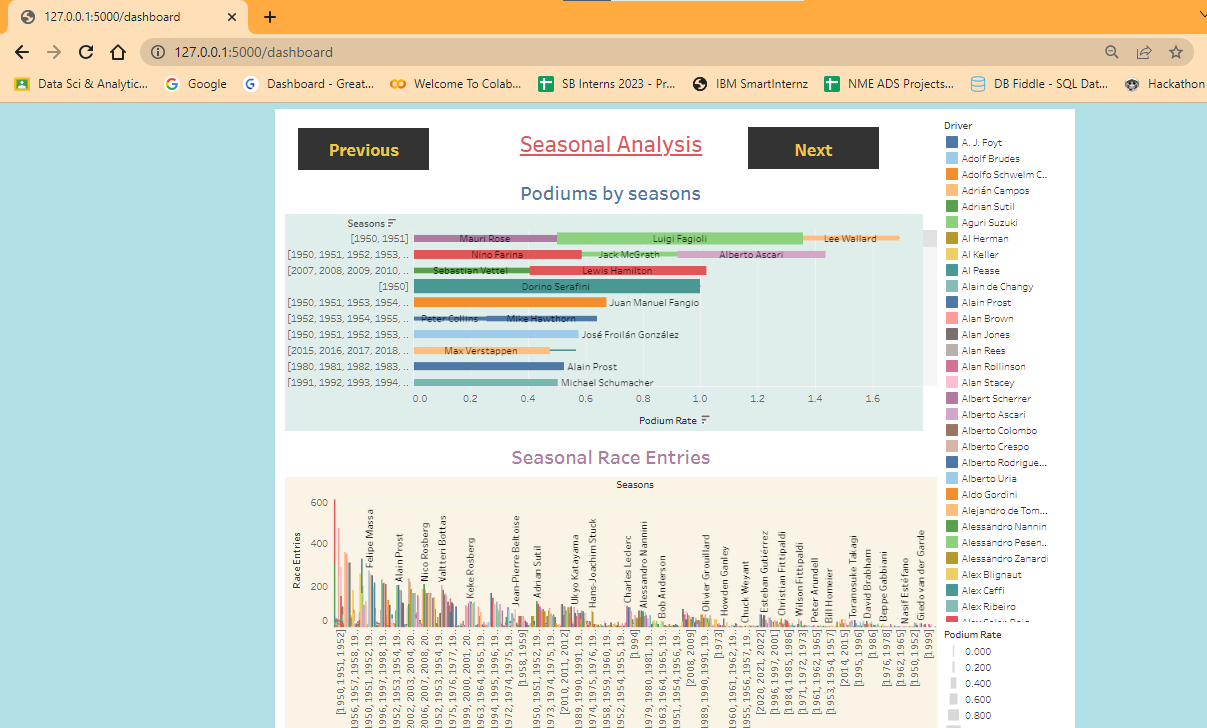
<https://drive.google.com/file/d/1nI6KHd-LUeGdDi_mZeIyLTvl00HB02JH/view?usp=drive_link>

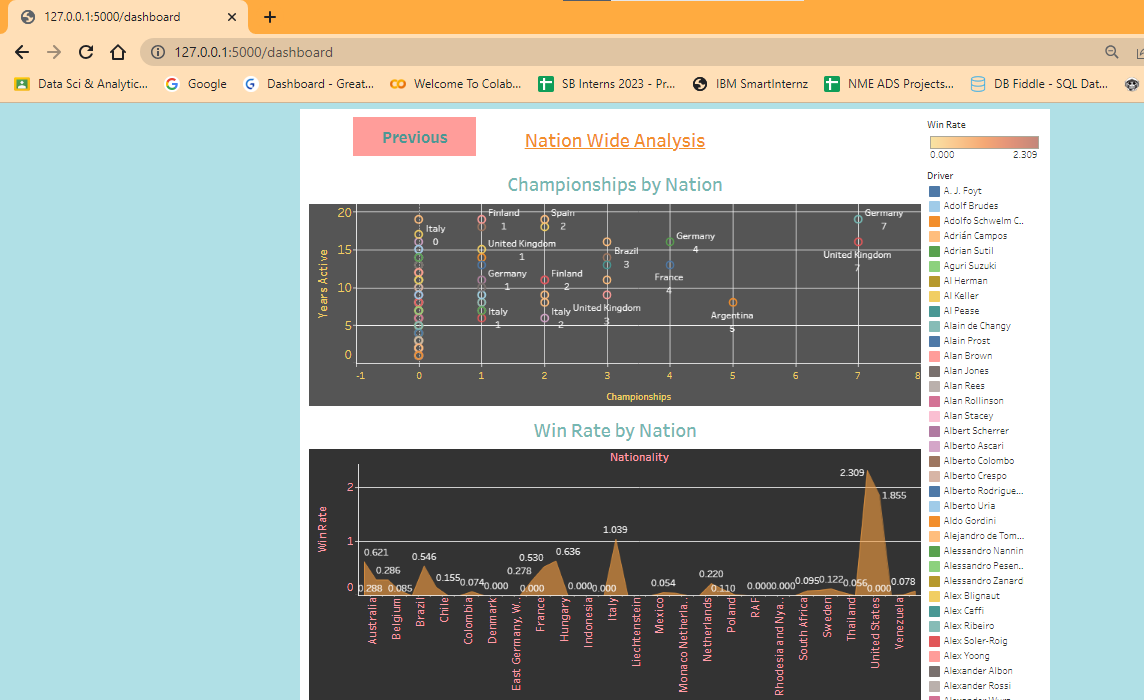
**Milestone 9: Using Bootstrap templates for Flask integration**











**Explanation video link:**

https://drive.google.com/file/d/1UWGVG5TwFm-UcjrQ3SZjIFW6yhSi40eO/view?usp=drive\_link

**Conclusion:**

In conclusion, the Formula 1 Drivers Analysis project offers a deep dive into the captivating realm of Formula 1 racing, unraveling the performance and achievements of its drivers. By leveraging interactive visualizations, comprehensive data analysis, and a compelling narrative, the project provides valuable insights into the factors that contribute to success on the race track. From geographic representations to comparative analyses and storytelling, it enables users to explore the journeys of these motor sport legends and appreciate the dedication and skill required to excel in Formula 1. By meeting the business requirements of comprehensive analysis, interactive visualizations, and user-friendly interfaces, the project delivers a rich and engaging experience, making it a valuable resource for motor sport enthusiasts, analysts, and stakeholders alike.

Sponsors and advertisers can leverage the project's visualizations and narratives to understand the popularity and marketability of drivers, enabling them to form effective partnerships. Media outlets and journalists can utilize the project's data-driven insights to enhance their coverage of Formula 1 races, adding depth and context to their stories. Overall, the project's comprehensive analysis and engaging visualizations make it a valuable tool for decision-makers, stakeholders, and enthusiasts looking to explore the fascinating world of Formula 1 drivers.