CSE 564 - Visualization & Visual Analytics

Project Proposal: Visualizing the Evolution of Formula 1 Through Interactive Dashboards

Background

Formula 1 racing, with its rich history spanning over seventy years, is not just about high-speed cars but about the evolution of technology, strategy, and human endeavor in motorsport. The Formula 1 World Championship dataset, covering from 1950 to 2023, provides an extensive record of races, drivers, lap times, standings, and more. This wealth of data offers a unique lens to examine aspects of performance, safety, technology, and team dynamics within the sport.

Source of Dataset:

The "Formula 1 World Championship (1950-2023)" dataset is a comprehensive collection of files detailing the various aspects of F1 races over a span of 70 years. This dataset includes a wide range of information about circuits, race results, driver standings, lap times, and more. Each file within the dataset is in CSV format, facilitating easy access and manipulation for data analysis purposes.

Link to Dataset: https:/formula-1-world-championship-1950-2023

Files and Their Contents

- **circuits.csv** - Contains information about the racing circuits, including location details and circuit statistics.
- **constructor_results.csv** - Provides results data specific to constructors (teams) for each race.
- **constructor_standings.csv** - Standings of constructors throughout the seasons, highlighting their competitive positions.
- **constructors.csv** - Details about the teams/constructors, including their names, nationalities, and foundational information.
- **driver_standings.csv** - Standings of individual drivers over the seasons, featuring points and ranking data.
- **drivers.csv** - Information on F1 drivers, including personal details, nationality, and code.
- **lap_times.csv** - Recorded lap times for drivers in each race, offering insights into performance during races.
- **pit_stops.csv** - Data on pit stops during races, detailing the strategy and efficiency of teams.
- **qualifying.csv** - Qualifying round details, which determine the starting positions of drivers for the races.
- **races.csv** - Information on each race event, including date, round, and location details.
- **results.csv** - Comprehensive results of races, featuring finishing positions, time differences, and points.
- **seasons.csv** - A list of seasons, typically linking to more detailed season reviews.
- **sprint_results.csv** - Results of sprint qualifying races, a recent addition to some race weekends.
- **status.csv** - Explanations for race retirements or classifications that denote the status of a driver's race completion.

Problem Statement

Despite the extensive data available on Formula 1, there lacks a comprehensive, interactive visualization that allows fans, analysts, and historians to explore the data dynamically to uncover trends,

patterns, and insights. Particularly, the relationships between driver safety, vehicle technology, team performance, and their evolution over decades are areas ripe for deeper analysis. Understanding these elements can enhance viewers' engagement and deepen their appreciation of the sport's complexity and the factors driving its changes.

Approach

The project aims to develop an interactive dashboard that will:

- **Visualize Performance Trends:** Use line plots, area charts, and bar charts to show changes in race times, team performances, and championship standings over the years.
- **Analyze Safety Improvements:** Deploy scatter plots and time series analysis to demonstrate how technological advancements have influenced safety, correlating these with major regulation changes and their impacts on race outcomes.
- **Highlight Driver Careers:** Implement parallel coordinates and sankey diagrams to trace drivers' careers across different teams and their performance metrics.
- **Compare Team Strategies:** Use stream graphs and scatterplot matrices to compare strategies and their effectiveness across different weather conditions, circuits, and eras.
- **Interactive Features:** Employ linked brushing and dynamic filtering to allow users to interactively select specific races, seasons, drivers, or teams, and see the detailed data and analytics update in real-time.

Non-standard Visualizations

- **Parallel Coordinates**: For multi-dimensional analysis of driver statistics.
- **Sankey Diagram:** To visualize the flow of drivers between teams across seasons.
- **Stream Graph:** To illustrate the ebb and flow of team performances over time.

What Do datasets look like?

Circuits.csv dataset looks as following:

circuitId	circuitRef	name	location	country	lat	lng	alt	url
1	albert_park	Albert Park Grand Prix Circuit	Melbourne	Australia	-37.8497	144.968	10	http://en.wikipedia.org/wiki/Melbourne_Grand_Prix_Circuit
2	sepang	Sepang International Circuit	Kuala Lumpur	Malaysia	2.76083	101.738	18	http://en.wikipedia.org/wiki/Sepang_International_Circuit
3	bahrain	Bahrain International Circuit	Sakhir	Bahrain	26.0325	50.5106	7	http://en.wikipedia.org/wiki/Bahrain_International_Circuit
4	catalunya	Circuit de Barcelona-Catalunya	MontmelÃ³	Spain	41.57	2.26111	109	http://en.wikipedia.org/wiki/Circuit_de_Barcelona-Catalunya
5	istanbul	Istanbul Park	Istanbul	Turkey	40.9517	29.405	130	http://en.wikipedia.org/wiki/Istanbul_Park
6	monaco	Circuit de Monaco	Monte-Carlo	Monaco	43.7347	7.42056	7	http://en.wikipedia.org/wiki/Circuit_de_Monaco
7	villeneuve	Circuit Gilles Villeneuve	Montreal	Canada	45.5	-73.5228	13	http://en.wikipedia.org/wiki/Circuit_Gilles_Villeneuve
8	magny_cours	Circuit de Nevers Magny-Cours	Magny Cours	France	46.8642	3.16361	228	http://en.wikipedia.org/wiki/Circuit_de_Nevers_Magny-Cours
9	silverstone	Silverstone Circuit	Silverstone	UK	52.0786	-1.01694	153	http://en.wikipedia.org/wiki/Silverstone_Circuit
10	hockenheimring	Hockenheimring	Hockenheim	Germany	49.3278	8.56583	103	http://en.wikipedia.org/wiki/Hockenheimring
11	hungaroring	Hungaroring	Budapest	Hungary	47.5789	19.2486	264	http://en.wikipedia.org/wiki/Hungaroring
12	valencia	Valencia Street Circuit	Valencia	Spain	39.4589	-0.331667	4	http://en.wikipedia.org/wiki/Valencia_Street_Circuit
13	spa	Circuit de Spa-Francorchamps	Spa	Belgium	50.4372	5.97139	401	http://en.wikipedia.org/wiki/Circuit_de_Spa-Francorchamps
14	monza	Autodromo Nazionale di Monza	Monza	Italy	45.6156	9.28111	162	http://en.wikipedia.org/wiki/Autodromo_Nazionale_Monza
15	marina_bay	Marina Bay Street Circuit	Marina Bay	Singapore	1.2914	103.864	18	http://en.wikipedia.org/wiki/Marina_Bay_Street_Circuit
16	fuji	Fuji Speedway	Oyama	Japan	35.3717	138.927	583	http://en.wikipedia.org/wiki/Fuji_Speedway
17	shanghai	Shanghai International Circuit	Shanghai	China	31.3389	121.22	5	http://en.wikipedia.org/wiki/Shanghai_International_Circuit
18	interlagos	AutÃ³dromo JosÃ© Carlos Pace	SÃ£o Paulo	Brazil	-23.7036	-46.6997	785	http://en.wikipedia.org/wiki/Aut%C3%B3dromo_Jos%C3%A9_Carlos_Pace
19	indianapolis	Indianapolis Motor Speedway	Indianapolis	USA	39.795	-86.2347	223	http://en.wikipedia.org/wiki/Indianapolis_Motor_Speedway
20	nurburgring	NÃ¼rburgring	NÃ¼rburg	Germany	50.3356	6.9475	578	http://en.wikipedia.org/wiki/N%C3%BCrburgring
21	imola	Autodromo Enzo e Dino Ferrari	Imola	Italy	44.3439	11.7167	37	http://en.wikipedia.org/wiki/Autodromo_Enzo_e_Dino_Ferrari
22	suzuka	Suzuka Circuit	Suzuka	Japan	34.8431	136.541	45	http://en.wikipedia.org/wiki/Suzuka_Circuit
80	vegas	Las Vegas Strip Street Circuit	Las Vegas	United States	36.1147	-115.173	\N	https://en.wikipedia.org/wiki/Las_Vegas_Grand_Prix#Circuit

Drivers.csv dataset as follows:

driverId	driverRef	number	code	forename	surname	dob	nationality	url
1	hamilton	44	HAM	Lewis	Hamilton	1/7/1985	British	http://en.wikipedia.org/wiki/Lewis_Hamilton
3	rosberg	6	ROS	Nico	Rosberg	6/27/1985	German	http://en.wikipedia.org/wiki/Nico_Rosberg
4	alonso	14	ALO	Fernando	Alonso	7/29/1981	Spanish	http://en.wikipedia.org/wiki/Fernando_Alonso
8	raikkonen	7	RAI	Kimi	RÃ¤ikkÃ¶nen	10/17/1979	Finnish	http://en.wikipedia.org/wiki/Kimi_R%C3%A4ikk%C3%B6nen
9	kubica	88	KUB	Robert	Kubica	12/7/1984	Polish	http://en.wikipedia.org/wiki/Robert_Kubica
13	massa	19	MAS	Felipe	Massa	4/25/1981	Brazilian	http://en.wikipedia.org/wiki/Felipe_Massa
16	sutil	99	SUT	Adrian	Sutil	1/11/1983	German	http://en.wikipedia.org/wiki/Adrian_Sutil
18	button	22	BUT	Jenson	Button	1/19/1980	British	http://en.wikipedia.org/wiki/Jenson_Button
20	vettel	5	VET	Sebastian	Vettel	7/3/1987	German	http://en.wikipedia.org/wiki/Sebastian_Vettel
154	grosjean	8	GRO	Romain	Grosjean	4/17/1986	French	http://en.wikipedia.org/wiki/Romain_Grosjean
155	kobayashi	10	KOB	Kamui	Kobayashi	9/13/1986	Japanese	http://en.wikipedia.org/wiki/Kamui_Kobayashi
842	gasly	10	GAS	Pierre	Gasly	2/7/1996	French	http://en.wikipedia.org/wiki/Pierre_Gasly
807	hulkenberg	27	HUL	Nico	HÃ¼lkenberg	8/19/1987	German	http://en.wikipedia.org/wiki/Nico_H%C3%BClkenberg
813	maldonado	13	MAL	Pastor	Maldonado	3/9/1985	Venezuelan	http://en.wikipedia.org/wiki/Pastor_Maldonado
815	perez	11	PER	Sergio	PÃ©rez	1/26/1990	Mexican	http://en.wikipedia.org/wiki/Sergio_P%C3%A9rez
817	ricciardo	3	RIC	Daniel	Ricciardo	7/1/1989	Australian	http://en.wikipedia.org/wiki/Daniel_Ricciardo
818	vergne	25	VER	Jean-Ã‰ric	Vergne	4/25/1990	French	http://en.wikipedia.org/wiki/Jean-%C3%89ric_Vergne
820	chilton	4	CHI	Max	Chilton	4/21/1991	British	http://en.wikipedia.org/wiki/Max_Chilton
821	gutierrez	21	GUT	Esteban	GutiÃ©rrez	8/5/1991	Mexican	http://en.wikipedia.org/wiki/Esteban_Guti%C3%A9rrez
822	bottas	77	BOT	Valtteri	Bottas	8/28/1989	Finnish	http://en.wikipedia.org/wiki/Valtteri_Bottas
824	jules_bianchi	17	BIA	Jules	Bianchi	8/3/1989	French	http://en.wikipedia.org/wiki/Jules_Bianchi
825	kevin_magnussen	20	MAG	Kevin	Magnussen	10/5/1992	Danish	http://en.wikipedia.org/wiki/Kevin_Magnussen
826	kvyat	26	KVY	Daniil	Kvyat	4/26/1994	Russian	http://en.wikipedia.org/wiki/Daniil_Kvyat

Dataset Relationships:

- **Circuits to Races**: One-to-Many. Each circuit can host many races, but each race occurs at one circuit.
- **Races to Results**: One-to-Many. Each race has many results (one for each driver).
- **Drivers to Results**: One-to-Many. A driver can have many results across different races.
- **Constructors to Results**: One-to-Many. A constructor can have many results across different races.
- **Races to Qualifying**: One-to-Many. Each race has a set of qualifying results.
- **Drivers to Driver Standings**: One-to-Many. Drivers can appear in many entries in the standings across different races and seasons.
- **Constructors to Constructor Standings**: One-to-Many. Constructors can appear in many entries in the standings.
- **Races to Lap Times**: One-to-Many. Each race has multiple lap times.
- **Drivers to Lap Times**: One-to-Many. Drivers generate multiple lap times.
- **Races to Pit Stops**: One-to-Many. Each race can have multiple pit stops.
- **Drivers to Pit Stops**: One-to-Many. Drivers can have multiple pit stops in a race.
- **Races to Sprint Results**: One-to-Many. Races with sprint qualifying have multiple sprint results entries.
- **Results to Status**: Many-to-One. Multiple results can have the same status (e.g., finished, did not finish).

These interconnected datasets allow for a composite view of the Formula 1 world. For a visualization project, these relationships enable:

- Creation of rich, layered stories about drivers, teams, and the sport itself.
- Multifaceted analysis of how different factors such as technology, regulations, and strategies impact performance.
- Exploration of complex questions like the correlation between qualifying performance and race results, or the impact of pit strategy on race outcomes.
- Interactive visualizations where selections in one view dynamically update related data in another, providing a seamless and intuitive user experience.

Dashboard Overview:

1. Header Section:

- Title: "Decades of Formula 1: An Interactive Exploration."
- Subtitle:"Discover the history, analyze trends, and explore the evolution of F1 from 1950 to 2023."

2. Navigation Bar:

- Filters: Dropdown menus for the user to filter the data by year, race, driver, constructor, and circuit.
- Reset Button: To clear all selections and filters.
- Help Icon: When clicked, displays instructions on how to interact with the dashboard.

3. Main Dashboard Panels:

Left Panel - Driver and Constructor Standings:

- Driver Standings: A bar chart or line graph showing the top drivers by season.
- Constructor Standings: A parallel bar chart displaying team standings over selected seasons.
- Interactive feature: Hovering over a driver or team highlights their performance across other visualizations.

Central Panel - Race Analysis:

- Circuit Performance Map: A geo-map with circuits highlighted; selecting a circuit displays performance data.
- Race Results Overview: A line plot or area chart to display the performance trends across races in a season.
- Safety and Technology Timeline: A timeline that marks significant safety and technological advancements in F1.
- Interactive feature: Clicking on a year or race on the timeline updates the central panel to reflect data from that period.

Right Panel - Detailed Race Data:

- Lap Time Analysis: A line chart showing average or fastest lap times per race or per season.
- Pit Stop Strategy: Scatter plots correlating pit stop times with final race positions.
- Interactive feature: Selecting a driver or race in the lap time analysis automatically filters the pit stop strategy data for related insights.

Bottom Panel - Career and Historical Analysis:

- Driver Career Trajectories: A sankey diagram to trace the movement of drivers between teams over the years.
- Team Performance Stream: A stream graph illustrating the rise and fall of team performances through time.
- Qualifying vs. Race Outcome: A scatter plot matrix showing the correlation between qualifying positions and race results.
- Interactive feature: Linked brushing allows selecting elements in one visualization to highlight related data in all others.