

CS 6630: Project Proposal - Formula 1 Stats

Team Members: Harshini Keerthi Vasan, Vinu Sreenivasan, Sanjana Aravindan

Basic Info

Project Title

Formula 1 Stats

Team Members

1. Sanjana Aravindan, u1135274, u1135274@utah.edu
2. Harshini Keerthi Vasan, u1136693, u1136693@utah.edu
3. Vinu Sreenivasan, u1138972, vinu.sreenivasan@utah.edu

Project Repository

<https://github.com/HarshiniVasu/Formula-1-Stats>

Background and Motivation

We, as a team are ardent fans of Formula 1 racing. This racing is a perfect blend of technology and the automobile industry. The mere adrenaline rush that the viewers get watching this sport, is something that gets us excited about working on this project. We are inspired enough to visualize various interesting correlations between top players of the league and eminent teams like Ferrari, Mercedes and McLaren.

The dataset we are planning to use for this project has instigated us with diverse ideas of interactive visualizations between racers, cars, race tracks and constructors (car teams).

Project Objectives

We would like to use the knowledge we have in d3 to visualize the formula 1 data set. This would help us understand interesting correlations between winning teams and the most prominent racers in the game. We will use concepts covered in class, assignments and also a few other d3 based visualizations to analyze the dataset and

build an effective user interface with interesting visualizations of formula 1's records, historic races, prominent players and the major-league teams.

Data

We did a significant amount of research and found a set of multiple csv files related to formula 1 racing from Ergast Developer API. These csv files provide mapping of various sets of data using primary keys. There is a circuit (race tracks) data file which comprises of a unique key id, the track location of the races with its latitude and longitude. The constructors file (f1 teams) has fields like id, nationality of the team and their standings game-wise and year-wise (1950-2016). The racer's data file includes the history of their games (wins and losses), the team they raced for, their standings in global level year wise as well as match wise. There is another interesting csv file which provides the status of the racers with respect to every game denoting the number of laps, damages, collisions, etc.

Data Processing

We are planning to do a bit of data cleansing and, re-organize the data according to our visualization specification. We intend to perform some normalization over the scoring system such that the data is in the right format to visualize various charts and graphs. We also plan to segregate the data and accumulate everything into a smaller set of csv files while ensuring that the reduced dataset will be effective enough to visualize the interesting relations in the dataset.

Visualization Design

Bar Charts

Individual Driver statistics

A shift in the position of an individual racer over the years can be visualized using bar chart.

Team statistics

A similar bar chart can be used for team stats where for each team, its position over the years is recorded and plotted.

We also plan to plot a bar graph between the number of races and number of wins for each team that participated over years from 1950-2017.

Cons

Sometimes chart space will be a limiting factor for comparing multiple drivers. If multiple drivers are selected, it will be difficult to plot the bar for each one of them.

Line Charts

Individual Driver Statistics

Line charts are a good way of depicting the relationship between two different attributes.

The first line chart below provides an illustration about the shift in the position of the individual racer over different locations of the F1 racing track. The second chart unfolds similar relation as that of the bar chart with an indication of the points in the y-axis.

Team Statistics

The line chart for team statistics depicts the shift of points over the years for each team.

Cons

Comparing performances using line chart can be strenuous sometimes if multiple teams or players are picked for comparison. Its hard to follow the trend if line charts are crossing each other. Hence, to distinguish the lines we can use different colors.

Bubble chart

Individual Driver Statistics

This chart uncovers the number of laps successfully completed by an individual racer over different years. The diameter of the bubble indicates the number of laps completed.

Cons

The size of the bubbles are hard to compare, therefore making it hard to understand the correlation.

Sankey Diagram

Individual Driver Statistics

This is a very interesting visualization that groups and illustrates the formula 1 winners over the years from 1950 to 2017.

Team Statistics

We also intend to visualize the relationship between various teams and players over the years using this diagram.

Geo-spatial maps

Race statistics

The location of the race tracks (circuits) of all the races for a particular year is depicted on the map. Along with it, we plan to add extra effects such as a hover-effect where hovering over the location will display the details of the racer and the team that won the race. This map may reveal many relationships between the different race tracks and might also illustrate about the performance of the racer who won over these tracks.

Comparison between drivers for each year

A geo-spatial map can be used for the comparison of performance of the drivers for a particular year over different race tracks. This will help us understand about the location where the racers generally perform really well and win most of the times.

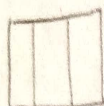
Collapsible Tree

We plan on building a collapsible tree whose root node is 'Teams' (constructor). Clicking on the node will expand all the teams that have participated in the Formula 1 racing. Further, clicking on one of the teams expands the tree further, displaying the details of the racers/drivers associated with the team. In a similar way, clicking on one of the racers will expand the statistical records such as finishing time, position, points. Clicking on a node will expand and collapse the tree thus making it a very interactive interface for visualizing the teams, the drivers/racers and their statistical records.

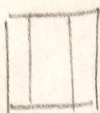
Champions - National flags

We plan to extract the flag of the nations associated with every team that won the race over the years. Going further, we intend to map the respective flag image with every year's record to portray the teams that won the race in that particular year.

Driver Champions:



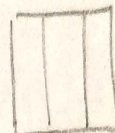
1990



1991



1992



1993



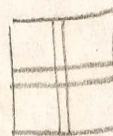
1994



1995



1996



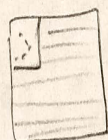
1997



1998



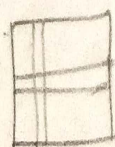
1999



2000



2001



2002



2003

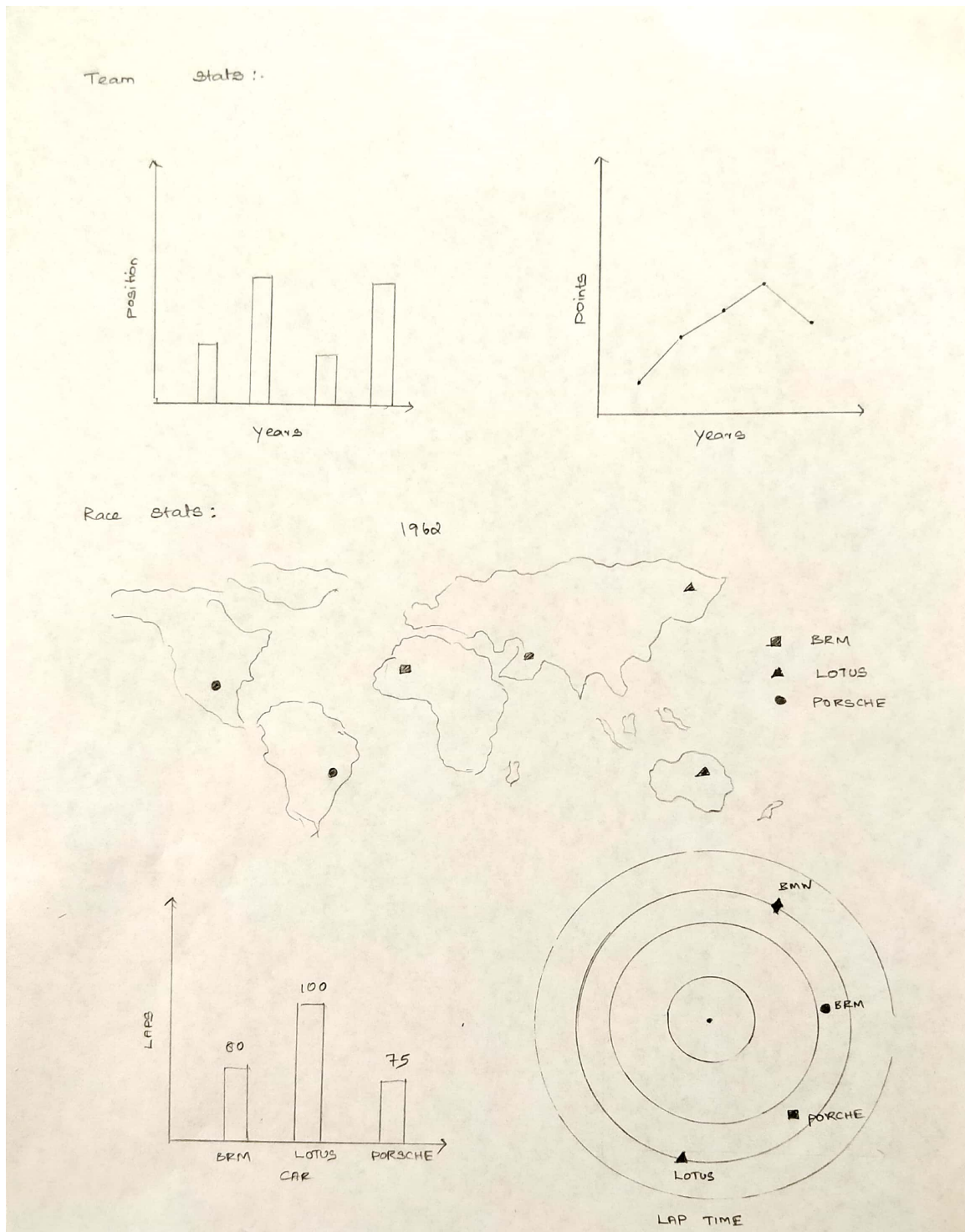


2004

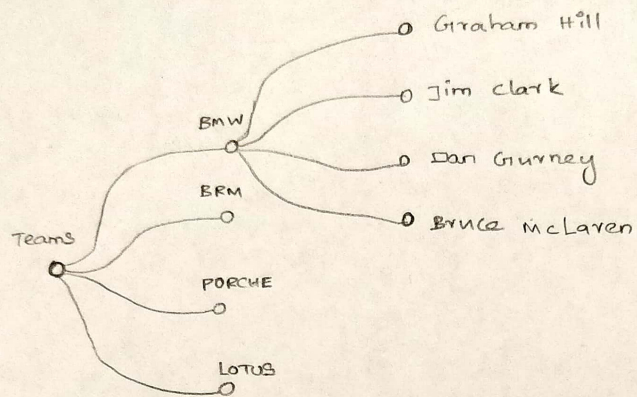


2005

Team Statistics Images

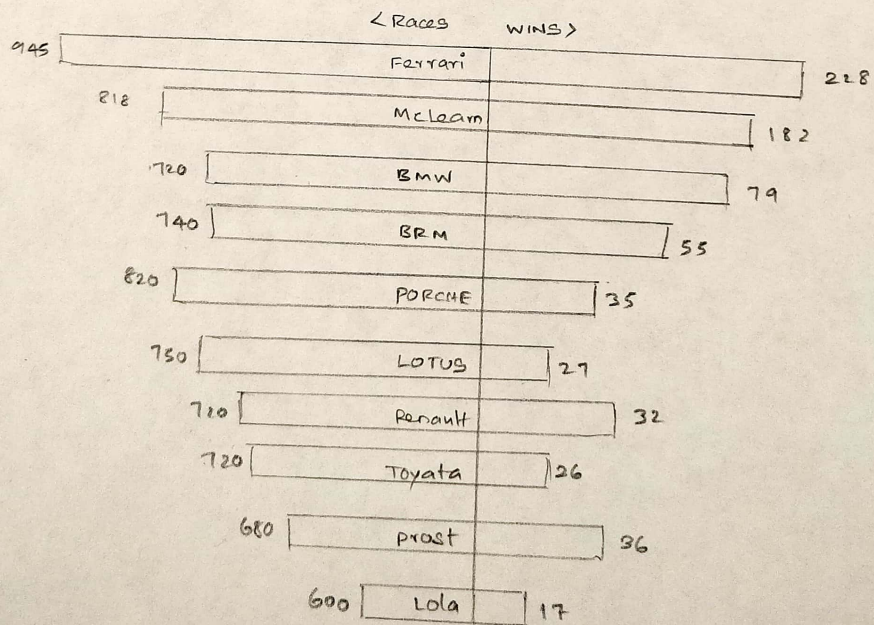


Team info:



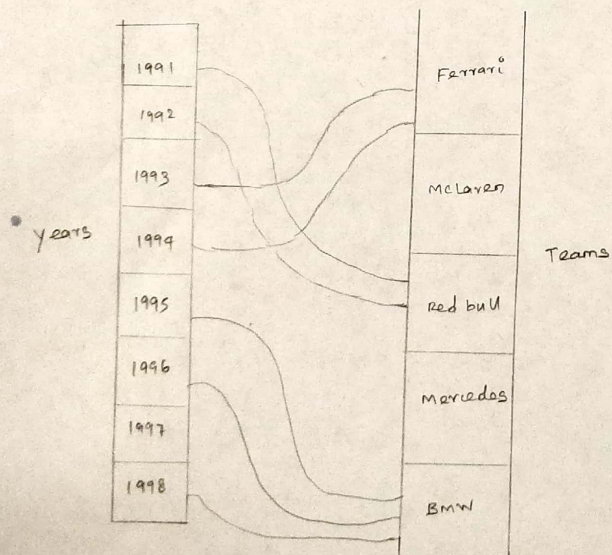
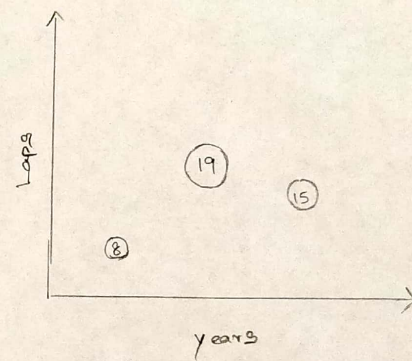
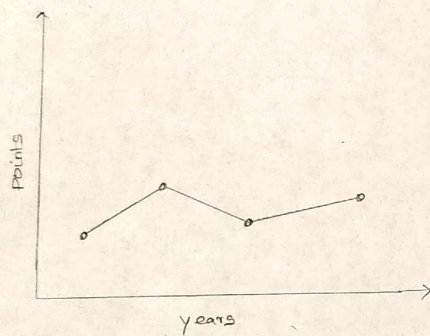
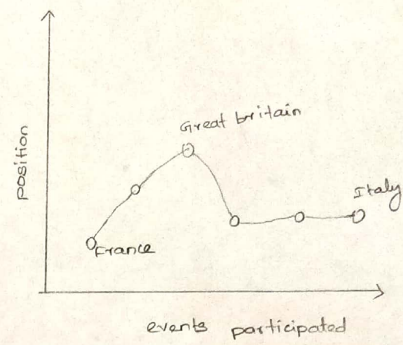
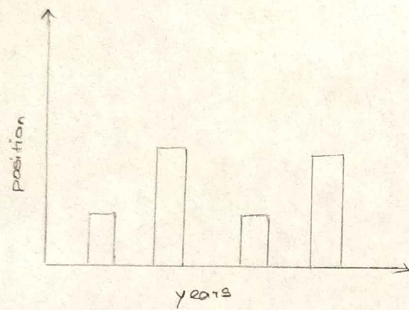
Treemap for teams

Team Stats: Races/Wins

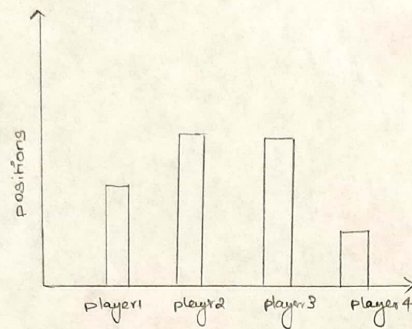


Individual Driver Statistics Images

Individual driver stats:

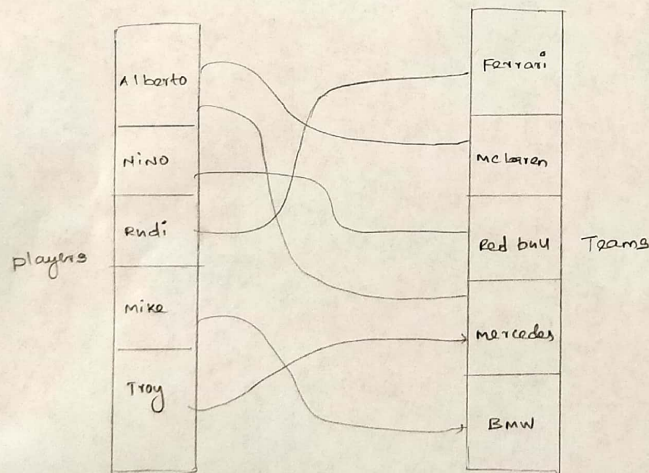
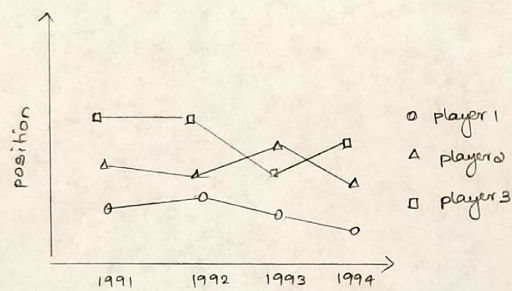


comparison of drivers per year



events around the world

- player 1
- △ player 2
- player 3



Must-Have Features

1. Individual racer statistics - Name, Team name, race finish timings and points.
2. Team statistics - Team name, Racers associated with the team, points, position.
3. Track statistics - number of locations

Optional Features

1. Individual racer statistics - Nationality, laps, number of stops for a race
2. Team statistics - Nationality
3. Status - Potential obstacles that can occur during a race (like collision, clutch, overheating, oil leak problems), overall number of laps.

Project Schedule

Week 1

1. Clean the data and determine the best layout for visualization. Survey visualizations with similar purpose.

Team members involved - Sanjana Aravindan, Harshini Keerthi Vasan

2. Find the right data structure to store the data and start working on developing the visualizations.

Team members involved - Vinu Sreenivasan

Week 2

1. Visualize the following stats:
 - (a) Racers and their associated team statistics and outstanding records
 - (b) Top teams and top players - globally and locally.

Team members involved - Harshini Keerthi Vasan, Vinu Sreenivasan

2. Compare the records of the players year-wise.

Team members involved - Sanjana Aravindan

Week 3

1. Improve the normal views of visualization and add extra views for dynamic visualization.

Team members involved - Vinu Sreenivasan, Sanjana Aravindan

2. Build a user interface to publicize the visualizations.

Team members involved - Harshini Keerthi Vasan

Week 4

1. Perform some additional interactivity in the data and meet out the edge cases.
2. Elaborate our project in the process book.

Team members involved - Vinu Sreenivasan, Sanjana Aravindan, Harshini Keerthi Vasan

References

- [1] Ergast Developer API : <https://ergast.com/mrd/db/#csv>
- [2] Formula 1 : <https://www.formula1.com/>