The history and evolution of Formula One

1 Objectives and Background

1.1 Introduction and History

The world of motorsport dates all the way back to the end of 19th century, with the first documented race being held in 1895. A race from Paris to Bordeaux, France, and back. A distance of 1170 kilometers, or 730 miles, where the average driver on four wheels was clocking in a speed of 24kph, or 14mph. But as the years went by, the race cars were able to reach speeds of up to 80kmh, or 49mph. This led to the Paris Bordeaux race being brought to a halt in 1903 due to the large number of accidents taking place. As the speed of the cars went up, so did the danger of drivers and spectators losing their lives to race accidents. [17]

This led to the birth of racing organizations that would dictate where and when these races would take place, the type and size of vehicles that would be racing against each other, and other aspects of racing that would later come to define the world of motorsport. The racing organization would also place rules and regulations that the drivers and teams had to abide by, to not only preserve the safety of the drivers and spectators, but also govern the technology that was developed and implemented during these races.

The largest and most famous of all racing organizations, is the ***Fédération Internationale de l'Automobile***, or **FIA**, which was established in 1904. Of the many things that the FIA has done since its inception, its most prominent role today is in the licensing and sanctioning of major races around the globe. And of those major races, the greatest of them all, is *Formula One*.

Formula One, or **F1**, is the highest form of international racing for open-wheel single-seater formula racing cars and the most popular form of motorsport racing in the world. Its first season took place in May 1950 and has been going on ever since. The world of F1 is far more than just fancy cars zooming past iconic landmarks across various countries and weekend parties, but it brings together some of the greatest minds, backed by billion dollar companies, to develop cutting edge technology that pushes the boundaries of engineering to new heights.

In [1]:

```
from IPython import display
display.Image("https://www.formula1.com/content/dam/fom-website/manual/Misc/2021manual/2021BritishManualAdds/2022CarImag
```

Out[1]:

The 2022 Formula 1 car, along with the 10 teams and their respective pairing of drivers that compete to win the F1 World Championship.[12]

1.2 Aims

The reason I have chosen this field, is that I wanted to explore the factors that have contributed to the technological marvel of the modern day F1 car, and the relative data that shows not only the progression of the cars themselves, but the controversial statistics that follow alongside them. The scope of covering a report about F1 is immense, thus, it will be scaled down to several areas.

These areas are:

- **Lap Times**
  - A modern day F1 season will host anywhere from 18 to 24 grands prix, with earlier seasons in the sixties and seventies having around 10 to 14 races. Each race takes places at iconic circuits around the world that tests the car's ability to drive under various conditions, from steep hills to high altitudes with windy conditions. A lap time is the time it takes for a car to complete one lap around the circuit, without exceeding track limits. We will see how the lap times have changed over the years. Have the cars gotten faster every single year without hiccups? Is it a steady linear progression?
- **Engine**
  - What type of engines were used? Were they regular engines seen in commercial cars, or are the certain technology used that are specific to the world of Formula 1? We will take a look at how they have changed over the years, and how these changes relates back to the lap times.
- **Environmental Impact**
  - F1 has been under scrutiny over the recent years over the impact their races have had towards the environment. Articles have been published stating the harm and waste of using fossil fuels for motorsport[16]. We will take a look at articles that various media bodies have published and gauge the media's sentiment towards the issue using natural language processing.

Why this field is relevant

In 2021, F1 generated a total of US$2.14 billion in revenue[1]. Today, a grand prix is more than just a racing event, it is an entertainment weekend. Drawing crowds from all over the world, a grand prix significantly boosts a country's tourism, bringing in a huge stream of income from various avenues. From the logistics, to food and beverages, to entertainment, a F1 Grand Prix brings with it job openings. Companies spend an absurd amount of money on advertising, in order to tap into the vast demographic of fans that F1 boasts. Teams are forced to employ creative and revolutionary tactics and ideas into their machines, in order to gain a competitive advantage over their rivals. All this translates to revenue to those who are able to pioneer new technology in order to dominate the sport and the market, therefore, it is not only a race on the track, but a race in the factories as well.

1.3 Data

Origin of data

The data that will be used in this report to explore laptimes and plot them was acquired from a web service, Ergast Developer API [2]. It contains CSV files of historical record of motor racing data for non-commercial purposes.

Limitations and constraints of the data

Timeframe

Because the data spans all the way back to 1950, there is simply too much data that can be reviewed in this report. Therefore, only certain areas will be hand picked to be viewed and discussed. These areas have been mentioned above(refer to 1.2). Throughout the evolution of F1, the FIA have imposed strict rules and regulations. The reason this is important is because the regulations dictates what can and cannot be done by the teams and their respective manufacturers. This includes the banning of certain technologies and strategies. Because the sporting, technical and financial regulations themselves change periodically, we will not be delving deeply into them, but only major decisions that pertain to the aforementioned subjects of F1. In regards to the articles, they were published in 2021,and 2019 respectively. This coverage only pertains to the modern era of F1, and therefore, does not reflect the sentiment of the public and media back in the sixties and seventies.

Publicly Available Data

There have been many private and public companies that have been sponsors of F1. This makes tracking down data difficult, as many of it is confidential and not publicly available. The roles these companies play range from logisitics, software and hardware, to the construction of the car parts themselves. An example would be Pirelli, the current tyre provider for F1. The rubber compound that are made to create the tyres used for racing, is an engineering feat on its own. Tyres are one of several key factors that affect one of the areas of research in this report, the lap times. However, it is impossible to retrieve any data about the tyres, its production, performance, source of rubber, etc, without it being provided from the actual company themselves. Thus, the data being used in this report will not cover all areas, but only what is publicly available.

1.4 Ethical Considerations

Use of logo and company images

Images of F1 cars, circuits, and engines for visualization purposes will be used in this report. Under the legal notice of the official F1 website, it is stated that, *"students can incorporate material from Formula1.com into educational projects, provided that the projects are not reproduced and are not for commercial purposes"* [3].

Use of website data

Regarding web scraping and accessing data directly from a website, F1 makes no specific mention of web scraping or web crawling. However, regarding the use of the materials available on their website for educational purposes, students may incorporate materials into their projects as long as credit and source is given to the official Formula 1 website[3].

ESPN is owned jointly by The Walt Disney Company, and under Disney's terms of use, it is stated under the "Usage Rules" section that, Disney products or intellectual property may not be used for commercial or business related purposes. This report is neither, and is for personal use only, which does not violate their terms.[9]

The Guardian's terms of service states that the use of material appearing on their website is for personal and non-commercial use only. It makes no mention of web scraping or anything of that nature. Since this report is made by a student and is non-commercial, no violation has been made.[10]

Mclaren racing makes no mention of using their website data and content, whether it be for commercial or non-commercial purposes. But for the purposes of this non-commercial, student made report, we can assume no violation has been made.

Potential negative impacts of this report

This report is purely for educational purposes. Written, coded and edited by a student. In no way is this report intended to portray a positive or negative image towards Formula 1, nor influence existing staff and/or fan's decisions pertaining to the sport. This report has been assessed by me in a fashion that the data and viewpoints raised will be unbiased. This report is purely meant to display and deduce publicly available data.

The conclusion of analyzing the articles given later in this report will not claim to be a representative analysis of the editor's or writer's journalistic style or view. Limitations of the data used have been mentioned above. Limitations of techniques used will be discussed throughout the report.

2 Exploring the lap times

In [2]:

```
#Importing libraries that will be used in this report

import pandas as pd
from bs4 import BeautifulSoup
import requests
from nltk.tokenize import sent_tokenize
from textblob import TextBlob
import matplotlib.pyplot as plt
import re
import datetime
import time
```

Structure of the data

The CSV files that will be used in this report are below. We will be using the data in each of these files to extract data regarding lap times. We will see later on how these files are labelled and sorted, but in general, the files and their relevant data are marked with an ID, which we will need to know in order to reference to other data files.

For example, in order to find the lap time for a particular grand prix, we will need to know the ID of the race and circuit in order to locate and sort the data.

"lap_times.csv" contains the lap times for the documented grand prixs. It has the lap time for each driver that participated in the given race.

"circuits.csv" contains all the names and locations of circuits that have been used to host a Formula 1 Grand Prix.

"races.csv" contains all the documented grand prix and the year in which the grand prix was held in.

Format for analysis

Because the data is stored in CSV files with unique identification columns, manipulating the data through the use of Pandas data frame is ideal as it allows easy sorting and pulling of data. The data frame will also be relatively easy to plot as well.

In [3]:

```
#Loading the CSV files that will be used in this report
laptimes_file = pd.read_csv('data_src/lap_times.csv')
circuits_file = pd.read_csv('data_src/circuits.csv')
races_file = pd.read_csv('data_src/races.csv')
```

The CSV files each have unique ID that is used to identify what the cells represent.

In [4]:

```
laptimes_file.head(3)
```

Out[4]:

	raceId	driverId	lap	position	time	milliseconds
0	841	20	1	1	1:38.109	98109
1	841	20	2	1	1:33.006	93006
2	841	20	3	1	1:32.713	92713

The lap times for a given circuit is classified under a "raceId" and "driverId". The same is true if we refer to the races file.

In [5]:

```
races_file.head(3)
```

Out[5]:

	raceId	year	round	circuitId	name	date	time	url	fp1_date	fp1_time	fp2_da
0	1	2009	1	1	Australian Grand Prix	2009-03-29	06:00:00	http://en.wikipedia.org/wiki/2009_Australian_G...	\N	\N	
1	2	2009	2	2	Malaysian Grand Prix	2009-04-05	09:00:00	http://en.wikipedia.org/wiki/2009_Malaysian_Gr...	\N	\N	
2	3	2009	3	17	Chinese Grand Prix	2009-04-19	07:00:00	http://en.wikipedia.org/wiki/2009_Chinese_Gran...	\N	\N	

We can see that the data is not sorted and contains columns that are not necessarily needed. We already have a separate CSV file that contains the lap times for each documented race. We can go ahead and clean the data for the races_file, and sort the race ID from ascending order.

In [6]:

```
#sort the raceId column with ascending true, then create new file as backup

sorted_laptimes = laptimes_file.sort_values(by=['raceId'],ascending = True)
sorted_laptimes.to_csv('data_src/sorted_laptimes.csv',index = False)

sorted_laptimes_file = pd.read_csv('data_src/sorted_laptimes.csv')
#print first 5 to see if its sorted
sorted_laptimes_file.head(5)
```

Out[6]:

	raceId	driverId	lap	position	time	milliseconds
0	1	1	1	13	1:49.088	109088
1	1	20	22	2	2:38.375	158375
2	1	20	23	2	2:31.909	151909
3	1	20	24	2	2:22.185	142185
4	1	20	25	2	1:33.553	93553

We will sort and remove unwanted data from the races file as well.

In [7]:

```
#cleaning the races.csv file, delete all columns after column "URL",
#since we do not need the empty cells of laptimes in another file

for column in races_file.columns:
    if column != 'raceId' and column != 'year' and column != 'round' and column != 'circuitId' and column != 'name' and
        races_file = races_file.drop([column],axis='columns')

races_file.to_csv('data_src/sorted_races.csv',index = False)
sorted_races_file = pd.read_csv('data_src/sorted_races.csv')
sorted_races_file.head(5)
```

Out[7]:

	raceId	year	round	circuitId	name	date	time
0	1	2009	1	1	Australian Grand Prix	2009-03-29	06:00:00
1	2	2009	2	2	Malaysian Grand Prix	2009-04-05	09:00:00
2	3	2009	3	17	Chinese Grand Prix	2009-04-19	07:00:00
3	4	2009	4	3	Bahrain Grand Prix	2009-04-26	12:00:00
4	5	2009	5	4	Spanish Grand Prix	2009-05-10	12:00:00

2.1 Selecting a circuit

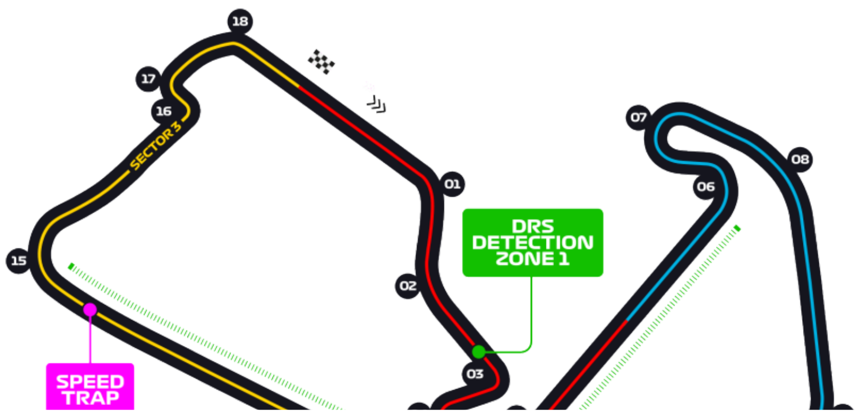
Now that we have a base point for our data, let us start to selectively extract some data to analyze. In order to see the reduction in lap times, we should first choose a particular circuit to delve into. For that, I have chosen Silverstone, an iconic English race track that is home to the British Grand Prix. The first F1 British Grand Prix was held in 1950[4].

Not all British Grand Prix were held at Silverstone, some were in Brands Hatch, another circuit based in England. Therefore certain grand prix will not be extracted from the file, but nevertheless, this will provide good data that we can track.

In [8]:

```
display.Image("https://www.formula1.com/content/dam/fom-website/2018-redesign-assets/Circuit%20maps%2016x9/Great_Britair
```

Out[8]:



Track map of SilverStone [13]

We will need to know the circuit ID for silverstone, in order to find the lap times associated with it.

In [9]:

```
#first, we find the circuit ID for silverstone
silverstoneID = circuits_file[circuits_file['circuitRef'].str.contains('silverstone')]

silverstoneID
```

Out[9]:

	circuitId	circuitRef	name	location	country	lat	lng	alt	url
8	9	silverstone	Silverstone Circuit	Silverstone	UK	52.0786	-1.01694	153	http://en.wikipedia.org/wiki/Silverstone_Circuit

We now find all race data that is associated with the British Grand Prix

In [10]:

```
#second, we find the grand prix ID for silverstone, the British Grand Prix
BritishGPObj = sorted_races_file[sorted_races_file['name'].str.contains('British Grand Prix')]

BritishGPObj
```

Out[10]:

	raceId	year	round	circuitId	name	date	time
7	8	2009	8	9	British Grand Prix	2009-06-21	12:00:00
25	26	2008	9	9	British Grand Prix	2008-07-06	12:00:00
43	44	2007	9	9	British Grand Prix	2007-07-08	12:00:00
59	60	2006	8	9	British Grand Prix	2006-06-11	12:00:00
80	81	2005	11	9	British Grand Prix	2005-07-10	14:00:00
...	...	...	...	...	...	...	...
985	998	2018	10	9	British Grand Prix	2018-07-08	13:10:00
1006	1019	2019	10	9	British Grand Prix	2019-07-14	13:10:00
1021	1034	2020	4	9	British Grand Prix	2020-08-02	13:10:00
1046	1061	2021	10	9	British Grand Prix	2021-07-18	14:00:00
1066	1083	2022	10	9	British Grand Prix	2022-07-03	14:00:00

73 rows × 7 columns

Now we know the ID of the circuit and the grand prix, and the circuit CSV file has been sorted, let us extract all the Silverstone lap times into its own unique table, so that we may plot it later.

In [11]:

```
#store race id into list
BritishGP_ListID = BritishGPObj['raceId'].values.tolist()

#check for british grand prix laptimes
BritishGP_Laptimes = sorted_laptimes_file[sorted_laptimes_file['raceId'].isin(BritishGP_ListID)]
BritishGP_Laptimes.to_csv('data_src/BritishGP_Laptimes.csv',index = False)

#csv file that has all British SilverStone GP
BritishGP_Silverstone = pd.read_csv('data_src/BritishGP_Laptimes.csv')

#Getting the fastest laptimes from drivers that raced in silverstone, and removing the rest
SilverStone_Laptimes = BritishGP_Silverstone.sort_values(by = ['raceId','time','milliseconds'], ascending = True)
SilverStone_Laptimes.drop_duplicates('raceId', keep="first", inplace=True)

#we do not need the columns pertaining to the info about drivers or their lap positions
finalx = SilverStone_Laptimes.drop(['driverId','lap','position'],axis=1)

print(finalx)
```

```
       raceId        time  milliseconds
63          8    1:20.735         80735
1833       26    1:32.150         92150
2414       44   10:32.179        632179
3517       60    1:21.599         81599
4438       81    1:20.502         80502
6079      100    1:18.739         78739
7280      118    1:22.236         82236
8065      133    1:23.083         83083
9018      151    1:23.405         83405
10245     161    1:26.217         86217
11378     182    1:28.309         88309
12433     199    1:35.704         95704
13205     215    1:24.475         84475
14158     233    1:29.288         89288
15656     346    1:30.874         90874
16264     849    1:33.391         93391
17716     868    1:34.608         94608
17919     887    1:33.401         93401
19431     908  1:01:07.384       3667384
19919     934    1:36.610         96610
21291     957    1:35.548         95548
22167     978    1:30.621         90621
23060     998    1:30.696         90696
24022    1019    1:27.369         87369
24753    1034    1:27.097         87097
25872    1061    1:28.617         88617
26639    1083    1:30.510         90510
```

We have the dataframe that contains the laptimes for the documented SilverStone British Grand Prixs in our sourced data set. But to better understand the data when we plot it, we replace the raceId to the appropriate year that the race was held in.

In [12]:

```
#get list of raceid in SilverStone df
check = SilverStone_Laptimes['raceId'].values.tolist()
#empty array to push sorted years
years = []
#for all rows in silverstone, find the year of the raceId, then push
for index,row in races_file.iterrows():
    for x in check:
        if row['raceId'] == x:
            years.append(row[1])
```

In [13]:

```
#new silverstone df, with raceId replaced with the year of the race
finalx = finalx.drop(['raceId'],axis=1)
finalx.insert(0,'year',years,True)
finalx = finalx.sort_values(by = ['year'], ascending = True)

#DROP DIRTY DATA
id44 = finalx[((finalx.year == 2007))].index
id908 = finalx[(finalx.year == 2014)].index

df = finalx.drop(id44,axis=0)
df = df.drop(id908,axis=0)

print(df)
```

```
       year      time  milliseconds
14158  1996  1:29.288         89288
13205  1997  1:24.475         84475
12433  1998  1:35.704         95704
11378  1999  1:28.309         88309
10245  2000  1:26.217         86217
9018   2001  1:23.405         83405
8065   2002  1:23.083         83083
7280   2003  1:22.236         82236
6079   2004  1:18.739         78739
4438   2005  1:20.502         80502
3517   2006  1:21.599         81599
1833   2008  1:32.150         92150
63     2009  1:20.735         80735
15656  2010  1:30.874         90874
16264  2011  1:33.391         93391
17716  2012  1:34.608         94608
17919  2013  1:33.401         93401
19919  2015  1:36.610         96610
```

Checking the legitimacy of the dataframe

Now that we have compiled our list of SilverStone Grand Prix fastest laptimes, let us check them against the F1 archives to see if the data is indeed correct.

We will use the 2000 British Grand Prix fastest laptime of 1 minute 26.217 seconds, to fact-check the data. We will web scrape the F1 archive to achieve this.

In [14]:

```
#begin webscraping of F1 archive page
def getwebsite(url):
    r = requests.get(url)
    #checking if website can be accessed
    response = requests.get(url)
    if response.status_code != 200:
        soupError = 'error'
        return soupError
    return r.text

htmldata = getwebsite("https://www.formula1.com/en/results.html/2000/races/50/great-britain/fastest-laps.html")
F1soup = BeautifulSoup(htmldata,'html.parser')
```

In [15]:

```
Table = F1soup.find("div", {"class":"resultsarchive-wrapper"})

print(Table.text)
```

```
1
1

Mika
Hakkinen
HAK

McLaren Mercedes
56
1:26.217
214.663





2
```

We extracted the table containing the fastest laps of the 2000 Silverstone Circuit, Foster's British Grand Prix, and at the top of the table we can see the fastest lap time was set by Mika Hakkinen of Mclaren Mercedes. A time of 1 minute, 26.217 seconds.

It matches the timing and date of our dataframe, which proves the data is accurate.

Lets visualize the change in laptimes

Now that we have a dataframe of sorted SilverStone Laptimes according to year, we can plot and visualize the change using matplotlib.

The "time" data in our data frame is a string. We need to replace the character in the string, so that we can convert the strings to integers. Only then can we plot the data

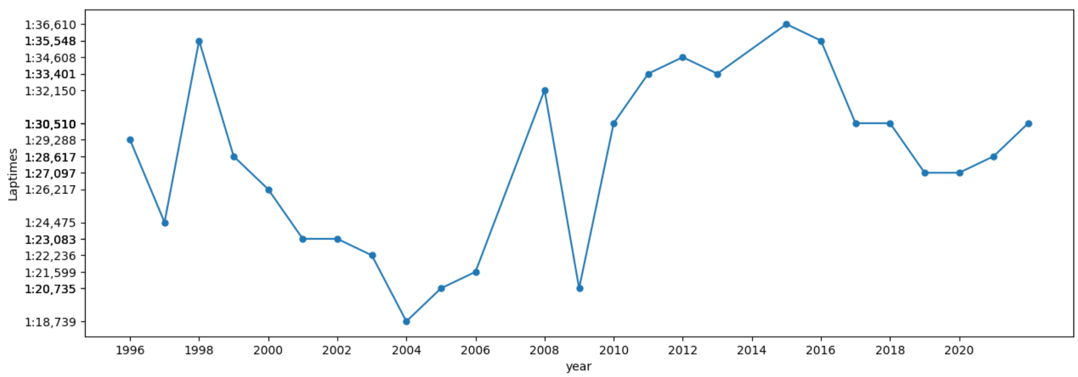
In [16]:

```
df['time'] = df['time'].str.replace('.',',')
```

```
C:\Users\shana\AppData\Local\Temp\ipykernel_6204\2289111245.py:1: FutureWarning: The default value of reg
ex will change from True to False in a future version. In addition, single character regular expressions
will*not* be treated as literal strings when regex=True.
  df['time'] = df['time'].str.replace('.',',')
```

In [17]:

```
plotMMSSList = df['time'].values.tolist()

#converting string to seconds
plotTime = []
for index,row in df.iterrows():
    x = time.strptime(row['time'].split(',')[0],'%M:%S')
    y = datetime.timedelta(minutes=x.tm_min,seconds=x.tm_sec).total_seconds()
    plotTime.append(y)


plt.figure(figsize = (15,5),dpi = 100)
plt.xlabel('year')
plt.xticks(range(1996,2022,2))
plt.yticks(plotTime,plotMMSSList)
plt.ylabel('Laptimes')

plt.plot(df['year'],plotTime,marker = '.', markersize = 10)

plt.show()
```



2.2 Summary

Through our data cleaning,sorting, and extraction, we can see that the lap times for the SilverStone British Grand Prixs, have a somewhat irregular trend to it. Initially, I was expecting to see a downward linear progression in terms of lap times. As the years go by, the cars would be able to clock in faster and faster lap times. We can see this isn't the case. The fastest lap time for SilverStone in the above graph is in 2004, which was set by the legendary Michael Schumacher. An impressive 1 minute, 18.739 seconds. Our modern 2022 F1 car is about 12 seconds slower.

It has been noted earlier in the report that the rules and regulations imposed by the **FIA** have impacted the performance of the cars (refer to 1.3). Though we will not delve into the rules change that produce such a dramatic shift in lap times, we can look at the changes in the engine.

3 Exploring the engine

3.1 Web scraping and article extracting

The list of URL that will be used for web scraping are referenced at the end of this report. We will be web scraping the official Mclaren website first, as it contains a well written article that explains the changes F1 engines have undergone over the years. Understanding the changes made to the engine will better explain the changes we have seen in F1 laptimes.

Brief explanation of an internal combustion engine

Combustion, also known as burning, is the basic chemical process of releasing energy from a fuel and air mixture. In an internal combustion engine (ICE), the ignition and combustion of the fuel occurs within the engine itself. The engine then partially converts the energy from the combustion to work. The engine consists of a fixed cylinder and a moving piston. The expanding combustion gases push the piston, which in turn rotates the crankshaft. Ultimately, through a system of gears in the powertrain, this motion drives the vehicle's wheels. [6]

Accessing the Mclaren webpage [5]

In [18]:

```
#begin webscraping
def getwebsite(url):
    r = requests.get(url)
    #checking if website can be accessed
    response = requests.get(url)
    if response.status_code != 200:
        soupError = 'error'
        return soupError
    return r.text

htmldata = getwebsite("https://www.mclaren.com/racing/car/history-of-the-f1-engine/")
soup = BeautifulSoup(htmldata,'html.parser')
```

We will narrow down to 5 major turning points in engine development in this article.

In [19]:

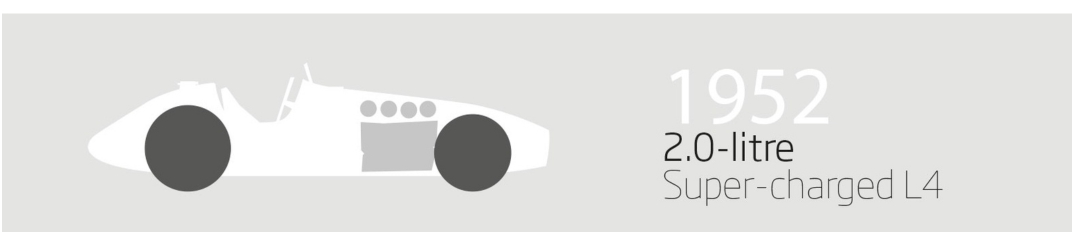
```
#function to find the header and paragraph to print
def printHeader_Article (h,p):
    Header = soup.find_all(re.compile('^h[2]'))[h] #finding header titles
    Para = soup.find_all('p')[p].text

    print(Header.get_text(),'\n')
    print(" ".join(Para.split()))

def printArticle (p):
    Para = soup.find_all('p')[p].text
    print("\n")
    print(" ".join(Para.split()))

printHeader_Article(0,6)
```

```
1950-1953: Formula 1 begins: the super-charger years

At the inception of the Formula 1 World Championship, teams had a choice of a super-charged engine of up
to 1.5l capacity or a normally-aspirated one of up to – monstrous by modern standards – 4.5l. Talbot ran
a 4.5l in-line six cylinder (L6) engine but the other manufacturer teams went down the super-charged rout
e. Borrowing aerospace technology, the lighter, super-charged engines held the early power advantage but
the combination of a small engine and enormous super-charger made for a very thirsty package, consuming p
erhaps five times as much fuel as the normally-aspirated cars. When more powerful, normally-aspirated eng
ines appeared, designed specifically for F1, the super-charged units went out of fashion. The super-charg
ed engines were an eclectic mix: Alfa ran an L8, Maserati an L4, while Ferrari had a V12. At the far extr
emes there was JAP V-twin, and a BRM V16.
```

In [20]:

```
display.Image("https://media-cdn.mclaren.com/media/images/galleries/1952.jpg")
```

Out[20]:



Teams could choose between super-charged [7] engines or normal aspirated ones, using technology from the aerospace sector. Some teams had different pistons configuration. Engine varied from team to team and were not standardized to a particular configuration.

In [21]:

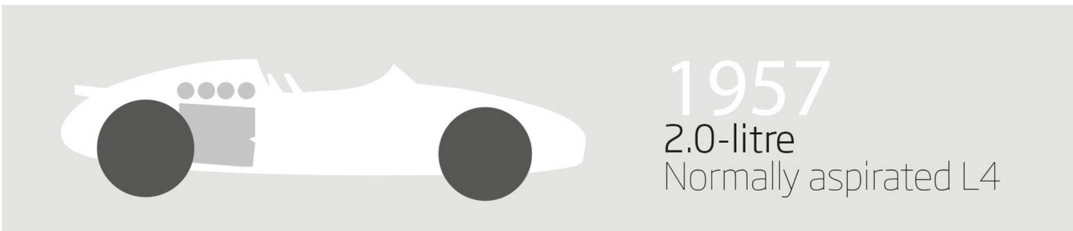
```
printHeader_Article(3,14)
```

```
1957: The dawn of the mid-engine formula

Mounting the engine behind the driver created a better-balanced car, less prone to the enormous understee
r that afflicted the front-engine cars. Auto Union’s pre-war effects had created cars that were prone to
oversteer and difficult to handle – but two decades of suspension development, notably the introduction o
f double wishbones made that less of an issue.
```

In [22]:

```
display.Image("https://media-cdn.mclaren.com/media/images/galleries/1957.jpg")
```

Out[22]:



The term understeer is defined as the car not following the steering input registered by the driver during a turn. See reference link for better understanding of understeer and oversteer [11].

By shifting the engine placement to the back, better balance was achieved, allowing drivers to perform better during turns, improving laptimes.

In [23]:

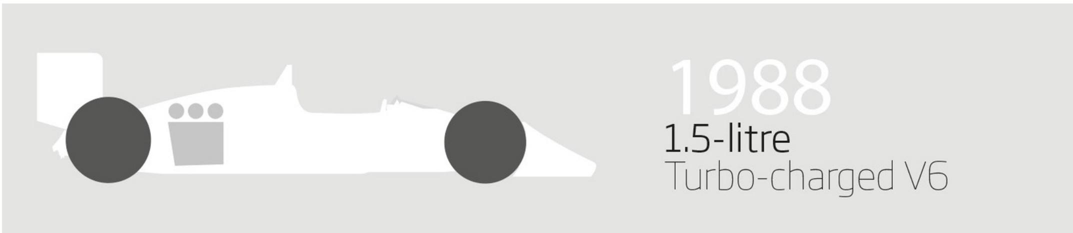
```
printHeader_Article(8,35)
printArticle(37)
printArticle(39)
printArticle(42)
```

```
1977-1988: The swift rise of the turbo

The regulations introduced for 1966 allowed for a normally-aspirated engine of up to 3.0L displacement or
a forced induction engine of up to 1.5l. The second part of the regulation was roundly ignored until 1977
when Renault debuted the RS01, powered by a 1.5l Renault-Gordini EF1 V6 turbo.


Renault continued to tweak the EF1. Electronic ignition was added in 1982 and water injection in 1983 – t
he later was a method of cooling ‘hot-spots’ in the combustion chamber to prevent premature detonation. I
t was retired at the end of 1983, having won 15 grands prix. In 1977 it was producing around 500hp; by 19
83 that figure was about 700hp.


Renault’s innovation had changed the sport: by 1983 turbos were the dominant force. By 1986 normally-aspi
rated engines had disappeared entirely from F1.


While turbo-charging wasn’t invented for F1, the sport carried the technology forward. In the plus colum
n, turbo engines produced prodigious power – but at a cost of great unreliability, poor driveability and
terrible fuel economy. The headlines of the 1980s were of turbo engines generating more and more power –
but in the background the engine makers were also busy solving the inherent weaknesses of the technology,
advancing combustion technology and metallurgy to a point where F1 transitioned from 500hp to 800hp in a
few short years. F1 went turbo-free in 1989 but the legacy of the era was that reliable, driveable and ac
ceptably economical turbo engines entered the automotive mainstream in the 1990s.
```

In [24]:

```
display.Image("https://media-cdn.mclaren.com/media/images/galleries/1988.jpg")
```

Out[24]:



Through Renault,turbo chargers [8] were introduced, along with electric ignition and water injection. Large increase in performance, beginning of the "Turbo Era". Notice how the chassis have changed as well. It has been mentioned at the start of the report that there are many factors that contribute to how well a Formula 1 car performs, we have only decided to delve into the engines, but from the images, I would like to simply address the changes in chassis were to provide better aerodynamic performance, resulting in things like increased top speed and better braking

In [25]:

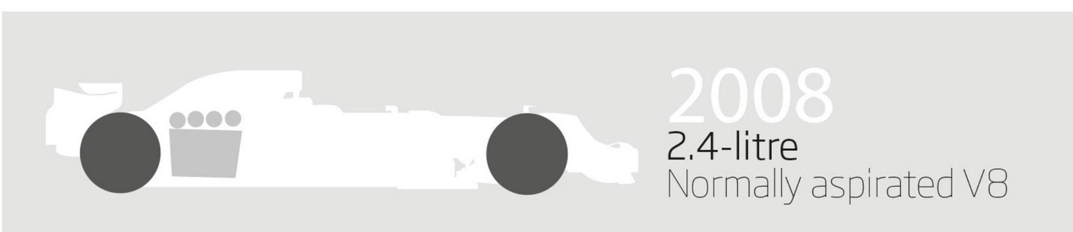
```
printHeader_Article(10,47)
printArticle(51)
```

```
1995-2013: Downsize, restrict, freeze

Engine size was restricted to 3.0l for 1995 as part of a raft of measures to control the speeds being rea
ched in F1. The modern trend for using regulation to retard speed really begins here. Lighter engines cou
ld run at higher speeds, and thus manufacturers continued to improve power output by increasing rpm, incr
easingly using exotic alloys to cut weight. A side effect of the lighter engines was that teams started r
unning significant amounts of ballast to make the minimum weight.


For 2006 the 3.0l V10 was replaced with a 2.4l V8. Exotic materials were banned in an effort to contain c
osts. Power output dropped back to 700-750hp for 2006 but Cosworth’s CA-series became F1’s first mainstre
am 20,000rpm race engine.
```

In [26]:

```
display.Image("https://media-cdn.mclaren.com/media/images/galleries/2008.jpg")
```

Out[26]:



Because of the rules imposed by the FIA to reduce costs, we can see there was a DECREASE in performance, even though the teams could push the car further. Focus was shifted to developing the longevity of F1 engines

In [27]:

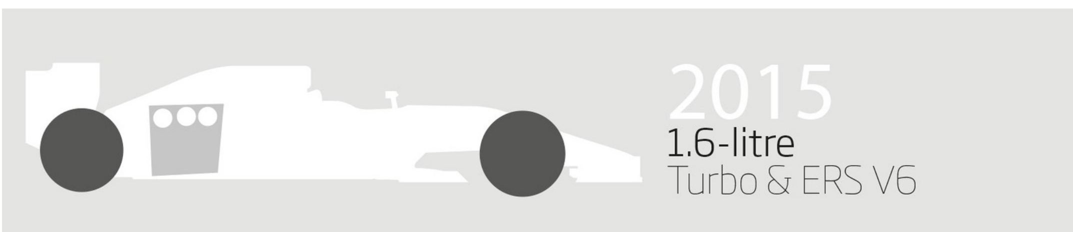
```
printHeader_Article(11,54)
print("\n")
printHeader_Article(12,56)
printArticle(58)
```

```
KERS

While the V8 engine hovered around the 750hp mark, in 2009 F1 cars gained – in short bursts – an extra 80
hp through the introduction of KERS. The Kinetic Energy Recovery System was a form of mild hybridisation
introduced with a view to making F1 more road-relevant. It recovered energy under braking, stored it in a
battery and allowed it to be restored to the powertrain at the driver's initiation.


Hybrid

Keen to push hybrid technology further, game-changing specifications were introduced in 2014. The most co
mprehensive overhaul in the history of F1came with a complex specification but a straightforward aim: the
new spec would consume one-third less fuel than the V8 with no corresponding drop in performance.


The manufacturers developed the specification for a 1.6l direct-injection V6 with a single turbo-charger.
It features two energy recovery systems: an uprated successor to KERS to recover energy from braking and
a turbo compounding loop recovering waste heat from the exhaust. The combined Energy Recovery System (ER
S) can add 120kW at the crank and deploy 4MJ per lap – essentially twice as much power for five times as
long as KERS. The power unit (it's a bit more than an 'engine' today) is either recovering or deploying e
nergy for most of the lap, mapped into the engine mode rather than controlled by the driver.
```

In [28]:

```
display.Image("https://media-cdn.mclaren.com/media/images/galleries/2015.jpg")
```

Out[28]:



From 2014 to date, began the turbo-hybrid era. Introduction of hybrid energy recovery systems to provide extra performance boost on top of the turbo charged internal combustion engine.

3.2 Summary

We have now seen the changes made to Formula 1 engines over the years, and how the performance boost and various technologies used such as the energy recovery system, lead to changes in grand prix laptimes. I was also right in my initial assumption that the rules imposed by the FIA affected the performance of the F1 cars. However with these changes, came an increase in fuel consumption. We will now explore the media's sentiment regarding F1's fuel consumption.

4 Environmental Impacts

4.1 Scraping and Analyzing The Guardian

Using the BeautifulSoup import library, we request for the website[14], access it, then locate the article text in the HTML body so that we may analyze it using natural language processing.

In [29]:

```
#requesting the url and getting website content
guardian = requests.get("https://www.theguardian.com/sport/2021/nov/26/climate-emergency-accelerates-f1-efforts-to-clea

soupG = BeautifulSoup(guardian.content,'lxml')
```

In [30]:

```
#article text
article = soupG.find('div', {'id': 'maincontent'}).find_all("p")

#code reference made to project example provided on coursera
article_text = ''
for content in article:
    article_text = article_text + ' ' + content.get_text().lstrip().rstrip()


article_text = article_text.replace('“', '"')
article_text = article_text.replace('”', '"')

article_text = article_text.replace('`', "'")
article_text = article_text.replace('’', "'")
#code reference made to project example provided on coursera
```

In [31]:

```
#function to analyze the tokenized text
def analyze_sentiment(text):
    analyze = TextBlob(text)
    if analyze.sentiment.polarity > 0:
        return 'Positive'
    if analyze.sentiment.polarity == 0:
        return 'Neutral'
    else:
        return 'Negative'
```

The polarity refers to analyzing textual data, where it will return a value of either -1 for negative sentiment or +1 for positive sentiment.

In [32]:

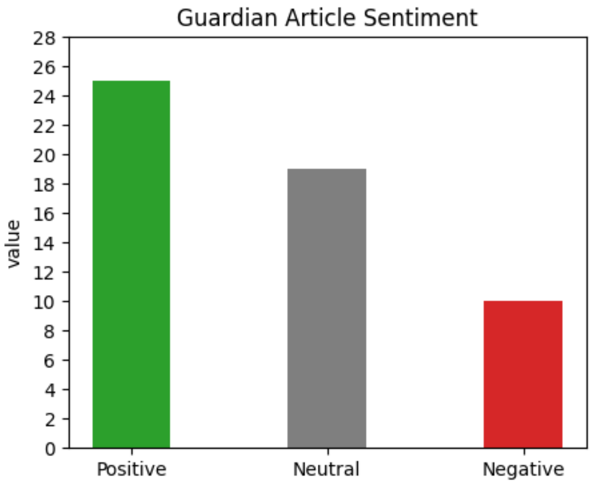
```
sentence = sent_tokenize(article_text)
sentence = pd.DataFrame(sentence)
sentence.columns = ['sentence']
sentence['Sentiment'] = [str(analyze_sentiment(text)) for text in sentence.sentence]

values = sentence.Sentiment.value_counts()

guardianArticleSent = pd.DataFrame(values)
print(guardianArticleSent)
```

```
          Sentiment
Positive         25
Neutral          19
Negative         10
```

This shows us that the overall sentiment of the Guardian article, leans towards a positive sentiment

In [33]:

```
#plotting the sentiment of the article
emotion = ['Positive','Neutral','Negative']
bar_colors = ['tab:green','tab:grey','tab:red']

#resize graph
plt.figure(figsize = (5, 4),dpi=100)

# creating the bar plot
plt.bar(emotion, values, color =bar_colors,
        width = 0.4)

plt.ylabel('value')
plt.yticks(range(0,30,2))
plt.title('Guardian Article Sentiment')
plt.show()
```



4.2 Scraping and Analyzing ESPN

We now perform the same web scraping and extracting for the ESPN website that contains the article[15].

In [34]:

```
#requesting the url and getting website content
ESPN = requests.get("https://www.espn.com/f1/story/_/id/28395224/the-environment-pose-f1-biggest-challenge-2020s")

soupE = BeautifulSoup(ESPN.content,'lxml')
```

In [35]:

```
#article text
articleESPN = soupE.find('div', {'class': 'article-body'}).find_all("p")

article_text_ESPN = ''
for content in articleESPN:
    article_text_ESPN = article_text + ' ' + content.get_text().lstrip().rstrip()

#code reference made to project example provided on coursera
article_text_ESPN = article_text_ESPN.replace('“', '"')
article_text_ESPN = article_text_ESPN.replace('”', '"')

article_text_ESPN = article_text_ESPN.replace('`', "'")
article_text_ESPN = article_text_ESPN.replace('’', "'")
article_text_ESPN = re.sub('[0-9]',"",article_text_ESPN)
#code reference made to project example provided on coursera
```

In [36]:

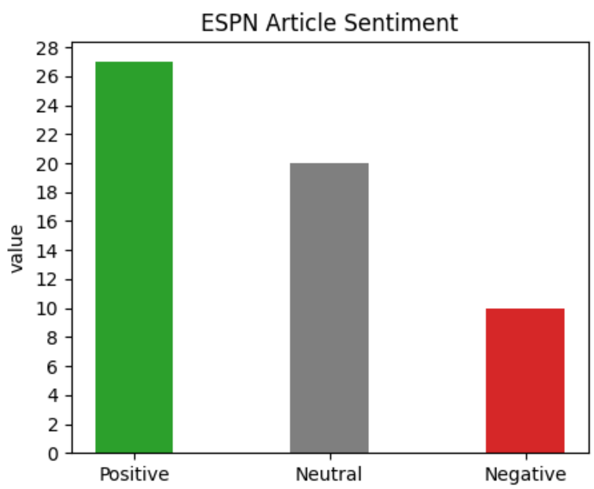
```
sentenceESPN = sent_tokenize(article_text_ESPN)
sentenceESPN = pd.DataFrame(sentenceESPN)
sentenceESPN.columns = ['sentence']
sentenceESPN['Sentiment'] = [str(analyze_sentiment(text)) for text in sentenceESPN.sentence]

ESPNvalues = sentenceESPN.Sentiment.value_counts()

ESPNArticleSent = pd.DataFrame(ESPNvalues)
print(ESPNArticleSent)
```

```
          Sentiment
Positive         27
Neutral          20
Negative         10
```

This shows us that the overall sentiment of the ESPN article, leans towards a positive sentiment

In [37]:

```
#plotting the sentiment of the article

#resize graph
plt.figure(figsize = (5, 4),dpi=100)

# creating the bar plot
plt.bar(emotion, ESPNvalues, color =bar_colors,
        width = 0.4)

plt.ylabel('value')
plt.yticks(range(0,30,2))
plt.title('ESPN Article Sentiment')
plt.show()
```



4.3 Summary

Through the web scraping and analysis of the two articles above, we can come to the conclusion that, regarding certain major media bodies such as The Guardian and ESPN, have a fairly positive sentiment regarding the environmental impact that Formula 1 poses. Formula 1 have publicized that they have taken a stand to reduce their carbon footprint and we can see that the news bodies we have analyzed has received their statements positively.

5 Conclusion

Through the exploratory data analysis, we have been able to see the changes Formula 1 cars has had throughout the years. From the speed of the cars and their performance on track, to the technology that has implemented into their engines, and the media's sentiment towards the sport and its impact on the environment. Though the approach undertaken in this report did not cover all aspects of the topic, the methodology used and the information gained did answer the question I set out to answer. I wanted see explore the engine changes, and the lap time changes, and I have successfully presented the answer and evidence in a comprehensible manner.

Within the project, it would be useful to explore sentiment analysis of more news articles from various media bodies, as it will help give us a more data and evidence to support our claim about the media's sentiment towards Formula 1.

10 References

[1] F1 2021 revenue up 87% to US$2.14bn (no date) Black Book Motorsport. Available at: https://www.blackbookmotorsport.com/news/f1-financial-results-q4-2021-revenue-income-liberty-media (https://www.blackbookmotorsport.com/news/f1-financial-results-q4-2021-revenue-income-liberty-media)

[2] Ergast developer API. Available at: http://ergast.com/mrd/ (http://ergast.com/mrd/)

[3] F1 legal notices: Official Formula 1 website (no date) Formula 1® - The Official F1® Website. Available at: https://www.formula1.com/en/toolbar/legal-notices.html (https://www.formula1.com/en/toolbar/legal-notices.html)

[4] Wikipedia. (2022). 1950 British Grand Prix. [online] Available at: https://en.wikipedia.org/wiki/1950_British_Grand_Prix (https://en.wikipedia.org/wiki/1950_British_Grand_Prix)

[5] McLaren (2020) History of the F1 engine, McLaren Racing. Available at: https://www.mclaren.com/racing/car/history-of-the-f1-engine/ (https://www.mclaren.com/racing/car/history-of-the-f1-engine/)

[6] Internal Combustion Engine Basics (no date) Energy.gov. Available at: https://www.energy.gov/eere/vehicles/articles/internal-combustion-engine-basics (https://www.energy.gov/eere/vehicles/articles/internal-combustion-engine-basics)

[7] Supercharger (2022) Wikipedia. Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/Supercharger (https://en.wikipedia.org/wiki/Supercharger).

[8] Turbocharger (2022) Wikipedia. Wikimedia Foundation. Available at: https://en.wikipedia.org/wiki/Turbocharger (https://en.wikipedia.org/wiki/Turbocharger).

[9] English – disney terms of use – United States (no date) Disney Terms Of Use. Available at: https://disneytermsofuse.com/english/ (https://disneytermsofuse.com/english/).

[10] Terms and conditions (2010) The Guardian. Guardian News and Media. Available at: https://www.theguardian.com/help/terms-of-service (https://www.theguardian.com/help/terms-of-service).

[11] Understeer vs oversteer – 3 simple & accurate definitions (2022) Your Data Driven. Available at: https://www.yourdatadriven.com/racing-talk-understeer-vs-oversteer-basic-explanation/ (https://www.yourdatadriven.com/racing-talk-understeer-vs-oversteer-basic-explanation/).

[12] F1 (2021) Gallery: A first look at the life-size 2022 F1 car, after drivers get up close at Silverstone: Formula 1®, Formula 1. Formula 1. Available at: https://www.formula1.com/en/latest/article.gallery-a-first-look-at-the-life-size-2022-f1-car-after-drivers-get-up-close.4rJgGsK2iEWcY6jjKWVjv.html (https://www.formula1.com/en/latest/article.gallery-a-first-look-at-the-life-size-2022-f1-car-after-drivers-get-up-close.4rJgGsK2iEWcY6jjKWVjv.html).

[13] F1 (2020) British Grand Prix - F1 race - Silverstone Circuit: Formula 1®, Formula 1. Formula 1. Available at: https://www.formula1.com/en/information.great-britain-silverstone-circuit.2DtFVI1FjkYgLVdGhbAIv0.html (https://www.formula1.com/en/information.great-britain-silverstone-circuit.2DtFVI1FjkYgLVdGhbAIv0.html).

[14] Climate emergency accelerates F1's efforts to clean up its image | Giles Richards (2021) The Guardian. Guardian News and Media. Available at: https://www.theguardian.com/sport/2021/nov/26/climate-emergency-accelerates-f1-efforts-to-clean-up-image (https://www.theguardian.com/sport/2021/nov/26/climate-emergency-accelerates-f1-efforts-to-clean-up-image).

[15] Edmondson, L. (2019) The environment will pose F1's biggest challenge in the 2020s, ESPN. ESPN Internet Ventures. Available at: https://www.espn.com/f1/story/_/id/28395224/the-environment-pose-f1-biggest-challenge-2020s (https://www.espn.com/f1/story/_/id/28395224/the-environment-pose-f1-biggest-challenge-2020s).

[16] How bad is F1 raceing for the environment? (2021) Thinking Sustainably. Available at: https://www.thinkingsustainably.com/f1-raceing-for-the-environment/ (https://www.thinkingsustainably.com/f1-raceing-for-the-environment/).

[17] Automobile racing (no date) Encyclopædia Britannica. Encyclopædia Britannica, inc. Available at: https://www.britannica.com/sports/automobile-racing (https://www.britannica.com/sports/automobile-racing).