Building influential sponsorship and celebrity endorsement with the right team and worthy driver in Formula 1

1. Background

Formula One (F1) is the global highest level racing project owned by the Formula One Group and supervised by The International Automobile Federation. It holds 20 Grand Prix each year across 20 different countries, which is one of the 10 most popular sports in the world ranked in the 7th place (Pledge sports, 2017). Because F1 contains elements of fast speed, excitement, competitiveness and embedded state-of-the-art in F1 racing, the popularity of F1 keeps increasing in recent years. In the United States, the average viewer per race was close to 1 million on ESPN, and the total number of worldwide television audiences in 2020 was around 433 million (Medland, 2021; Lange, 2021).

In this highly exposed racing event, businesses can increase their marketing exposure and brand visibility internationally by taking advantage of the popularity of leading teams and top performance drivers. Companies could either achieve sponsorship with F1 constructors or invite famous and popular drivers to endorse products. So, the choices of which team and which driver must be made, if corporates want to show up in the camera of this eye-catching international racing matches and let fans aware of their brands.

2. Business Question and the Rationale behind it

This report will demonstrate the business question from two perspectives: in order to increase brand awareness and international visibility of multinational companies, which F1 team should they sponsor, and which driver could be the right person endorsing their products?

According to Rust and Oliver (1994, as cited in Cornwell & Kwon, 2019), television advertising's position has already been shifted from 'traditional media' to so-called 'legacy media', which might eventually diminish in the long run. However, there are many substitutes arising such as indirect marketing approach sponsorship. Cornwell and Kwon (2019) pointed out sponsorship refers to paying amounts of cash or equivalent property in return for obtaining the commercial potential that can be developed by that activity such as sports and arts. In this case, some companies may want to make use of the popularity of F1 events.

With respect to celebrity endorsement, this is a commonly used advertising strategy in recent years, where 25% of the United States ads contain a celebrity. For example, Nike held $9.4 billion endorsement contracts in total in 2016 which was roughly 25% of its annual revenue at that time (Euromonitor International, 2014; McIntyre, 2017, as cited in Carrillat & Ilicic, 2019). Out scope of the teams, F1 drivers are also well recognized by global businesses regarding celebrity endorsement. Essentially Sports (2021) reported that F1 star Lewis Hamilton has endorsements from Monster Energy, Puma, Sony, Bose, etc., where his total endorsements accounted for $12 million in 2020. Meanwhile, his latest base salary paid by Mercedes is $55 million in the 2021 season.

Therefore, for some multinational companies who are new to the F1 industry and eager to increase their global popularity by reaching cooperation with either famous constructors or outstanding drivers, the problem is what factors should they consider maximizing the business influence and value when choosing teams and drivers. The simple answer is to decide the long-term top performers among constructors and drivers, as well as potential great players in the near future.

3. Methodology and Hypothesis

This report will unpack the question into two parts for evaluating the constructor performance of the top 10 constructors and driver performance of the top 10 drivers in the F1 2021 season. To evaluate constructor performance, this research considered the following performance indicators: position – average position per race, wins – average wins per race or the win rate, points – average points earned per race, times – how many times did the team present in games, current – 2021 team standing from F1 official site. For driver performance, except position, wins, points, times and current, the age is also considered to compare the experiences or career stage among drivers.

This research assumed that these selected factors are related to the frequency of constructor and driver appearances during F1 races, interviews, documentaries, etc. Under this hypothesis, a higher frequency of appearances will lead to higher market exposure, global visibility and brand awareness, and an eventual larger business value. Thus, a smart choice should be made to filter the suitable F1 team or driver, because business aims to optimize how they spend the money. They could decide to sponsor which team and invite which player to endorse their products based on selecting top performers with the help of radar charts and histograms produced in the next few procedures.

4. Data preparation

In this report, all the datasets are all provided by the teaching team, the following datasets are used:
'races.csv',
'constructors.csv', 'constructors_standings.csv',
'drivers.csv', 'driver_standings.csv'.

- '**races.csv**': because the research objects are the current top 10 constructors and drivers, so the time should be narrower than the original timeline contained in the dataset. I selected races in recent 10 years from Australian Grand Prix in March 2011 to Austrian Grand Prix in July 2021, which has 206 races in total. No missing value or duplicates are detected in the selected period.

Selected Timeline

First race: 841 - Australian Grand Prix 2011-03-27
Last race: 1060 - Austrian Grand Prix 2021-07-04

- '**constructors.csv**': according to Formula 1 Official Website (2021), the current top 10 teams on October 24th, 2021 (see **Appendix**) are Mercedes, Red Bull Racing, McLaren, Ferrari, Alpine, AlphaTauri, Aston Martin, Williams, Alfa Romeo Racing and Haas. The constructor names are slightly different from the names listed on the F1 official website because the most recent team names contained the power unit constructor's name at the end. This file was used to identify the unique key – 'constructorId', which can be used for later processes.

Constructor List

Constructor ID	Constructor Name
131	Mercedes
9	Red Bull
1	McLaren
6	Ferrari
214	Alpine F1 Team
213	AlphaTauri
117	Aston Martin
3	Williams
51	Alfa Romeo
210	Haas F1 Team

- '**constructors_standings.csv**' & 'constructors_results.csv': these two files have similar information, however 'constructor_standings.csv' has more information such as wins and position for each race. After filter the dataset by using conditions computed in previous steps, the remaining data is related to the 2021 top 10 constructors' performance within 2011 and 2021. Importantly, data cleaning is necessary before finalizing the key performance indicators. Then order the data by constructor id and race id. There are details must aware where columns like 'wins' and 'points' have implicit meanings that are accumulated winning and points for each team in a specific year, which differ from the meaning of winning per race and points per race. Therefore, these two columns should be recovered as original and required format, which are wining per race and points per race. Afterwards, Position, Wins, Points per race are calculated, and information about how many times the team is qualified for races and the current standings on F1 official website are added.

Summarized Table for Constructors Performance

Constructors	Position	Wins	Points	Times	Current
Mercedes	2.049	0.505	26.650	206	1
Red Bull	2.427	0.272	21.646	206	2
McLaren	5.252	0.058	8.408	206	3
Ferrari	2.752	0.107	18.587	206	4
Alpine F1 Team	6.200	0.000	3.400	10	5
AlphaTauri	6.444	0.037	5.778	27	6
Aston Martin	5.800	0.000	4.800	10	7
Williams	7.291	0.005	4.141	206	8
Alfa Romeo	7.500	0.000	1.271	48	9
Haas F1 Team	7.736	0.000	1.673	110	10

For Position and Current, if the average position and current standing are smaller, the better the constructor performance. For Wins, it can be understood in two ways, which are the average win rate or win time per race. Taking Mercedes as example, it refers to 0.505 winning time per race over the 206 total races. And the Times column indicates how frequent the team is qualified for each season. In this case, 206 represents no absent in past ten years events, and there are five teams met this standard, which are Mercedes, Red Bull, McLaren, Ferrari and Williams.

- '**drivers.csv**': according to Formula 1 Official Website (2021), the current top 10 drivers on October 24th, 2021 (see **Appendix**) are Max Verstappen, Lewis Hamilton, Valtteri Bottas, Lando Norris, Sergio Perez, Carlos Sainz, Charles Leclerc, Daniel Ricciardo, Pierre Gasly and Fernando Alonso. This file is similarly used as 'constructors.csv', by using drivers' forename and surname to identify the unique key - 'driverId'. However, Sergio Perez is a Mexican driver, so his surname is recorded as 'Pérez' with Latin letter.

Driver List

Driver ID	Driver Full Name
830	Max Verstappen
1	Lewis Hamilton
822	Valtteri Bottas
846	Lando Norris
815	Sergio Pérez
832	Carlos Sainz
844	Charles Leclerc
817	Daniel Ricciardo
842	Pierre Gasly
4	Fernando Alonso

- '**driver_standings.csv**': this file has no missing values but one duplicate, after handled them, filter the information of points, position and wins with respect to the top 10 drivers within the given time from 2011 to 2021. Similarly, after order the data by driver id and race id, the information provided by 'wins', 'points', 'times' columns should be edited like the previous steps in processing 'constructors_standings.csv'. On the other hand, merge the 'drivers.csv' date of birth column into the modified data, and calculate the age for each driver. Furthermore, adding the current driver standings according to F1 official website. Lastly, the average position, average wins and average points are calculated based on different drivers in the recent 206 races.

Summarized Table for Driver Performance

Drivers	Position	Wins	Points	Times	Age	Current
Max Verstappen	5.977	0.124	10.349	129	24	1
Lewis Hamilton	2.141	0.403	16.015	206	36	2
Valtteri Bottas	5.796	0.048	9.228	167	32	3
Lando Norris	7.479	0.000	5.021	48	21	4
Sergio Pérez	10.146	0.010	3.912	205	31	5
Carlos Sainz	10.419	0.000	3.264	129	27	6
Charles Leclerc	7.913	0.029	6.391	69	24	7
Daniel Ricciardo	9.510	0.035	5.894	198	32	8
Pierre Gasly	10.707	0.013	3.093	75	25	9
Fernando Alonso	8.551	0.036	6.168	167	40	10

From this table, we can see that Lewis Hamilton's overall performance is extraordinary compared to other drivers, especially the average position and average wins per race. At the same time, the times of present may not be a fair performance indicator because some of the drivers are very young who joined in F1 in later years. Such as Lando Norris who is only 21 years old, but ranked at 4th place in the 2021 season, although his overall performance in the average position, wins and points are not very flourishing in recent events, his main competitiveness lies in his age.

5. Visualizations

5.1 Evaluating Formula One Constructor Performance using Radar chart

- **Data manipulation before visualization**

  Before passing the data in the python visualization package Plotly, it is essential to rescale the data into the same range, which is helpful for creating an interpretable and meaningful radar chart. If using the average position per race, average wins per race, average points, times (number of entries) and current standing directly in their original scale, there will be some inconvenience for the audience, even worse, some misunderstandings because these columns have different ranges. For instance, if the driver's current standing is higher, the number is smaller. This rule is applied for average position per race as well. Hence, theses two columns are converted into negative number.

  For Position and Current: $\boldsymbol{X' = X * (-1)}$

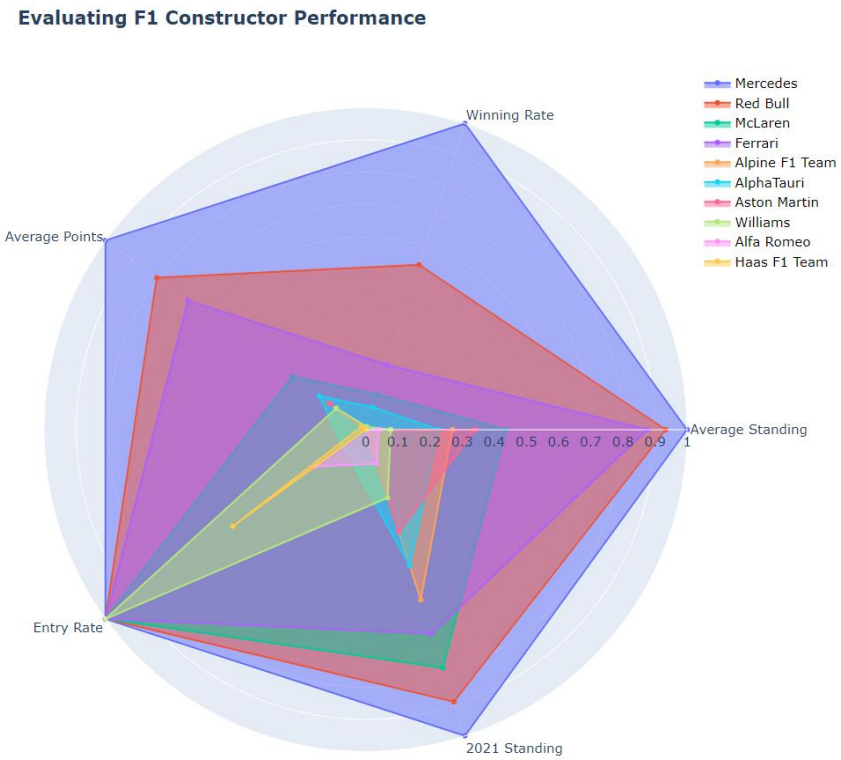
  After that, all five columns align with the rule '**the larger the number, the better**'. The next step is to rescale the data to the same range (0,1) by using MinMaxScaler from Scikit-learn package.

  For all five columns: $\boldsymbol{X' = \frac{X - X_{min}}{X_{max} - X_{min}}}$

  Rescaled Table for Constructors Performance

Constructors	Position	Wins	Points	Times	Current
Mercedes	1.000	1.000	1.000	1.000	1.000
Red Bull	0.934	0.539	0.803	1.000	0.889
McLaren	0.437	0.115	0.281	1.000	0.778
Ferrari	0.876	0.212	0.682	1.000	0.667
Alpine F1 Team	0.270	0.000	0.084	0.000	0.556
AlphaTauri	0.227	0.073	0.178	0.087	0.444
Aston Martin	0.340	0.000	0.139	0.000	0.333
Williams	0.078	0.010	0.113	1.000	0.222
Alfa Romeo	0.041	0.000	0.000	0.194	0.111
Haas F1 Team	0.000	0.000	0.016	0.510	0.000

After the data manipulation, the numbers' meaning also changed a little due to the rescale. Especially for 0 and 1 value, they do not represent none and full mark on a particular performance metric, however, 0 and 1 are illustrating the lower bound and upper bound for each column. Such as the column 'Times', the value 1 means the constructor attend all series during the selected period from 2011 to 2021. On the other hand, Haas F1 Team has two zeros for 'Position' and 'Wins', the meaning here is Haas has the lowest average position per race and lowest average wins per race among 2021 top 10 constructors.



- **Is Radar Chart suitable for this evaluation?**

  A radar chart is a multi-axis plot that displays multivariate data in the form of a two-dimensional graph and contains multiple quantitative variables' information on the same scale axis. It has many aliases such as polar chart, spider chart, circular parallel coordinate chart, etc. (Porter & Niksiar, 2018). This visualization works best when you are making decision according to different features simultaneously. Meanwhile, it uses different color for each polygon, which provides convenience for audience to identify different groups (QBUS6860, 2021).

In this research, the radar chart is a very suitable and appropriate visualization for evaluating F1 constructor performance based on 2021 standing, entry rate (how frequent they present in races), average standing, average win rate and average points per race from 2011 to 2021 seasons. Multinational businesses can identify the better-performed constructors by comparing all teams directly, and then, negotiating the sponsorship with valuable teams later. However, as mentioned in the previous data manipulation procedure, the highest value dost not equal 100%, but the upper boundary. Taking Mercedes as an example, it seems this team got a full mark on every performance indicator, but the more proper understanding is 'Mercedes is the top performance constructor in 2021 compared to other teams based on these five dimensions.' In addition, the radar chart created by Plotly is very user friendly to hide or show the team user wants to see, simply by clicking legends on the top right (for an audience who is interested, please follow the steps in the attached python code). As a result, direct comparisons among all teams or between two or three interested teams are easy to operate. Another advantage is radar chart allows the audience to decide which group is best solely based on the shaded area, which follows the rule 'the larger the shaded area, the better'. However, there are some potential issues in the radar chart, such as lacking dynamics (chronological information) and leading to potential misunderstanding if the data is not handled properly.

5.2 Evaluating Formula One Driver Performance using Radar chart

- **Data manipulation before visualization**

  Similarly, to evaluate driver performance with radar chart, the data preprocessing followed the same rule in last sector. Firstly, the data should be rescaled into 0 and 1 to ensure the data alignment of radar chart axes. Secondly, columns like 'Position', 'Current' should be inversed by multiplying negative one. Afterwards, all columns follow the rule 'a larger number indicates better performance'. However, the column 'Age' may differ slightly, it is related to driver's experiences, career time and future development potentials. Such as Fernando Alonso who is now 40 years old, he might have lower-level performances in win rate and points gained per race compared to young drivers, but he might be more experienced confronting emergency in a match.

  For Position and Current: $\boldsymbol{X' = X * (-1)}$

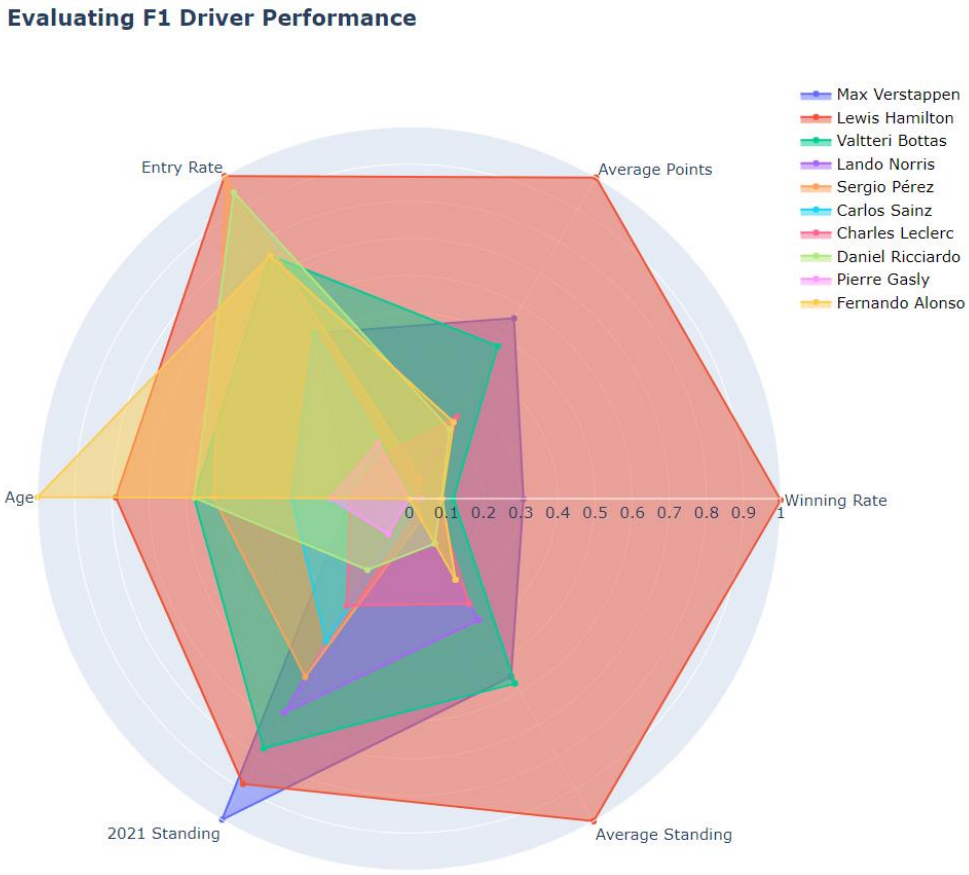
  Then, all six columns should align with the rule '**a larger number indicates better performance**'. The next step is to rescale the data to the same range (0,1) by using MinMaxScaler from Scikit-learn package.

  For all six columns: $\boldsymbol{X' = \frac{X - X_{min}}{X_{max} - X_{min}}}$

  Rescaled Table for Driver Performance

Drivers	Position	Wins	Points	Times	Age	Current
Max Verstappen	0.552	0.308	0.562	0.513	0.158	1.000
Lewis Hamilton	1.000	1.000	1.000	1.000	0.789	0.889
Valtteri Bottas	0.573	0.119	0.475	0.753	0.579	0.778
Lando Norris	0.377	0.000	0.149	0.000	0.000	0.667
Sergio Pérez	0.066	0.025	0.063	0.994	0.526	0.556
Carlos Sainz	0.034	0.000	0.013	0.513	0.316	0.444
Charles Leclerc	0.326	0.072	0.255	0.133	0.158	0.333
Daniel Ricciardo	0.140	0.087	0.217	0.949	0.579	0.222
Pierre Gasly	0.000	0.032	0.000	0.171	0.211	0.111
Fernando Alonso	0.252	0.089	0.238	0.753	1.000	0.000

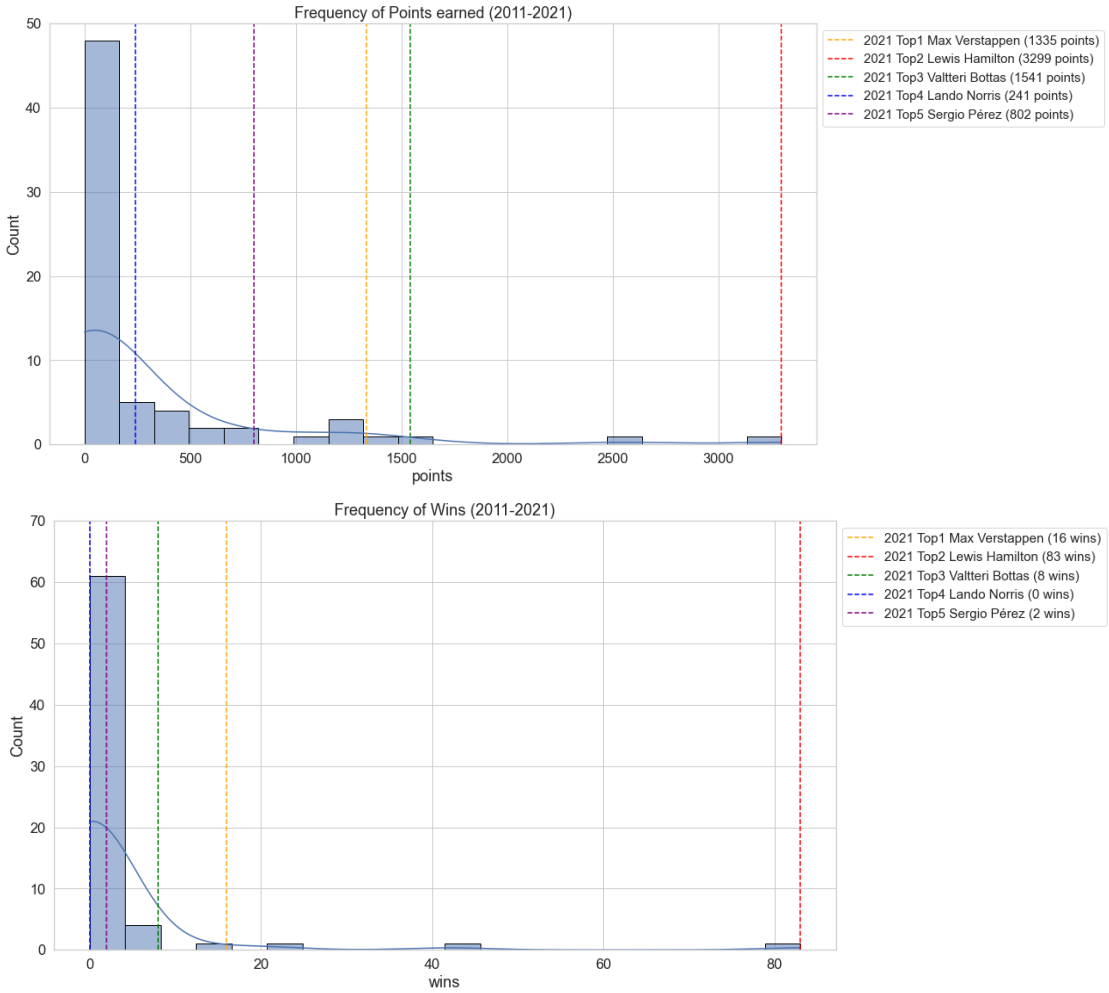
  In the rescaled driver performance table, the judging criteria are similar as the previous part for 'Position', 'Wins', 'Points', 'Times' and 'Current'. However, the target here is each individual driver rather than team.



- **Is Radar Chart suitable for this evaluation?**
  Because the methodology used in processing data for radar chart visualization is very similar to the last one, it is reasonable to conclude that radar chart is also suitable and good visualization for evaluating F1 driver performance from different metrics like average standing, points, entry rate, etc. This graph offers a comprehensive evaluation of different players, which is a very intuitive way to decide the potentially worthy driver and reach an endorsement contract with that player. And the original graph in the provided code is interactive, so it is handy to do an individual comparison for driver versus driver by the audience.

5.3 Comparing Formula One Driver wins and points using Histogram

On the other hand, there are some visualizations illustrate how rare and unusual are the top F1 drivers in 2021. This research uses histogram as the way to show the overall approximated distribution for total points and total wins respectively. During the selected period 2011 to 2021, there are 69 drivers participated in F1 races, and these two histograms are based on all recorded points and wins during this period.

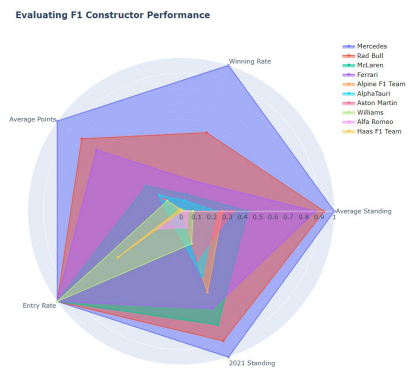


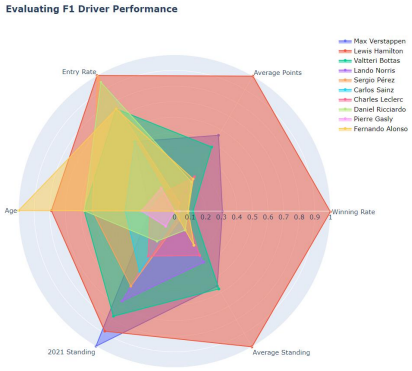
- **Is Histogram suitable for this evaluation?**

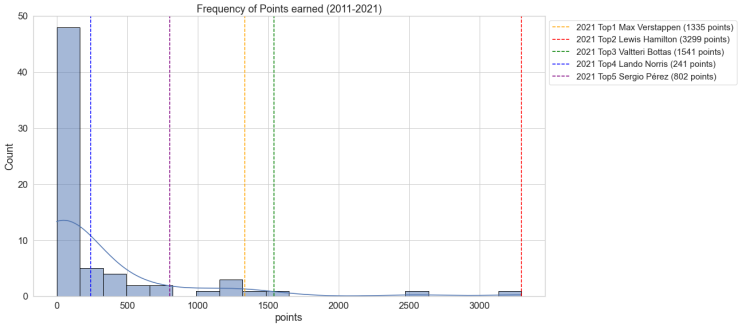
  Histograms are useful to show the shape of numerical data's distribution, and it is also a reliable way to visualize likelihoods (QBUS6860, 2021).

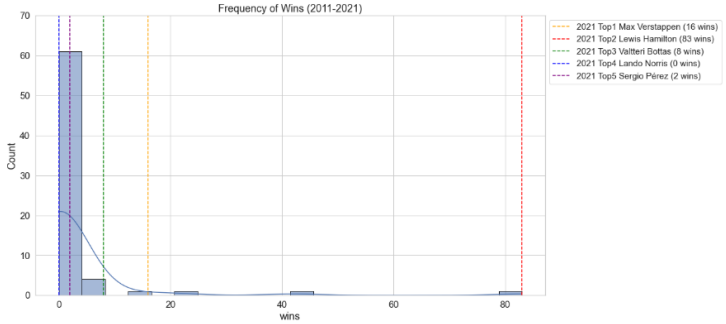
  These two histograms are not very ideal visualizations because they have some sparse categories, and the distributions are extremely highly right-skewed. But these histograms could be useful by adding some extra auxiliary lines to show how good the current top 5 drivers are (visualizing all top 10 drivers are feasible, but the graph may look chaotic) with respect to other drivers from 2011 to 2021 and where are they located in these distributions.

6. Results & Recommendations









6.1 Which constructer to sponsor?

Based on the radar chart of F1 constructor performance, it is obvious that Mercedes dominated (who occupied the upper limit in all five dimensions) F1 races in the past 10 years, followed by Red Bull, Ferrari and McLaren. At the same time, Williams has a high entry rate compared to other remaining teams, although this team has relatively lower performance on winning and scoring. However, there is evidence to prove that the current standing does not align with the overall performance (shaded area). So, there is still possibilities for lower entry rate team getting higher standing and increase their average performance indicators in the future. As a result, multinational enterprises that want to get high exposure with the help of F1 races should sponsor Mercedes or Red Bull for sure. Since these two teams are in the leading roles based on selected performance metrics, and their average performances in the past decade are relatively stable due to the larger shaded area on the radar chart.

6.2 Which driver endorse products?

According to the radar chart of F1 driver performance, it is very astonishing to see how much Lewis Hamilton overwhelmed other drivers in the past 206 races, where he played the dominant role from scoring, winning, standing and presenting perspectives. But the current top 1 driver Max Verstappen denied that in the 2021 season, who is 12 years old younger than Lewis Hamilton. And the youngest top 10 driver Lando Norris followed in 4th place, who is only 21 years old. From histograms, it is reasonable to conclude that top performance drivers are quite rare, and both winning and scoring are dominated by the leading player, therefore the majority falls in the first few categories (lower number of wins and points). Meanwhile, these two histograms produced the same conclusion for Lewis Hamilton, which is the top 1 driver in the past 10 years who defines the most wins and points. For new player Lando Norris, he still has a long journey to go. Overall, global businesses who want to build new celebrity endorsements with F1 drivers should try to negotiate with Lewis Hamilton for instant business cooperation, because he is already a legendary driver in F1 history. However, they should also consider achieving long-term cooperation with the new generation players like Max Verstappen and Lando Norris, because they have the potential to be the future 'Lewis Hamilton' in F1 competitions.

7. Limitations of the research

This report mainly relied on data-driven analysis, which lacks qualitative analysis and detailed business insights. For example, in the recent F1 fans survey (Formula 1 Official Website, 2021; Racing News 365, 2021), the top favorite team was McLaren with 29.5% of the vote, while Mercedes only gets 11.9% in the fourth. However, drivers for celebrity endorsement in this research is partially aligned since Max Verstappen and Lando Norris are popular among the overall vote and female audiences. This evidence justified that despite races performance, multinational companies also need to consider factors outside of the F1 competitions, such as constructor's reputation, driver's social media influence, etc. In addition, the cost of sponsorship and celebrity endorsement should also be taken into consideration. Simultaneously, the performance indicators should be further improved for both constructors and drivers, such as unequal career time that might be unfair for some newly joined next-generation drivers and produce biased performance comparison.

8. Reference list

Carrillat, F. A., & Ilicic, J. (2019). The Celebrity Capital Life Cycle: A Framework for Future Research Directions on Celebrity Endorsement. *Journal of Advertising, 48:1, 61-71*. https://doi.org/10.1080/00913367.2019.1579689

Cornwell, T. B., & Kwon, Y. (2020). Sponsorship-linked marketing: research surpluses and shortages. *Journal of the Academy of Marketing Science 48, 607-629*. https://doi.org/10.1007/s11747-019-00654-w

Essentially Sports. (2021). *Lewis Hamilton 2021: Net Worth, Salary and Endorsement*. https://www.essentiallysports.com/tag/lewis-hamilton/

Formula 1 Official Website. (2021). 2021 Driver Standings and Constructor Standings. Retrieved October 24, 2021, from https://www.formula1.com/en/results.html/2021/drivers.html; https://www.formula1.com/en/results.html/2021/team.html

Formula 1 Official Website. (2021). *New global fan survey reveals Formula 1's most popular team and driver*. https://www.formula1.com/en/latest/article.new-global-fan-survey-reveals-f1s-most-popular-team-and-driver.2evqTWEPsWYeBlw9vmSyBs.html

Lange, D. (2021). *F1 - Statistics & Facts*. Statista. https://www.statista.com/topics/3899/motor-sports/#dossierKeyfigures

Medland, C. (2021). *INSIGHT: What's driving F1's U.S. boom*. Racer.com. https://racer.com/2021/08/10/insight-whats-driving-f1s-u-s-boom/

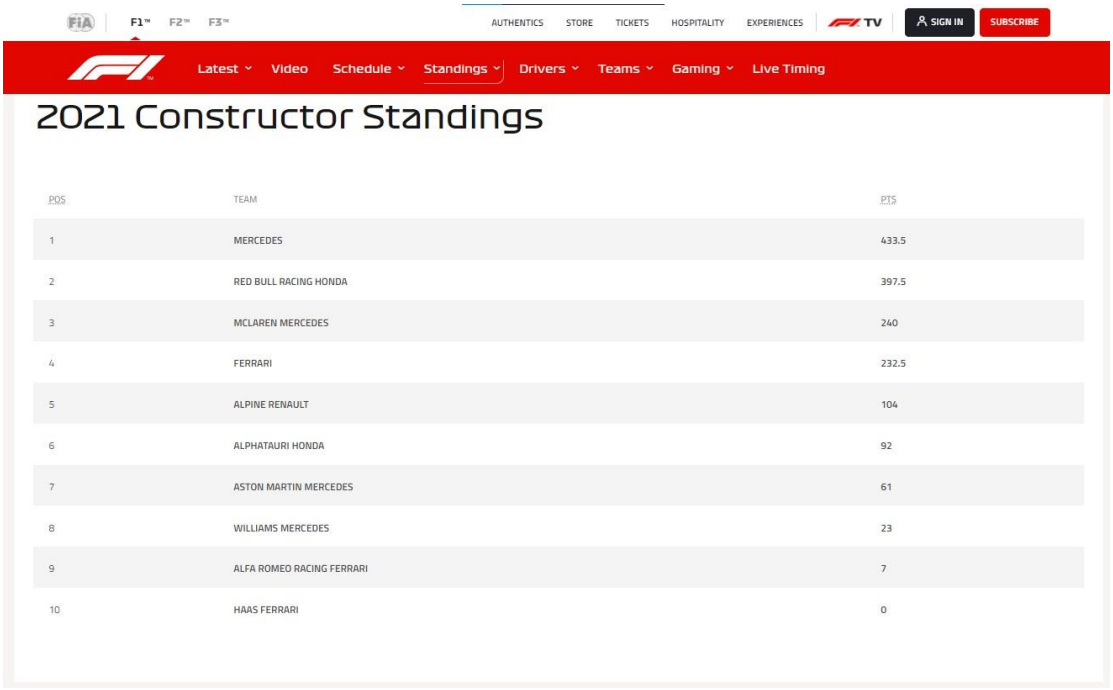
Pledge sports. (2017). *10 Most Popular Sports in The World*. https://www.pledgesports.org/2017/06/10-most-popular-sports-in-the-world/

Porter, M. M., & Niksiar, P. (2018). Multidimensional mechanics: Performance mapping of natural biological systems using permutated radar charts. PLoS ONE 13(9): e0204309.https://doi.org/10.1371/journal.pone.0204309

QBUS6860. (2021). *Visual Data Analytics - Theory Lecture 2: Data Types and Visualisation Types*, University of Sydney Learning Material, accessed at https://canvas.sydney.edu.au/courses/36335/pages/week-2-theory-lecture-and-materials-data-types-and-visualisation-types?module_item_id=1299789

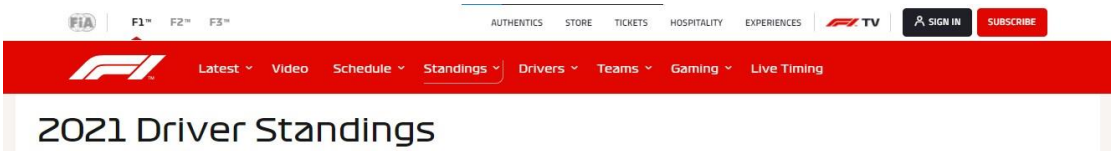
Racing News 365. (2021). *New F1 survey: Who are Formula 1's most popular drivers and teams*. https://racingnews365.com/new-f1-survey-who-are-formula-1s-most-popular-drivers-and-teams

Appendix



2021 Constructor Standings

POS	TEAM	PTS
1	MERCEDES	433.5
2	RED BULL RACING HONDA	397.5
3	MCLAREN MERCEDES	240
4	FERRARI	232.5
5	ALPINE RENAULT	104
6	ALPHATAURI HONDA	92
7	ASTON MARTIN MERCEDES	61
8	WILLIAMS MERCEDES	23
9	ALFA ROMEO RACING FERRARI	7
10	HAAS FERRARI	0



2021 Driver Standings

POS	DRIVER	NATIONALITY	CAR	PTS
1	Max Verstappen	NED	RED BULL RACING HONDA	262.5
2	Lewis Hamilton	GBR	MERCEDES	256.5
3	Valtteri Bottas	FIN	MERCEDES	177
4	Lando Norris	GBR	MCLAREN MERCEDES	145
5	Sergio Perez	MEX	RED BULL RACING HONDA	135
6	Carlos Sainz	ESP	FERRARI	116.5
7	Charles Leclerc	MON	FERRARI	116
8	Daniel Ricciardo	AUS	MCLAREN MERCEDES	95
9	Pierre Gasly	FRA	ALPHATAURI HONDA	74
10	Fernando Alonso	ESP	ALPINE RENAULT	58