

Project 2: Bahrain Gran Prix

Project 2: Bahrain Gran Prix

Data base

During this project we are studying three different datasets:

- **Before the race:** From the 22nd of March to the 25th of the same month.
- **During the race weekend:** All three days, 26th, 27th and 28th of March.
- After the race: Between the 29th to the 31st.

We have chosen this schedule because the first news of the gran prix started during race week, so on the 22nd. However, the last important new of the Bahrain gran prix was on the 30th, when it was discovered why Fernando Alonso retired, recording the last discussions about the gran prix on the 31st. Also, on the 01st of April it was April fools and it was the anniversary of the famous Bahrein gran prix of 2014 (when a battle between Rosberg and Hamilton took place) which created a different discussion which we did not wanted to record.

On the other hand, the database focuses on two languages, English and Spanish, both of them having a much greater presence in the three databases that we have downloaded. Another reason for this is due to our understanding of languages, being the knowledge of the members of the group fluent in both.

	lang	n		lang .	n		lang	n
	<chr>></chr>	<int></int>			<int></int>		<chr></chr>	<int></int>
1	en	<u>28</u> 589	1	en en	46386	1	en	<u>33</u> 420
2	es	<u>4</u> 041		es	<u>6</u> 556	2	es	<u>4</u> 616
3	fr	<u>2</u> 387		fr	<u>3</u> 865	3	it	<u>2</u> 457
4	it	<u>1</u> 443		it	<u>2</u> 316	4	ja	<u>2</u> 278
5	pt	639		ja	1823	5	fr	<u>2</u> 112
6	de	553		pt	1798	6	da	800
7	und	548		und	1675	7	und	786
8	ja	352		tr	1649	8	pt	760
9	tr	307		ar	437	9	tr	696
10	nl	197	10		354	10	n1	399

Objective: Emotional study

Our objective is to combine both, the r study about the main topics talked during the different timelines with the real events that happened on that weekend. Because of it, you will see a common structure in the explanation which will go like this:

- R study about the main topics: Cloud graphics.
- Events that happened at that time: Different news and media which will be cited at the end of the project.



Project 2: Bahrain Gran Prix

Before the race



During the first word cloud we can see mainly the excitement of having back the competition. Most of the expressions of exactment can be seen in:

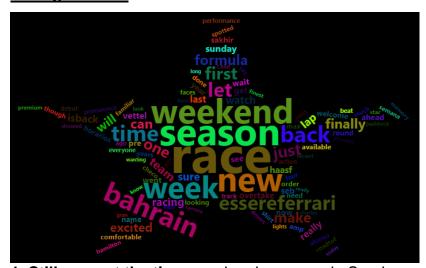
- **Place:** With Bahrain having more than 3.000 mentions in all the tweets in English and Spanish.
- Words like: Back, new, race, week, first... all of them put it

together showing that the sport is back in town.

However, the only curious thing we have in this study is that only one team seems to have a head start over other specific formula 1 words. Obviously, Ferrari being the only team which breaks with all the standards and general words in the word cloud. This fact is curious taking into account that Italian language was not considered in this study, being an Italian team with the *tifosi* (Italian fans) as its main asset, we were expecting another classical English team like Mclaren or Williams to have a higher presence.

Nevertheless, we cannot forget that Spanish is being considered in the study, which could have derived into a higher mention of Ferrari due to being the debut of Carlos Sainz in this team. A conclusion which has been drove by looking at the huge amount of presence of the word debut, with almost 1.000 mentions.

During the race



26th Between the the 28th and March, during the race, the audiences' sentiment did not change much. Most of the words used described feelings, and as we see in starwordcloud there different are expressions:

- 1. Still present the time: weekend, new week, Sunday, semana...
- **2. Sentiment**: back, excited, isback, performance, welcome, join, beat, premium...

Interesting during the race is how the **place** tag changed, if the #bahrain at the begging had more 3.000, the same tag during the race was used as often as

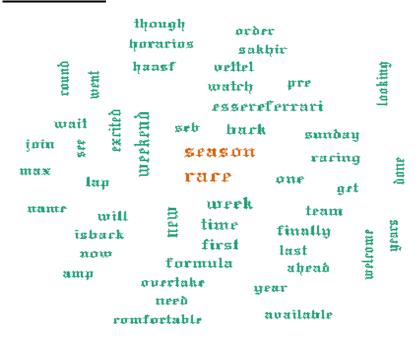


Project 2: Bahrain Gran Prix

before the race, but a more specific word #bahraingp has been used by 32.748 people.

Here we can see, how just adding 2 different letters (gp) would change the study. Thus, #bahraingp is the word most used, followed by #race with a frequency of 5947. And still, just 86 people used #joking to describe their sentiment in watching the race (the reason being a penalty given to Verstappen that made him lose the race win).

After the race



For the last wordcloud we decided to use a modern and unique design.

However, the most important feature of the graph is the limit on the words that it has, enable us to study closely new words have appeared.

Straight away we can see that still a lot of the words which already

appeared in the other two graphs are still on this one.

Nevertheless, some other words gain importance, like "seb", "Vettel" or "esserferrari". The reason for the increase in those words is due to the poor performance that the German driver had in the race, compared to the huge improvement its previous team.

Lastly, it is quite remarkable, that when comparing the English words to the Spanish ones we can see a change in the topics. While, the English words bring discussions and talk about the race, the only Spanish word that appears in the cloud is "horarios" probably showing the amount of relevance Formula 1 has in the Spanish country compared to the British one. That brings us to the conclusion that the British people came up with a larger amount of discussions about the race than Spanish folks, who were worried about the time when it was starting (typical Latin style).