How I tried to prove the Filter bubble Theory: Twitter Sentiment Analysis

Some weeks ago, I was in my house in Washington DC watching the Inauguration Day on TV. I realized that I was probably living one of the most important moments of the history of the next couple of years.

As many Americans, I was also surprised when I saw Donald J. Trump elected to be the next president of The United States. From my point of view, every newspaper was mentioning bad comments about him, all news on the internet were negative about him and all tweets were against him… However, the elections did reflect something different. So, what happened?

When I was trying to find an answer to that question I remembered the [Eli Pariser’s theory](https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=es) about the [**filter bubble**](http://www.thefilterbubble.com/) **problem**. This theory suggest that we live inside bubbles of information where all that you want to read appears on your internet’s screens. So, we wouldn’t be having objective perspectives about public things.

So while watching TV I was thinking that I was suffering the **filter bubble problem**, as Eli Pariser suggest in one of his talks. This theory

Some weeks ago, when I was watching the Inauguration’s Day on TV I was wondering what had happened to all the news that I had read, all the posts that I used to see on my twitter’s and Facebook’s walls.

Then I remembered the [Eli Pariser’s theory](https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=es) about the [**filter bubble**](http://www.thefilterbubble.com/) **problem**. This theory suggests that we live inside bubbles of information and we are not able to access to objective information. It

On that moment I thought to myself, if American citizens could vote today again, would we see the same results?

For that reason, I decided to apply **The Power of Big Data** to know what was people´s thoughts in that moment. So I took my best guns and started to work (code): Twitter, R sentiment analysis libraries and Carto.

?

So, coming back to the moment in the coach, I asked to myself: Has US citizens changed their mind since the day that they voted? I mean, I they had to vote again today, would the results have been the same?

To solve this questions I did the following simple steps:

## Get a feeling about people´s conversations: Twits.

To get a feeling about people´s conversation I get 36XXX from twitter containing the words Trump and #InaugurationDay. In order to being able to plot them in a map, the only requirement I asked for was that they should have a geo tag. Thanks to the Carto tool, they made my live easier!

In another post I will write how to use the tool!

## Run a text analysis to extract the sentiment

## Aggregate the results by state

## Plot in a map and compare result with the election´s results map

## [WHAT IS AMERICA THINKING ABOUT THE NEW PRESIDENT? (Through twitter sentiment analysis)]

[DATA SCIENTISTS = BUSINESS + PROGRAMMING + STATISTICS]

[STEPS]

1.- What is America thinking? (Business problem definition + maps)

2.- Step by step to data mining (Programming)

2.1.- Extract twits with Carto

2.2.- Clean dataset

2.3.- analyse sentiment of tweets (Tokenize + word sentiment)

3.- Hypothesis testing (T-test) (Statistics)

4.- Conclusions (Business)

Note: A data scientist is something blured in between a business person, a programmer and a scientist. We know something of everything but nothing of everything…

(Links to github and notebooks in R)