

Matemáticas en Twitter



Viviana Márquez

Konrad Lorenz Fundación Universitaria
Clase: Big Data — Profesor: Jorge Camargo

Noviembre 11, 2017

Objetivo inicial

Se recolectarán tweets que incluyan la palabra **matemáticas** en varias ciudades de Colombia, con el objetivo de posteriormente realizar un análisis de los tuits publicados. A través de este proyecto se busca determinar la actitud que los tuiteros colombianos tienen hacia la matemática.

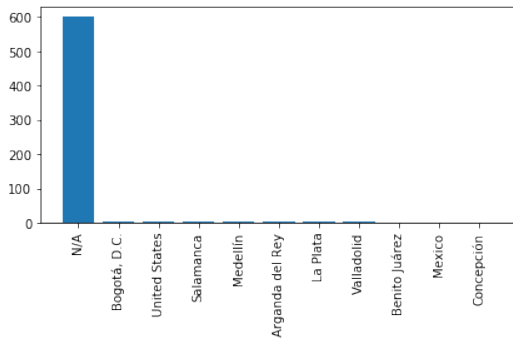
- Dada la falta de tuits con la palabra **matemáticas** que estuvieran geolocalizados en Colombia, se expandió la búsqueda al mundo entero.
- El patrón continuo en todo el mundo, así que el objetivo se modificó a estudiar la actitud de los tuiteros hacia la matemática en general.

Se realizó la búsqueda de los tuits con los siguientes parámetros:

```
In [3]: #Parámetros
latitude = 4.624335
longitude = -74.063644
max_range = 20000
num_results = 1000
quest = "matematicas"
outfile = "output_matematicas.csv"
```

```
In [4]: result_count = 0
last_id = None
while result_count < num_results:
    query = twitter.search.tweets(q = quest, count = 100, max_id = last_id)
```

Podemos ver que la mayoría de tuits no declaran su geolocalización:

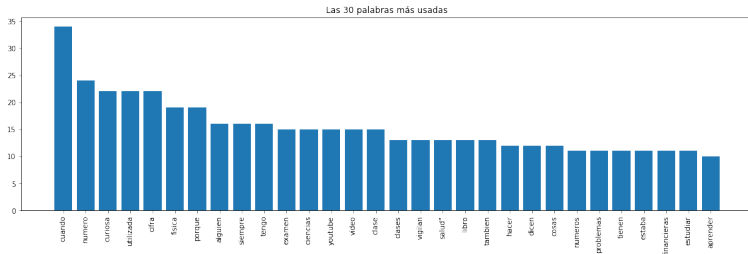


- Al hacer el estudio de palabras, se descartaron las que tienen un tamaño menor a 4 caracteres dado que estas son conectores y no brindan mucha información.
- También se removi6 la palabra matemáticas, pues estaba en todos los tuits.
- En adici6n, se decidi6 estudiar el conjunto de datos sin los RTs.

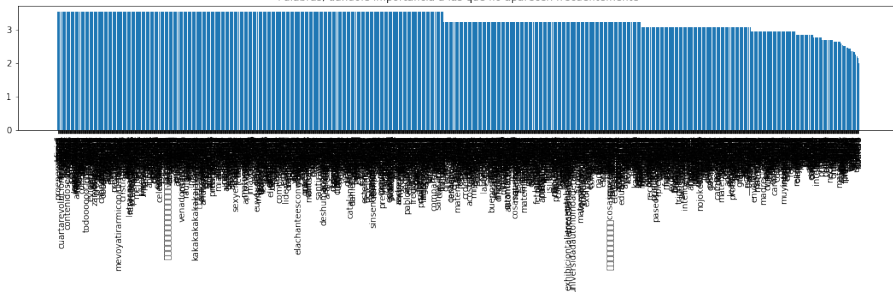
TOP 15 de palabras

- cuando
- numero
- curiosa
- utilizada
- cifra
- fisica
- porque
- alguien
- siempre
- tengo
- examen
- ciencias
- youtube
- tambien

Bolsa de palabras



Vemos que con d_f no se obtiene nueva información sobre los datos:



Algunos tuits que incluyen la palabra problema:

matematicascsl
acabo de ver las cuatro etapas del cerebro para resolver problemas matematicos haga clic ver tambien ➡ <http://stcopn-ddxy2rzf>

BrandoPk123
se suicido cuaderno de matematicas tenia muchos problemas 🙄🙄🙄🙄

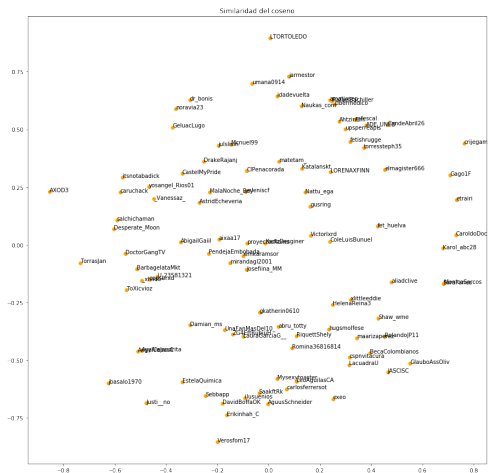
ColeLuisBunuel
resolviendo problemas de matematicas en 5°primaria coleluisbunuel <http://stco9ejccdez4>

nicolargacha
parece que siguen con "problemas de matematicas los muchachos" <http://stcosrr20qw5av>

Argenttty
la solucion de los problemas esta en el propio problema o eso dicen las matematicas muggles <http://stcojy2ynm6an4>

Similitudes

Similaridad del coseno:



Bajo la similaridad del coseno, tuits que están lejos vs cerca:

```
In [23]: print tuits["dr_bonis"]  
print ""  
print tuits["Verosfoml7"]
```

wrojewski seranodino que el sistema de ingreso a facultades de medicina seleccione principalmente a chicas empoll htt
pstcohxpxjjwndz

doelgomo estefanymena3 lauravignatti nadaldrama minorfanita alvaralai mimifdez9 anakaroulloa httpstcopexbuijscj

```
In [24]: print tuits["SaraFaries"]  
print ""  
print tuits["MonicaSarcos"]
```

numero pi una curiosa cifra utilizada para casi todo matematicas httpstcohvwjmepls7

numero pi una curiosa cifra utilizada para casi todo matematicas httpstcontbkqligpp

K-Means

