



Title:

10 Academy July 2020 Training - Weekly Challenge: Week 1

Topic:

African influencers: Twitter users segmentation

Goal:

Identify influencers rank position from Twitter data.

Method:

The following are the main steps to completing this challenge

1. Web scrape [100 most influential Twitter users in Africa](#) using Python or Bash to obtain the 100 African twitter influencers. This website uses three key metrics to find the top 100 influencers
 - a. **Popularity (Retweet Influence)**: measured by the number of Retweets and Likes users get
 - b. **Reach (Indegree Influence)**: measured by the size of their audience
 - c. **Relevance (Mentions Influence)**: measured by the relevancy of their content
2. Web scrape [African leaders respond to coronavirus... on Twitter](#) using Python or Bash to obtain the twitter account of African top government officials. [Top 18 African Heads of State on Twitter: it's a mixed bag](#) ranks African governments based on their followers and other metrics.
3. The key paper listed in the reference [Measuring User Influence in Twitter: The Million Follower Fallacy](#) gives a good understanding of identifying/segmenting influencers on twitter data. **You are required to read this paper.** Segmenting influencers on networked data, such as mobile money transaction, supermarket, telecommunication, etc. is important to shape business strategies.
4. Write a python code to search and download the tweets posted by all the 100 influencers and top government officials. You have already seen the core part of this coding during the pre-training week. A fully working error free notebook is posted [here](#).

5. By analysing their followers count, number of likes, number of retweets, number of people they follow, number of tweets they shared, score each of the influencers and top government officials as follows:
 - a. $\text{popularity_score} = \# \text{retweets} + \# \text{likes}$
 - b. $\text{reach_score} = \# \text{followers} - \# \text{they follow}$
 - c. $\text{relevance_score} = \# \text{comments} + \# \text{mentions in other tweets}$
6. Extract the top 5 unique hashtags these influencers and government officials used in their top 10 retweets. Group them based on their top 5 hashtags similarities.
7. Make a bar plot to show the fraction of influencers and top government officials by hashtag.
8. Depending on how far you managed to go, interpret your finding in a report that doesn't exceed 3 pages and an accompanying slide deck of no more than 10 slides.
9. Share your finding in your twitter account using #10AcademyCoronaResponse hashtag and tag @10acad as well.

Guidelines on your report & slides:

- Let's imagine that you are doing this work for a large multinational company (let's use Nike) that is planning a new digital campaign in Africa, and wants to understand who the social media influencers are in Africa and in which area fall into (e.g. politics, fashion, art, etc.). This is key to Nike as they don't want to be associated with the wrong influencers (ethnic and religious polarisation figures for example). This work is commissioned by a combination of their strategy and their marketing departments.
- They want to think about:
 - Could they try to partner with or hire some key influencers to amplify their planned digital campaign?
 - How are the influencers endorsed by key government officials?
 - Which topics are important for both influencers and key government officials?
- Your report and accompanying slides will be aimed at the heads of marketing and strategy for the [EMEA](#) area.
- Template suggestion (each user can iterate/edit this) - the same progression can be used for both the report and the slides
 - Your understanding of what the client is looking for
 - What analysis you performed. Be sure to highlight limitations of the analysis.
 - Key outcomes (incl. graphics) that are relevant to the questions the client wants to answer. This may include highlighting particular users, hashtags and/or trends that could be relevant to your client. This may also include topics to be avoided.
 - Bonus: Suggest a course of action - who to partner with and which hashtags they could align with. Even better would be to look for new trends and dimensions that may provide actionable insight in how to segment

influencers.

- Bonus: Sticking with the assumption of the client being Nike, what behaviour do you see from your analysis which they should be aware of when extending their digital campaigns to Africa.

Assumptions:

It is hard to determine the influence of users who have few tweets. Use the concept of “active users” from the traditional media research (Levy and Windhal 1985) and ensure all of the 100 influencers accounts are active users - they have a minimum level of activity. For this challenge we define 5 tweets in the search period to be the minimum level of activity.

Due Date and Submission:

Sunday 19 July 2020 at 1800 GMT.

Submit:

- PDF of your report
- PDF of your slides
- URL of your twitter post
- URL of Github repo to this project
- Upload your notebook. If you have multiple notebooks, zip the folder and upload the zipped file.

A submission mechanism will be provided.

Grading Rubric:

Please see weekly overview

Portfolio:

Be ready to post the following in your portfolio that will be shared with prospective employers:

- Your Github repository link for this project
- URL to report. If you can, convert your report as a blog entry in your github pages.

Support:

Your key tutor for this week is Jean-Henock. The rest of the team is available to support, of course.

Reading Material:

Key paper and blogs are to be used as major references.

1. **Key Paper** (pdf): [Measuring User Influence in Twitter: The Million Follower Fallacy](#)
2. **Key Data Blog**: [Top 18 African Heads of State on Twitter: it's a mixed bag](#)
3. **Key Data Blog**: [African leaders respond to coronavirus... on Twitter](#)
4. **Key Scrapping Blog**: [Practical Introduction to Web Scraping in Python](#)
5. Paper (pdf): Anger & Kittl (2011) [Measuring Influence on Twitter](#)
6. Blog: [How to calculate Twitter impressions and reach](#)
7. Blog: [A Beginner's Guide to learn web scraping with python](#) - using selenium
8. Blog: [Web Scraping using Python](#) - more complete tutorial
9. Blog: [Advanced Web Scraping using Scrapy](#) - advanced use