EPPS6356 Data Visualization

Data Visualization Project Proposal

Federico Ferrero

Area/Topic:

The #AlevelResults Case: Visualizing Social Networks and Text Mining on Twitter Data

Context of the research questions: The *#AlevelResults* case

Every year in the United Kingdom, "Level A" exams are taken by young people to determine which university they can aspire to. However, this year such tests were suspended due to the COVID-19 pandemic and the school lockdowns. Instead, high school teachers were asked to give a predicted grade for their pupils and then rank them in order within their class. Posteriorly, an algorithm was expected to only adjust and standardize results at the national level.

With the purpose of calculating the final grades, the algorithm of the Office of Qualifications and Examinations Regulation (named "Ofqual") used the historical performance of schools and the ranking of students within their own school.

The result was that 37% of estimated grades were lowered and 5.3% were raised by Ofqual's algorithm in comparison with teachers' predictions. The unexpected effect was that the algorithm ended up benefitting private schools over public schools. This meant that results favored those students who were enrolled in better schools even if their grades were worse, compared to those in schools with historically worse performance. As a result, in some cases, bright pupils in poorly performing schools saw their grades lowered because last year's cohort of pupils did not do well in their exams.

It is evident that such a situation can have an impact on the future life of students since it conditions the higher education institutions they can access and limits their opportunities for social mobility.

As soon as the results were known on past August 13th, the demonstrations of students, teachers, and educational institutions on social networks did not wait. Mainly on Twitter with the hashtag #AlevelResults, comments about lack of consistency, the erratic and disconcerting character of the results, and anomalies in the construction of the model used by Ofqual were just some of the main public manifestations.

As corollary, in the face of media pressure, the United Kingdom government ended up discarding the A levels results corrected by the algorithm and established a plan of future evaluations and appeals to mitigate the biased results.

Seen in perspective, the "#AlevelResults" case is one of the first times that public opinion has spoken out so strongly about results considered unfair coming from evaluations driven by algorithms in educational settings. Undoubtedly, the analysis of such data can be a great opportunity to explore the reaction and the arguments offered by students and teachers. To everyone's concern, it is expected that situations like these will become more and more frequent in educational organizations.

Research Questions

The general objective is to explore data from Twitter with the hashtag *#AlevelResults* in order to produce visualizations about the social network and about the content of the published tweets.

Specifically, the main questions that will be addressed are:

• How has the reaction around #AlevelResults on Twitter developed in time?

• From which geographic locations did users participate?

• What is the network or sub-community that most participates in the discussion (tweets with responses and interactions)?

• Who are the top users involved in the discussion around #AlevelResults on Twitter?

• What feelings are associated with the texts published in the tweets?

Data

78,839 tweets have been downloaded from Twitter from 8/14/2020 to 9/2/2020 (when the last tweet about the topic was recorder) using as a criterion the hashtag #AlevelResults. This hashtag was trending topic when the discussion exploded on past 14th of August. The database can be explored in the GitHub repository created for this visualization project: https://github.com/federico-jf/a_level_results_visualization_project

The variables that will be worked with are the following:

- Id_str
- From_user
- Text
- Created_at
- Time
- In_reply_to_user_id_str
- In_reply_to_screen_name
- From_user_id_str
- In_reply_to_status_id_str
- Profile_image_url
- User_followers_count
- User_friends_count
- User_location

Visualization method(s)

The visualization techniques to use include general charts to initially explore data. Then, Social Network Analysis will be implemented to visualize the network of interactions (influential users will be highlighted in size, among other graphics marks). Finally, if time is available, Text Mining techniques will be applied (specifically Sentiment Analysis) over Twitter content with hashtag #AlevelResults.

While there is some experience with Software Network Analysis using Gephi application, the challenge of this proposal will be the use of R to construct and analyze the social networks.

Therefore, R will be used with RStudio.

<u>Bibliography</u>

Abdelsadek, Y., Chelghoum, K., Herrmann, F., Kacem, I., & Otjacques, B. (2018). Community extraction and visualization in social networks applied to Twitter. *Information Sciences*, 424, 204-223.

Bruns, A. (2012). How Long is a Tweet? Mapping Dynamic Conversation Networks on Twitter Using Gawk and Gephi. *Information, Communication & Society*, 15:9, 1323-1351, DOI: 10.1080/1369118X.2011.635214

Ediger, D. et al., (2010). Massive Social Network Analysis: Mining Twitter for Social Good. *39th International Conference on Parallel Processing*, San Diego, CA, 583-593, DOI: 10.1109/ICPP.2010.66.

Kolaczyk, E. D., & Csárdi, G. (2014). *Statistical Analysis of Network Data with R*. New York: Springer.

Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. New York: O'Reilly Media.

<u>Journalistic References about "A level exams" Case</u>

Alex Hern (14 August 2020). Do the maths: why England's A-level grading system is unfair. The Guardian. Retrieved from: https://amp.theguardian.com/education/2020/aug/14/do-the-maths-why-englands-a-level-grading-system-is-unfair?__twitter_impression=true

Amit Katwala. (15 August 2020). An Algorithm Determined UK Students' Grades. Chaos Ensued. Wired. https://www.wired.com/story/an-algorithm-determined-uk-students-grades-chaos-ensued/

Dan Davies (15 August 2020). This year's A-level results are a fiasco – but the system was already broken. The Guardia. Retrieved from: https://www.theguardian.com/commentisfree/2020/aug/15/a-level-results-system-ofqual-england-exam-marking

Donna Ferguson and Michael Savage (16 August 2020) Autumn term chaos feared over exam resits and appeals. The Guardian. Retrieved from: https://www.theguardian.com/education/2020/aug/16/autumn-term-chaos-feared-over-exam-resits-and-appeals

Heather Stewart and Kate Proctor (17 August 2020)Penny Mordaunt becomes first minister to voice concerns about A-level crisis. The Guardian. Retrieved from: https://www.theguardian.com/education/2020/aug/17/delaying-a-year-wont-be-an-option-for-many-a-level-pupils-says-penny-mordaunt

Helen Pidd (14 August 2020). 'Punishment by statistics': the father who foresaw A-level algorithm flaws. The Guardian. Retrieved from: https://amp.theguardian.com/education/2020/aug/14/punishment-by-statistics-the-father-who-foresaw-a-level-algorithm-flaws?__twitter_impression=true

No specified author (13 August 2020). A-levels: 'Dreams ruined by an algorithm'. BBC News. Retrieved from: https://www.bbc.com/news/uk-northern-ireland-53767773

No specified author (13 August 2020). A-levels: Over a third of NI results lower than estimated grades. BBC News. Retrieved from: https://www.bbc.com/news/uk-northern-ireland-53757767

Ofqual (2020). Awarding GCSE, AS, A level, advanced extension awards and extended project qualifications in summer 2020: interim report. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/909368/6656-1_Awarding_GCSE__AS__A_level__advanced_extension_awards_and_extended_project_qualifications_in_summer_2020_-_interim_report.pdf

Sean Coughlan, Katherine Sellgren, Judith Burns (13 August 2020) A-levels: Anger over 'unfair' results this year. BBC News. Retrieved from: https://www.bbc.com/news/education-53759832