



UK Data Service - Changing perceptions on social media

Sima Aykin | BSocSc Sociology

Overview of the Data Fellowship

This Q-Step project on changing perceptions on social media at the UK Data Service gave me an opportunity to work closely with another intern over an 8-week period. We chose to explore attitudes towards veganism and plant-based foods and decided on Twitter as the social media platform.

At first, we conducted a literature review to learn more about our research topic, which helped us to develop our research questions and hypotheses. We then used the Twitter API function to scrape the data and this was possible by writing a code in Python. To analyse the data we gathered, we used Python to produce data visualisation tools (e.g., histograms, word clouds and heat maps).

At the end of the internship, we presented our findings to the UK Data Service team through a short presentation and a detailed written report.

A collaborative blog post in which I reflected on my experiences at the UK Data Service is published on the UKDS Impact blog.

Data Analysis

Our main aim was to access a diverse sample, hence we collected data across multiple cities in the UK during 'Veganuary 2019'. Veganuary is an annual challenge that encourages a vegan lifestyle for the month of January.

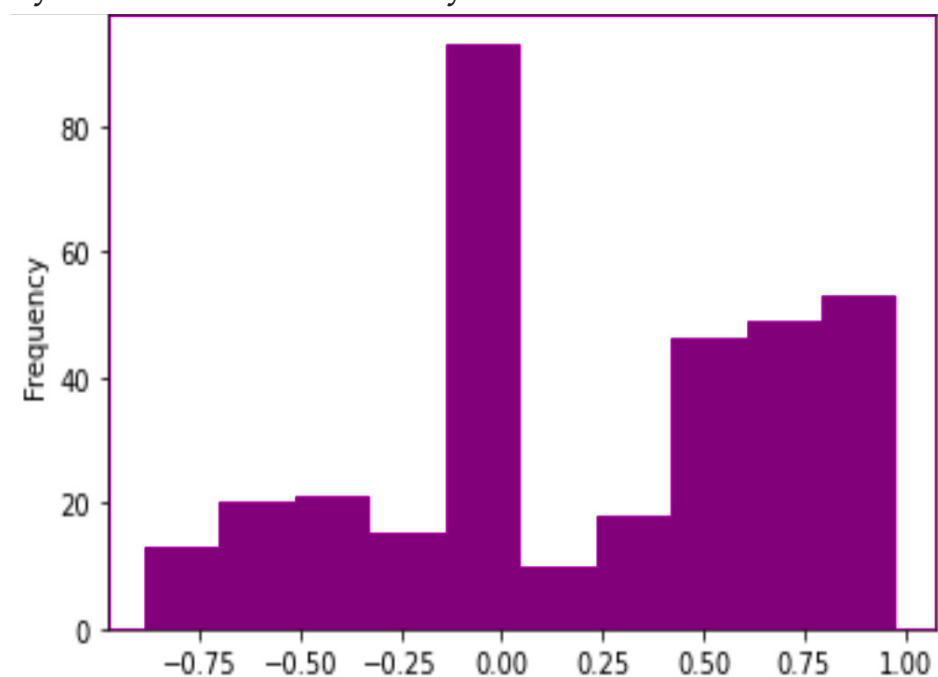


Figure 1: Histogram showing the frequency and sentiment of Manchester tweets

We searched the keywords "plant based" and "plant-based" when scraping the Twitter data. To avoid sampling bias, we used a random sampling method by selecting random tweets over the timeframe we have chosen. By using 'sentiment analysis', we measured attitudes of each tweet we gathered.



Sentiment analysis can detect changes (pos/neg) in attitudes towards plant-based foods from the Twitter data. We set our scale for the sentiment analysis as -1 (negative), 0 (neutral) and 1 (positive). We covered London, Manchester and Preston in our research to explore differences and similarities between regions. Prior to the data analysis, we believed that there would be a significant relationship between attitudes and geographical location.

Findings

- The sample size for London was bigger than the other cities. The analysis showed that the mean sentiment was 0.23 which reveals that the majority of tweets were positive and neutral.
- As seen in Figure 1, for Manchester, the mean sentiment was 0.22 which shows a similar result with London.
- The mean sentiment for Preston was 0.13, this implies that people in Preston are less positive when tweeting about plant-based foods compared to the two big cities we researched.

Key Skills Learnt

- Web scraping with Twitter API
- Familiarity with the Python programming language
- Data visualisation techniques (heatmaps, graphs)
- Collaboration with team members
- Managing workload efficiently under strict deadlines
- Data presentation skills