UK Data Service - Changing perceptions on social media

Rumeysa Piskin | Department of Social Sciences

Overview of the Data Fellowship

This project on changing perceptions on social media, for the UK Data Service, required me and another intern to pick a topic and create a report. We picked the topic of attitudes towards veganism/plant based foods because we were both interested in this area and chose the social media platform, Twitter. We used the Twitter API function to scrape the data we needed and used the programming language python to do this. Throughout we typed up our literature review and developed questions and hypotheses for the report. Once we had the datasets from Twitter we used python to analyse the data and create visualisations of our findings e.g. histograms, heat maps and word clouds. We then competed our report and developed a presentation for the rest of the UK Data Service interns and leads. We finished up the internship by presenting these findings and showed the results of our eight weeks of work with them.

We also collaborated on a blog post, which will be displayed in November 2021 on the UK Data Services website, discussing our experiences as Q-step interns.

Data Analysis

We decided to collect data across multiple cities in the UK, during "Veganuary 2019". We specified the keywords "plant-based" and "plant based" when scraping the twitter data. These keywords allowed us to have a more diverse sample. To avoid sampling bias, we used a random sam-

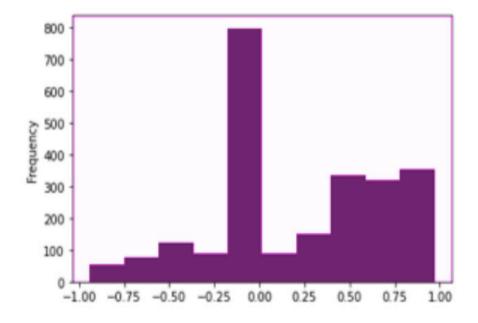


Figure 1: Histogram showing the frequency and sentiment of London tweets. pling method by selecting random tweets over the timeframe we chose. We also decided to use sentiment analysis. Twitter provides a variety of information, in which people post real time messages about their opinions on many topics, this created the challenge of summarising an overall sentiment for the data. Coding the sentiment analysis allowed us to detect these (pos/neg) changes in attitudes towards plant based foods from the



Manchester Tweets - Word Cloud

Twitter data we collected. Attitudes are often used together interchangeably with sentiment analysis and allowed us to make conclusions about our data. We set our scale for the sentiment analysis as -1 (neg), 0 (neu) and 1 (pos). We chose to cover Manchester, London and Preston in our project. The goal was that we had two or more regions to compare between that would show either similarities or differences in attitudes towards plant based foods. Initially, we expected to find a polarisation in attitudes.

Findings

- Analysis showed that people in Preston do not tweet as much about plant based foods as people who live in London and Manchester, possibly due to the varying population size between the locations.
- Tweets gathered from London showed a mean sentiment analysis of 0.23. Figure 1, shows that the majority of tweets in London are positive in sentiment, with few being negative, and then a peak of neutral tweets. This was an unexpected. Manchester showed a mean sentiment of 0.22 and Preston was 0.13. The outcomes from all three areas suggested that tweets are mainly neutral or positive in sentiment.
- We also coded a word cloud, using python, for each location. Manchester's wordcloud can be seen above. The biggest words show the most tweeted e.g., vegan, veganuary, and dairyfree.

Key Skills Learnt

- Learning the programming language (python)
- Web scraping, using the Twitter API function
- Collaboration with other team members at a professional work-level
- Managing and organising a heavy workload with strict deadlines