

DSC-65 Wales on social media

2018-06-29

Summary

This project aims to better understand what other countries talk about when they mention Wales on social media.

Team members

- Steven Hopkins

The need

Twitter has been used as the data source for this project where tweets are analysed for key topics or themes. The topics are then assessed on their sentiment which will give a basic indication of which topics are generally positive and which are negative.

Impact

Data is obtained through the Twitter API using the tweepy package. Tweets can be obtained by country by specifying a country code as a tag in the search criteria.

Currently, with a free twitter developer account, there are rate limits imposed on the number of possible searches and on the number of tweets that can be returned. Also, the free API only allows the return of tweets from the past 7 days.

A future phase of this work will be to develop a data collection program that automatically collects tweets and stores them in a database. This would provide a larger data set and make results more informative.

Data science

- Keras and word embeddings for sentiment
- R - real time twitter analysis using word embeddings

- PythonProgramming.net - great tutorials on python live twitter streaming API
- Textblob - how to build a classifier

Stakeholders

- National Assembly for Wales

Code and outputs

- GitHub repo

Related and existing work

- Mining twitter data for sentiment analysis
- Twitter united airlines sentiment - scraped using hashtags

Further information

Please contact datasciencecampus@ons.gov.uk for more information.

Updates

- No updates yet.