

# Exposé: Charity survival and Twitter use

Group 1

## Purpose of research

This project looks at charity survival and use of Twitter, which is a widespread but not well understood activity undertaken by charitable organisations which may contribute to their success. The project uses three different sources of data on charities in England and Wales to answer several questions which could not be answered by a single source of data alone. The longitudinal aspect of the data is leveraged to determine charity survival.

## Research questions

Research question	Operationalisation	Variables used
1. Can source of funding be used to predict the success or failure of a charity?	- Success measured by not dying & by growth of income	- Funding source (a) - Survived? (b) Growth of income (itotal (a) and updated_income (b))
2. How is government funding related to charity use of Twitter?	- Having Twitter vs using Twitter	- Funding source (a) - Has Twitter binary, number of Tweets (c)
3. Do charities which seek to help the public engage more with Twitter to reach the public?	- Being followed on Twitter	- Benefit groups (b) - Followed on Twitter (c)
4. Does number of staff, rather than size, determine a charity's successful use of Twitter?	- Staff controlling for size (income) - Following on twitter (which reflects active use)	-Staff (b) -Size (updated_income) (b) -Following on Twitter (c)

Original data (a), Scrape data (b), Twitter data (c)

## Data sources

1. UKDA data<sup>i</sup>: Traditional data set containing detailed financial information for 2011-12 for 12,150 charities. This level of detail is not available for the more modern year.
2. Scrape data: Up-to-date web scraped data on income, staff, and benefit groups in 2017. Scraped from the Charity Commission website<sup>ii</sup> which is a complete record of charities.
3. Twitter data: Data on numbers of followers, following, and Tweets. Obtained through the API.

Data source 1 is the backbone of the project and forms the primary sample of 12,150 charities which more up to date data will then be collected for. The second source of data scrapes up-to-date and new information for these charities which allows for the calculation of growth, survival and adds accurate data on staff and benefit groups (who the charity helps). Data on staff and benefit groups is not available in the original data. The final source, the Twitter data, adds some basic metrics which reflect how active and popular each of the charities is on Twitter (and if they have an account at all). Charity number is used to link the data as it is unique for each organisation; the specific charities are the common factors between all the data sources.

## Limitations

- The sample is based on charities active in 2011-12, any charity founded after this date is not included. The UKDA data is also a stratified random sample and not a complete record of charities.
- Staff count may not be a reliable variable (unknown if FTE or headcount).
- Twitter handles have been linked using fuzzy matching and there will be errors (both false positive and negatives) in the results. This error should be small enough as to not unduly affect the analysis.

<sup>i</sup> <http://reshare.ukdataservice.ac.uk/850933/>

<sup>ii</sup> <https://www.gov.uk/government/organisations/charity-commission>