

Charity Data Hack Methodology

CharityHack

* on Sat, 26th of July, 2014

* Dublin, t-cube

* Organized by Open Knowledge Ireland

* 13 participants who made it happen: Dave Corley, Tracey P. Lauriault, Patrick Killalea, Adrian Corcoran, Allen Thomas Varghese, Ingo Keck, Helen Nic Giolla Rua, Adrian O'Flynn, Flora Fleischer, Denis Parfenov, Chris Garde, Grazia D'Aversa, Salua Nassabay

Description of the data gathering process

The day started with a welcome speech by our conveners and hosts Denis Parfenov and Flora Fleischer of Open Knowledge Ireland. A [short presentation](#) by Flora Fleischer.

Adrian O'Flynn, our charities subject matter specialist, and the person who inspired the event, introduced the topic of spending for charities in Ireland, recent issues portrayed in the media and why it would be useful for the public to be able to compare charities based on their financial reports. Here is [his presentation](#).

Adrian Corcoran, the event's project manager, provided detailed instructions on how we were to work together for the day. He followed this outline (<http://openknowledge.ie/chy-01-charity/>)

Three main datasets were used for the CharityHack:

1. An overview working document listing basic data about the charities: (https://docs.google.com/spreadsheets/d/1sIH9NKBkpQMFMnt_0sYW9B8DSyv839EbbtLmibdml1s/pubhtml).
2. An excel spreadsheet ([CharityFinancialDataFinal.xlsx](#)), that includes detailed data for a number of charities, which were initially extracted from the annual reports by Adrian O'Flynn.
3. A legend document which includes the types of codes used in the document above containing the basic charity data (e.g. expenditure codes) (https://docs.google.com/spreadsheets/d/1sIH9NKBkpQMFMnt_0sYW9B8DSyv839EbbtLmibdml1s/edit#gid=861039018)

Participants were separated into three teams "red", "green", "blue", following the team label of the charities in the overview document.

Participants from each team picked a charity from the overview document. Adrian Corcoran provided instructions on how to label the data, to quality check the numbers within and to verify the data with the annual reports. These data were then assembled into one shared document. Data quality issues were reported and then corrected by Adrian Corcoran.

Problems were reported during this process and more serious problems were shelved for later.

Once this work was completed, only data that passed data quality processes were included.

The charity data used for this CharityHack exercise are as follows:

1. Data were extracted from the annual financial reports of 24 of the largest charities in Ireland (1Mil+€). NOTE This is a small sample of the thousands of charities in Ireland.
2. These are charities renowned by their public brand awareness for the fiscal year 2012 (see the [definition](#)).

3. ONLY charities where funding represents charitable funds collected voluntarily from the public were selected. This excludes charities that primarily rely on large institutional funds (e.g., Irish Aid Grants, HSE Grants).
4. Only charities that have been independently audited and who have published their financial statements on a publicly accessible websites were used.
5. Only charities who followed the Statement by the Accounting Standards Board on the SORP Accounting and Reporting by Charities: [Statement of Recommended Practice](#) were used (Section **B: Resources Expended**). Not all organizations in Ireland follow this standard. It is only because of this standard that it was possible to derive standardized financial data from the financial statements in the annual reports. The data here reflect resources spent on Governance, Fundraising and Charity.
6. The list of charities was derived from the survey conducted by the [Irish Charity Engagement Monitor](#) (ICEM).
7. This final dataset was quality checked by the participants at the 2014 Charity Hackday on July 26.
8. The original dataset was created by Adrian O'Flynn.

It is a very small sample, but this illustrates the power of open data.

Finally a [first draft of a future website](#) was created, where charity data will be displayed in an easy to understand fashion.

All of the Datasets are available for download from our [Charity Datasets](#) page and we hope to expand the quantity of data, details to follow.

Published by [Google Drive](#) – [Report Abuse](#) – Updated automatically every 5 minutes
