Media Technology Single Project

TE3980

"Data Journalism"

By Heather Amanda List

SCHOOL OF JOURNALISM AND DIGITAL COMMUNICATIONS

University of Central Lancashire

8th May 2015

Acknowledgements

Acknowledgements are extended to the final year project supervisor's and lecturers within the School of Journalism and Media; University of Central Lancashire, for their assistance in proof reading and useful general comments on this report.

Acknowledgements are also extended to Johnsons Press who provided the foundation for this research project during collaborative work in 2014 and permitting further co-operation again in 2015 alongside the Uclan Data Journalist students.

Finally gratitude is expressed to…

- Megan Knight - Data Journalism module leader.
- Ruby Kitchen - Senior reporter at Johnston Press.
- Uclan Data Journalist students.

Summary

What is Data Journalism? A brief outline will be presented with references to professional opinions and observations of published reports?

Data Journalism at the University of Central Lancashire is in its infancy; 2015 being the first year it has run as a fully fledge module. Two years ago it was conducted as an element within a current Journalism module and one year ago it was added as a sidebar independent study component. It was this proving ground and the work by the students with organisations in the industry such Johnsons Press that convinced Uclan course curators to add Data Journalism as a module on the curriculum.

The first section of this report is a first-hand case study of its inception as tutored by Megan Knight the course leader and a documented account of the Lancashire Evening Post's collaboration with the students in creating data driven reports on the 2015 General Election within their readership coverage area.

The second section of this report is a look at the skillsets requested by media industry leaders seeking to employ a data journalist along with what tools of the trade current data journalist use.

The final section is an analysis of the software programs predominantly used by a data journalist to create visualisations with regards to the skill level required to operate them.

Contents

Introduction

The majority of the data stored around the world is in the form of factual descriptions, definitions, and historically written conclusions of events, conversely there are also equally as many databases with numerical statistics that match the facts. Data Journalism is a process of gathering and combining this data then using software tools and techniques to clarify the information and produce a visual representation that can tell you more than the original pages of data. An example of this is *Elisabetta Tola's* '*map of ignorance'* (Appendix A.v) an interactive map of the seismic safety of Italian public schools created by merging gathered data. (datadrivenjournalism.net 2013)

There are seasoned journalists that consider data journalism to be the equivalent of a graphic novel as compared to the works of Shakespeare. It is not taken seriously for its true journalistic content in the same way the story's told via graphic artists are not readily recognised for their literature content. Before the internet this was probably a realistic perception, however since technology began pushing newsprint to news-pixels it has become a more attractive way of telling the story of the day especially if it is data driven. *Nick Bilton* a technologist for *The Times* said "Paper is dying, but it's just a device, replacing it with pixels is a better experience." (newyorkdailysun.com 2015)

Off the shelf software programmes such as Google Fusion Tables, (google.com) which is a powerful tool to merge repetitive tables of information into one visualisation and Datawrapper (datawrapper.de) a customisable chart builder, are widely available to assist journalists in designing graphs and charts for print and interactive web elements to accompany online news articles. An experienced front end web designer however can hand code geo location data such as the latitude and longitude of landmarks and buildings into an interactive map just as quickly as Google map engine (Appendix A.vii) and provide bespoke representations that exemplify the information so are these tools necessary.

The question is what technological skillsets do data journalists need to keep up with this visual information revolution. *Tim Berners-Lee*, founder of the World Wide Web said, "data-driven journalism is the future, journalists need to be data-savvy" (The Data Journalism Handbook, 2012, p.16) this project will explore if they also need to be code-savvy.

Section 1- Data Journalism

Case Study

Overview

This case study will follow a group of *Uclan* Journalism students as they work in collaboration with (JP) *Johnsons Press* and the (LEP) *Lancashire Evening Post* to take the data driven stories behind the general election 2015 from conception to completion.

The scope of the stories will primarily focus on the constituency's surrounding the LEP's readership coverage (Appendix A.i) with a view to them featuring as data driven news articles in the paper during the run up to the election.

The intention is to examine the technological skillsets the students cultivate to gather, clean and merge the data behind their chosen stories and what data visualisation techniques they develop to portray the information. By following their journalistic endeavours and analysing each stage of their progress it will be possible to see how these future data journalists turn words and numbers in a spreadsheet into a data visualisation that reveals the true story.

Tuesday 20th January

The data journalism module project as outlined by *Megan Knight* (2015) *University of Central Lancashire* module leader.

> *"This module combines two main outcomes: the development of your skills in data journalism and working on a publication project with the Lancashire evening post, based on those skills.*
> *Data journalism is often presented as an entirely new form of journalism, but the fact is that journalism has always involved the use of complex data in the process, but the existence of large tranches of public data and ready access to the tools needed to analyze it is relatively new."*

The intention of the module is to teach the journalist students key skills in data analysis and visualisation so they might produce data driven news stories. (Google Ref G.i)

What is Data Journalism?

Tuesday 27th January

It is only recently that 'data journalism' was ascribed the title, it is more commonly known as 'computer assisted reporting' and *The National Institute for Computer-Assisted Reporting* (NICAR) has been supporting data journalist since 1989 with is extensive database library's and data analysis training. (ire.org 2015)

However all 'journalism' is 'data journalism' if you ascribe the term 'data' to mean 'information' as the news media industry exists to present information. In the context of this project the term data journalism will be refer to stories that have been derived from collated data sets or have been reinforced by visualizations generated from the data.

> *"Telegraph journalist Holly Watt had set up a spreadsheet which listed MP's, names, the addresses of properties on which they had claimed, whether each was rented or mortgaged and notes of other points about their claims. By the middle of the second week ... she had a potential breakthrough. As she typed in the address of the second home of Ian Cawsey, an obscure Labour MP, a box popped up on her screen with the rest of the address. This 'auto-complete-function' meant she had already typed in the address for another MP.*
> *"Bingo" Watt shouted across the office.*
> *Scrolling back up the spreadsheet, Watt found that Cawsey's London address was shared with (another MP) ... In other words, the taxpayer was being billed twice for the same property."*
>
> *No Expenses Spared*
> *Robert Winnet & Gordon Rayner, 2009, p.220*

You cannot get a more perfect example of data journalism than the stories that emerged from the analysis of parliamentary expense claims forms leaked to *The Telegraph* in 2009. The documentation had been requested pry to the leak but access turned down; access which under the Freedom of Information Act, that came into force on 1st January 2005 would have been freely available.

Data journalism in this form is relatively new and has piggybacked the transformation of the news industry to the web, allowing journalists to portray more evidently the myriad datasets of information available to them as visualization's that not only clarify but expose.

The *BBC* in The Data Journalism Handbook (2012, p.30) described the purpose of data journalism with three rules that a reporter should consider before adding a data visualization to their story.

Does the visual…

- *Enable a reader to discover information that is personally relevant.*
- *Reveal a story that is remarkable and previously unknown.*
- *Help the reader to better understand a complex issue.*

The Washington Posts "Faces of the Fallen" a collection of interactive searchable data visualisations of the U.S. service members that died in the Iraqi conflicts, is a worthy example of all three rules. (washingtonpost.com 2013)

The best description of data Journalism is probably is by *Simon Rogers* (The Data Journalism Handbook, 2012, p.38) who built the Guardian data blogs from a small side line to an award-winning online data visualisation and free database repository.

> *"It's not that long ago journalists were the gatekeepers to official data. We would write stories about the numbers and release them to a grateful public, who were not interested in the raw statistics. The idea of us allowing raw information into our newspapers was anathema.*
> *Now that dynamic has changed beyond recognition. Our role is becoming interpreters; helping people understand the data, and even just publishing it because it's interesting in itself.*
> *But numbers without analysis are just numbers, which is where we fit in. When Britain's prime minister claims the riots in August 201 1 were not about poverty, we were able to map the addresses of the rioters with poverty indicators to show the truth behind the claim."*
>
> *Simon Rogers, 2012*

These descriptions suggest there is a degree of interpretation involved in producing data driven stories and their associated visuals, some of this interpretation is founded in the journalist's skills but how much is revealed when it is processed by technology?

First Meeting Lancashire Evening Post

Wednesday 28th January

Johnston Press plc is a multimedia company, based in London and is the second-largest publisher of local newspapers in the UK. *The Lancashire Evening Post* is one of the newspapers they publish daily; its head office in Preston was the location of development meetings.

This initial gathering was to define an agenda for the project and a path for the students to follow in regards to the subject matter which was to be the general election in May 2015. The LEP outlined a scope for data driven stories based on the results of a *Johnson Press* reader survey they conducted into the main issues that would sway a voter.

The top six of the results of the survey are…

- *Health 79%*
- *Immigration 68%*
- *Europe 56%*
- *Welfare 53%*
- *Economic Policy 49%*
- *Support for the elderly 46%*

Editor *Mike Hill* was keen to make the stories relevant to local readers' comments on the election. (Google Ref G.vii) He finished the meeting by asking how fast we could get a jpeg visualization of the election results to the paper for them to run alongside the main story.

Speed to print is vital for the newspaper industry especially headline breaking news such as election results so an editor would require a data journalist who could accurately gather the data, format it and produce visualizations very quickly. The trade-off is between this speed to press and the skill of the journalist as to whether they produce a chart created from an off the shelf piece of software and on the editors desk in say thirty minutes or if they go for a hand coded custom diagram that takes the same time. (Appendix A.xvii)

Spreadsheets

Tuesday 3rd February

Since the introduction of the freedom of information act literally millions of data sets and records are now free for the public to download or request. This task has been made easier with the ability to access computer programmes such Microsoft Excel and compatible software like OpenOffice.org to open and examine the data files. It is online accessible data sets from sources such as the Office of National Statistics (ONS) (ons.gov.uk) that modern journalist scour for hidden stories or use to back up their own investigative reports.

These extensive spreadsheets of data however do not always come in a human readable style let alone one that a software programme can understand:

> *"Those spreadsheets often have to be seriously tidied up-all those extraneous columns and weirdly merged cells really don't help. And that's assuming it's not a PDF, the worst format for data known to humankind."*
>
> *Simon Rogers, The Data Journalism Handbook, 2012, p.39*

Using pre-existing xlsx files on cities populations (Google Ref G.ii) in Abode Excel the data journalism module tutor explained to her students how information should be correctly organised in a spread sheet so that when the file is upload to a program such as Tableau's desktop data visualisation software, (tableau.com) the results are not obscure or inaccurate. Cleaning your data is very important The Online Journalism Handbook (Bradshaw and Rohumaa, 2011) provides a few examples of how 'dirty' data can mess up your visualisations and suggest using 'Google Refine' a spreadsheet analysis tool now branded as "Open Refine", that can help you filter and sort large data sets that could take hours of scrutiny to clean by hand. (openrefine.org)

The module tutor also demonstrated for the students how to perform filters and sorts using the Excel spreadsheets tools but did not refer to any online software applications that could do this for them and the lecture only skimmed the surface how you can manipulate data in spreadsheets. Understanding spreadsheets, operating them and performing formulated calculations is a subject that is taught separately at both high school and college level under the current curriculum.

However journalists that graduated from their education prior to the advent of computers and information technology lessons would have to seek an education in this discipline as

advised by *Cynthia O'Murchu*, of *The Financial Times*. (The Data Journalism Handbook, 2012, p.131) However some of the students in the lecture did not have the prerequisite Excel skills to perform even simple mathematical calculations.

Maps and mapping

Tuesday 10th February

Google maps are the most famous mapping software and have a developer playground with dedicated Api's to plot everything from lampposts to the latest traffic jam, so long as you have a geo-location reference point i.e. name of a city, country, continent, post/zip code, latitude and longitude coordinates, or a combination zipped into a Keyhole Markup Language (KML) file.

The tutor introduced her students to maps and mapping with a Google Map Engine Lite lesson plan. (Google Ref G.iii) As a starting point Google Map is a respectable piece of software to teach mapping visualization skills and given enough time to play with layers, custom icons and uploading kml area plotting files you can produce effect maps. The tutor demonstrated this by creating a colour coded map of the general election results 2010 that make up the LEP readership distribution area. (Appendix A.ii / A.iii / A.iv)

These 'off the shelf' produced maps showing a breakdown of the major political parties in the distribution area, could be used as base a comparison of the upcoming 2015 election results and took at most thirty minutes to produce once the data had been cleaned. A post 2015 election map would probably take even less time now that a template has been established something *Helsingin Sanomat* a *Finnish* reporter believes all data journalist should a stash of. (Open Knowledge Blog, 2013)

> *"When there is a breaking news event, we have about five hours to come up with an idea. In most cases it would be too slow to start from scratch, to overcome this problem, we have been creating kind of a Style Book for data journalists. The Style Book contains a set of News App templates we can modify and publish very fast."*
>
> *Helsingin Sanomat 2013*

Google Maps though are not the only online mapping software; the CartoDB map (cartodb.com 2015) (Appendix A.xiv) of the *University of Central Lancashire* campus highlighting crimes committed in the area during November 2014 was created from an

extensive Excel spreadsheet downloaded from the data police website. (data.police.uk 2015) The same map could not be duplicated in Google Map Engine because the file was too large for the software to handle, however neither CartoDB nor any of the other open source mapping tools were demonstrated to the students or even mentioned in the lectures.

Election data planning and development

Tuesday 17th February

The Lancashire Evening Post asked for the Data Journalist students to produce five stories based on their top six survey responses, to facilitate this the subjects were broken down into research topics and the students split into pairs.

The tutor asked the students to consider two things when formulating their stories and possible visualisations.

- *What data do we need?*
- *Where can we get it?*

They were then requested to find these data sources and present the results to the LEP the following day. (Google Ref G.xi) At this point the plan as an observer was to see how the student's journalist skills merged with those required of a data journalist to find the information essential to produce data visualisations.

Second meeting with Lancashire Evening Post

Wednesday 18th February

The purpose of this meeting was to find out what the students were considering in regards to story ideas and whether they met with *The Lancashire Evening Posts* expectations. Help and advice as to story direction was given by *Mike Hill* and alternative suggestions of where the data could be sourced.

Offers of utilising the LEP as a foundation to process freedom of information requests (FOI) were forthcoming however none of the students took them up on this offer. It was also recommended that the students read the stories that the paper was currently printing on the election campaign and issues around it; the editor wanted the stories to remain focused on local issues for their localised readership, only a few of the students bothered to take a copy of the newspaper home.

Data Journalist Students Progress

Friday 20th February

After the meeting at *The Lancashire Evening Post* the module leader *Megan knight* sent an email arranging to meet each group of students separately to discuss how far they had got and what they needed to do to progress further. (Google Ref G.iv)

Sitting in on one of these meetings it was obvious that some of the students were struggling to find the appropriate data. The two students working on 'The Cost of Living' story were at a loss as to how to phrase keywords to search the Office of National Statics for data tables and they failed to come up with anything new since the last lecture. This could probably be more of a cultural issue than a skills issue as the two girls in question are international students from China.

Subsequent to this meeting *Megan* sent an email outlining the students' progress stating that she thought they had done some good work. (Google Ref G.v / G.vi)

Wednesday 25th February

At this point 'The Cost of Living' students still had not made any progress and were floundering. The problem the students were facing was that they would not find a single document with all the data they needed and they would have to scrape data from a variety of sites and collate the information for themselves this would require some analytic skills that so far they had not demonstrated. It was only when the tutor created a spread sheet with headings and sample data that they seemed to comprehend what they had to do.

Friday 27th February

The lecture had been billed as a show and tell so the tutor could get a sense of their story developments. 'The Cost of Living' group were still striving to make sense of ONS data files and had to exploit the tutor as a search assistant to complete the task. Once they had the numbers though due to their lack of knowledge with Excel they had to lean on the tutor's knowledge to perform mathematical formulas to convert the data so they could compare the yearly percentage values.

The third year and post graduate students appeared to be working independently and with better results. They had found, cleaned, and converted large data sets to extract the

information they required; in fact they had too much information and had planned on using the editor of the LEP as a sounding board for the direction they should take next.

When asked what kind of visualisation they had in mind for the data each group responded with the same answer of a 'line graph'. The tutor asked them to think carefully about whether this choice of visualisation would work for the story behind the data and to consider if it would help the reader understand the story. *Gestalt* a psychologist in 1912 came up with eight *Gestalt Laws* of pattern recognition indicating how colour, shape and perception can influence the way people make sense of visual data, an explanation of these 'laws' might have helped the students with visualizing the patterns in the data. (Interactive Visualization, 2013, p.113)

Furthermore it was suggested that the NHS data could be displayed as a timeline or that the migration data be shown as an interactive layered map; neither of these suggestion went down very well with the non-technological students and were not followed up on.

Wednesday 4th March

Lancashire Evening Post meeting with 'cost of living' and 'health' group. The purpose of this meeting was for the students to show the editors the data they had found and how they were going to use to use it to tell a story. Unfortunately due to dearth of confidence the two students covering 'cost of living' hardly spoke and allowed the tutor to exhibit their work. The 'Health Story' students were more able to explain their data, asked relevant questions, and arranged to speak with the LEP's health correspondent.

Wednesday 11th March

Lancashire Evening Post meeting with 'environment and 'migration' group. Again the purpose of this meeting was for the students to show the editors the data they had found to support their election issue story. The tutor was unable to attend this meeting with the students and the only LEP reporter present was *Jenny Simpson* conversely though the two groups of students demonstrating their work were the strongest and the most organised, they had gathered relevant data independently, and considered what would make a good story.

They were though still unsure as to how to present their data visually, undoubtedly due to their lack of exposure to data visualization software techniques. Both groups left the

decision to the editor to explain what they could possibly do for a visual and what kind of news article they could produce, however she only talked about print options and did not discuss to how the data could be visualized for online.

Tableau & Graph design workshop

Tuesday 3rd March

As an introduction to the desktop data visualisation software Tableau the students were instructed in how to upload an Excel file on city population figures and drag and drop the relevant rows and columns of data to build a map that could be filtered by density and population. (Appendix A.vi) This was only achievable because the data in the file had been cleaned and structured in a way that was compatible with the software's data recognition methods.

It was proposed to the module leader that she introduced the students to other online chart design software such as DataWrapper (datawrapper.de) which is popular with many national newspaper groups; she commented that it was more for online use than print. Very few news media groups do not have online publications, for that reason if the data journalism module is purporting to teach the skills required of a data journalist it should include some visualisation techniques for web.

Tableau can produce some good data visualizations however it requires more than a brief introduction if you really want to work with it. Some of the students were due to attend a Tableau workshop and as the tutor has training in the software it is believed she will be encouraging her students to produce their visualisations with it.

Data scraping and cleaning

All the students have strived to locate data and extract relevant information from the sea of material they have downloaded. Experienced data journalist use online tools such OpenRefine and Data Wrangler (vis.stanford.edu) to clean and transform tables of information into workable files that can be uploaded to software programmes such as Tableau. A firm favourite to convert those elusive pdf documents to spreadsheets is

ScraperWiki, (scraperwiki.com) even so the students were not informed of these tools and never mind instructed in how to use them.

> *"ScraperWiki has helped journalists become data journalists ever since we were founded in 2009."*
>
> *scraperwiki.com, 2015*

The students were also very myopic in their data sources, as journalist they should be able to research and find information to back up and support their articles from a variety of creditable sources. Conversely the cost of living students spent hours searching the Office for National Statistics (ONS) for information even when other sources were suggested to them such as the databases of The Guardian Blogs, (theguardian.com).

The migration team establish all their data from the 2011 census which was possibly the most accurate but was also sourced from the ONS. The only team that looked elsewhere for their data was the environment team but only so far as to direct their search to the datasets of the gov.uk.

This deficiency of investigative skills is what the collaboration with the newspaper was supposed to overcome however the experiment has possibly not gone quiet to plan.

It might be worth them reading two books written by *Paul Bradshaw, Scraping for Journalists (2013)* and *Finding Stories in Spreadsheets(2014),* highlighted in an article on the Journlsim.co.uk website *'Two reasons why every journalist should know about scraping*

> *"Scraping is faster than FOI, can provide more detailed and structured results than a PR request - and allows you to grab data that organisations would rather you didn't have. If information is a journalist's lifeblood, then scraping is becoming an increasingly key tool to get the answers that a journalist needs, not just the story that someone else wants to tell."*
>
> *Paul Bradshaw, 2012*

Data Journalist Students Work

Regardless of their skirmishes with data, meagre analysis skills and visualisation techniques the students did produced finished articles; some of which will be published in the *Lancashire Evening Post* over the weeks leading up to the general election.

One of the hardest things for the students to grasp was how they would visualise their data, as students of words and pictures considering graphs, charts and maps did not come to them easy and in most cases had to be pressed into their choice by the editor at the LEP.

The environment team developed maps in both Tableau and Google, yet after all its hype when submitting their final story they choose to use the Google map over the Tableau one. (Appendix A.x / A.ix) Conceivably due to their customisation of the development areas with Google Map polygraph lines, something they may not have accomplished in Tableau without the right tutorial.

The migration team of students have possibly established the most comprehensive collection of data and graphics to demonstrate their story. (Google Ref G.xii) They have used a map created in Tableau, (Appendix A.xi) but it should be noted that these two girls went on the Tableau workshop so they have skillset advantage.

They also produced illustrated infographics, in the book Interactive Visualization (Ferster and Shneiderman 2013 p139) there are 'Five rules of information communication' one of them is "Do not beautify, clarify" The images the team created to accompany the data on migration figures exemplifies this rule, by enhancing the understanding visually what would have been harder to comprehend in words. (Appendix A.xii / A.xiii)

The same could not be said for the visualisations produced by 'the cost of living' group; they could be forgiven for having one of the hardest data sets to scrape and configure however their portrayal of the data is poor quality and girls failed to explore any other avenue than Excel line and bar charts as visuals. (Appendix A.xv / A.xvi)

When graphing data it is vital to choose the correct form; in Excel there are nine ways of charting data (Ferster and Shneiderman 2013 p139) and there are five methods of organising information: by *location, alphabetically, liner time, category* and *hierarchy*. The cost of living girls tried to merge all of these concepts in each chart resulting in confusing visuals that are hard to understand. They should take a lesson from *Duncan Clark* and his comment in an interview with the Interhacktives.

> *"One golden rule is to let the information speak for itself. There is no point making a pretty visualization if it doesn't make the data clearer to understand and easier to interrogate."*
>
> *Duncan Clark, 2014*

The students had time in their favour however they did not use it to experiment with a wider range of data visualisation software, they stayed within the comfort zone of their skillset instead of seeking to develop it further.

Data Journalist Students Survey

Back in 2011 a survey was conducted by the *European Journalism Centre* (EJC) (ejc.net) into what training was required for future and current journalists, the most prominent result of the survey as reported on datadrivenjournalism.net was:

> *"While 70% of respondents unsurprising felt that data journalism was 'vary important', an even higher proportion felt that a 'lack of adequate knowledge' was the biggest hurdle to using data."*

So a survey was organised for to gather some of data Journalist student's opinions on the module and their work, the results revealed:

- *All the students said they took the module to add to their skills or improve them.*
- *80% said the module met with their expectations.*
- *Even with the Tableau workshop trip only 30% said they would use it to design data visuals for print however 60% said they would use it for online visuals.*
- *Surprisingly 30% said they would use Google Fusion Tables for online visuals.*
- *Not surprisingly the over whelming majority of the students chose Excel and Google maps as their go-to tool for print and web.*
    - *The main reasons for choosing these tools were ease of use and knowledge of software.*
    - *Speed of delivery was not considered overall important.*

Only 50% of the students said that would consider a career as a data journalist perhaps this is also due to a 'lack of adequate knowledge'.

Case Study Conclusion

If the objective of data journalism is to tell a story or reveal a hidden story through visualization techniques, we can mark the module and case study a success. The students did indeed create data driven stories and visualisations for the *Lancashire Evening Post* based on election issues raised in their survey and in some cases clarified fact from fiction that would have been harder to distinguish in words.

However it is the feeling of this project writer that the visualisation techniques imparted to the students lacked the nuances of what can be accomplished with more versatile software

such as Google Fusion Tables and CartoDB mapping, which from testing have been established to be within the abilities of the students with no knowledge of coding languages.

As part of the module they have been asked to analysis real journalist's data visualizations and critically challenge their choice of graph, chart or map when barely a month ago they were stuttering over the question of how they would portray their own stories data with the words 'I think a line graph' hardly making them candidates for judging other peoples work.

The level of data analysis covered by the module also falls short of that which would be required by a journalist working for a national newspaper. The basic Excel skills demonstrated to the students have in no way prepared them for the role of a data journalist. Their success was more down to their current skills in research and data analysis than any they acquired during the lectures.

The University of Central Lancashire is famed for producing media reporters of merit and note however if it is to produce an award winning data journalist the content of the data journalism module may require a rethink. Perhaps taking a lesson or two from the Interhacktives who recently ran a webinar on how to clean and visualise election data using the web mapping tool CartoDB. (interhacktives.com 2015)

Section 2 - Data Journalism in Practice

The Job of a Data Journalist

The *BBC* news media describes data journalism as a team effort with everyone pooling their individual talents together to create the data story and its relevant visuals. They say there is no 'unicorn' no one "specifically identified as a data journalist" (The Data Journalism Handbook, 2012, p.33) they talk of their staff being proficient in the use of spreadsheet applications and whether they are a journalist, designer or developer they are expected to know what the other person job entails so they can work more proficiently together.

Skills they call upon their team to demonstrate include analysing data with Google Docs and Fusion Tables, developing geographical visualisations with Google, and Esri's ArcMap and design infographics around the data with Adobe Illustrator, After Effects, and Photoshop. Some of their hard core hackers use MySQL, Python, and Perl to pick apart larger data sets and then use JavaScript and JQuery libraries to publish their findings as interactive visuals. This would suggest they require a high level of technical knowledge from their journalist as well as that all important 'nose' for a good story.

The Guardian also describes the role of a data journalist as agile requiring more than one head or pair of hands. (Appendix A.vii) *Simon Rogers* who edited *The Guardian Blog* for many years defines data journalism as a process that is changing as new techniques emerge and advocates the use of free tools.

> *"Our top tools are the free ones that we can produce something quickly with. The more sophisticated graphics are produced by our dev team. This means that we commonly use Google charts for small line graphs and pies, or Google Fusion Tables to create maps quickly and easily."*
>
> *(The Data Journalism Handbook, 2012, p.38)*

This is an example of the trade-off between speed to press and the skill of the journalist.

In April 2015 the *BBC* careers website was advertising the role of a 'Data Journalist, Visual Journalism' they were asking candidates to be "brimming with original ideas on how to best apply data-driven journalism." (careerssearch.bbc.co.uk 2015)
They wanted the potential journalist to have:

- *A good understanding of statistics and statistical analysis.*
- *A strong grasp of how clean, parse and query data as well as database management.*
- *Demonstrable experience of visualising data and using visualisation tools such as Tableau, Refine and Fusion Tables.*
- *Demonstrable experience of working to tight deadlines.*
- *A good knowledge of Microsoft Excel*

They were also keen for them to have:

- *Knowledge of several of the following packages; SPAA, SAS, tableau, Refine and Fusion Tables.*
- *A Knowledge basic scripting and HTML, as it might pertain to data visualisation or data analysis.*

Obviously the *BBC* is a major player in the field of journalism and this is reflected in the expertise they ask for when recruiting new staff. Conversely when *Mike Hill* editor of the *Lancashire Evening Post* a small local daily newspaper was asked what essential and desired skills he would seek in a potential data journalist he replied:

> *Essential; Maths, Journalism, data analysis, excel, google Fusion.*
> *Desirable: OpenRefine, Html, JavaScript, Python, DataWrapper, Tableau.*

So even at the lower end of the spectrum editors are looking for journalist with online reporting skills perhaps even that elusive 'unicorn' data journalist. He also suggested an important ability would be *"A nose for spotting a story in the data and a sack-full of ideas".* (Mike Hill, 2015) (Google Ref G.ix)

Interview with a Data Journalist

At the first meeting with the Lancashire Evening Post editorial staff we were introduced to *Ruby Kitchen* an investigative reporter and data journalist for the Harrogate Advertiser and she agreed to an email based interview. The questions sent to *Ruby* were modelled on two online articles based on interviews with data journalists.
(radar.oreilly.com 2012) (interhacktives.com 2013).

The two journalists interviewed were *Ben Welsh* from the *Los Angeles Times* (palewi.re/who-is-ben-welsh/ 2013) who describes himself as:

> *"A reporter, albeit an unconventional one. I specialize in what some people call computer-assisted reporting, and some others call data journalism. Journalists in the past were often hack novelists. I'm a hack computer programmer. I write code to collect, organize, analyse, and present large amounts of information."*

and *James Ball* one of the UK's leading data journalists for *The Guardian* who has worked on notorious stories such as *Guantanamo Files* and *Reading the Riots*. (Gennady Kolker, the Guardian 2013)

Perhaps it is a bit unfair to compare *Ruby Kitchen* to these two national reporters with backgrounds in computer assisted reporting but they each '*fell into'* the role with no formal education or instruction in how to compile data driven stories or produce data visualizations. Conceivably because when they set out to follow a career in journalism the

term data journalism was not a widely known concept and formal education in the subject matter was not considered, *Ruby Kitchen* admitted as much in her interview.

> *"I didn't even know it was called data journalism until very recently, it was just the way I sourced stories."*

James Ball when explaining how he got into data journalism said *'I had data skills basically," Ben Walsh* polished his skills while working for the *National Institute for Computer-Assisted Reporting* (NICAR) *" Decades before anyone called it 'data journalism'."*

Ruby analyses spreadsheet data from government and public sources to either find or back up her stories of daily life in Harrogate, the same starting point for the majority of *James and Bens* reporting accomplishments. In fact *James Ball* recommends starting small with Excel.

> *"If I could only ever pick one tool it would always be Microsoft Excel, because you've got to start at the most basic level."*

Both *Ben* and *Ruby* use third party software to help analyses databases; *Ben* uses PostgreSQL (postgresql.org) a powerful, open source object-relational database system, while *Ruby* uses OpenRefine a tool for cleaning messy data, the former being a more 'techie' option while the latter is a playground for many data analysts.

She may not have had to lay low because the data she carried on a memory stick was sensitive but like the journalism students *Ruby* has struggled to find data, she explained how she had to scrape information from multiple sources for her most famous story about 'legal highs' because there was no official record at that time on the increasing dilemma. (harrogateadvertiser.co.uk 2015)

When Ruby was asked which comes first the story or the data her response was:

> *"Both. Neither. Sometimes I'll scan some data publication, and find an interesting story. Sometimes I'll look at a topic, and find the data. On a national scale, or a wider scale, I would say the data comes first. There's a clear indication of a story. On a local level, you need to know what the story is before you find the data or the data might not apply. E.g. Poverty figures have come out this week. Nationally, the data creates a story. Locally, it doesn't, as Harrogate doesn't have a poverty problem."*

James Balls reply to the same question was, *"it's always better to have the story in mind but if the data doesn't back your story you drop the story".* It would seem the answer is more complicated than the question and is where the experience, training, and skillsets of a good reporter come into practice.

The final question posed to *Ruby* and James was 'What three pieces of advice would you give to an aspiring data journalist?' One of James answers was, *"You have got to have skills that other people in the newsroom don't."* Which is good advice and so was *Ruby's:*

- *Every story needs a victim. As interesting as the figures are, they are dull without demonstrating the impact to real people.*
- *Don't live by the data. If it's not showing what you expect, use your gut reaction to dig deeper.*
- *Data needs a pretty picture. An infographic can do in three seconds what 1,000 words can do. It's far more understandable, readable, and interesting.*

The front page was once the domain of the press photographer not anymore; it is progressively becoming the province of the data journalist? The headlines on May the 8th will be based on data journalism and the public will be looking for a visualisation that demonstrates clearly how the election was won.

The complete interview with reporter *Ruby Kitchen* can be found on the Google drive. (Google Ref G.x)

Section 3 - Software Analysis

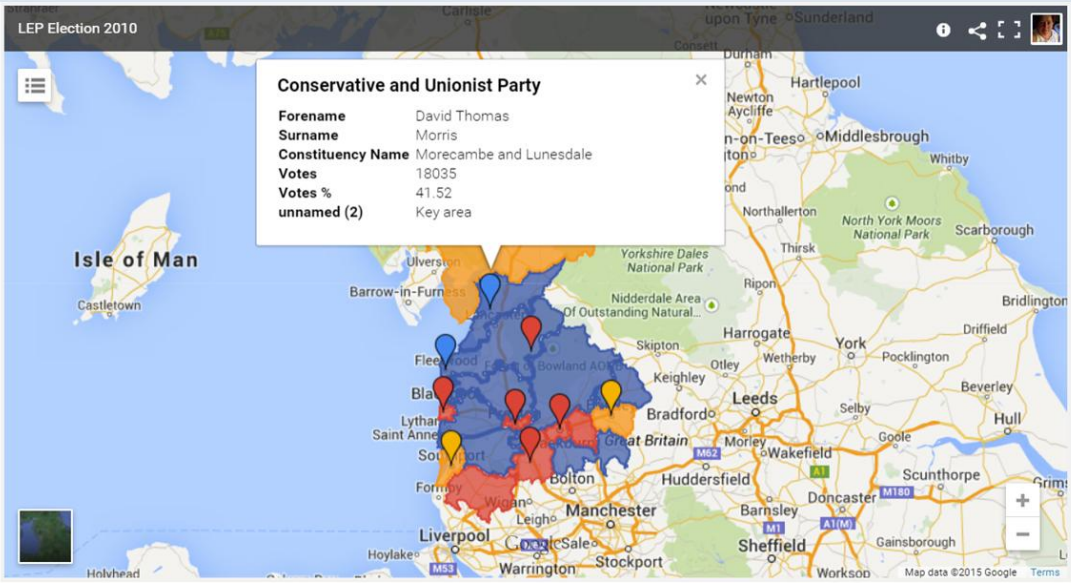
Tools of the trade

"Excel, Google Sheets, QGIS, CartoDB, HighCharts, Quartz Chartbuilder, Outwit Hub, Illustrator – each one has their advantages"

Billy Ehrenberg Interhacktives 2015

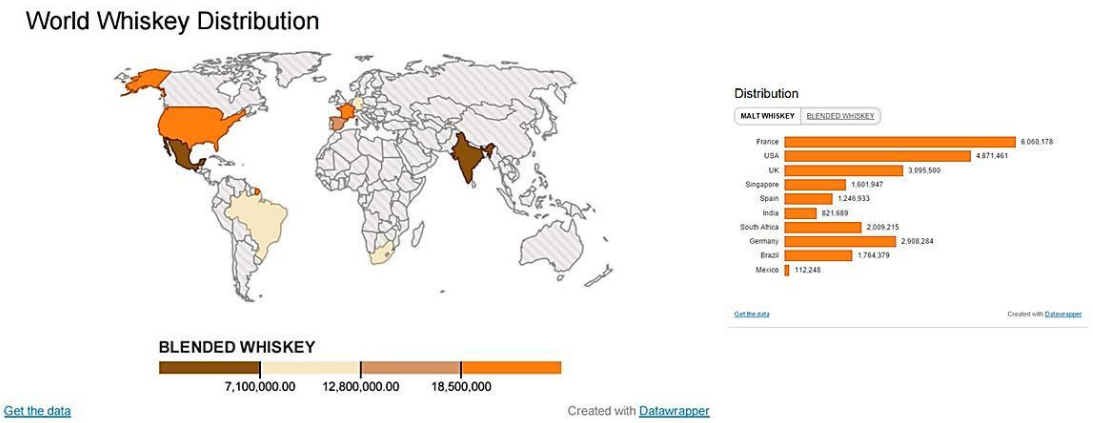
Journalists from the national award winning *Ben Welsh* to the local data wrangler *Ruby Kitchen* all have their personal favourite software programme. Each as *Billy Ehrenberg* states in the quote above have their merits so in accordance with the rules of data journalism here are some visualisations to clarify the merits along with an honest opinion from a coder.

Google Maps Lite	*Mapping*	*Beginner level* *No coding required*
Beginners playground for designing interactive maps		
Easy to learn and customisable with icons and overlays		



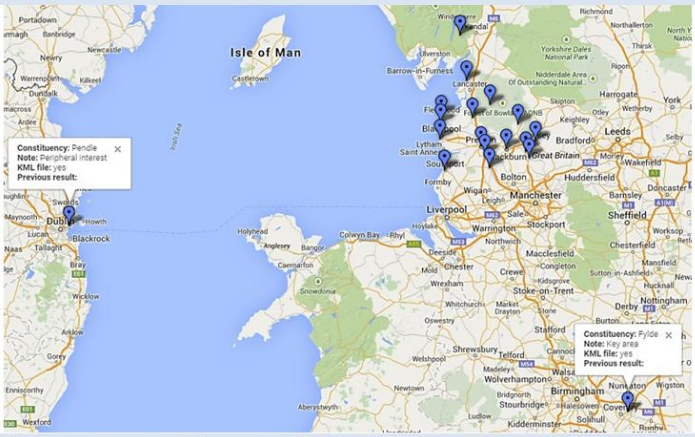
Lacks ability to add searchable data		
Cannot handle large data sets		

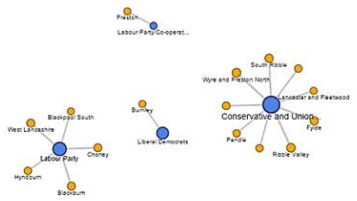
DataWrapper	*Mapping, Charts & Graphs*	*Intermediate skills* *Some coding but not essential*
Able to create charts and maps in four steps reducing design time from hours to minutes		
Can produce a variety of charts quickly provides both static and embedded files		



Unless you pay the option to export static image or embed file is denied		
So far unable to get the zipped files to open either the map or the chart in a web browser		

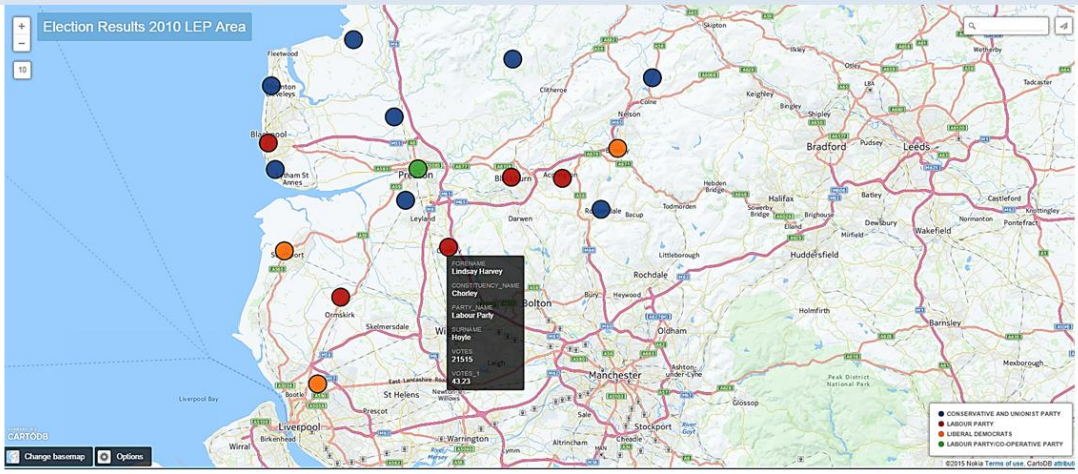
Google Fusion Tables	*Mapping, Charts & Graphs*	*Beginner level* *No coding required*
Works with Google Drive to create data visualizations and interactive web applications		
Can be published as a link for sharing or an iframe for embedding in webpage		





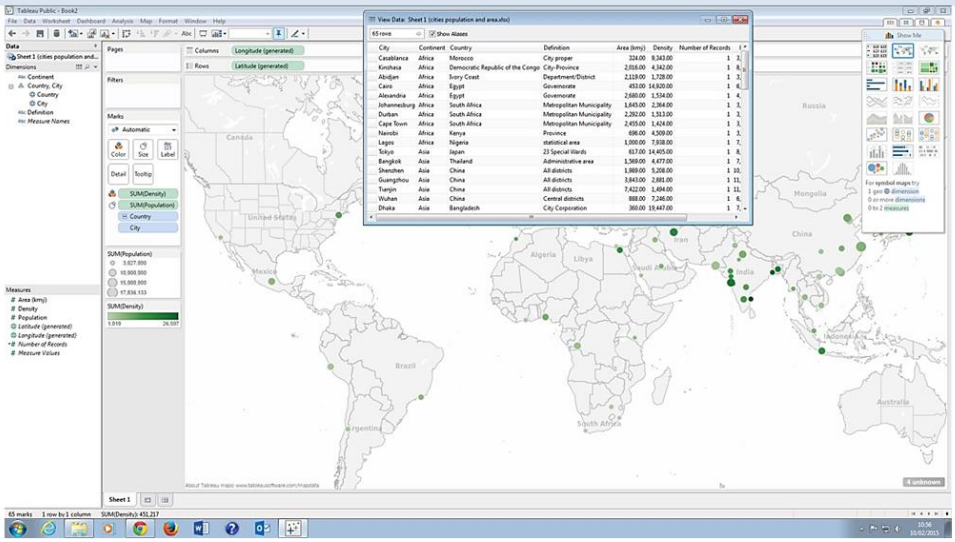
Can automatically geocode locations		
Unfortunately it thinks Pendle is in Dublin and the Fylde coast is in Coventry		

CartoDB	*Mapping*	*Beginner level* *No coding required*
Can design fully publishable interactive maps very quickly		
Extremely customisable and can handle large data sets		



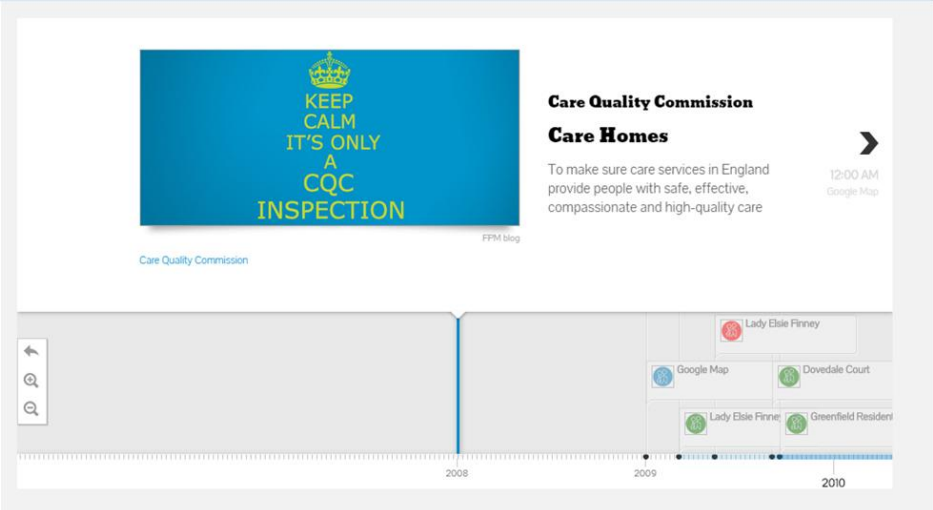
Only produces maps and has Limited layers unless you pay		
Had some problems changing the data headings it displayed the symbol '%' as '-1'		

Tableau	*Mapping*	*Training Required*
If you can afford the training it is a powerful tool for data visualization		
Requires the ability to organise spreadsheet data accurately or software cannot read it		



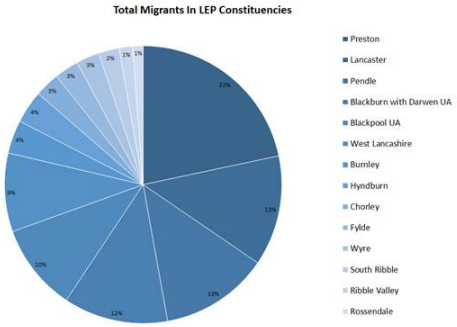
Two versions Tableau public and Tableau Desk Top / Desk Top needs an activation code	
Charts designed in Tableau Desk Top will not open in the free Tableau public	

Timelinejs	*Mapping*	*Intermediate* *Some coding but not essential*
Open source tool enabling you to build interactive timelines from google spreadsheets		
Can be customised with images and links to build embeddable timelines		



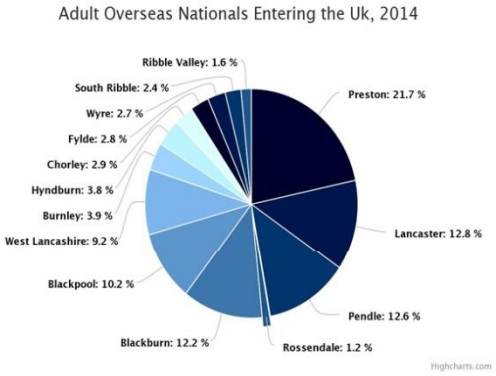
Restricted to templates	
If you require anything more you need to fork code from git hub and hand code it yourself	

Excel	*Charts & Graphs*	*Beginner* *No coding required*
Easy tool if you require a quick chart or graph to demonstrate the trend or comparison of figures		
Data can be linked to Google spreadsheets to make it updatable if embedded in a webpage		



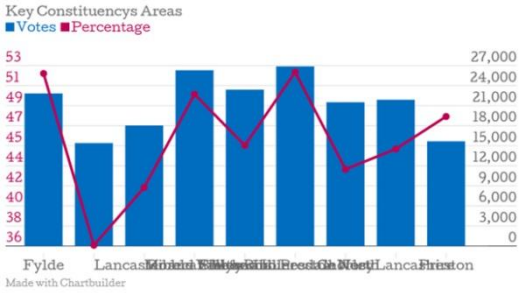
Speed over quality
Charts are not interactive and depend on accurate data for best results

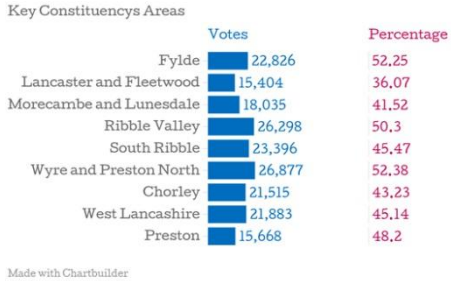
HighCharts	*Mapping, Charts & Graphs*	*Intermediate* *Professional Coder*
A powerful desktop map and chart builder for non-profit use		
"Huge Variety" of Javascript based charts and graphs with Json files fully editable in jsfiddle		



Requires a working knowledge of html and Javascript and google Api key for maps
Needs a jsfiddle account and a GitHub account to save templates of charts and maps

Quartz Chartbuilder	*Charts & Graphs*	*Intermediate Professional Coder*
Paste the data into the tool and then pick how you want to visualise each data series/		
Open source code, A developers playground for customising charts and graphs		





Speed to press may be an issue and it only builds line and bar charts
Not perfect would take a professional coder to get the best results from this chart builder

If you are serious about becoming a data journalist you need to immerse yourself in the myriad software programmes, experiment with customisations, and downloadable formats. The trick is to find the tools that you feel the most comfortable working with or update your skills by taking a course and learn how to use the software professionally; especially if you want to impress the editor with quickly produced quality visuals that look nerdy but really only took a couple of mouse clicks.

> *"As a developer, you have a few options. You can flex your engineering muscle by building a gorgeous map using custom software. Or you can use existing tools like Google Fusion Tables or open source mapping libraries and finish the job in a couple hours. The first option will give you a better app; but the second might give you more time to build something else with a better chance of having a lasting impact."*
>
> *(Chase Davis, Data Journalism Handbook, p.152)*

Conversely the best data visualisation tool could be the one that is currently being designed. Good sources of what is new in data journalism and data visualisation are journalism.co.uk and the Interhacktives.com, who publish some incisive interviews with data journalists and information on data visualisation tutorials.

Section 4 - In Conclusion

Investigative Analysis

> *"You don't have to be a programmer. You can become a top coder if you want. But the bigger task is to think about the data like a journalist, rather than an analyst. What's interesting about these numbers? What's new? What would happen if I mashed it up with something else? Answering those questions is more important than anything else."*
>
> *Simon Rogers, The Guardian 2012*

It was the intention of this project to explore whether a data journalist needed to be 'code-savvy' the conclusion is that they can fake the code with off the shelf technologies, they can even fake their ability to clean, sort and analysis data with programmes like OpenRefine, they cannot though fake the talent of being able to see the wood from the trees or the skill to create a visualisation that allows us to clearly see the path through the forest.

Whether you are a 'clickbait' visualizer as described by *Dr James Cheshire* (Interhacktives 2014) or an awarding winning data journalist you do undeniably as *Tim Berners-Lee* said, *"Need to be data-savvy"* (The Data Journalism Handbook, 2012, p.16) The most essential talent a data journalist requires is the ability to understand, sort, filter and manipulate a spreadsheet of data; however it would not hurt to be a little *'code-savvy'*.

The projects investigation into the skillsets required to produce data journalism and its visualisations revealed a two tier system, which are accurately laid out by investigate data visualizer *Duncan Clark* (Interhacktives 2014)

> *"It depends what kind of data journalist you want to be. If you're mainly interested in breaking stories, then getting acquainted with how to get unexplored data via Freedom of Information requests might be a good idea. If you're more interested in interactives and visualization's then learning to code can't hurt: access to good developers is always a bottleneck for journalists, so being able to do at least some of the coding yourself is a huge advantage."*
>
> *Duncan Clark, Kiln.it 2014*

At the top of this tier are the multiple skilled journalists that come from a data analysis background or have educated themselves in the skillsets they require to create the data visualisations they wish to portray. They have built up a playground of tools such as Highcharts and Tableau; they understand how to tweak JavaScript code and css styling to produce the most effective visualisation and probably could build a widget to forecast the

next five years rainfall. These journalists are more likely to be investigative reporters with time to examine data, compare, and merge it until they extract the truth from the figures.

Further down the ladder you have weekly reporters who need to produce graphs and charts on the fly from data they have sourced from freedom of information acts and government statistic files. These journalists do not have to time to play in 'jsfiddle's' coding playground (jsfiddle.net) or build illuminating infographics in Illustrator; they have to feed the data into simple efficient off the shelf visualisation builders such as DataWrapper and Google Fusion Tables. When push comes to shove and a deadline is looming they may even resort to the old Excel files and export a quick jpeg of a bar chart to the editor then ask the newsroom's dev team if they could build an interactive one for the website.

The Kiln.it website have a moto "Data + Journalism + technology + Design = kiln" If you replace 'Kiln' with 'News' you get the equation for the skillset that a data journalist requires.

Data + Journalism + Technology + Design = NEWS

Part of this project was a study of the University of Central Lancashire Data Journalism module and whether it is tutoring the students in the skillsets they will require as a data journalist. Based on the observations of lectures, the students work, conversations with data journalists and research into skillsets request by prospective employers, the consensus is that the module lacks an open minded view in relation to the difference in skills required for print data journalism and online data journalism and therefore is not comprehensively preparing the students for such a role.

The relevance of the online news industry to a data journalists target audience is marginalized by the data journalism module and may be disadvantageous to current and future students in a world where pulling an iPad out on the morning commute to check the news is becoming more common place than the rustle of a broadsheet.

During the British enquiry into the Phone Hacking Scandal Rupert Murdoch (nytimes.com, April 26, 2012) predicted the end of the print news media industry:

"The day will come when the news business would be 'purely Electronic' in five, 10 or 20 years"

Bibliography

1. Jonathan Gray, Lucy Chambers, Liliana Bounegru (July 2012) The Data Journalism Handbook, United Kingdom: O'Reilly Media.
2. Bill Ferster (2013). Interactive Visualization. 10th ed. Cambridge: Massachusetts Institute of Technology. p4.
3. Paul Bradshaw, Liisa Rohumaa (16 Jun 2011). The Online Journalism Handbook. New York: Routledge.
4. James C. Foust, 2004. Online Journalism: Principles And Practices Of News For The Web. Edition. Holcomb Hathaway Pubs.

Online References

1. #safeschools – A Data-driven Investigation into Seismic Safety Assessment of Italian Public Schools | Featured Projects | Data Driven Journalism. 2015.*#safeschools – A Data-driven Investigation into Seismic Safety Assessment of Italian Public Schools | Featured Projects | Data Driven Journalism*. [ONLINE] Available at: http://datadrivenjournalism.net/featured_projects/safeschools_investigation_seismic_safety_assessment_Italian_Schools. [Accessed 28 April 2015].
2. #scuolesicure | Controlla se la tua scuola è a rischio sismico. 2015.*#scuolesicure | Controlla se la tua scuola è a rischio sismico*. [ONLINE] Available at: http://old.voglioilruolo.it/scuolesicure. [Accessed 28 April 2015].
3. University of Central Lancashire. (2014). Journalism, BA (Hons). Available: http://www.uclan.ac.uk/courses/ba_hons_journalism.php. Last accessed 7th December 2014.
4. contact us - Lancashire Evening Post. 2015. *contact us - Lancashire Evening Post*. [ONLINE] Available at: http://www.lep.co.uk/contact-us. [Accessed 22 February 2015]. (2015. [ONLINE] Available at: https://www.google.com/maps/d/.[Accessed 22 February 2015].)
5. Data Journalist, Visual Journalism | Jobs and careers with BBC. 2015. *Data Journalist, Visual Journalism | Jobs and careers with BBC*. [ONLINE] Available at: http://careerssearch.bbc.co.uk/jobs/job/Data-Journalist-Visual-Journalism/7374. [Accessed 09 April 2015].

6. who is ben welsh? . palewire. 2015. who is ben welsh? . palewire. [ONLINE] Available at: http://palewi.re/who-is-ben-welsh/. [Accessed 11 April 2015].
7. Interview: James Ball data journalist at the Guardian - #Interhacktives. 2015. *Interview: James Ball data journalist at the Guardian - #Interhacktives*. [ONLINE] Available at:http://www.interhacktives.com/2013/01/20/interview-james-ball-data-journalist-at-the-guardian/. [Accessed 11 April 2015].
8. Profile of the Data Journalist: The Human Algorithm - O'Reilly Radar. 2015.*Profile of the Data Journalist: The Human Algorithm - O'Reilly Radar*. [ONLINE] Available at: http://radar.oreilly.com/2012/03/profile-of-the-data-journalist-2.html. [Accessed 11 April 2015].
9. Data journalism survey: a mixed picture | News & Analysis | Data Driven Journalism. 2015. *Data journalism survey: a mixed picture | News & Analysis | Data Driven Journalism*. [ONLINE] Available at:http://datadrivenjournalism.net/news_and_analysis/data_journalism_survey_analysis. [Accessed 14 April 2015].
10. Two reasons why every journalist should know about scraping (cross-posted) | Online Journalism Blog. 2015. *Two reasons why every journalist should know about scraping (cross-posted) | Online Journalism Blog*. [ONLINE] Available at: http://onlinejournalismblog.com/2012/08/09/two-reasons-why-every-journalist-should-know-about-scraping-cross-posted/#more-16643. [Accessed 21 April 2015].
11. Interviews Archives - #Interhacktives. 2015. *Interviews Archives - #Interhacktives*. [ONLINE] Available at:http://www.interhacktives.com/category/interview-2/. [Accessed 22 April 2015].
12. Data downloads | data.police.uk. 2015. *Data downloads | data.police.uk*. [ONLINE] Available at: http://data.police.uk/data/. [Accessed 22 March 2015].
13. KILN | Data Visualization • Digital Journalism • Maps • Interactives • Training. 2015. *KILN | Data Visualization • Digital Journalism • Maps • Interactives • Training*. [ONLINE] Available at: http://www.kiln.it/. [Accessed 27 April 2015].
14. Faces of the Fallen - The Washington Post. 2015. *Faces of the Fallen - The Washington Post*. [ONLINE] Available at:http://apps.washingtonpost.com/national/fallen/. [Accessed 28 April 2015].
15. Use Templates to Make News Apps Quickly | Open Knowledge Blog . 2015. *Use Templates to Make News Apps Quickly | Open Knowledge Blog* . [ONLINE]

Available at: http://blog.okfn.org/2013/04/03/use-templates-to-make-news-apps-quickly/. [Accessed 28 April 2015].

16. The Future of Newspapers? | New York Daily Sun - The Trusted New York Daily Broadsheet. 2015. *The Future of Newspapers? | New York Daily Sun - The Trusted New York Daily Broadsheet*. [ONLINE] Available at:http://www.newyorkdailysun.com/the-future-of-newspapers/. [Accessed 28 April 2015].
17. Data journalism at the Guardian: what is it and how do we do it? | News | The Guardian. 2015. *Data journalism at the Guardian: what is it and how do we do it? | News | The Guardian*. [ONLINE] Available at:http://www.theguardian.com/news/datablog/2011/jul/28/data-journalism. [Accessed 27 April 2015].
18. Log In - The New York Times. 2015. *Log In - The New York Times*. [ONLINE] Available at:http://www.nytimes.com/2012/04/27/world/europe/rupert-murdoch-testimony-leveson-inquiry-day-2.html?. [Accessed 27 April 2015].
19. DRUGS INVESTIGATION - legal highs in Harrogate - Harrogate Advertiser. 2015. *DRUGS INVESTIGATION - legal highs in Harrogate - Harrogate Advertiser*. [ONLINE] Available at:http://www.harrogateadvertiser.co.uk/news/crime/drugs-investigation-legal-highs-in-harrogate-1-7165782. [Accessed 14 April 2015].
20. James Ball to Join Guardian's US Newsroom | Guardian US press office | The Guardian. 2015. James Ball to Join Guardian's US Newsroom | Guardian US press office | The Guardian. [ONLINE] Available at:http://www.theguardian.com/guardian-us-press-office/james-ball-joins-guardian-us-office. [Accessed 11 April 2015].
21. Use Templates to Make News Apps Quickly | Open Knowledge Blog . 2015. *Use Templates to Make News Apps Quickly | Open Knowledge Blog* . [ONLINE] Available at: http://blog.okfn.org/2013/04/03/use-templates-to-make-news-apps-quickly/. [Accessed 30 April 2015].
22. Using Data Journalism to Generate Content Ideas. 2015. *Using Data Journalism to Generate Content Ideas*. [ONLINE] Available at: http://www.sitepoint.com/generate-content-ideas-data-journalism/. [Accessed 30 April 2015].

23. Investigative Reporters and Editors | NICAR. 2015. *Investigative Reporters and Editors | NICAR*. [ONLINE] Available at: http://www.ire.org/nicar/. [Accessed 30 April 2015].
24. Data Wrangler. 2015. *Data Wrangler*. [ONLINE] Available at: http://vis.stanford.edu/wrangler/. [Accessed 30 April 2015].
25. Create a new fiddle - JSFiddle. 2015. *Create a new fiddle - JSFiddle*. [ONLINE] Available at: https://jsfiddle.net/. [Accessed 30 April 2015].
26. Scene-setting survey. 2015. *Scene-setting survey*. [ONLINE] Available at:http://spacebabyuk.co.uk/Graduation/DataBank/LancashireEveningPost.htm. [Accessed 22 March 2015].
27. Data journalism survey: a mixed picture | News & Analysis | Data Driven Journalism. 2015. *Data journalism survey: a mixed picture | News & Analysis | Data Driven Journalism*. [ONLINE] Available at: http://datadrivenjournalism.net/news_and_analysis/data_journalism_survey_analysis. [Accessed 30 April 2015].
28. Data Journalism Students Survey. 2015. *Data Journalism Students Survey*. [ONLINE] Available at: https://www.surveymonkey.com/s/PYF5F3F. [Accessed 14 April 2015].
29. OpenRefine. 2015. *OpenRefine*. [ONLINE] Available at:http://openrefine.org/. [Accessed 22 March 2015].
30. Datawrapper. 2015. *Datawrapper*. [ONLINE] Available at:https://datawrapper.de/. [Accessed 22 March 2015].
31. cartodb Maps. 2015. *cartodb Maps*. [ONLINE] Available at:http://spacebabyuk.co.uk/Graduation/DataBank/cartodb.htm. [Accessed 22 March 2015].
32. ScraperWiki. 2015. *ScraperWiki*. [ONLINE] Available at:https://scraperwiki.com/. [Accessed 21 April 2015].
33. Chartbuilder. 2015. *Chartbuilder*. [ONLINE] Available at:http://quartz.github.io/Chartbuilder/. [Accessed 28 April 2015].
34. Interactive JavaScript charts for your webpage. 2015. *Interactive JavaScript charts for your webpage*. [ONLINE] Available at:http://www.highcharts.com/. [Accessed 28 April 2015].
35. Timeline JS - Beautifully crafted timelines that are easy, and intuitive to use.. 2015. *Timeline JS - Beautifully crafted timelines that are easy, and intuitive to use.*. [ONLINE] Available at: http://timeline.knightlab.com/. [Accessed 28 April 2015].

36. Tableau Desktop | Tableau Software. 2015. *Tableau Desktop | Tableau Software*. [ONLINE] Available at: http://www.tableau.com/products/desktop. [Accessed 28 April 2015].
37. . 2015. . [ONLINE] Available at: https://www.google.com/maps/d/. [Accessed 28 April 2015].

Google Drive Links

G.i. JN3111/JC4005 Data Journalism MIP
https://drive.google.com/file/d/0ByTm8YTrwNuoSGZsN1REUTFMMTQ/view?usp=sharing

G.ii. City Population density and area xlsx file
https://drive.google.com/file/d/0ByTm8YTrwNuoN1hUUU16ZnhSYVE/view?usp=sharing

G.iii. Google Map Engine Lite lesson plan
https://drive.google.com/file/d/0ByTm8YTrwNuoaW5RSUU1MV9DMFE/view?usp=sharing

G.iv. Email 18th February 2015 Megan Knight to Journalist Students / meetings
https://drive.google.com/file/d/0ByTm8YTrwNuocm94bFhuR0Y0OUU/view?usp=sharing

G.v. Email 22nd February 2015 Megan Knight to Journalist Students / meetings
https://drive.google.com/file/d/0ByTm8YTrwNuoSVh6aVMwb2pnTTg/view?usp=sharing

G.vi. Email 22nd February 2015 Megan Knight to Journalist Students / Update
https://drive.google.com/file/d/0ByTm8YTrwNuoM0lyTXpoZW1rV1U/view?usp=sharing

G.vii. LEP Survey readers' comments
https://drive.google.com/file/d/0ByTm8YTrwNuoNWllYWFRdXdoZzA/view?usp=sharing

G.viii. United Kingdom election rules and regulation PowerPoint
https://drive.google.com/file/d/0ByTm8YTrwNuobkpvNkN6TzdrVlE/view?usp=sharing

G.ix. Mike Hill email regarding data journalist skills
https://drive.google.com/file/d/0ByTm8YTrwNuod0dLWk9fZlR3bFE/view?usp=sharing

G.x. Ruby Kitchen Email Interview
https://drive.google.com/file/d/0ByTm8YTrwNuoVXhIT05sNHdLZlE/view?usp=sharing

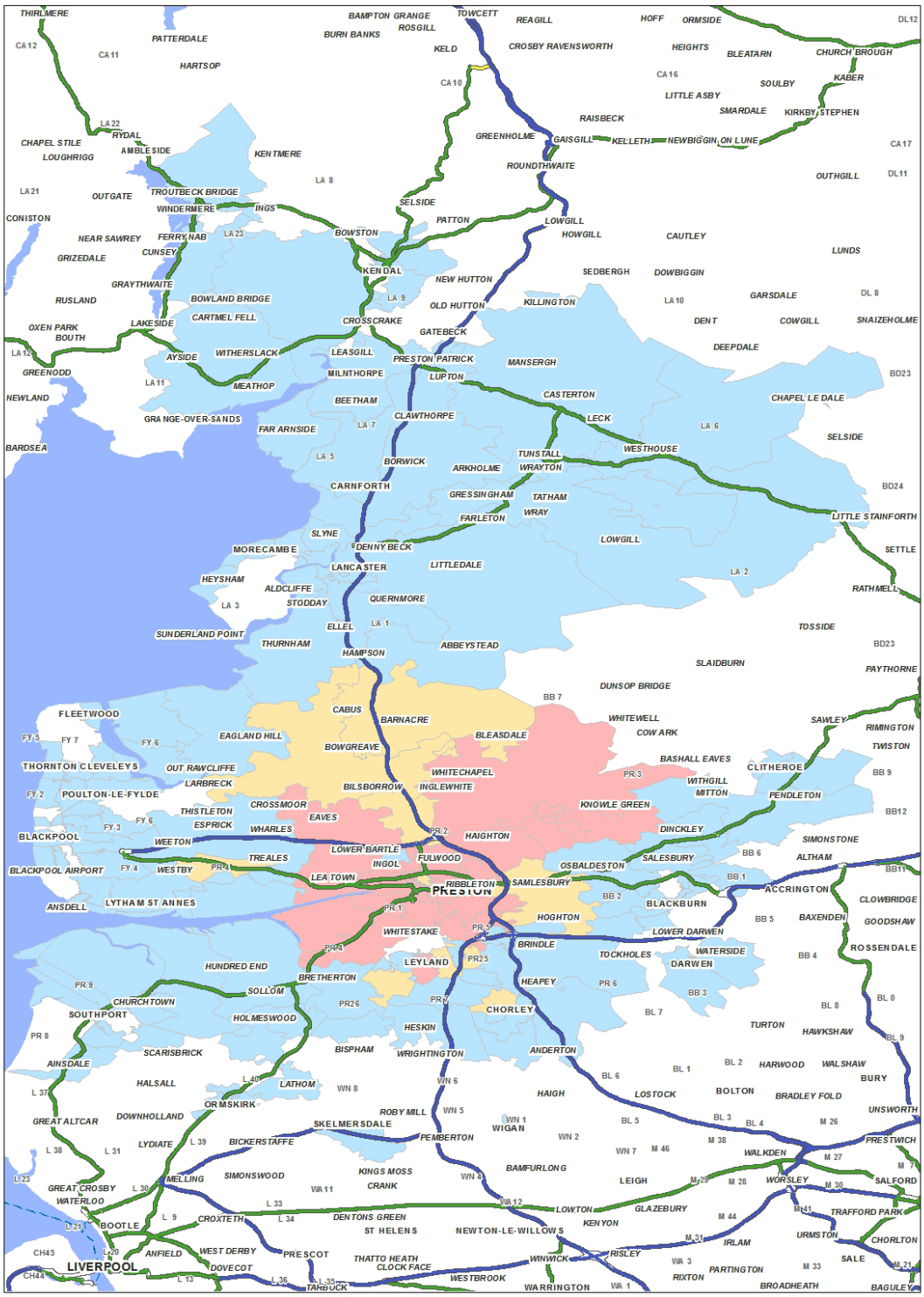
G.xi. Story planning
https://docs.google.com/document/d/1xAlVaPkb49acb_6NBOC5CNvIAoXEKvEVhkfsnptr9Ck/edit?usp=sharing

G.xii. Migration story development folder
https://drive.google.com/folderview?id=0ByTm8YTrwNuofnp6R2ktSGVwZkl0NUNsWkw3Yk1EdkRadWNGSlRPRVUyaHF1bVVIRVJzSjQ&usp=sharing

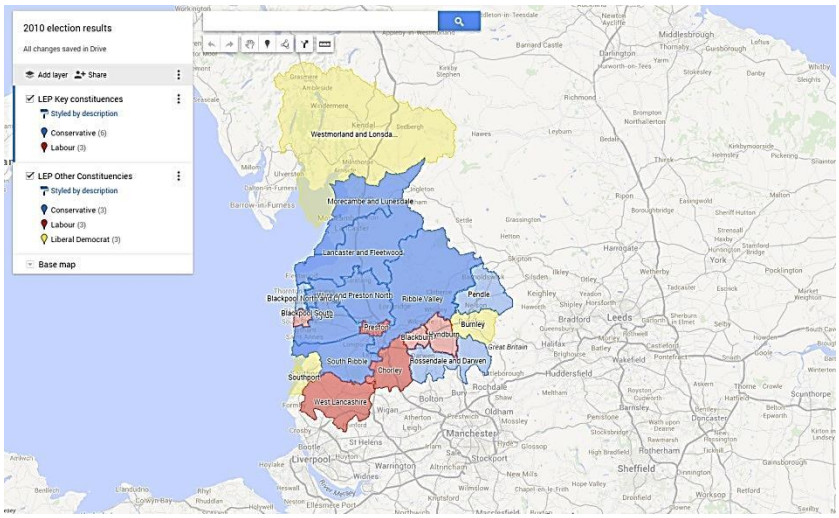
Appendix

A.i

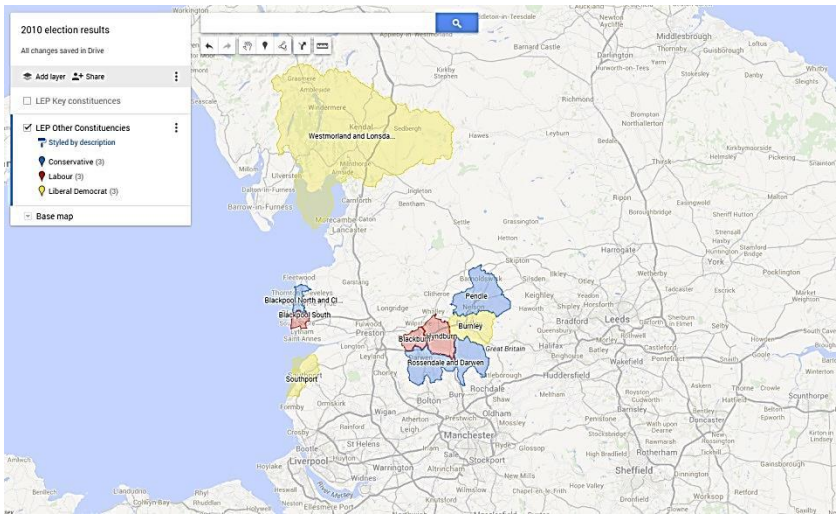
Lancashire Evening Post Readership Coverage Map



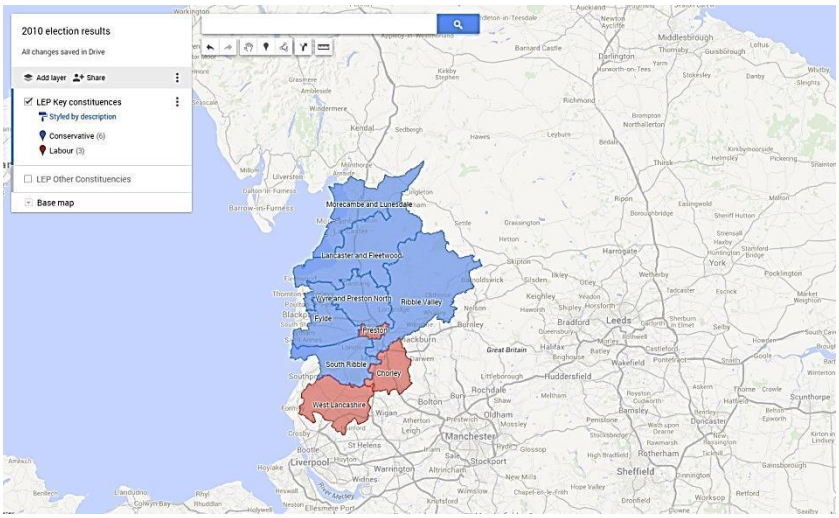
A.ii


2010 election results
All changes saved in Drive
Add layer
Share
LEP Key constituences
Styled by description
Conservative (6)
Labour (3)
LEP Other Constituencies
Styled by description
Conservative (3)
Labour (3)
Liberal Democrat (3)
Base map
Westmorland and Lonsda...
Morecambe and Lunesdale
Lancaster and Fleetwood
Blackpool North and Preston North
Ribble Valley
Pendle
Blackpool South
Preston
Burnley
Blackburn
Great Britain
South Ribble
Chorley
Rossendale and Darwen
Southport
West Lancashire

A.iii


2010 election results
All changes saved in Drive
Add layer
Share
LEP Key constituences
LEP Other Constituencies
Styled by description
Conservative (3)
Labour (3)
Liberal Democrat (3)
Base map
Westmorland and Lonsda...
Pendle
Burnley
Great Britain
Rossendale and Darwen
Southport

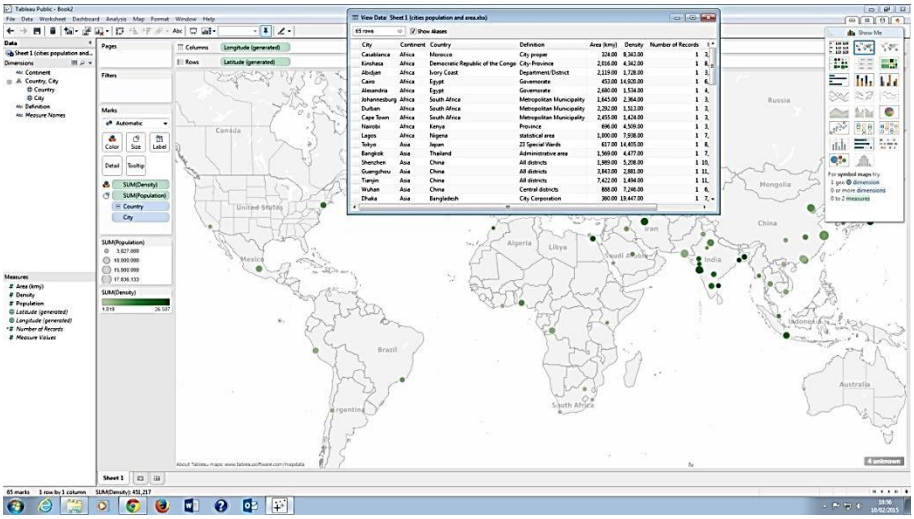
A.iv


2010 election results
All changes saved in Drive
Add layer
Share
LEP Key constituences
Styled by description
Conservative (6)
Labour (3)
LEP Other Constituencies
Base map
Morecambe and Lunesdale
Lancaster and Fleetwood
Ribble Valley
Preston
Great Britain
South Ribble
Chorley
West Lancashire

A.v

Elisabetta Tola's 'map of ignorance'

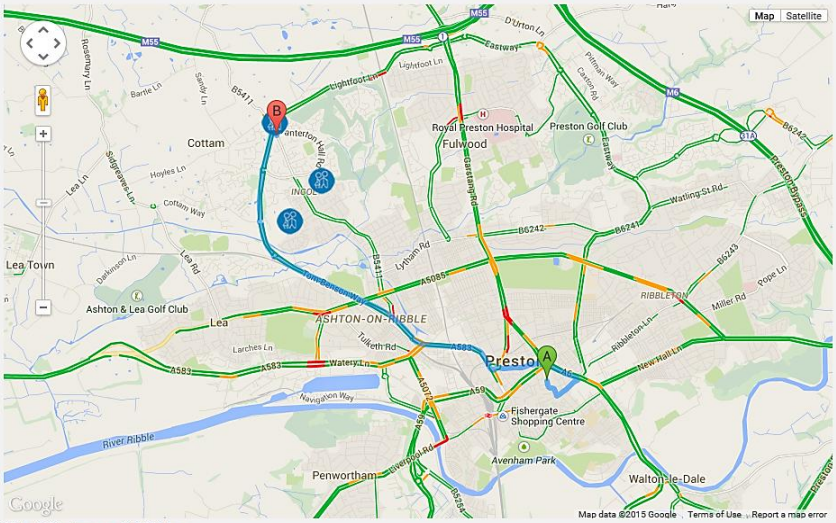
A.vi



A.vii

Welcome to the 'Last Stop' because 'We Care'

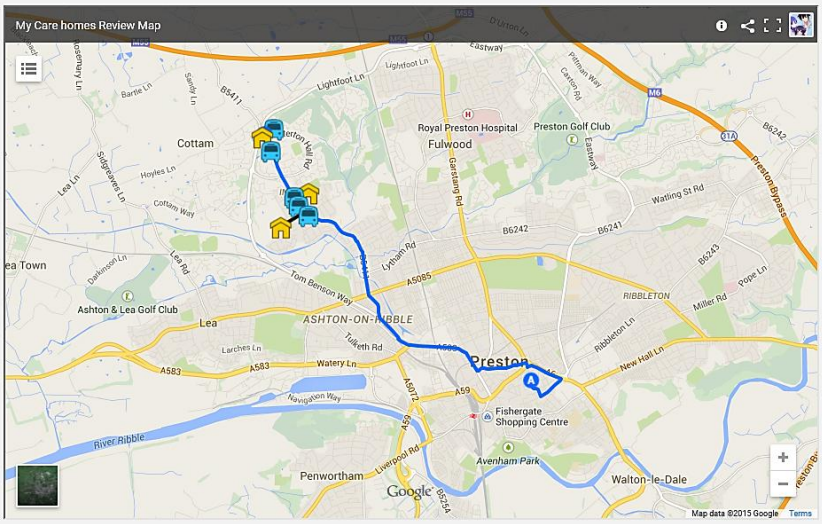
Google Map hand coded using an API key Location of 3 comparison care homes with information, images, traffic overlay and directions from city bus station.



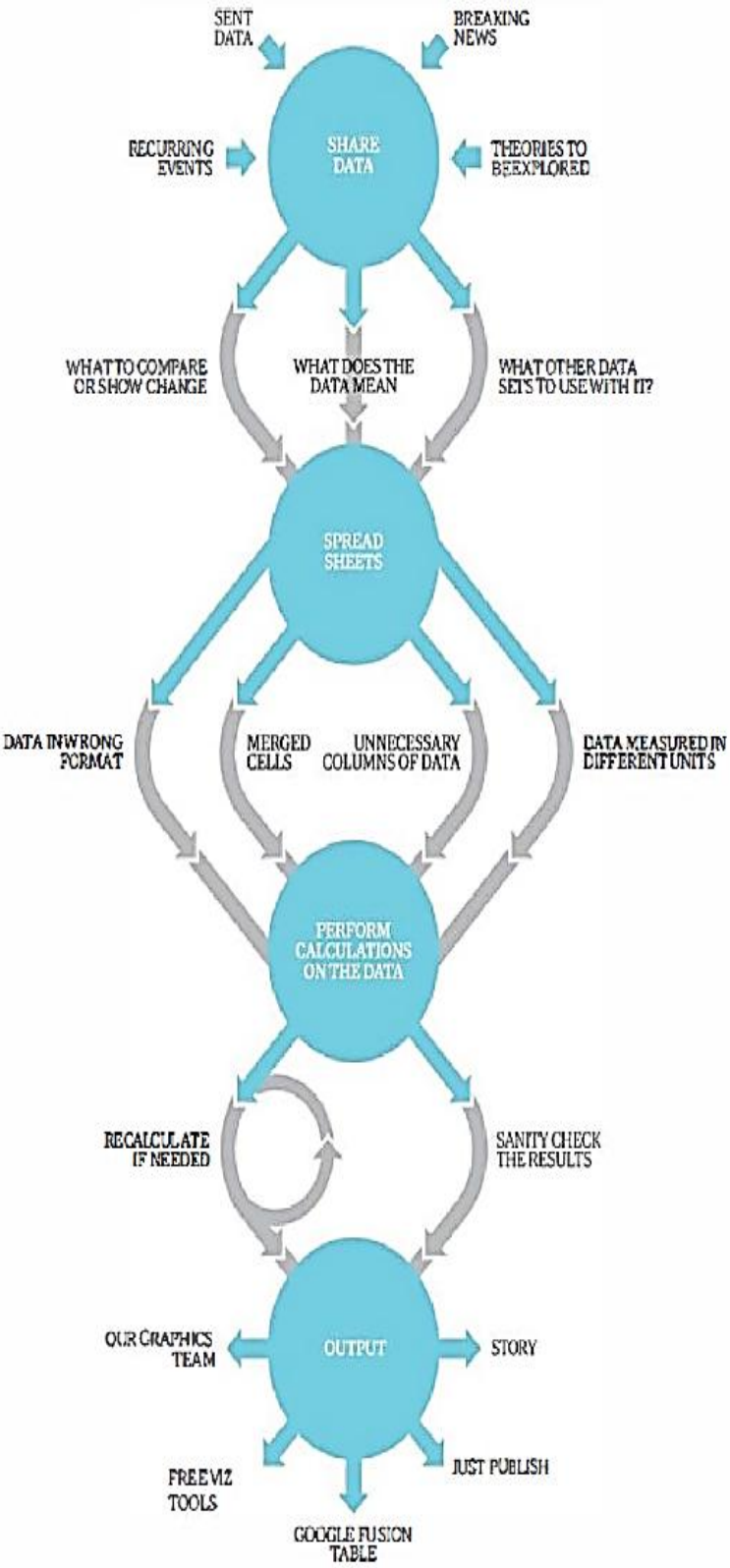
Total Distance: 6.878 km

Welcome to the 'Last Stop' because 'We Care'

Map created using Google Map Lite Location of 3 comparison care homes with information, images, and directions from city bus station.

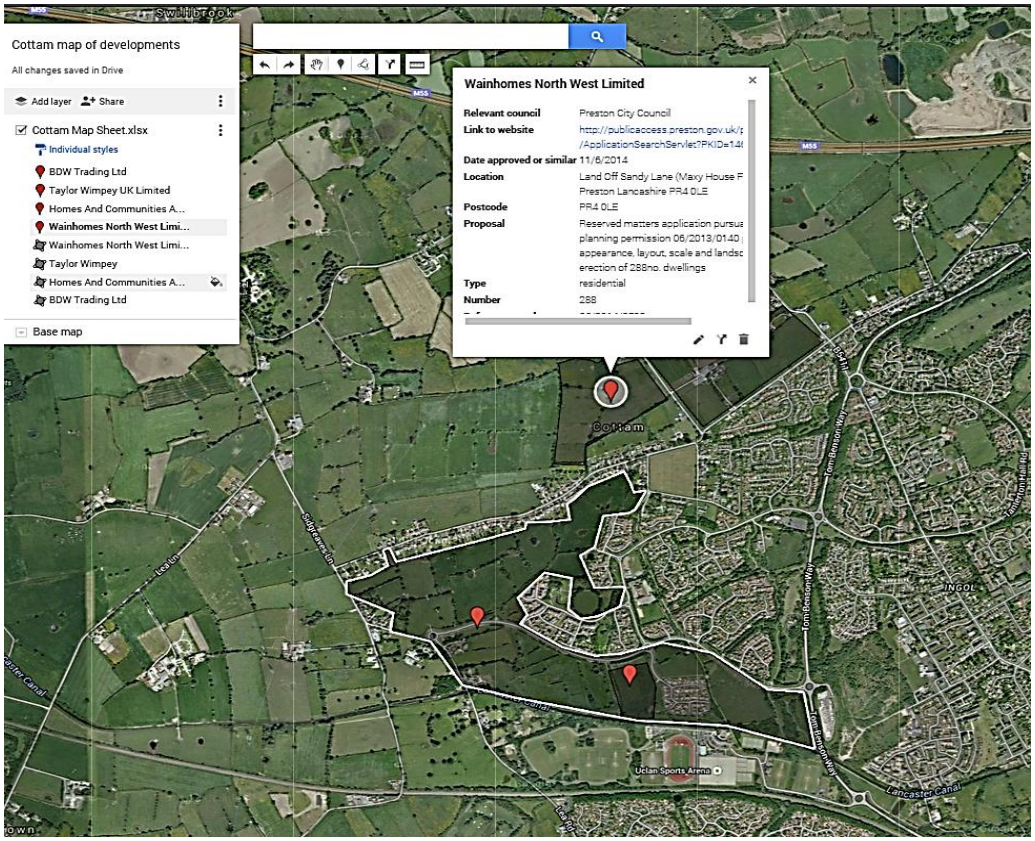


A.viii

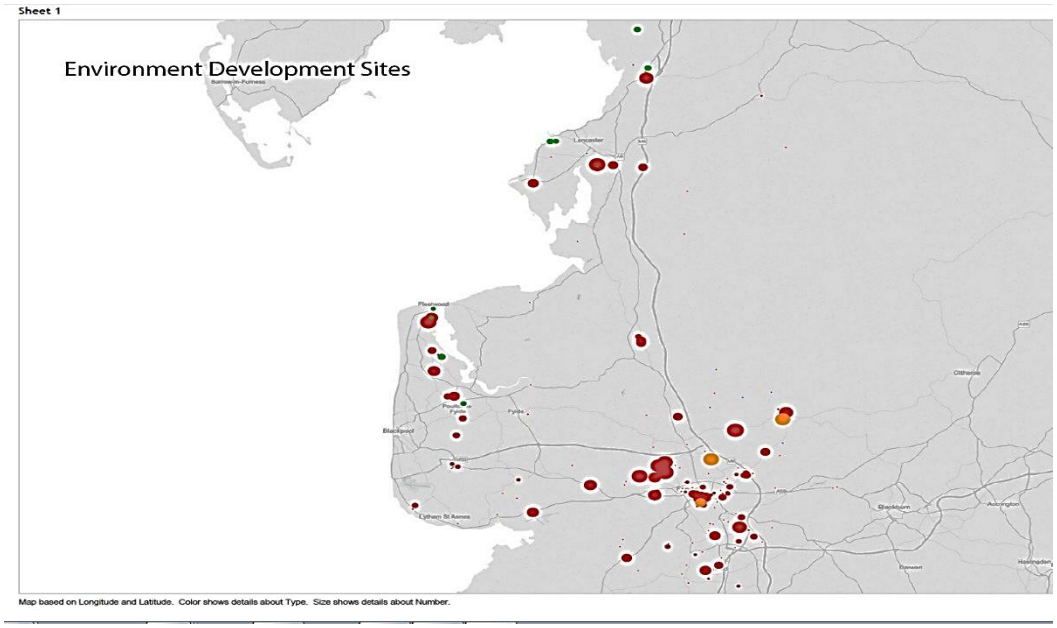


The Guardian Datablog production process visualized (the Guardian)

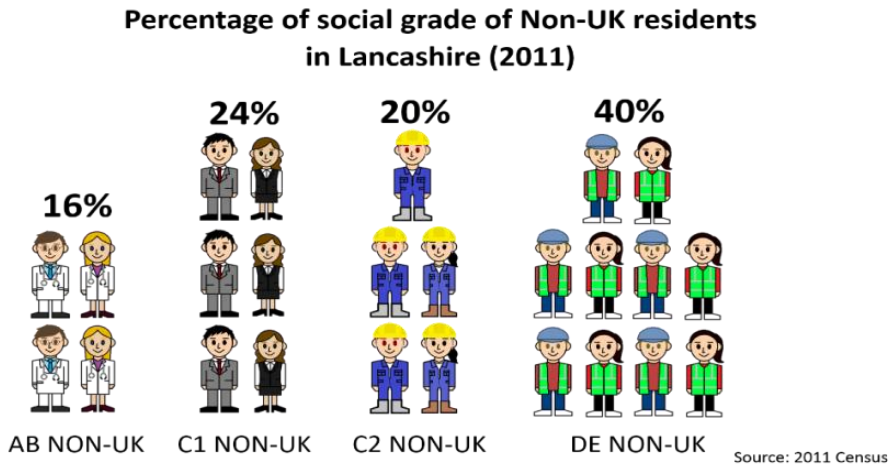
A.ix


Cottam map of developments
All changes saved in Drive
Add layer
Share
Cottam Map Sheet.xlsx
Individual styles
BDW Trading Ltd
Taylor Wimpey UK Limited
Homes And Communities A...
Wainhomes North West Limi...
Wainhomes North West Limi...
Taylor Wimpey
Homes And Communities A...
BDW Trading Ltd
Base map
Wainhomes North West Limited
Relevant council
Preston City Council
Link to website
Date approved or similar 11/6/2014
Location
Land Off Sandy Lane (Maxy House F
Preston Lancashire PR4 0LE
Postcode
PR4 0LE
Proposal
Reserved matters application pursua
planning permission 06/2013/0140
appearance, layout, scale and landsc
erection of 288no. dwellings
Type
residential
Number
288
Cottam
INGOL
Uclan Sports Arena
Lancaster Canal

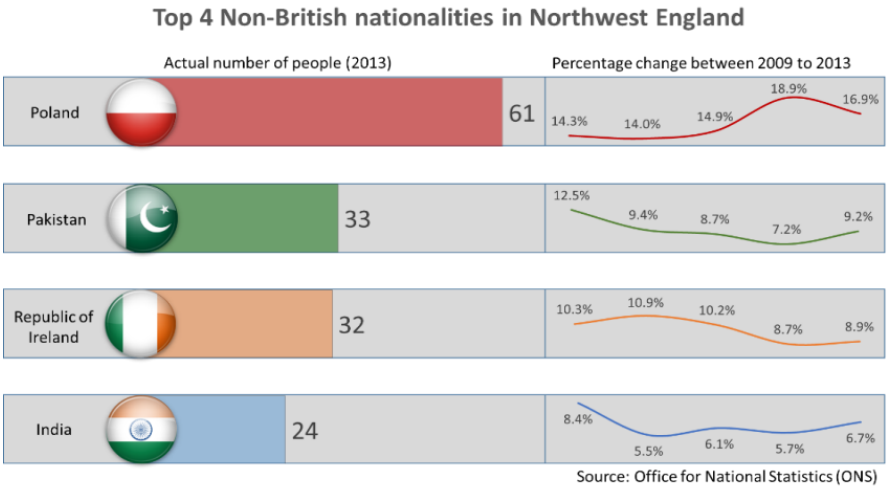
A.x


Sheet 1
Environment Development Sites
Map based on Longitude and Latitude. Color shows details about Type. Size shows details about Number.

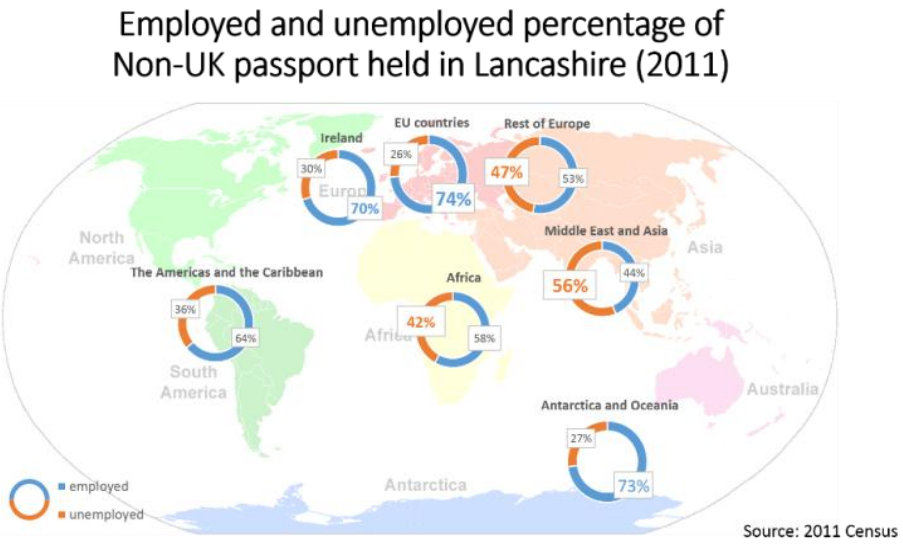
A.xi



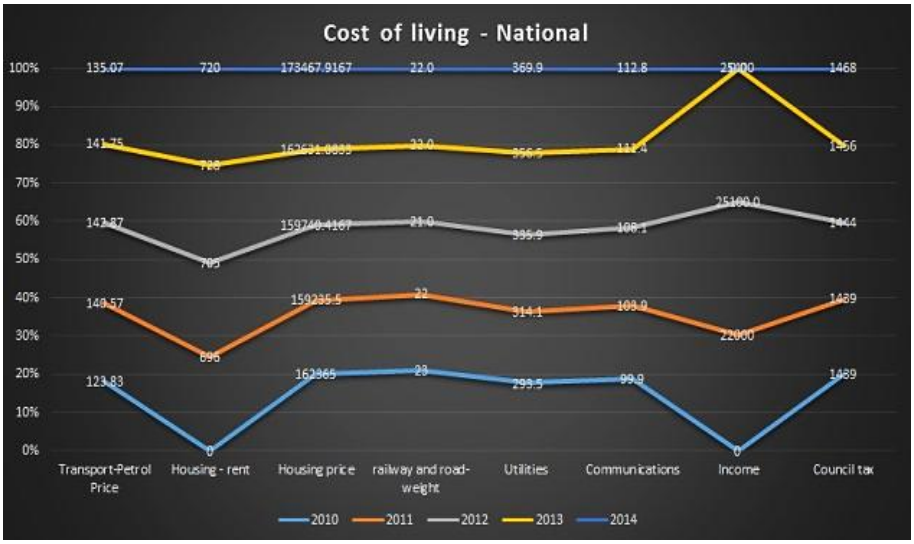
A.xii



A.xiii



A.xv


Cost of living - National
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Transport-Petrol Price
Housing - rent
Housing price
railway and road-weight
Utilities
Communications
Income
Council tax
2010
2011
2012
2013
2014

A.xvi


Cost of living - National weigths
2010
2011
2012
2013
2014
130
125
120
115
110
105
100
95
90
Transport Petrol Price
Housing - rent
Housing price
Cost of food
railway and road
Utilities
Council tax
Communications
Income

A.xiv (http://spacebabyuk.co.uk/Graduation/DataBank/cartodb.htm)

A.xvii