

# *FIT5147 DATA VISUALISATION PROJECT*

A Sentiment Analysis on Donald Trump's Twitter

*Isobel Rowe*

30042585

## Table of Contents

|                             |           |
|-----------------------------|-----------|
| <b>1. INTRODUCTION</b>      | <b>3</b>  |
| <b>2. DESIGN</b>            | <b>4</b>  |
| <b>2.1 5DS SHEET 1</b>      | <b>4</b>  |
| <b>2.2 5DS SHEETS 2/3/4</b> | <b>4</b>  |
| SHEET 2                     | 4         |
| SHEET 3                     | 5         |
| SHEET 4                     | 5         |
| <b>2.3 5DS SHEET 5</b>      | <b>5</b>  |
| <b>3. IMPLEMENTATION</b>    | <b>6</b>  |
| SECTION 1                   | 7         |
| SECTION 2                   | 8         |
| SECTION 2/3/4               | 8         |
| SECTION 6                   | 9         |
| SECTION 7                   | 9         |
| SECTION 8                   | 10        |
| <b>4. USER GUIDE</b>        | <b>10</b> |
| SECTION 1                   | 10        |
| SECTION 2                   | 11        |
| SECTION 3                   | 11        |
| SECTION 4                   | 12        |
| SECTION 5                   | 12        |
| SECTION 6                   | 13        |
| SECTION 7                   | 13        |
| SECTION 8                   | 14        |
| <b>5. CONCLUSION</b>        | <b>14</b> |
| <b>6. REFERENCES</b>        | <b>15</b> |
| <b>7. APPENDIX</b>          | <b>16</b> |

## 1. Introduction

Donald Trump's singular experience of Twitter is different from that of the ordinary user. Because he is one of the most well-known men in the world, and his twitter is one of the most notorious on the platform, they're guaranteed to get noticed and discussed. Trump's tweets don't just amass thousands of favourites, replies and retweets; they appear, often within minutes of being posted, in high-definition blow-ups on news breakfast shows; they're read verbatim, by TV and radio newscasters and personalities; they're embedded in stories and articles on news websites; and they're praised, attacked and analysed both at home and abroad. Whether it's sharing far-right extremist content (Pearce, 2017), attacking federal judges (Phillips, 2017), or jostling with various world leaders (Johnson, 2017) - when Donald Trump tweets, the world listens.

There is so much talk around the content of his tweets, but I want to dive a little bit deeper, and use this project to analyse and break-down exactly what Donald Trump's Twitter looks like and how he uses it. In the exploration project, I discovered a lot of different things from the content-rich dataset I gathered. It's so rich in information, that it would be impossible to break down each and every theme. So, for this visualisation, I will aim to provide an overall view of what his Twitter feed has looked like throughout his presidency so far through the lens of sentiment analysis.

So, who benefits or could use this information? Well, it's hard to identify a particular group of individuals, because, there's so many: his political party, the opposing political party, his supporters or enemies, his 370 million twitter followers - the world at large. Because, not only is his twitter actually *entertaining*, but, whether we like it or not, politicians and political parties play a huge role in society and in our everyday lives.

This is why I think it's important to understand the way in which the nature of communication between the world leaders and the public is changing. This new wave of communication between politicians and their supporters via social media has provided an unprecedented view into their thoughts and psyche, and perhaps paved way for a big change. Because, after all, you could chalk his prolific twitter use up to a one-of-a-kind celebrity president fond of those 280 characters, or you could call it a step into an uncertain future.

## 2. Design

In order to design an effective and practical visualisation, I used the Five Design Sheet methodology to create quick ideation of low-fidelity models. These sheets will be explored in the proceeding sections and can be viewed in Appendix 1.

### 2.1 5DS Sheet 1

The first design sheet is used to sketch various ideas and different possibilities to think them through and consider their potential effectiveness as visualisation methods. I created preliminary outlines of various visualisations, including joyplots, interactive donut charts, bar charts, word clouds and beeswarms. I thought about linking sentiment ratings to different colours, and having mouseover options to display the content of the tweet to the corresponding data point. Eventually, I settled on three options to explore in a more in-depth sense in the following sheets.

### 2.2 5DS Sheets 2/3/4

Next, I created three principal design sheets, detailed in the following sections.

#### Sheet 2

In Sheet 2, I investigated the option of an interactive beeswarm in D3.js. This visualisation would have various interactive capabilities, including:

- A zoom-able axis, allowing users to zoom in on certain sections of the axis, which displays a 'favourite' count. This is helpful as, because there are so many data points due to the sheer number of tweets, data can become compacted on the axis – zooming in and out would help to display data points in a clearer way.
- Colour linking to sentiment, where each data point would be coloured based on sentiment – red for negative, and green for positive, providing an easy-to-understand visual indicator for users.
- Click-for-tweet, which allows users to click on a data point of interest to display the tweet content.

I think that this is a really good visualisation technique. It's effective, provides links back to the data, and allows for various user interactions.

### Sheet 3

In Sheet 3, I explored an interactive word cloud Shiny app. This would use a menu system, allowing users to vary the following inputs:

- Choose a sentiment: controls what sentiment is displayed, for example 'negative', 'positive', 'angry' etc.
- Maximum words: controls the maximum number of words displayed in the word cloud.
- Minimum frequency: controls the minimum frequency in Trump's Twitter corpus for each word in the word cloud.

This option is really easy to implement, but offers little insight into the data findings other than frequency of words.

### Sheet 4

In Sheet 4, I assessed a data visualisation scroller with D3.js. This operates on a 'scroll to trigger' function where users are guided through narrated sections with graphs. The act of scrolling to the next section indicates the users' readiness to ingest the next piece of content. This is, in fact, the main interaction that this visualisation comprises. However, there is also room for added interaction within sections, with things like mouseover functions to display added data.

This is by far the most difficult visualisation option for me as I am only just becoming familiar with the elements of D3.js – HTML, CSS, and JavaScript. However, it is also the visualisation option that makes the most sense for the data – it uses the data to tell a story in a dynamic way, providing a lot of interesting information along the way.

## 2.3 5DS Sheet 5

Finally, I decided to choose Option 3: Scrollytelling with D3.js. Ultimately, this decision came down to accessibility – my target audience is broad, and I wanted to create a visualisation that would be

approachable to everyone. It was important to make it easy, not hard, to use and understand the webpage.

As a form of interaction, scrolling is ubiquitous – those who use the internet do it constantly, which makes this interaction type almost effortless for the user. There are other options, such as using steppers, where the user has to click to open the next sections, but this is a comparatively complex option compared to scrollers.

This is what most of my following design decisions have come down to, as well, because I don't necessarily know what the user's knowledge of or familiarity with both the internet and data visualisation, but I'm pretty sure most people will know how to scroll down a webpage, and be able to understand simple grids and bar graphs.

In order to converge on a final realisation design to be implemented, I honed in on a few different graph options, including beeswarms, bar graphs, or even a grid, which is divided into small squares that each represent a tweet, that can be highlighted according to sentiment value. Other than graphs, sections may include just words or an embedded tweet to really emphasise or focus on a piece of information or data point.

### 3. Implementation

Although it was an uphill battle in learning, I decided to use D3.js for the visualisation. This is because it offers a level of versatility that Shiny apps cannot. It was important to me that the final visualisation be visually effective, and tailored to fit my exact requirements. Moreover, D3 was required if I was to do the 'scrollytelling' or 'data journalism' option.

To start, I decided to pre-process the data a little bit more to create a dataset that would be easy to use in D3. I used the Donald Trump's Twitter dataset that I gathered and wrangled in the Exploration Project, but I added a few columns for the sentiment values – for instance, I added a Boolean column for both 'positive' and 'negative' which had the values of either 1 for 'true' and 0 for 'false'. Beyond this, I mainly kept everything the same, but cleaned it up so that it only contained the data necessary for the visualisation, which meant removing most columns.

Next, to implement the design created in Sheet 5, I created three files – a HTML, CSS, and JavaScript file. These are shown in the code submission, with comprehensive comments. In addition, I used a JavaScript file created by Jim Valladingham which is a standalone, reusable scroll detection function used to create the scrolling animations – essentially handling the details of figuring out which section of the visualisation the user is currently scrolled to.

When it came to actually implementing the design set out in Design Sheet 5, I realised quickly that there were two problems:

- I was limited by my lack of experience. Having no previous practical knowledge of JavaScript, I found the scroll functionality really hard to implement with respect to the graphs as there is so many little things to get right in a lot of different sections across different files - there was a lot of iterations to get it to the point that it was functional.
- The design on paper didn't necessarily translate very well to the computer. I had to make sure the visualisation itself worked well on screen and that the scrolling usability factor was high, with the end user always in mind.

So, in the end, the final design strayed a small amount from the Design Sheet 5, but still maintained the same integrity and feel. The visualisation itself is split up into sections based on where the user is currently located - as you scroll, the current section fades out, and the next section fades in. These sections are detailed in the following sections:

## Section 1

Upon first open of the webpage, users are greeted with an introductory section. This section is short and to-the-point.

Details:

1. Intro text: I thought it was important to use this introduction to set the scene, and give the user a little bit of background information. However, I don't necessarily want the narrative visualisation to have a lot of text and I don't want the user to have to rely on reading a lot of content to understand the section. The bulk of the communication will come from the

graphs themselves that change as the user scrolls, and will be complimented by a very little amount of text that guides you through the story.

2. Animated eyes: This sets a mood – it's fun and commands attention, enticing users to keep reading.
3. Animated 'keep scrolling' icon: Suggests that there is more information. Depending on the users' screen resolution, without the scroll icon, they might not realise they can scroll. The icon is a visual indication or explicit prompt of additional content, avoiding leaving the user stranded on the introduction section.
4. Colour choice: The background blue is #1DA1F2 – the colour of Twitter's logo. I wanted to keep the same colour palette as a way of linking the content back to the source.
5. Font choice: I chose to use Helvetica, as it is a common typeface that is accessible by everyone, since it's a web font. It's also really legible, which means people won't have a problem reading it on any background – especially the blue of the introduction section.

## Section 2

This section is a secondary opener, encompassing just simple text, drawing attention to the information being presented.

1. Hierarchy of information: Here, I used visual characteristics to help influence and direct the user's perception of information on the page. I wanted users to first see the number, then the tweets, and finally the block text to the left. I did this by using visual characteristics of: size, colour and contrast to draw attention and highlight certain aspect is succession of one another.
2. Text placement: In Western society, we read left to right, so information should be on the left. Once again, there is not a lot of text – just a digestible amount of supplementary information/explanation of the graphs.

## Section 2/3/4

These sections are linked as they use the same visualisation, but are broken up between sections to hone in on the points.

1. Colour choice: Although the concept of keeping the blue theme was great, in practice, it became clear that sticking to blue, and only blue, was letting the visualisation down. It just didn't display enough contrast, nor map to 'positive' and 'negative' in the way that red and



blue could. It's important to have these clear identifiers for the user, ensuring digestibility and comprehensibility. It was a shame to break this theme, but I think the end result turned out a lot clearer, and sometimes you have to break consistency in favour of clarity.

2. Grid choice: I chose this graphing technique as it is an easy way to visualise each data point, showing both the depth and scale of the data. I also found that it was novel and engaging, as the more the user scrolls, the more information they gradually receive across each section.

## Section 6

1. Bar chart: I found in the exploration project that bar charts were a surprisingly efficient form of visualisation methods. They may not be the most interesting, but they get the information across succinctly. So, in this section I decided to use a horizontal bar chart, where the colours are again mapped to sentiment. The bars are also ordered from the highest to lowest value, as the bars represent unordered categories. All of these design choices are linked back to the fact that this is the easiest way to read a bar chart.

## Section 7

As per the Design Sheet 5, I really wanted to include a beeswarm chart, that mapped tweets and their sentiment values to the amount of 'likes' or 'favourites' that it received. However, I had a lot of trouble actually implementing it in D3 – I managed to make it perfectly in its own, separate file, but had a lot of trouble whilst trying to integrate it into the scrollytelling files. And ultimately, I was unable to add it to the page without inevitably breaking the other graphs. I'm not sure exactly why this happened, but I came to the conclusion that it was outside of the scope of my JavaScript abilities. Hence, it is not included in the final visualisation files, although I have added it in the user guide to give a visual idea of what it would have looked like.

In any case, the reasoning for this chart is that it showed each of the data points in an interesting way, introducing a new variable whilst sticking to the theme of sentiment analysis. I decided to just include the top 200 most 'liked' tweets, as including all 5,000+ tweets on one beeswarm chart was just not feasible. Again, I stuck with the same colour pallet, and included just a little bit of text explaining the chart.

## Section 8

Finally, the outro section. Originally, I hadn't thought this section was necessary, but I thought it was a good way of signalling the end of the visualisation, without just abruptly ending it. It was a good way of circling back and wrapping up. I used the same style as the intro section – the same font, text alignment, background colour etc.

1. Embedded tweet: I embedded this tweet to link back to the previous section (the beeswarm chart), and the to reconnect to original data source.
2. Data source: I also thought it was important to pay homage to the data source that made this project possible, so I added in a quick data source section, that links the user to the Trump Twitter Archive website if they want to find out any more.

## 4. User Guide

The user guide for this visualisation is simple – just scroll down the webpage! Upon scrolling, new sections will fade in and open up. There is reversible scrolling, meaning users can scroll in either direction, and the functionality should still work. There is a screenshot of each section as follows:

### Section 1



## Section 2

5,632

## Tweets

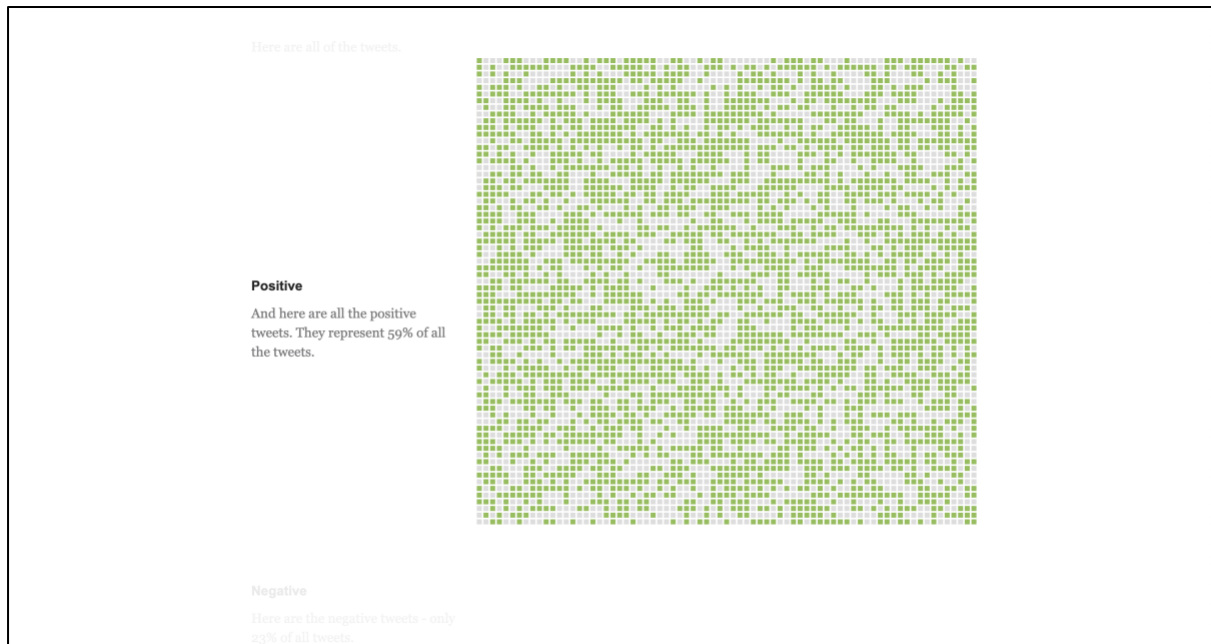
## Tweets

Since the start of his presidency, Donald Trump has pressed post on over 5,362 unique tweets (not including retweets).

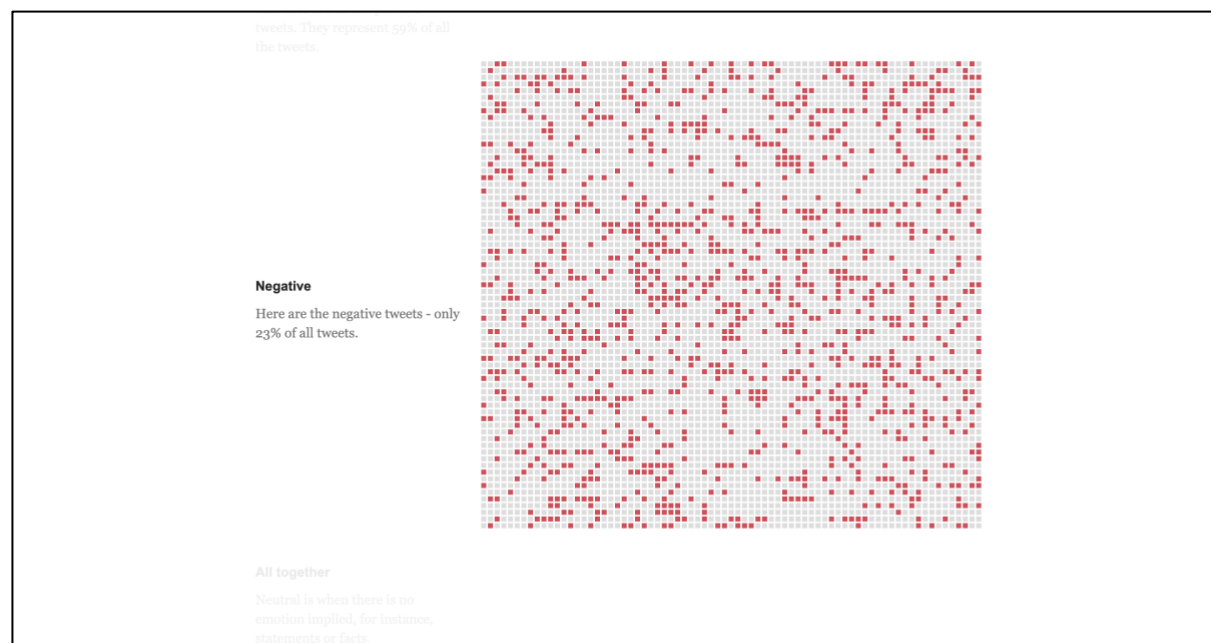
### Section 3

Here are all of the tweets.

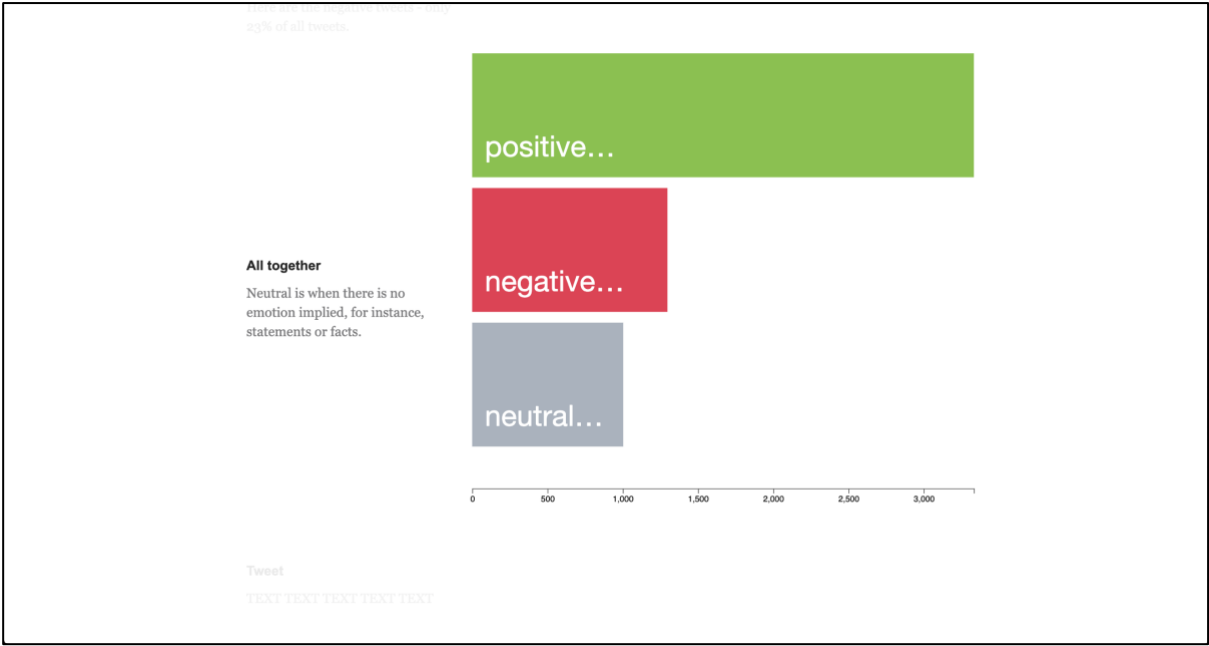
## Section 4



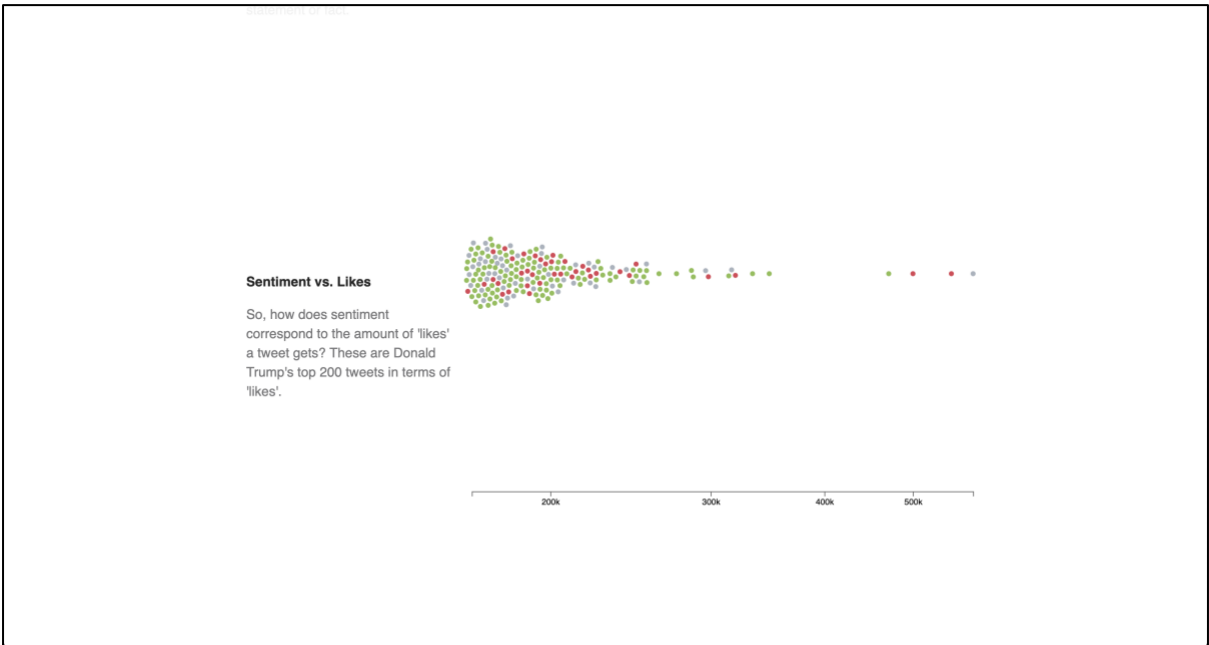
## Section 5



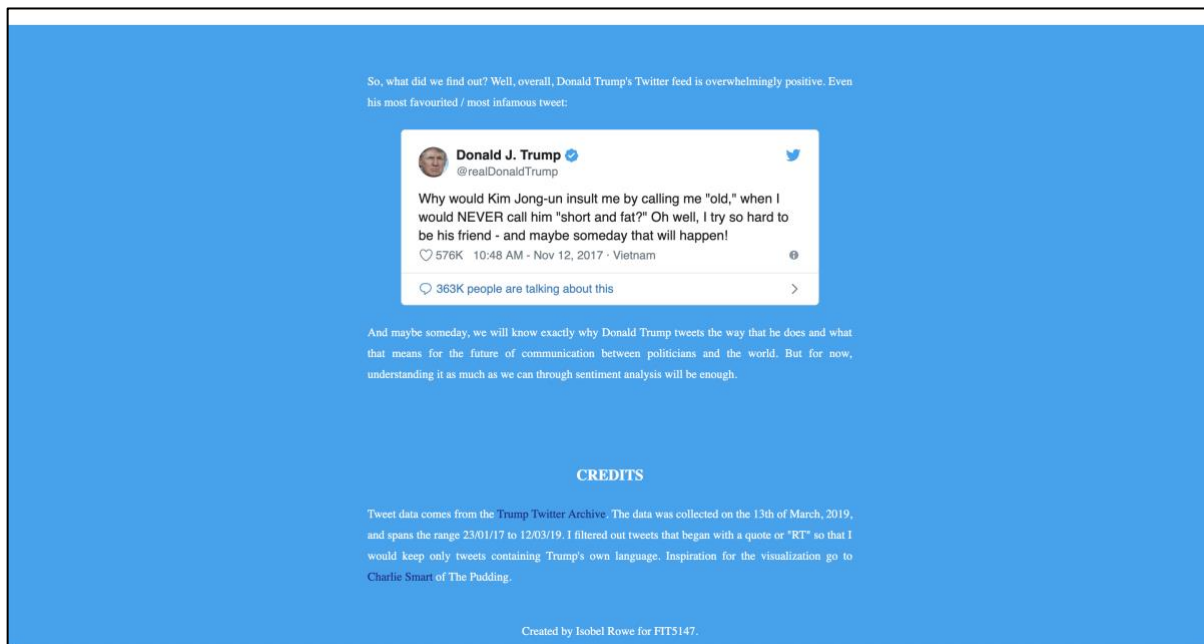
Section 6



Section 7



## Section 8



## 5. Conclusion

In this project, I managed to create an appealing visualisation underpinned by the notion that it should be accessible to my target audience. I used the insight found through data exploration to successfully communicate the findings to my target audience through an interactive story.

In terms of programming, I learnt *a lot*. I gained a deeper understanding in HTML, CSS, and JavaScript, and of the technicalities of how D3.js works. And discovered a lot about interactive visualisations and the process of design that goes into making them. I am excited to ameliorate the new skills I've gained, and put them into action in future projects.

In hindsight, I would have tried to keep the graph on the right-hand side the absolute same, but changed the inputs in each section, or on each 'scroll'. This essentially filters the data to show different findings, and changes the scales on the axis, or adds different colours etc. For instance, using a beeswarm chart that I could change based on various variables. I think this is a really effective visualisation method, and it's something I would have loved to do.

## 6. References

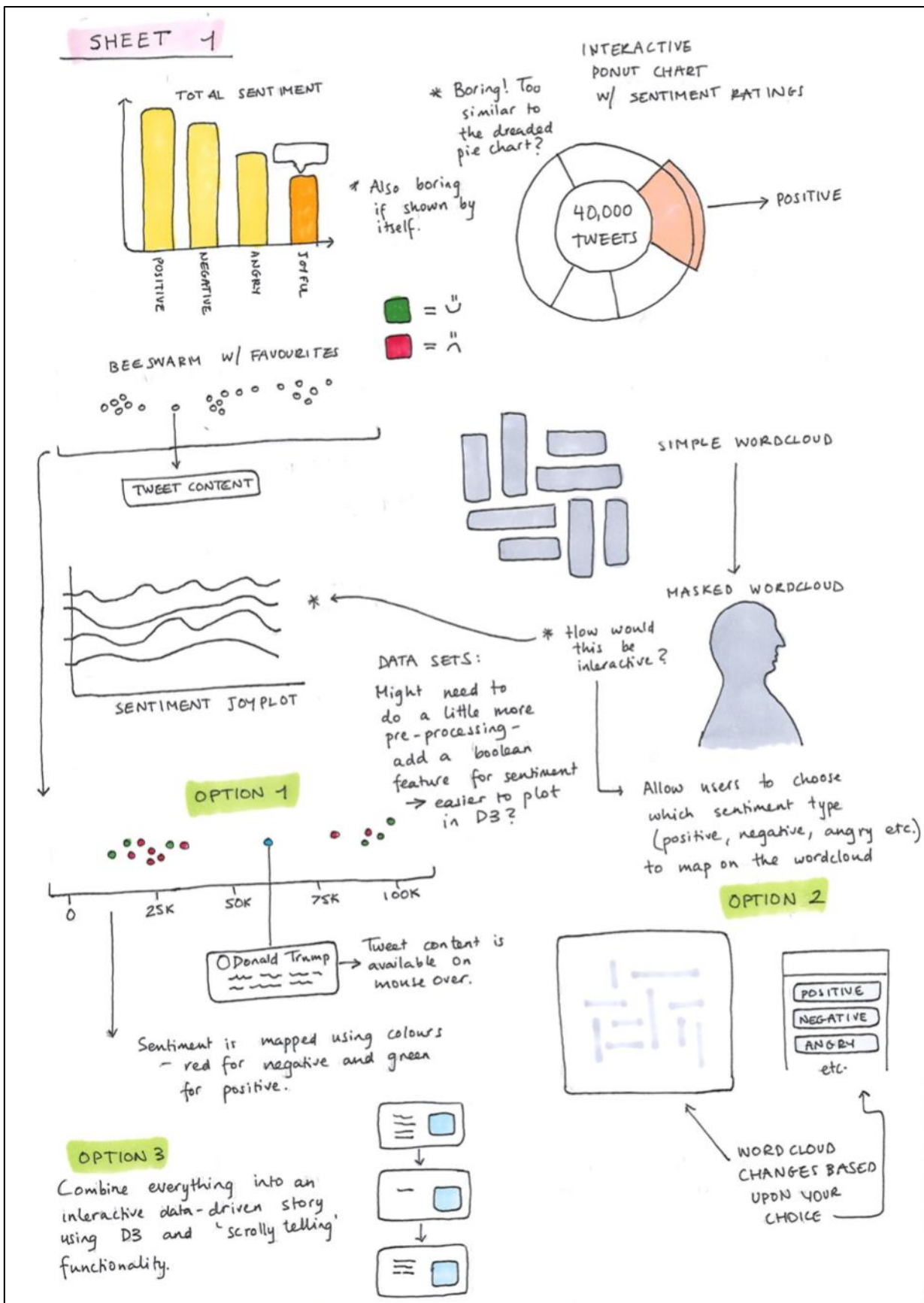
Bostock, M. (2014, November 3). How To Scroll. Retrieved from <https://bost.ocks.org/mike/scroll/>

Pearce, M. (2017, July 3). Once again, Trump tweets a meme linked to a brazen racist. *The Los Angeles Times*. Retrieved from <https://www.latimes.com/>

Phillips, K. (2017, April 26). All the times Trump personally attacked judges — and why his tirades are 'worse than wrong'. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/>

Johnson, A. (2017, September 26). Twitter Promises to Clarify Rules in Wake of Trump North Korea Threat Tweets. *NBC News*. Retrieved from <https://www.nbcnews.com/>

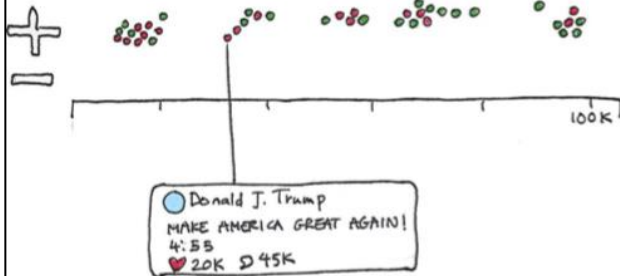
## 7. Appendix





## SHEET 2

### LAYOUT



### FOCUS



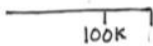
#### ZOOMABLE AXIS

- Allows users to zoom in on certain sections as data is compacted on scale due to sheer size of datapoints.

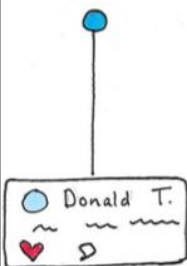


#### COLOURED POINTS BASED ON SENTIMENT

- Good visual indicator - easy to understand.



#### SCALE INDICATES NUMBER OF FAVOURITES



#### CLICK FOR TWEET

- Tweets embedded in mouse-over function.

TITLE: BEESWARM

AUTHOR: ISOBEL

DATE: 25/5/19 SHEET: 2

### OPERATION

- User has initial view of beeswarm - perhaps an explanation of how to navigate the visualisation is needed.
- Can zoom in and out on axis at whatever point they wish.
- Can click on any datapoint, which will bring up the embedded tweet that the datapoint corresponds to.

### DISCUSSION

- + Relatively simple to implement.
- + Simple to understand and use.
- + Shows the distribution of the datapoints, not just the measured values.
- Misses a lot of the interesting findings from the exploration.
- Gets boring quickly for user as there isn't much interaction involved!

## SHEET 3

### LAYOUT

### TRUMP'S TWITTER IN WORDCLOUDS

CHOOSE A SENTIMENT

CHOICES

MAXIMUM WORDS

0 250

MINIMUM FREQUENCY

25 25

TERRIBLE HOAX POLITICS  
 PHONY COURT  
 CRAZY FIGHTING  
 CRIMINAL MONEY  
 ABUSE

TITLE: INTERACTIVE WORDCLOUD

AUTHOR: ISOBEL

DATE: 25/5/19 SHEET: 3

### OPERATION

- Created as a R Shiny application.
- Uses a menu system to allow users to vary the inputs to create various wordclouds. This is detailed in the 'FOCUS' section.

### FOCUS

Colour changes depending on which sentiment is chosen to map.

### DISCUSSION

- + Fast and engaging for users of the app.
- + Really easy to implement.
- + Shows overall theme of Trump's Twitter.
- Offers little insight into data findings other than frequency.
- Way too simple for such an important project.
- Doesn't provide context.

CHOOSE A SENTIMENT

POSITIVE

MAXIMUM WORDS:

1 150 250

MINIMUM FREQUENCY:

1 10 25

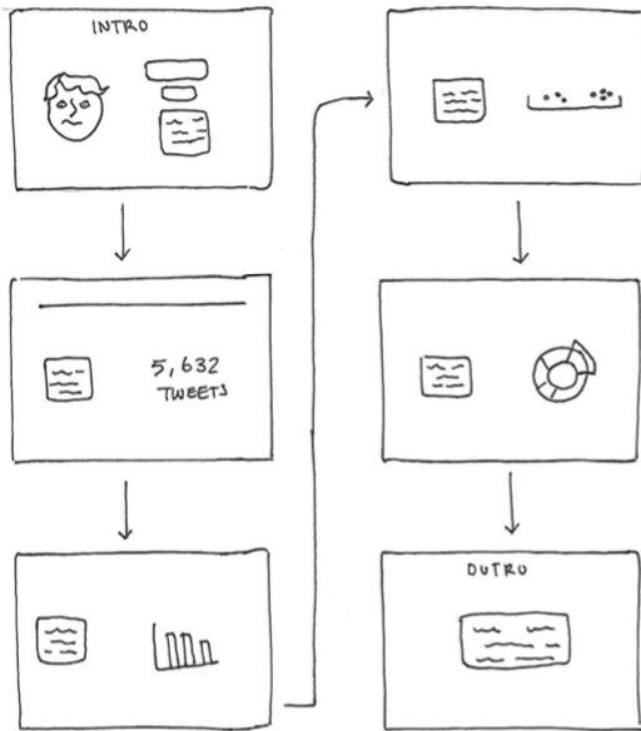
Controls what sentiment will be displayed

POSITIVE  
 NEGATIVE  
 POSITIVE  
 ANGRY  
 etc.

Controls the minimum frequency in the Twitter corpus for each word in the wordcloud.

Controls the maximum number of words in the wordcloud.

## SHEET 4



TITLE: DATA VIZ SCROLLER

AUTHOR: ISOBEL

DATE: 26/5/19 SHEET: 4

### OPERATION

BROADLY:

- Users use a 'scroll to trigger' functionality where they are guided through narrated graphs/plots.

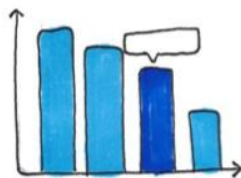
WITHIN SECTIONS:

- Users can mouseover/click certain points on the plot to bring up more information.

### DISCUSSION

- + Tells a story
- + Provides a lot of information on Trump's Twitter in a dynamic way.

- Main interaction is out of the users hands as it is dictated by scrolling.
- Difficult to implement with a lot of aspects that I haven't used before.



MOUSEOVER:  
o highlights bar and provides exact quantity.

Possibly use the blue from the Twitter logo as inspiration for the colour palette?

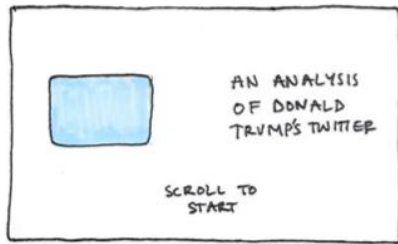


- The main interaction that the visualisation has with the user is the 'scroll to trigger' functionality. However, some of the sections will have mouseover functions for added interaction

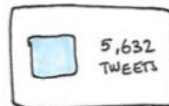
## SHEET 5

### LAYOUT

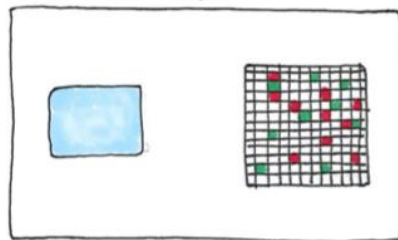
### FOCUS



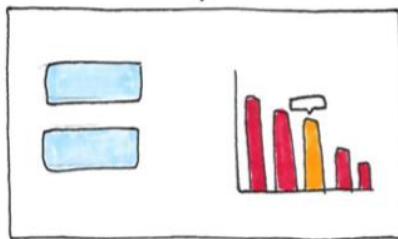
An introduction page with a brief explanation



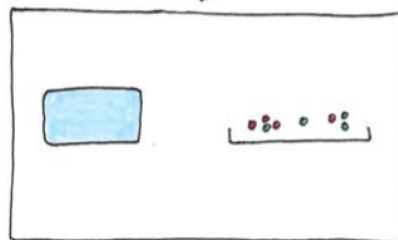
OPTIONAL



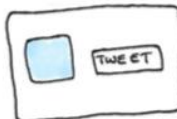
Split into two sections:  
1. A grid with squares representing all 5,000 tweets.  
2. Same grid but squares coloured based on sentiment.



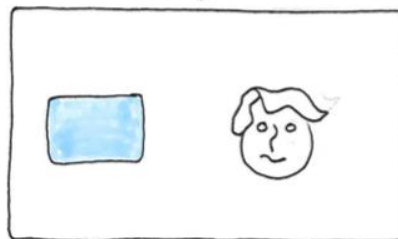
Total counts for each sentiment type. Chart has mouse over interaction - changes colour of bar and shows value.



How these sentiments correspond to the amount of 'favourites', shown in the form of a beeswarm.



OPTIONAL



An outro page with a quick summary.

\* = narrative text

TITLE: DATA VIZ SCROLLER

AUTHOR: ISOBEL

DATE: 30/5/19 SHEET: 5

### OPERATION

- o This visualisation relies on a 'scroll as a trigger' functionality where the act of scrolling through the page starts the next visualisation or piece of information.
- o The act of scrolling to a part of the page indicates the users readiness to ingest the next piece of content.
- o Within the inner sections, some visualisations will have additional interactivity e.g. mouse over to bring up more data. - these are detailed next to each page.

### DETAILS

- o The data sources are Donald Trump's Twitter and approval ratings.
- o The implementation platform is D3.js.
- o The estimated time to complete this visualisation is approx. one week.
- o Data sets may need some additional pre-processing to make suitable.
- o Colour scheme is based on the 'Twitter blue' + different hues of blue.