Assignment 3

Group Report

Wanting Chen, Alexandra Feix & Mark Williams

Context of our project

In the run-up to the Brexit referendum a curious thing happened. Instead of the issue being played out through party politics, the vote instead divided the Conservative Party in two, with the Labour leader Jeremy Corbyn being almost absent from the whole ordeal.

In the vacuum left by Corbyn's absence the two main players leading the opposing sides of the debate were the Prime Minister David Cameron, and one of his closest allies in parliament; Boris Johnson. The pair had both been to Eton together and it seemed as though their old schoolyard rivalry was being re-awoken.

When the results of the vote came in, many were left in shock. Boris and the Vote Leave side had been victorious, and David Cameron was swift to resign the next day. But, unlike the predictions, Boris didn't immediately snatch the job up. Instead there was a period of chaos where nobody knew what was going to happen. Culminating in Theresa May becoming the Prime Minister and Boris Johnson becoming the new Foreign Secretary in her cabinet.

Audience

We wanted to market our design towards the average web-news reader. Many of the UK's big newspapers such as; The Guardian, Independent, and The Times have large online presences, with interactive features becoming more and more prevalent as more of their readership views their content on tablets where the technology and now the expectation is there to make moving & interactive visualisations.

We wanted our visualisation to be a full screen interactive application which would act like a two-page spread in a digital newspaper, for our audience to click through and choose which data set they wanted to view.

Data Cleaning

Our data holder presented us with a dataset containing every BBC news podcast from the years 2016 and 2017. The data was given to us in the form of JSON files with information related to each word per speaker per podcast. There was more data such as time of start and end, word confidence, etc, which we ignored for our purposes.

One could argue that the word confidence should have taken a much more significant role in our process but we found that the confidence value was a very unreliable measure of how accurately the word was being predicted. Many words were being given a 1.00 confidence (the maximum value) but were obviously being transcribed incorrectly. These errors became clear as soon as we attempted to read some of the text. After globbing all relevant files together and creating a timeframe with the information, we could now focus on getting the right words.

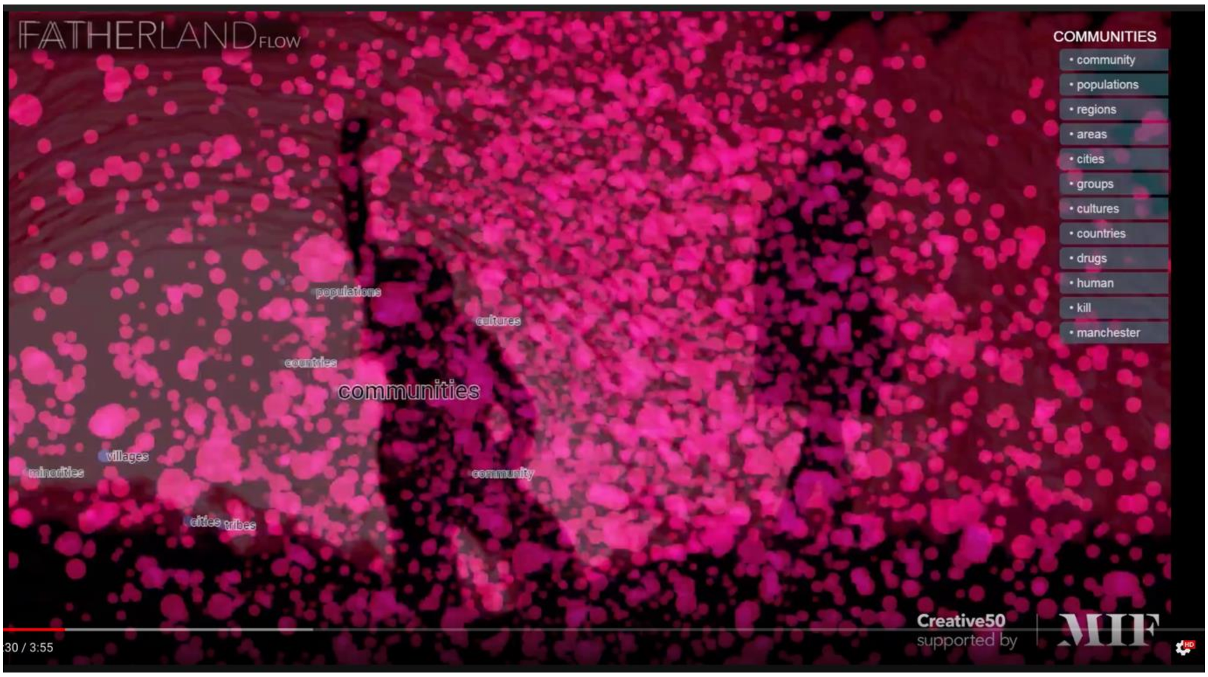
In order to filter only those words which are associated with a certain person, we used an iteration to search for a combination of words in the data which could be interpreted as Boris Johnson and David Cameron. Since the data was based on spoken words, we had to define a range of 30 words before and after their name in order to get a basic representation of the associated words.

The speech recognition software used by our data holder turned out to be outdated and couldn't identify words like "Brexit" which were only coined during the political debates of the past years. Therefore, we had to scan the data for representations of these words and the change them accordingly. Combinations we found were e.g. "brecht's" + "it" or "breaks" + "it".

For filtering (unnecessary) words, we made use of the natural language toolkit (nltk) for python. We labelled all the words we have previously filtered and selected only those words which had a label indicating they were either nouns, adjectives or verbs. In a final step, we had to manually create a blacklist of words such as "it", "he", "she", etc. in order to display only the most meaningful ones.

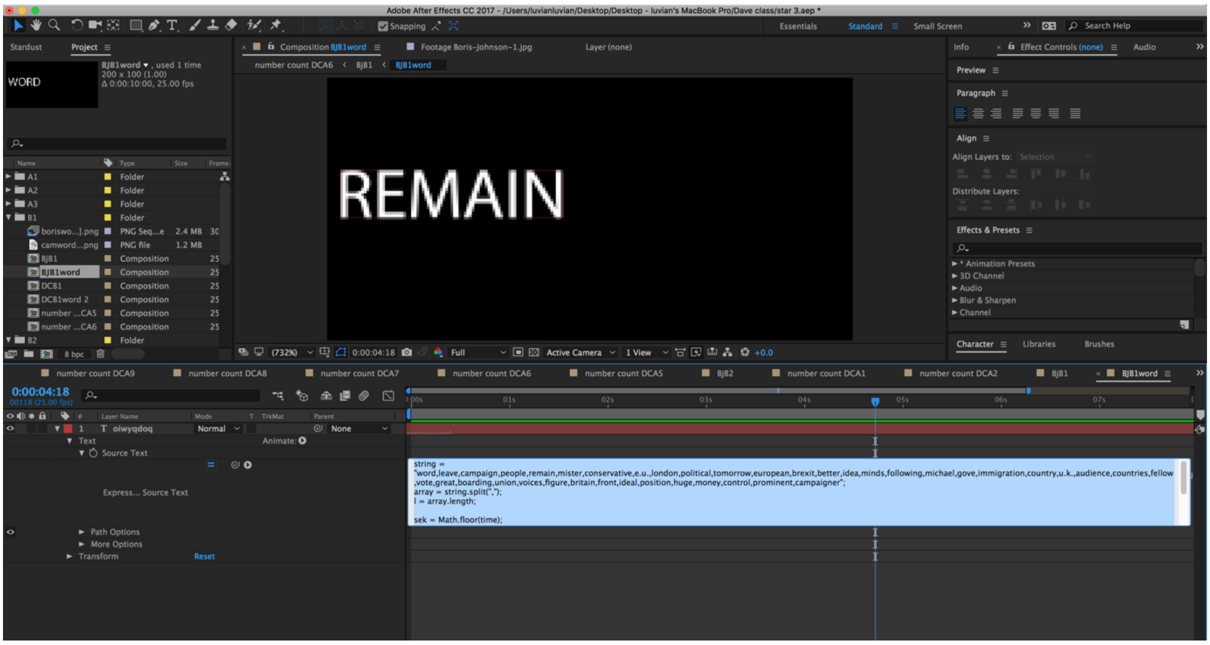
Making our Visualisation

Before we started to work on our visualization, we first discussed about what we wanted to show to our audience. Since we focused on the comparison of two politicians, we decided to use world cloud to make it easier for the audience to read the information. We got the inspiration for a video from an example we saw on Youtube and we really wanted to emulate that style.



In the sketch, we had two persons surrounded by word and beneath them is a time line to enable the users to interact with the data. In order to make the visualization look more futuristic and modern, we used a stardust tool to create a universe which consisted of words. At first, we wanted to make the words flow through the air and surround the politicians.

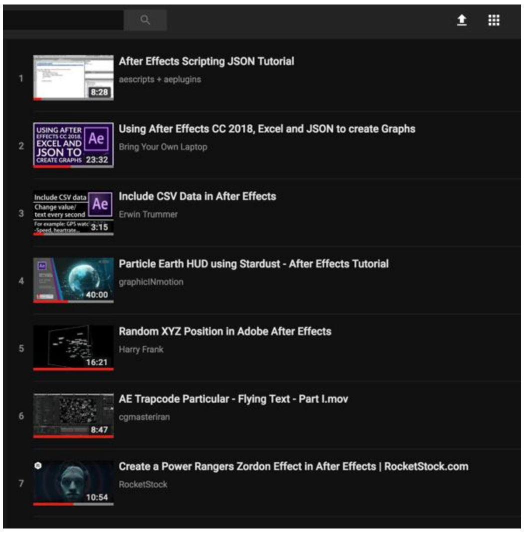
The visualization we created just looks superficially like a word cloud in a 3D space. Unfortunately, it wasn’t produced using any code inputs, so there was nowhere to put our data. We hadn’t done any 3D effects with AE before, so we searched for tutorial on Youtube to help us.


Adobe After Effects CC 2017 - /Users/luvianluvian/Desktop/Desktop - luvian's MacBook Pro/Dave class/star 3.aep *
REMAIN
0:00:04:18

```
string =
"word,leave,campaign,people,remain,mister,conservative,e.u.,london,political,tomorrow,european,brexit,better,idea,minds,following,michael,gove,immigration,country,u.k.,audience,countries,fellow
,vote,great,boarding,union,voices,figure,britain,front,ideal,position,huge,money,control,prominent,campaigner";
array = string.split(",");
l = array.length;

sek = Math.floor(time);
```

These are all the most important sources that we used to make the visualization.



1 After Effects Scripting JSON Tutorial
aescripts + aeplugins
8:28

2 Using After Effects CC 2018, Excel and JSON to create Graphs
Bring Your Own Laptop
23:32

3 Include CSV Data in After Effects
Erwin Trummer
3:15

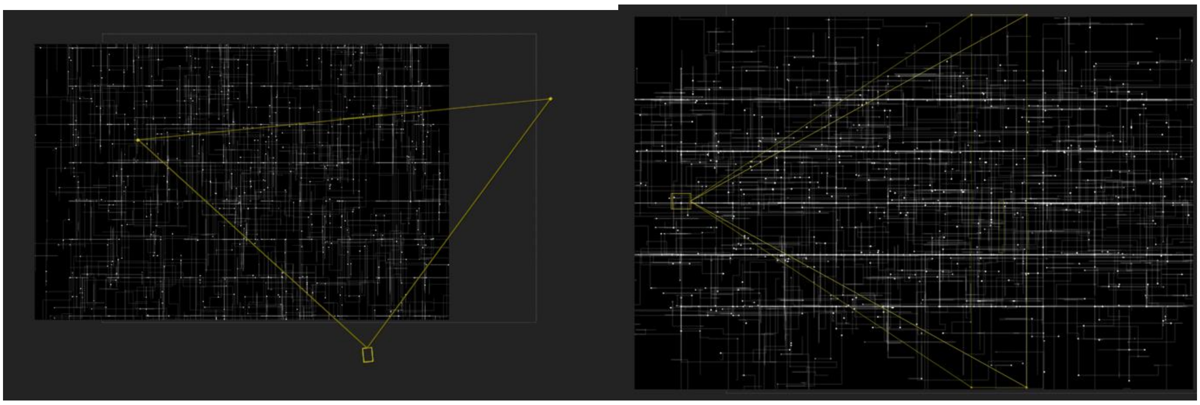
4 Particle Earth HUD using Stardust - After Effects Tutorial
graphicINmotion
40:00

5 Random XYZ Position in Adobe After Effects
Harry Frank
16:21

6 AE Trapcode Particular - Flying Text - Part I.mov
cgmasteriran
8:47

7 Create a Power Rangers Zordon Effect in After Effects | RocketStock.com
RocketStock
10:54

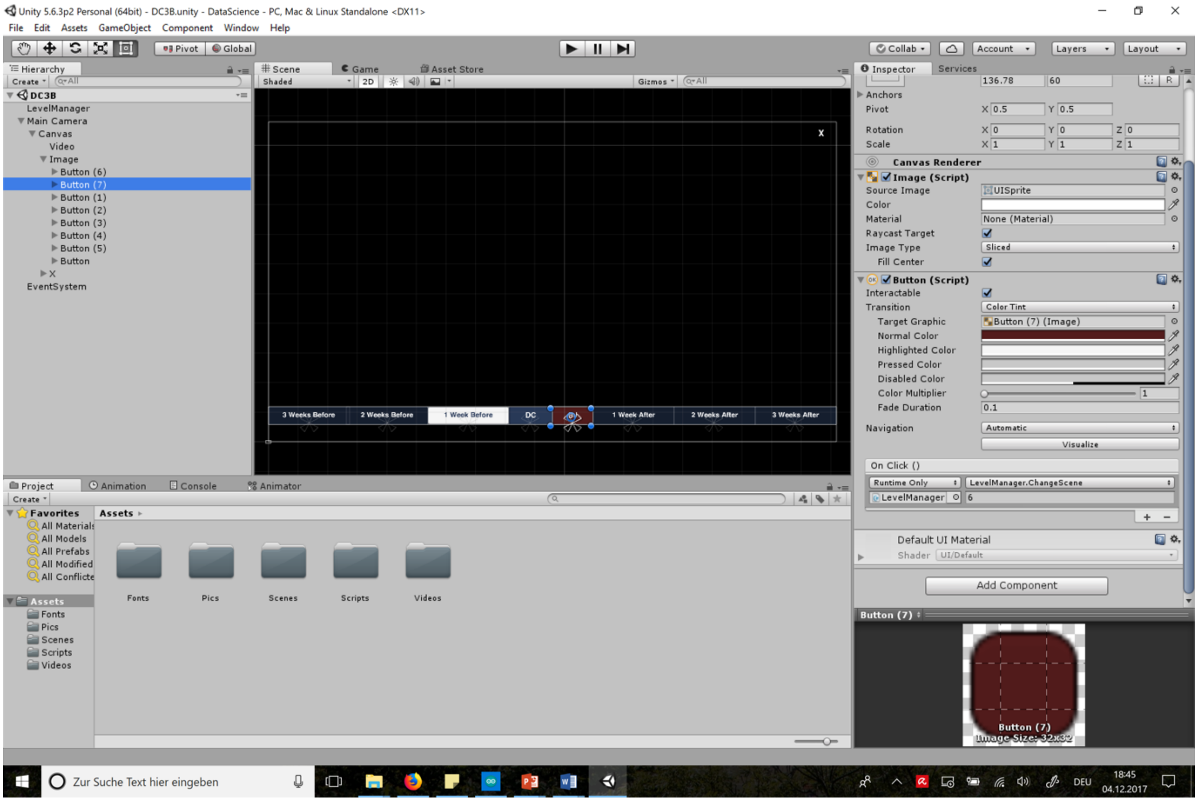
So after creating the word cube, we noticed that the cube is actually a layer. A big problem for us as there’s no way to put something in the middle, such as our two politicians! So we discussed about what to do and decided to keep the effect and use other ways to show the data count of the top words. We then chose to use the zone in and out tools to show each word as a workaround.



We output 6 videos for both David Cameron and Boris Johnson, each of them have 3 videos showing 3 weeks before the referendum and 3 weeks after the referendum.

After creating the videos, we needed a way to put them together in an interactive application. We used Unity 3D in order to offer an application which can be run from basically anywhere. Another reason is the fact that our group members weren't really knowledgeable in JavaScript, so we had to come up with another solution.

Since Unity 5 introduced a new video Game object, we could create a scene with each video and play it instantly on loading. We created a task line with buttons to load each scene, so the user could choose the word cloud for a specific week. The buttons had a On Click () function and a script attached to them which enabled the loading of a specific scene when pressed by the user.



Unity 5.6.3p2 Personal (64bit) - DC3B.unity - DataScience - PC, Mac & Linux Standalone <DX11>

```
DC3B
  LevelManager
  Main Camera
    Canvas
      Video
      Image
        Button (6)
        Button (7)
        Button (1)
        Button (2)
        Button (3)
        Button (4)
        Button (5)
        Button
      X
  EventSystem
```

```
136.78    60
Anchors
Pivot             X 0.5   Y 0.5
Rotation          X 0     Y 0     Z 0
Scale             X 1     Y 1     Z 1
Canvas Renderer
Image (Script)
Source Image      UISprite
Color
Material          None (Material)
Raycast Target    ☑
Image Type        Sliced
  Fill Center     ☑
Button (Script)
Interactable      ☑
Transition        Color Tint
  Target Graphic  Button (7) (Image)
  Normal Color
  Highlighted Color
  Pressed Color
  Disabled Color
  Color Multiplier  1
  Fade Duration   0.1
Navigation        Automatic
On Click ()
Runtime Only      LevelManager.ChangeScene
LevelManager      6
Default UI Material
Shader            UI/Default
```

Button (7)
Image Size: 32x32

Conclusion

In summary, our project was hampered by many compromises. Our lack of ability to make the visualisations more interactive was down to a skills gap in our team. But considering the difficulties we had, out outcome was still something that we are very proud of. We learnt many new skills to unpack all of the data from Catherine, and had to overcome many challenges both in the coding of our python notebooks and in the production of our interactive visualisation.