INTERACTIVE MEDIA DESIGN

Continuous Assessment 1 - Tableau

Daniel Buckley

X00109141

Contents

Summary

This is my document for my continuous assessment for Interactive Media Design. This document will outline how I came up with my dashboard visualisation and the steps in how I arrived at them. This continuous assessment was elapsed over three weeks. In my opinion, I felt the hardest part of the continuous assessment was gaining the data I wanted to use. I searched for a few hours on multiple occasions to find the datasets that suited me. Finally, I found one set at http://www.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=MUM01 . This gave me all the data I needed for multiple months in a half decent format which helped me to clean it and use it for myself to use in Tableau. In this project, I wanted to show the correlation between unemployment and the recession which occurred in 2008. I was trying to show the unemployment epidemic happening in Ireland at this current moment in time compared to before and after the Celtic Tiger era which impacted this country greatly. I wanted to cross reference this with data which I gained from the Road Safety Authority's website. I wanted to see if there was any links with the rise or fall of unemployment on the issuing of the number of penalty points.

Background

I became interested in the topic of unemployment when I was in secondary school a few years ago. I had an English teacher who would talk about current situations in the country and the recession seemed to come up a lot in discussion. At first I was not so interested but towards the end of my time in secondary school I began to see and feel the impact of the recession with family members and family friends losing jobs and homes first hand. I realised that everything my teacher had discussed with us in class was coming through. At first it was slowly, but month after month it gathered momentum.

This is what gave me the idea for this project. One aim of this project is to see just how long after events such as recessions and crashes are the effects felt by citizens within Ireland. I would like to examine this data along with data on the issuing of penalty points in the Republic of Ireland and see if there are any links between the pair.

I hope that this will be different as I will be using this report and my visualisations to convey a powerful message to the reader on events regarding unemployment and penalty point issuing. From looking at my visualisations I want the user to have a good grasp on the message Ii am trying to put across on these particular topics.

Unemployment Data

Acquire

My research began looking at worldwide unemployment rate datasets which I accessed through some sites such as Eurostats.com and data.worldbank.org. From doing research on this I realised that I was finding data on such a large scale and I thought completing the project on something closer to home would be of much more benefit to the public which I had a stronger connection with. This is the reason why I chose my home nation, Ireland. I felt the terrible inflictions of the recession from 2008 first hand, as did several of my friends and family. This is the reason why I thought starting my research in this direction would be more interesting and of more benefit to me as I would see trends which I have experienced first-hand rather than the world.

I continued my research looking for datasets on the topic of unemployment within Ireland. At first I found a considerable number of datasets. I looked at these sets and thought that they could be used for comparisons more so down the road but they were not an exact match that I was wanting at that time. I saved these datasets and kept moving on towards the relevant data that I needed to show my message.

After days of frustration looking for the one right dataset I stumbled across the Central Statistics Office website and data portal. I searched through all categories getting datasets which were of no use to me and my goals that I set out at the start which I was determined to prove. The dataset was of the unemployment rate in Ireland from the year 1998 – 2016, month by month, and by genders. This was exactly what I was looking for. It included male and female information also which I will use to create some great visualisations for easy comparisons.

However, reviewing the data within the excel spreadsheet which the Central Statistic had put together was a mess. Trying to make sense of it was hard at first but from doing this a few times I gained a better grasp and understanding of the data that was on hand to me. Each time I learned more and more about the characteristics of the dataset that I had found. However, I still had found no trend by looking at the excel spreadsheet.

Parsing

At this point in time I was happy with the dataset I had gained. The next stage I needed to do was to clean it and get it ready to be imported into Tableau. The dataset was a mess and it would take a lot of time and effort to clean this data manually. To combat against this massive time consumption totally manual process I decided to write a script in C#. This script took about an hour to perfect and I enjoyed doing it at the same time. In the dataset, the month and year was in one string value e.g. suppose we had January 2016, this was shown as "2016M01" which was not very useful to us in this format. My first step was to split these into a more usable format for Tableau. I decided to split the string into month and year columns in my new excel spreadsheet. When I ran the script, it produced an output for the months and years 225 times which was the exact number of times I needed them. Next all I had to do is to copy this data into the excel spreadsheet. I had no problems importing this, the columns were inserted perfect. These scripts can be found on my GitHub repository at https://github.com/DanoBuck/InteractiveMediaDesignCA1/tree/master/ParseDataScript .

One problem with parsing the data was trying to complete the data with age groups all along the spreadsheet. If I dragged the cell downwards it would go down in jumps of years. This was not what I wanted. I wrote this extra piece of functionality into my script to produce the age group I needed and how many times it was needed. Next I copied this output from the console and it inserted with no problem.

A very important piece of data was the last part to go into the spreadsheet. This was the number of Irish citizens who were unemployed. If I was to copy this information from the acquire phase to the parse phase it would be inserted into one cell as a long string value. This is not the expected outcome I was hoping for. I needed to prevent this from happening as all my previous work would have been a waste of time and I would have to enter the data one cell at a time. Luckily I found a solution. I opened a Microsoft word document, copied the text I needed, and converted it into a table. It was a simple approach for a big win on this project. It saved me an enormous amount of time if I had to go another route as in doing it manually. I repeated this approach for another row in the acquire stage. My dataset was now taking the shape I wanted and could move to the next stage.

Filter

This stage involved me removing data I found irrelevant for my data visualisations which will come. In this dataset, I only had to remove one column. This was the percentage of people currently unemployed. I thought I would allow Tableau to calculate the percentages for me as part of my data mining step.

Data Mining

This step was carried out within Tableau to show some more complex tasks within the software. I looked at certain data by adding extra table calculations such as the average and sums to name a few options. This allowed me to see the entire sum of people unemployed within any given month.

Represent

This stage involves me visualising the data which I have cleaned up and imported into Tableau for further analytic work. This was the most interesting part of the project as I seen trends that I did not notice before I started inserting the data into charts.

My first visualisation is something which opened my eyes to this problem. It is a horizontal bar chart with the number of unemployed people on the X-Axis and years on the Y-Axis. From this I can see a lot of detail already from a simple chart. From investigating this chart, I see that there was a massive rise in unemployment from 2008 – 2015. This began getting increasingly worse at the year of 2011 where there was a big jump in the unemployment rate. From glancing at the chart, I see that this year was the worst year in the data I have for unemployment. I would like to investigate this more and see what other than the recession could have effected this big change. I have shown the years on the Y-Axis as 2016 at the top and working its way down in order of year. I done this as it did not make sense to me to show the year with the highest unemployment number at the top followed by the next and so on. The visualisation can be seen below:
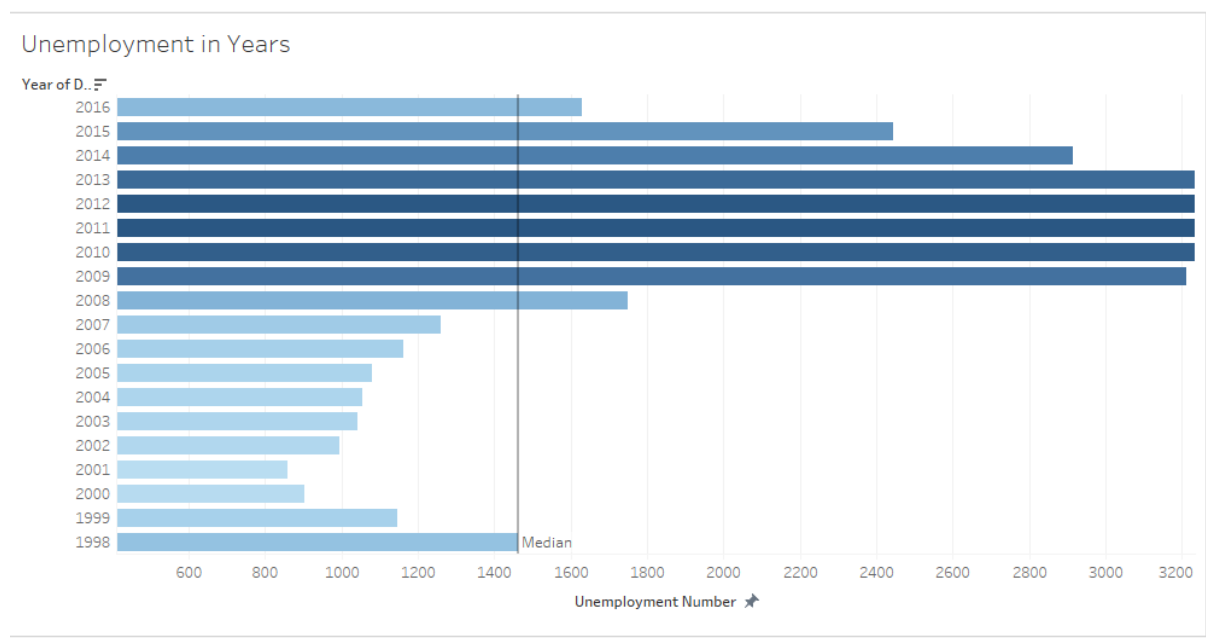
My second visualisation revealed another unknown pattern to me which I had no prior knowledge about. From looking at the chart I created I can see that there is a big difference in the age groups and genders. In saying this, I can reveal that there is not a big gap between 15 – 24-year-old males and females in relation to unemployment. However, if we look at the next group and gender i.e. the 25 – 74 - year-olds there is a great difference in unemployment. It seems to me that males have a higher rate of unemployment than their female counter parts. This is an odd pattern and I would like to know why there is a greater number of males unemployed than women. Another great feature of Tableau is the use of being able to forecast predictions for the coming years. I predicted the next three years of unemployment and on each occasion, it is showing that unemployment is rising towards levels like that of 2014. This visualisation can be seen below:

Forecast of Unemployment
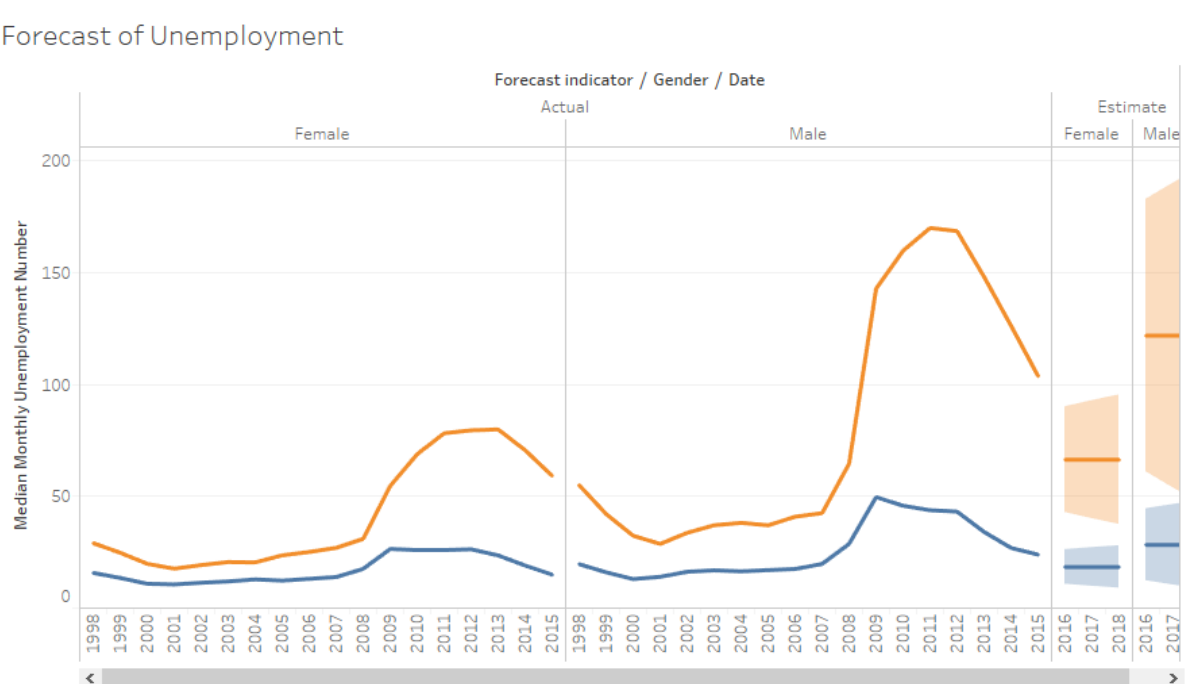
My third visualisation created on this dataset is a simple table. This simple table shows the number of unemployment for age group and gender. It also adds up everything and shows it in a column underneath the corresponding information. This table can be seen below:

Unemployment by Gender and Age Group

| Age Group	Year of ..	Female	Male
15 - 24 years	1998	201	247
	1999	162	195
	2000	133	152
	2001	123	169
	2002	139	198
	2003	143	204
	2004	153	197
	2005	149	201
	2006	160	209
	2007	166	238
	2008	205	358
	2009	308	588
	2010	329	553
	2011	315	523
	2012	313	505
	2013	286	398
	2014	229	334
	2015	182	291
	2016	122	189
25 - 74 years	1998	360	654
	1999	292	498

Refine

To improve my first visualisation I have added in more colour. The colour code is low unemployment is light blue as the lower that rate is the better and dark blue for higher unemployment as that is bad. There are also different shades within the charts to show the transition between lowness and highness with regards to unemployment. The user can also click on whatever bar they want and they are provided with extra information about the bar they have chosen.

To improve the second visualisation I have added labels and used the mean average to show the line graph. The visualisation is divided on genders and age groups allowing the user to compare easily. The X-Axis is over the years while the Y-Axis is the average number of unemployment for that year.

To improve my third visualisation I added in an extra field to show the total number of unemployment for that row giving the user all the numbers that they need for comparisons.

The Dashboard

I then inserted these visualisations into my new dashboard. The dashboard consisted of 3 separate visualisations. One was the bar chart I have mentioned above, another was a line graph, and the final one was a table. The line graph showed the before and after effects of the financial crash that hit in 2008. This dashboard can be seen below:
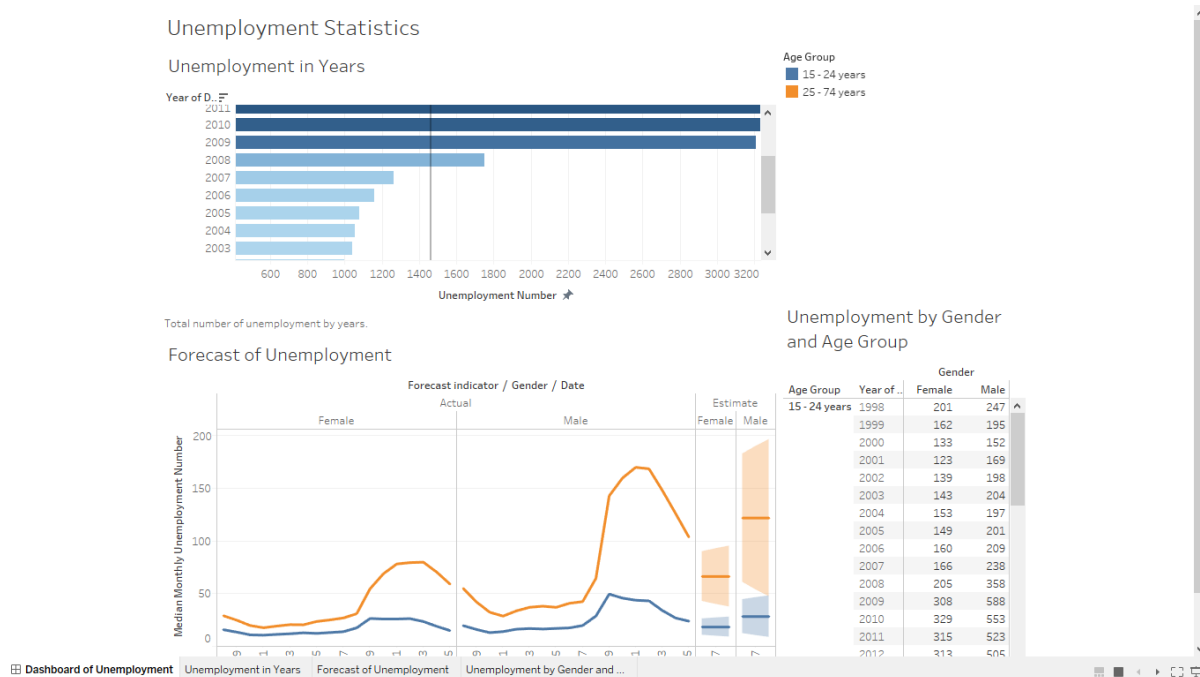
The dashboard is very interactive as the user can click on the age group filter and everything will filter on the dashboard. This allows the user to gather all the relevant data that they need. The user can also select a year on the bar chart and the view will filter accordingly to their selection. The last statement is correct to the table also.

Interact

In visualisation 1, the user can click on the bar which they want to see more information about. Once clicked a tool tip is presented to them This shows some more information about this bar to the user.

To manipulate the data in visualisation 2, the user is given a menu on the right-hand side of the screen. On this side of the screen the user can click on the group of values which ever they wish to view. By doing this, the user is shown the corresponding values only. This gives the user some more power over the visualisation as they can see whatever they wish to see.

The dashboard which I developed was interactive. All charts were linked up to one another. If something was clicked in one chart this was seen the other 2. If a user clicked a specific year, the table would highlight that row and give the user to compare the data with ease. The user could also use multiple years for comparisons. This is a nice feature of using Tableau. This has also been mentioned above.

Problems and solutions

At the start of the project, getting the data into the format that would allow me import it into Tableau provided a great challenge to me. The solution to this problem was writing a script in C# which would output the chosen text into the format that I needed to be able to import the excel spreadsheet into Tableau. This was splitting out the months and the years into two separate string values.

Likewise, with the numeric values. When I was copying the long string values into the parse stage from the acquire stage I had problems getting them to fill the columns which corresponded to their values. When I would paste them over they would go into one very long cell which is not what I wanted. The solution to this problem was by copying the chosen text into a Microsoft word document and converting that text into a table. This allowed me then to copy the whole table into the excel spreadsheet the way I needed it to be.

Some of my visualisations did not work in conveying the message I wish to show to the users. One chart which completely failed on me was the line chart. This chart is below:
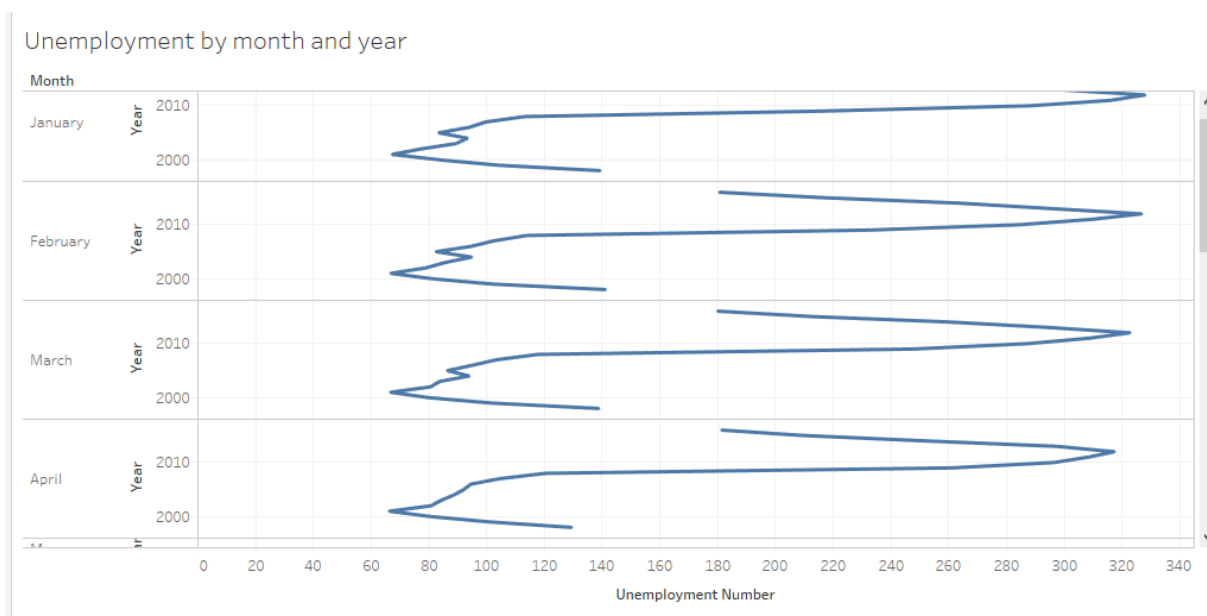
There is a lot of information on this chart which would confuse the user. The year starts at 2000 on the Y-Axis and as the axis goes up the year increases. This is not visible to the user. There are also 12 separate charts on this which is a lot of information to the user.

When creating my visualisations in this I realised that I could not complete them correctly with the month and year format I decided to go with. I needed to change this to give me more selection over the data. I wrote a function in the C# script which allowed me to convert this easily. Again, I copied and pasted the output into the excel spreadsheet resulting in a lot of saved time.

Penalty Point Data

Acquire

To acquire data regarding the issuing of penalty points within Ireland I first searched on the Central Statistics Office website with no luck. I then navigated through Google until I found a suitable site for me to the gain the data I needed. Soon I landed on the Road Safety Authority's website. I found a vast amount of data regarding the subject on this website. However, it was not in a useable format for me and sometimes there was data was missing for a few months. The dataset I acquired listed all offence types on the road and the total penalty points that had been issued for that offence by counties.

The dataset did not come in one perfect format going back years. Instead I had to go month by month downloading the excel data file that had been put together over that month. I then had to copy this data and insert it into my new excel file which I would use for parsing, filtering, and data mining later. I repeated this tedious process until I had a worthy dataset which I could apply the seven stages on. The dataset had around 70 columns and over 2000 rows. It was my biggest dataset but it was cut down for the import to Tableau.

Parse

The dataset was in such bad shape it made cleaning it a tremendously hard job for me. This took a matter of days to complete. Luckily, I took measures to reduce the manual work needed to clean this dataset. I wrote another script in C# to help me with the cleaning of this data. I failed on 3 separate occasions to get my data parsing script to work properly. My first few attempts were very frustrating. I could get the script to output text in a format I almost needed, however, this was not good enough and still would require a lengthy manual process of cleaning it up some more as the script would not do all. These scripts were starting to feel like lost causes. I contemplated completing the full process manually, but I knew this would take many more hours, if not days. If I done this I would had to cut down the dataset by a vast amount. Luckily it did not come to this.

At this stage I was quite fed up and was getting close to calling it quits and doing it fully manually but I gave it one more go. The big problem with the data was that it was formatted horizontally instead of vertically. Tableau needs it in this format to work to the best of the tools ability. Fortunately, this time I was successful and the script ran almost perfectly after several more failed attempts. The script would then produce a new csv file which was in the format I needed for this stage. The only job I needed to do was to copy it back into the excel data file. Unfortunately, I only got the script to work on a single month's data and I had to remove all the data such as the date, county, and offence, only leaving the actual numeric data. This meant I had to run the script 29 separate times, each time copying and pasting the produced data into my the excel file I was using. This proved very tedious but not as much as it would have been if I had of done it the way as mentioned above. This was only part 1 on the parsing stage.

Part 2 consisted of writing another function in the script to remove hyphens within the datafile produced previously. Compared to the above challenge this was one I was very optimistic about getting to work. I got it to work and the hyphens were removed from the data and replaced with "0"'s as they were number columns needed for processing. This completed part 2 of the parsing stage.

Next I proceeded onto part 3, which turned out to be the final parsing stage. In this stage I did not require any scripts to help with the parsing of data. Instead I used tools inbuilt in excel. The tool was the "Find and Select" tool in the Insert tab. I needed this to replace all the blank cells left in the dataset. Previously I had tried to work around this with a script but because I was processing the file as a csv, whenever I tried to remove empty cells, the output file was heavily affected and would not import into the excel file the right way. This is what lead me to using the inbuilt tools. This tool allowed me to highlight the empty cells and then enter a value of "0" in all of them. It completed this in a matter of seconds without any side effects to the data in question.

Filter

When filtering the data, I noticed there was a lot of overlap in the header information. I knew that this data would be useless to me when I imported it into Tableau. After identifying the columns that Ii did not need, I highlighted these. Some columns though I had the option of merging them as they were all sub categories but this will be explained in the data mining stage. To decide what could be merged I colour coded the columns, green and yellow meant they were merged into one. However, the presence of the colour red meant that the column was removed from the data file. The reason I removed this was that I see it as irrelevant.

Data Mining

A good bit of the data mining in this data set was done in the excel spreadsheet. As mentioned above I wanted to merge sub category columns into one. I did this by using the SUM function and dragging the bottom right hand corner down to the end of the spreadsheet. This populated the columns with the sum of all the merged columns. Exactly what I needed and it didn't take too much time to complete. This reduced the column number from 67 to just over 20. A perfect number of columns which would allow for great visualisations and comparisons which are spoken about later in this document. This was four of the seven stages completed using C# and excel. Now it was time to move onto the implementation of Tableau.

While carrying out this section I realised that there was a small problem in my dataset. This was due to the way the date was formatted. I split out the date into two separate columns i.e. one column for month and another for the year. This was overkill for Tableau. I went back to my script and wrote a function which output the date in a useable format e.g. the format 1/1/2001 was perfect, this is what I made the script output and I copied this output into the excel spreadsheet.

Represent

My first visualisation of this dataset was of the whole island of Ireland. In this dataset, I showed all the counties and the percentage of penalty points present within the county. I created a calculated field in Tableau which allowed me to compare the percentages of the amount of issued penalty points between all counties. The county with a high percentage of penalty points was shaded a dark blue. As the number of penalty points decreased in counties, the shade of blue faded. If the percentage got larger, then the colour of the county turned a darker blue. From this visualisation I was shocked to learn that Waterford was the county which had the second highest percentage of penalty points within Ireland. I would have suspected that Cork would have had the largest percentage due to the large city population. I was totally wrong and I would not have known this if it was not for the visualisation I created. Another reason I enjoyed working with Tableau. This visualisation can be seen below:

Percentage of Penalty Points in the Republic of Ireland
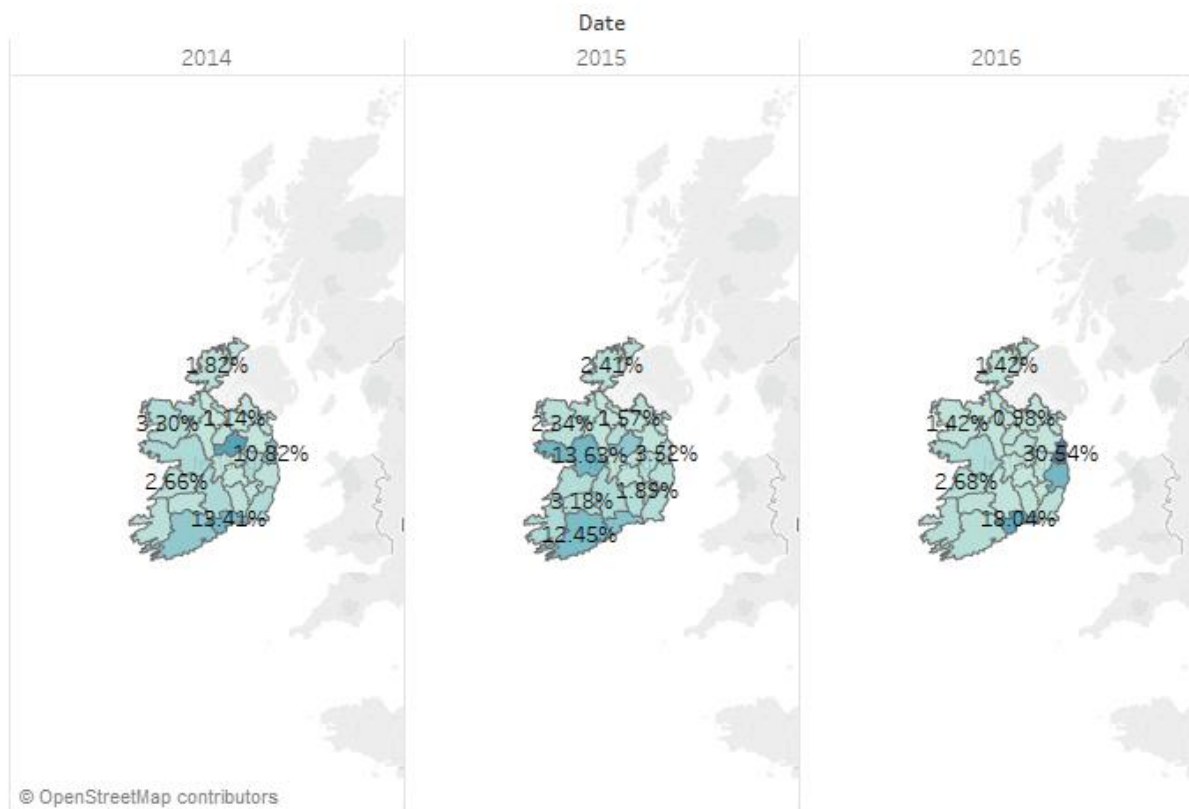
The next visualisation I completed was in as a bar chart. This visualisation portrayed the total number of penalty points for any given offence in the dataset. I added grid lines to this to allow the user to locate the amount with ease. I also gave them something to compare against. This was a bar across the grid which represented the average of penalty points. This can be seen below:
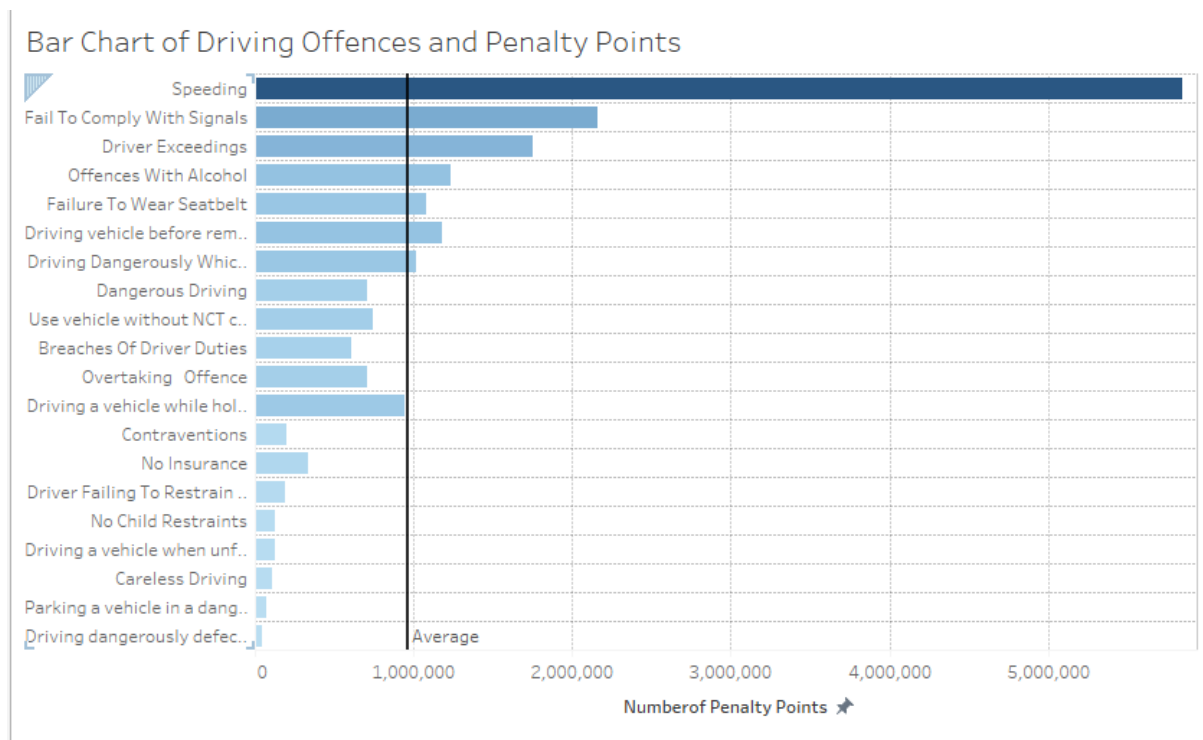
Another visualisation I created was of a simple table. This table showed all offence types and the sum of penalty points for the years in the dataset. After the years, I then showed the grand total for the offences being viewed. This can be seen below:

Table of Totals for Offences

	Date			
	2014	2015	2016	Grand Total
Speeding	1,456,757	17,316	4,375,597	5,849,670
Fail To Comply With Signa..	769,329	1,135,462	257,328	2,162,119
Driver Exceedings	594,246	700,770	457,219	1,752,235
Offences With Alcohol	588,280	589,649	59,081	1,237,010
Failure To Wear Seatbelt	154,665	601,221	320,692	1,076,578
Driving vehicle before re..	2,687	964,943	213,006	1,180,636
Driving Dangerously Whic..	567,861	153,892	293,243	1,014,996
Dangerous Driving	59,823	606,452	42,345	708,620
Use vehicle without NCT c..	427,196	269,481	47,876	744,553
Breaches Of Driver Duties	580,179	12,641	11,024	603,844
Overtaking Offence	42,362	20,561	645,197	708,120
Driving a vehicle while hol..	339,922	5,941	593,851	939,714
Contraventions	31,921	56,069	115,409	203,399
No Insurance	101,866	109,766	128,409	340,041
Driver Failing To Restrain ..	50,145	128,558	13,349	192,052
No Child Restraints	15,958	105,467	7,617	129,042
Driving a vehicle when un..	8,130	111,077	7,197	126,404
Careless Driving	46,232	27,651	36,757	110,640
Parking a vehicle in a dan..	41,062	34,538	1,019	76,619
Driving dangerously defec..	7,351	18,801	21,492	47,644

I then went on to creating another bar chart visualisation. However, this time I included all the counties and years in the dataset for this visualisation. I believe this gave it more value to the user. This visualisation allows for comparisons very easily. The visualisation is below:
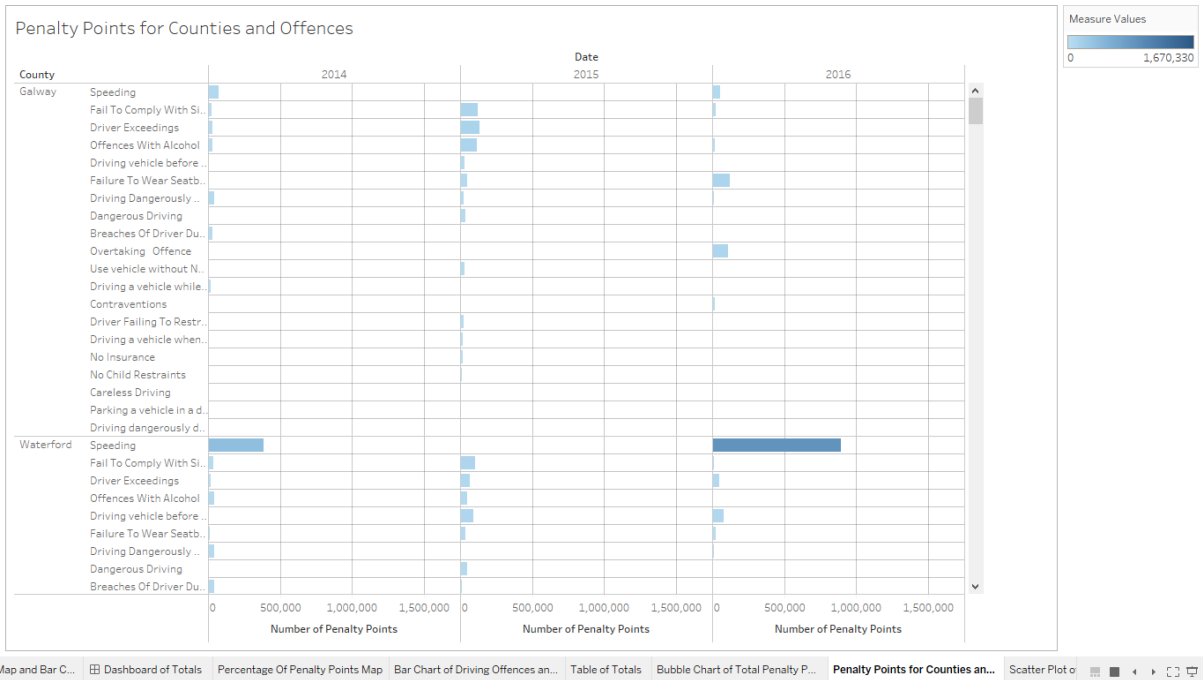
My last visualisation on this dataset was of a scatter plot showing the correlation between Speeding and Dangerous Driving in every county, for every year in the dataset. This showed the relationship between the two offences and can be seen below:

Scatter Plot of Speeding and Careless Driving
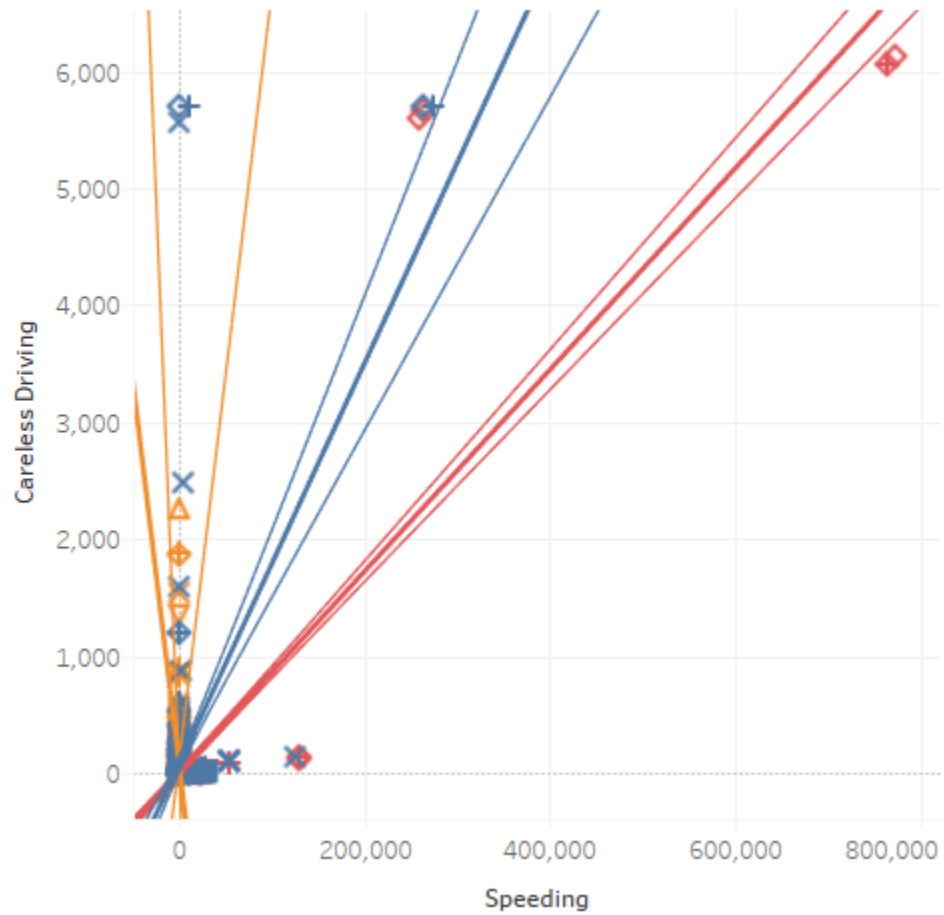
Refine

To refine everything, I inserted my visualisations into a new dashboard. In this dashboard, I inserted my map visualisation alongside a bar chart which showed the number of penalty points per county and offence. I added in cool colours which appeal to the user and do not give a misperception to them. This is done by using different shades of blue. The darker the blue in the map and the bar chart, the higher the penalty points. In the map the penalty points are out of the total percentage. In the bar chart, the penalty points are represented as full figures for that county and offence type. This gives the user the ability to quickly glance at it and receive the vital information they require at that moment in time. This dashboard can be seen below:
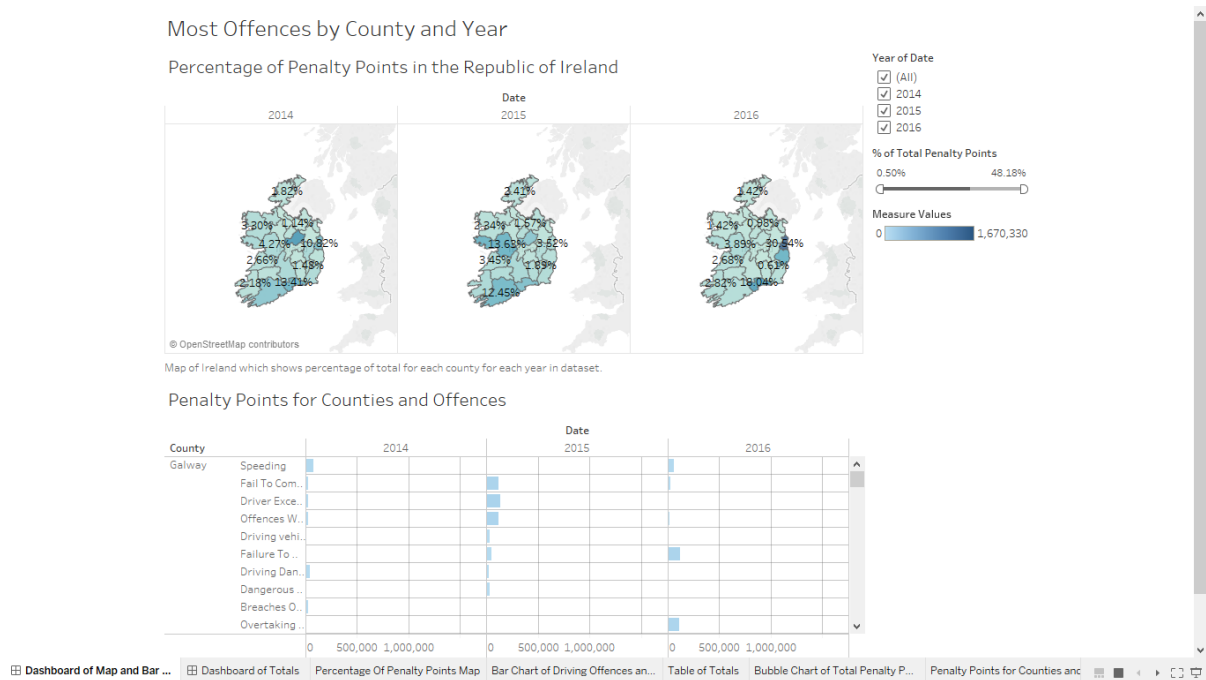
I made another dashboard. This time it was a dashboard of totals concerning penalty points. The dashboard has a bar chart on it with all the offence types on the Y-Axis and the number of penalty points on the X-Axis. This information sums up all the penalty points over the years in the data. Underneath this visualisation, there is two separate visualisations. The one on the left of the screen is a table of the total number of penalty points split by year with a grand total. To the right of this, there is a scatter plot which relates to Speeding and Careless Driving.
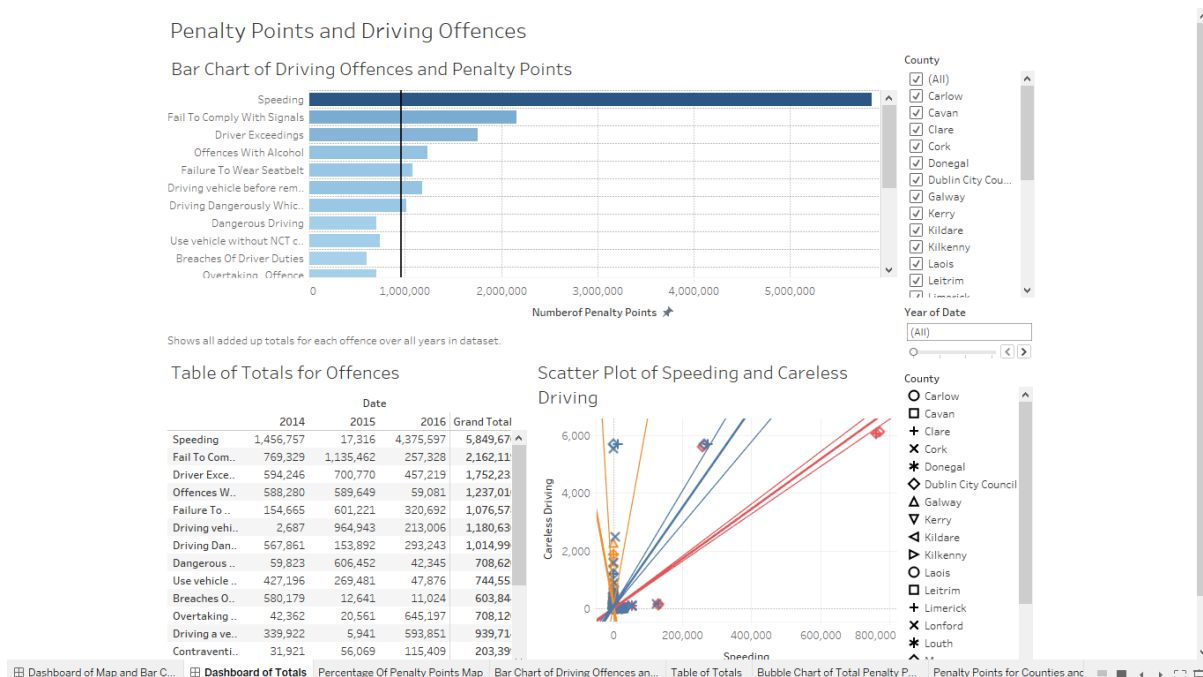
Interact

Interaction is a big part of the first dashboard. The user can interact in three separate ways with the map. They can choose to hover on a county and they will be shown more information about that county i.e. the percentage of the total of penalty points that county possesses. This is the same on the bar chart except it shows the total number of penalty points rather than the percentage. This is a great feature for filtering through the map and bar chart. My favourite part of this dashboard is the interaction between the two visualisations. If the user clicks on a county, the bar chart automatically adjusts to these criteria. The user can also select as many counties as they wish to compare, the bar chart automatically adheres to these criteria again. The user can also select and compare different years. The user chooses a year on the right-hand side and can pick a county and the bar chart will adjust to that criteria.

In the second dashboard, I linked up all the visualisations. These visualisations proved tricky at first to link up, but again it got easier before the end. If the user clicked on one of the attributes in the top bar chart, the system would then highlight the corresponding row in the table beneath. This allowed the user to view the raw data if necessary. In this table, if the user clicked on a specific year, these changes would filter into the above bar chart. These changes also were seen across in the scatter plot. The scatter plot would show the relevant year along with all its attributes in the visualisation.

Problems and solutions

When attempting my first visualisation, I noticed that I could not make it very interactive. This was due to the fact I had the month and year in separate columns at first. I needed to change this so my map could tell the user a story about the evolution of penalty points throughout the years. I went back into my datasheet and identified the problem. It would have been a lot of manual work if I changed the dates into formats I needed. I decided to write a function in my script which would produce this. This was very helpful and was complete in a matter of minutes. I copied the data that the function produced and pasted it into the columns corresponding to the date. I was now able to make my first visualisation much more interactive after completing this step.

Another problem I faced was that Tableau did not recognise some Irish counties in my dataset. This was down to the fact that some of them were entered in as County Councils or City Councils. This was not a big problem. I decided to leave my dataset the way it was and to manually insert the counties in Tableau by choosing from a drop-down list.

Linking up separate visualisations proved difficult at the start, but it began to get easier. At the start, I realised the two visualisations had to be using the same measure or dimension to link together. This took me a while to figure out. I searched the internet and could not find much help so eventually I realised what the error was saying and checked the two charts. This allowed me to get to the bottom of it. In my penalty point dashboard, I implemented one "Action" which carried out the linking mentioned previously.

One visualisation I implemented on this dataset was a bubble chart. After developing this chart, I realised that it was no use and did not show the message I was trying to convey. It was a good way of showing the difference between counties but even saying that is a push. The visualisation can be found below. A better way of doing this would have been by a line graph or even a bar chart.
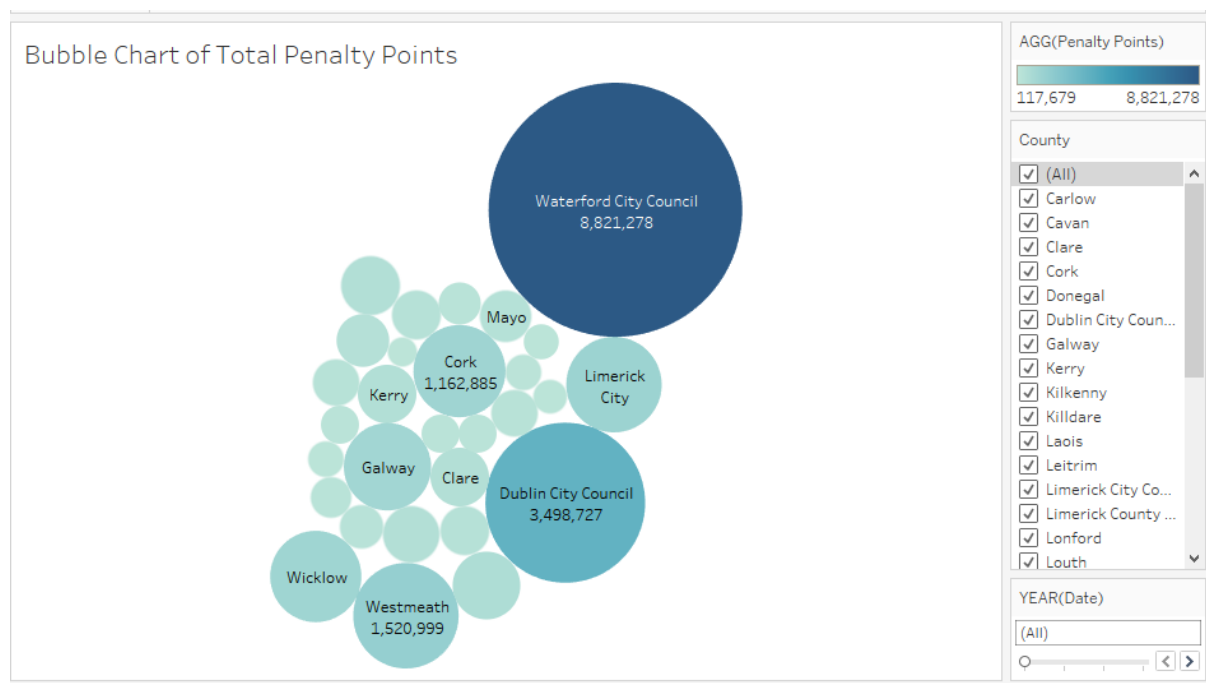
Bringing it all together

In this step, I decided to merge the two datasets i.e. the unemployment dataset with the Penalty Point dataset and then import it into Tableau. The first visualisation I created was two line charts. One was for the number of unemployed persons within the Republic of Ireland and the other was for the amount of penalty points received for driving with no insurance. I chose this example as I know insuring a car is expensive and if a person is not working it would make paying this quite difficult. At first I tried to visualise the two selections on the same chart as a line graph. This was unsuccessful. The problem with doing this was the vast number of differences. The unemployment dataset is in thousands while the penalty point dataset is in hundreds of thousands. This was giving me two Y-Axis on the chart. It was even confusing to me. The result can be seen below.
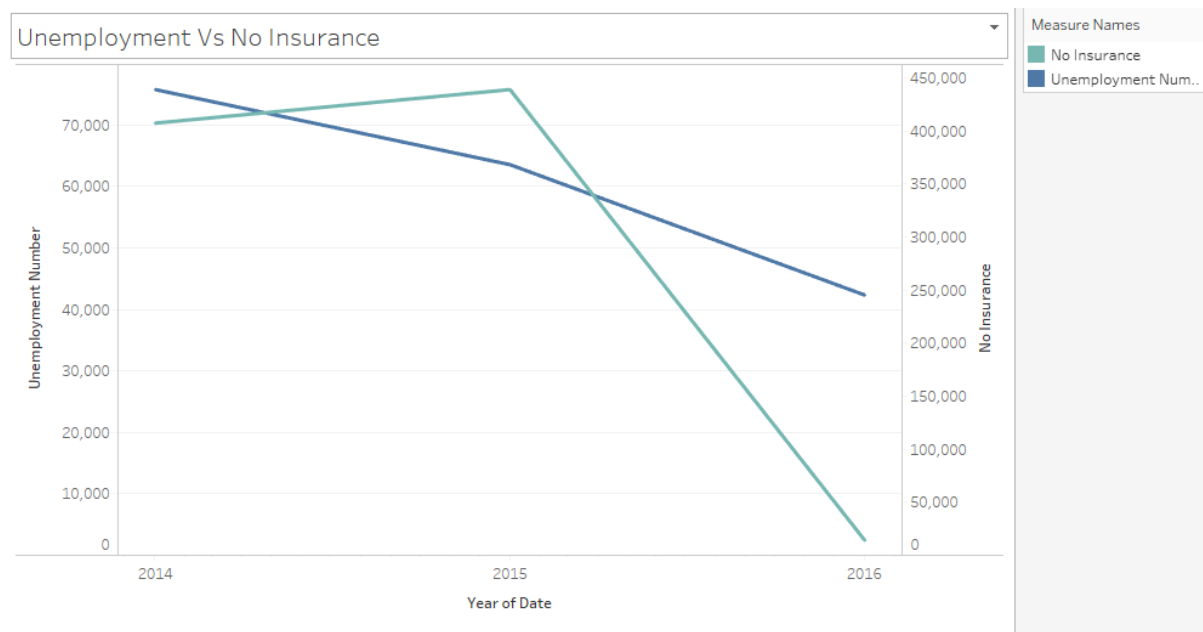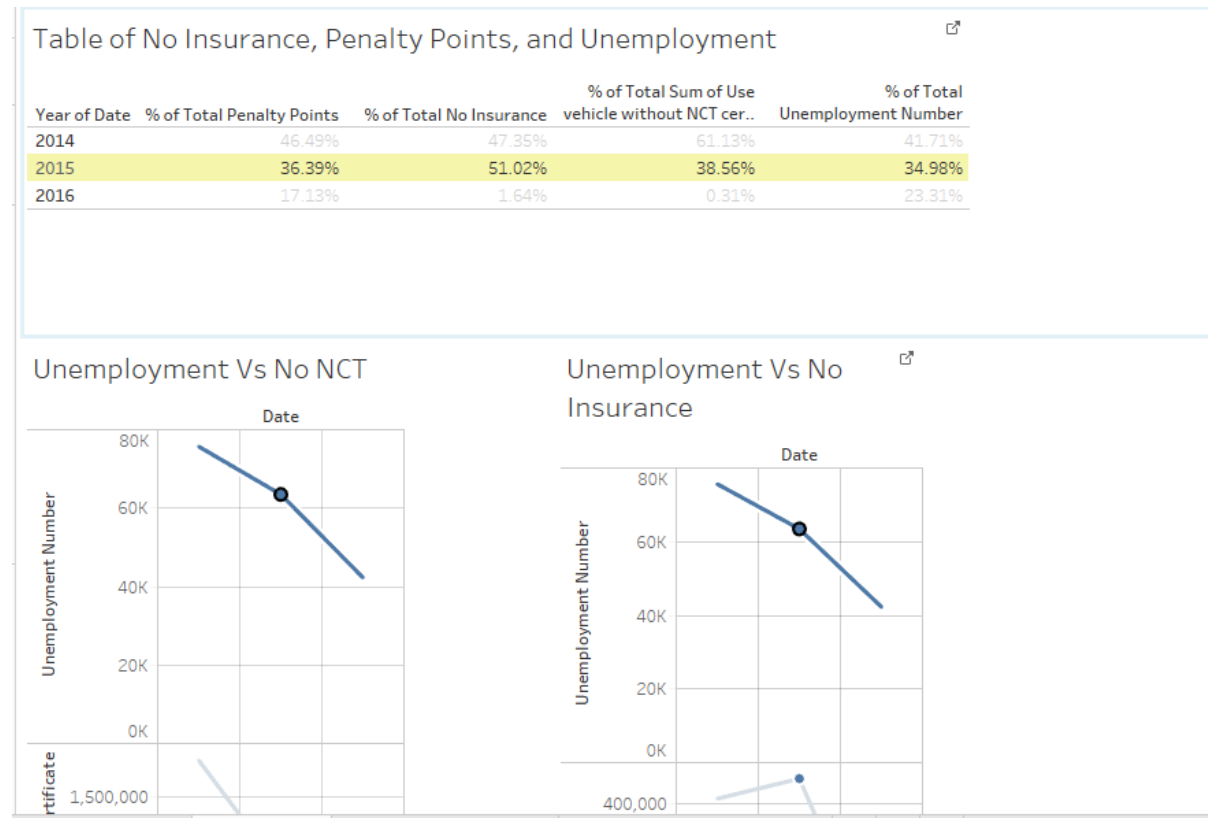
I scrapped this idea and went with the option of using separate line graphs to convey my message. This worked better for me. I felt the message was much clearer than the above example. This can be seen below.

Unemployment Vs No Insurance
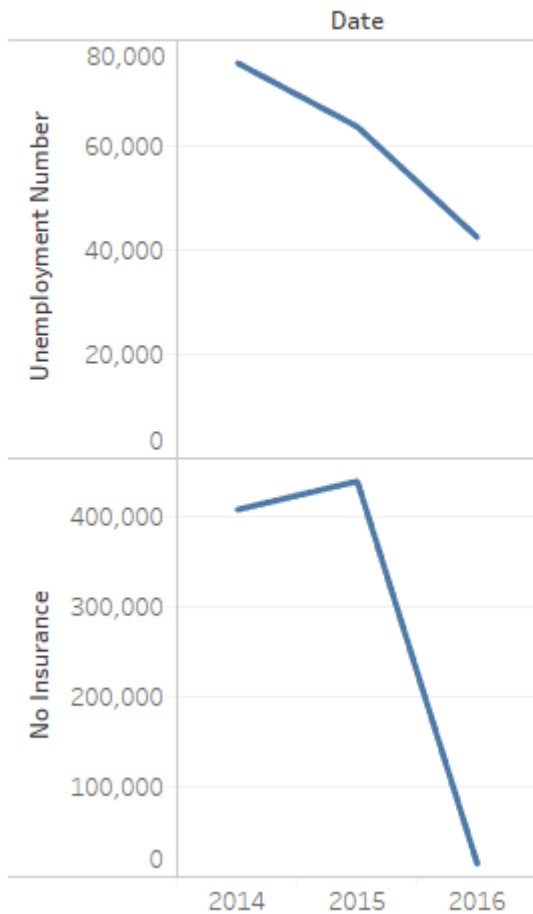
From creating this visualisation, it revealed that the number of drivers driving without insurance plummets when the unemployment rate drops, i.e. people going into employment. This was an interesting find.

From this discovery, I pursued investigating my data to see if there were any other hidden patterns. Another pattern I found was that when the unemployment rate is falling, so does the amount of penalty points being given out for driving a car without an NCT certificate. My completed dashboard can be found below.

Table of No Insurance, Penalty Points, and Unemployment

Year of Date	% of Total Penalty Points	% of Total No Insurance	% of Total Sum of Use vehicle without NCT cer..	% of Total Unemployment Number
2014	46.49%	47.35%	61.13%	41.71%
2015	36.39%	51.02%	38.56%	34.98%
2016	17.13%	1.64%	0.31%	23.31%

Unemployment Vs No NCT

Unemployment Vs No Insurance

Conclusion

From completing on this project, I realised that is extremely difficult to compare two different datasets like I have tried to do in this project. However, comparisons internally between the datasets was much more manageable and beneficial. I have learned a lot over the course or this project. The main thing being perceptions can be very wrong. By saying this, I refer to the Penalty Point dataset. From visualising this data, I discovered that Dublin was the worst offending county overall although it was not for one of the years which surprised me. Before completing this, I would have thought that the likes of Galway, Cork, or Donegal would have been next but no I was wrong. The second worst county was Waterford, and after that Westmeath. This was a shock to me as the populations here are smaller than what they are in Cork or Galway City.

Another thing I learned while investigating my dataset is that the issuing of penalty points in Dublin is increasing year after year except for a single. Dublin are way ahead of other counties when it comes to the rise in penalty points. Westmeath in 2014 had the highest rate of penalty points but in the following two years it successfully reduced this vastly. Dublin dipped in 2015 dramatically. This did not last for long and it was back to the highest in the country the following year.

With regards to the Unemployment dataset I also learned a lot. The most important of which is that within the Republic of Ireland, the highest group of unemployment exists in the age group of 25 -74 years old males. But this figure is dropping. It peaked to high levels in 2011 but it is now steadily dropping all though I have predicted a small rise in this using Tableau.

I learned a lot in this project about data analytics and just how important it is to identify certain trends in societies as well as businesses. I enjoyed acquiring my datasets and parsing them through some C# scripts although it took some time to get some what working. Tableau is an amazing tool for business intelligence and it performs superb. I look forward to continuing my studies with Tableau.

What makes my work stand out is the interactivity of the dashboards that I have created. I have found work like my own. However, it was not interactive at all and only showed a single visualisation. This was very limited beforehand and my project builds on top of that.

If I had more time to investigate these subjects I have chosen I would like to gather bigger datasets in these fields. The unemployment dataset was big but limited. It only showed the groups of unemployment within the Republic of Ireland and the number that exists for that group. It also broke down the groups by genders. I would like to know more about these groups and genders. I would like to find more data about their employment before their unemployment. This would allow me to know what employment sectors are worse hit by financial crashes.

The penalty point dataset was the largest one I used. However, it was limited to 3 years. I would like to gather all penalty points issued for each offence in the Republic of Ireland going back to the creation of the penalty point system. This would give my data a better meaning and allow for more accurate forecasting over the next few years.

In the Project Upload

In my upload, you will find several folders. These are:

- DataSets
  - PenaltyPointsData – Where the seven stages excel spreadsheet exists.
  - UnemploymentDatt – Same as above.
- Document – This contains the document I wrote alongside the development of the project.
- ParseDataScript – Where my C# scripts live.
- Visualisations
  - BringingTheDataTogetherVisualisations – This is where I made my attempt at a dashboard by merging two datasets together and doing some analysis on them.
  - PenaltyPointsVisualisations – Where my penalty point visualisations live alongside the imported spreadsheet.
  - UnemploymentVisualisations – Where my unemployment visualisations live alongside the imported spreadsheet.

Published Dashboards

I have published my dashboards to Tableau Public. These can be found at the following URL's:

- https://public.tableau.com/profile/daniel.buckley#!/vizhome/PenaltyPointsMapVisualisations3/DashboardofMapandBarChart
- https://public.tableau.com/profile/daniel.buckley#!/vizhome/PenaltyPointsVisualisations3/DashboardofTotals
- https://public.tableau.com/profile/daniel.buckley#!/vizhome/UnemploymentStatisticsDashboard/DashboardofUnemployment
- https://public.tableau.com/profile/daniel.buckley#!/vizhome/ComparingDataSets2/DashboardofUnemploymentandPenaltyPoints