Data Visualization Coursework:

1. Research Question(s):

Retail has become a huge part of our lives but apart from the everyday shopping trip and occasional news article, by choice we don't really choose to research Retail information unless it is part of our day job. The main questions I am trying to answer using Data Visualization are:

- What is the current state of the retail market?
- How are other retail sectors broken down? What are they?
- How have the top companies performed both now and in the past?
- What other retail opportunities are there outside of the UK
- Who are the biggest retailers in the UK and how are they performing?

These are just some of the questions that I feel are valid and useful questions in which the answers could be both insightful and useful for us.

2. Data Source(s):

I have retrieved data from various sources ranging from online open data sources, specialist subscription areas and magazines. I have categorised each set of data into its on section stating where the data was retrieved.

- Grocer 33 Data
    - Availability – This data was manually scraped from the premium area of the Grocer website. The way in which I gained access was through my employer who has a subscription. I manually entered the data myself into an excel spreadsheet and saved it as a CSV. This covered the last 20 weeks
    - Cheapest Basket Data - This data was manually scraped from the premium area of the Grocer website. The way in which I gained access was through my employer who has a subscription. I manually entered the data myself into an excel spreadsheet and saved it as a CSV. This covered the last 40 weeks
    - Top 100 Brands – This data was manually entered from the March issue of the Grocer magazine. This involved manually entering the data into excel and saving to a CSV.
    - Top 150 Suppliers – This data was given to me by my Director who had scanned in a page from a previous issue of the Grocer magazine and used Excel to clean it up.

- 2013 GRDI Data - http://www.atkearney.com/documents/10192/1315829/Global+Retailers-+Cautiously+Aggressive+or+Aggressively+Cautious.pdf/b8f528f4-cb6f-411d-8ae6-7afb7fbc7b7e (Page 4)

- UK Channels - http://www.igd.com/our-expertise/Retail/retail-outlook/3371/UK-Grocery-Retailing/

- Retail Price Index (ONS) - http://www.ons.gov.uk/ons/rel/rsi/retail-sales/february-2014/rft-rsi-poundsdata-february-2014.xls

3. User Instructions:

Brief instructions on how to use your data visualization. You may include screen shots if you wish.

This section will briefly describe how to use my data visualization. My visualization has been built to be used as more of a tool, an application of visualizations collectively giving insights. The basic way in which you navigate is by either using the TAB button, clicking on the tab you wish to go to or pressing number 0-4. I will briefly describe what each page includes:

Welcome Page – This is a basic page welcomes you to the tool and tells you how to navigate

Summary Page – This page is an overview of data. The first graph you'll see is the retail sales comparing overall retail sales with fuel sales. This is calculated in 2 graphs across the past year and also 2003-2013. Below that is a quick snapshot of how the big 5 retailers are doing in terms of availability of stock for the past 20 weeks. The table on the far right are the top 100 brands for 2014.

GRDI and Retail Sales Index – The first 2 pie charts show the division of retail sales by both value and volume. You can hover over a pie segment to see the actual value in a tooltip. The map shows you the values of different GRDI metrics across the top 30 countries, you can click on the other metrics at the top to switch between them. Below that is the table of figures

Brand and Supplier Analysis – Both sides of the page contain tables and range change metrics across 2 years for suppliers and brands. The graph in the middle visualizes the sales for the top 100 brand starting with number 1, hovering over a bar conceals its value and brand in a tooltip. The two pie charts are presented to give an idea of how UK retail channels are split up and also out of the top 150 suppliers, how they are broken down in terms of ownership type, final the two box plots show the distribution of turnover and profit for the suppliers.

Grocer 33 Analysis – the "Arc Pie" as I call it on the left shows the number of times each retailer has been the cheapest for the past 40 weeks, you are able to hover over the dot to get the value in a tooltip. The other metrics related to the last 40 weeks are shown in a pentagon square graph, you can switch between the different metrics using the buttons.

4. Design Justification:

4.1 Colour:

Colour was extremely important to me when doing my data visualizations as I know it can be very easy to stick so many different colours into a visualization to try and make it look pretty but it ends up looking a mess. I have generally chosen a consistently calm colour palette consisting of whites, greys and blues for the most part. Where appropriate I have chosen brighter colours. Examples of this are demonstrated in any graph related to the Top 5 retailers. Looking through each retailers logo and assigning the colours in my visualization to the correct retailer (i.e. Asda = Green, Tesco = Red, Sainsbury's = Orange etc). I have use interpolated colour scales where appropriate as well where I have mapped values to certain values. Associating with retailer values I stuck with light to dark signaling

low to high. This is also demonstrated in the mapping where for risk, a darker red is used to show high risk and light to show low. Most of my graphs use the blue, black, grey palette. Using different shades of blue, including interaction. I feel this helps keep everything consistent but also stand out at the same time as not everything is the same colour or same shade. Colour has also been used effectively to indicate active elements such as tabs or current data viewed (as shown in Grocer 33 Analysis). The only place where generic colour not associated with brand of layout is used if for the rank change tables where red bars show a decrease in rank whereas green shows an increase.

4.2 Layout:

I thought very carefully in terms of laying out my data visualizations. I was able to do this with the help of my containers class which enabled me to lay each element out neatly with a consistent title and colour scheme. The layout technique you could say I used was a kind of grid, each page is filled with containers. The top part of my tool has the navigation which is consistently in the same place whichever page is visited. Each graph, table and pie chart is laid out in a similar fashion again to keep consistency across the tool. It is important that everything is easily viewed and related elements are kept close together (i.e. page 3 the top 100 brands table and sales chart were next to each other on the same page. I felt it was a good idea to lay everything out in such a way that each page tells its own story and is grouped in a nice way.

4.3 Symbolisation:

Symobolisation is used a bit in my visualizations, for example on the summary page where each retailer is symbolized by a coloured circle related to their own logo colour. Also each graph is shown with a consistent bar type, each element shown using a blue bar. Where necessary symbolic legends have been used and in the case of the Grocer 33 analysis Pie Arc graph a small circle symbolizes the point at which the arc stops. Also the varying bars in the rank change graphs are another representation of how I have symbolized values (in this case by a bar that varies in length). I have chosen these symbols to give the user an instant insight into what they're looking at.

4.4 Interaction:

There are numerous interactions I have implemented as part of my data visualization. I will go though each one by page:

Common across all – The navigation pane which allows users to click and see which tab is highlighted (current page viewed). Users can also navigate using the TAB button on the keyboard as well as the number keys (0-4).

Summary Page – You are able to hover over of the bars of each graph which change colour when the mouse is over it and also a tooltip is shown with the value associated with that particular bar. You can navigate the table by clicking the up and down button on the page, or if you have a scroll mouse, if you are within the element you can scroll up and down.
GRDI and Retail Sales Index – Each pie segment when hovered over a certain point expands to signify that it is the one being focused on. A tooltip with the value associated with that segment is shown as well. This is valid for both pie charts. The map element

allows you to hover over a country and view the name with the help of the tooltip. You are also able to view the different metrics by click on the other metrics on the title bar of the map element. You can navigate the table by clicking the up and down button on the page, or if you have a scroll mouse, if you are within the element you can scroll up and down. There is also a help tooltip that you can view by hovering over the question mark in the table which will create a help box over that element describing what that element is.
Brand and Supplier Analysis – You can navigate both tables by clicking the up and down button on the page, or if you have a scroll mouse, if you are within the elements you can scroll up and down as well. The graph allows you to hover over an element and view the sales value for that bar using a tooltip with the brand name and sales value.
Grocer 33 Analysis – You can hover over the end points of each arc (the little circles) and view the amount of times each retailer was the cheapest. In the Pentagon square graph you are able to click on the button associated with a metric and view the updated values with the selected metric box highlighted as well.

5. Further Work:

I think I could have implemented interactive filtering if I had more data that could be joined. This would mean a person could click on a graph and update lets say a pie chart showing different metrics relating to the same subject. Also zooming on the map would have been useful, although its easy enough to see where each country is and with the table you can see which countries to click on. A zoom option I feel would have been a nice addition. There was significant slowdown on the GRDI page due to the map, I'm not sure why but the effects the interactions and causes a slight delay, if I could find a way of increasing the performance it would be better. If I had more time I would have implemented more help buttons for each graph. Perhaps explaining what each graph is trying to show.