COS30045 DATA VISUALISATION

DATA VISUALISATION PROJECT GUIDE HINTS AND TIPS

SWINBURNE UNIVERSITY OF TECHNOLOGY

DATA VIS PROJECT

SUBMISSION GUIDELINES

Weight: 70%

Due Date:

- check canvas

Items:

- Data Visualisation Process Book as a PDF (45%)
    - Design (25%)
    - Link to live Website (appears on title page of Process Book) (20%)
- Code (format TBA)
- Project Reflection (20% Individual PDFl)
- Stand-ups (5% team/individual Oral/PDF/Text)

This is an overview of requirements for the Data Visualisation Project.

In this overview we will go over the sections of the Project and show examples of past student work.

DATA VIS PROJECT

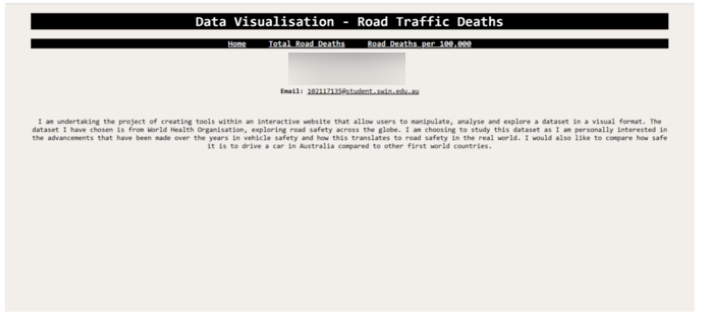
INTENDED LEARNING OUTCOMES

1. Critically evaluate data visualisations and propose improvements based on an understanding of human perception and cognition and data visualisation design principles.
2. Apply a structured design process to create effective visualisations.
3. Conceptualise and iterate data visualisation designs using sketching and low fidelity prototyping techniques.
4. Create web-based interactive data visualisations using a real-world data set.

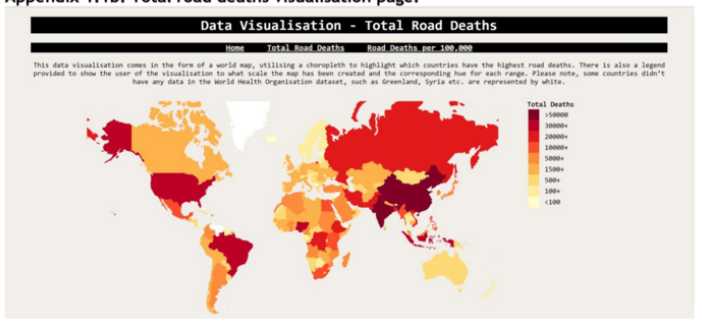
DATA VIS PROJECT

- Standup (5 team/ individual)
- Design Process Book (25 team)
- Website (20 team)
- Reflection (20 individual)

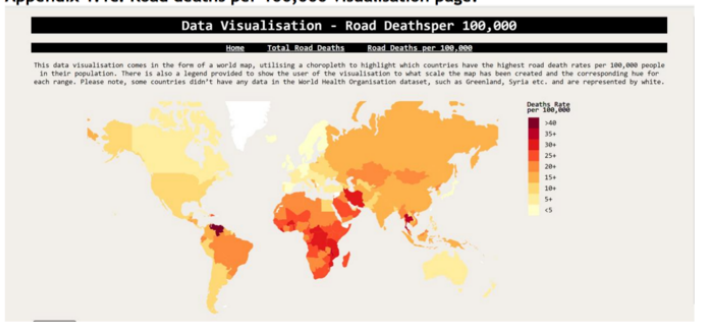
Appendix 1.1a: Index page:

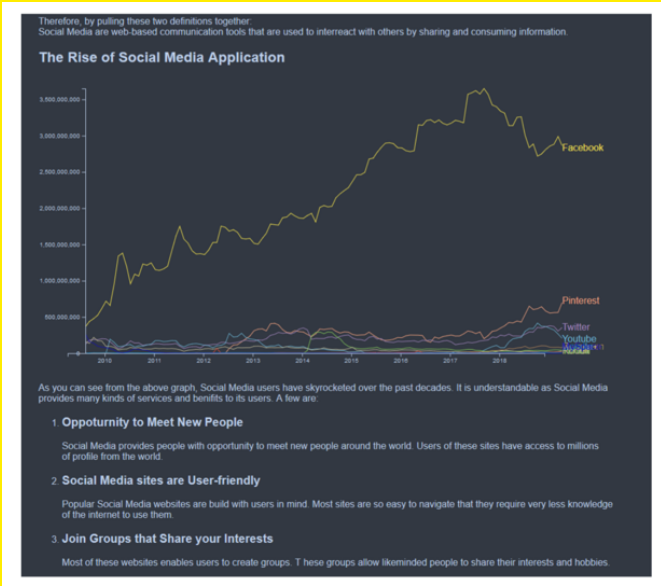


Appendix 1.1b: Total road deaths visualisation page:



Appendix 1.1c: Road deaths per 100,000 visualisation page:





VISUALISATION PROJECT TOPIC - SUSTAINABILITY: ENVIRONMENT IN CRISIS

https://www.theguardian.com/world/gallery/2013/sep/10/bushfires-sydney-suburbs-in-pictures

Swinburne is committed to helping individuals and society work towards a more sustainable future. In this Project you will be designing and building a visualisation to help us realise this ambition. In particular we would like you to focus on visualisations of recent and/or ongoing environmental issues. Your visualisation might be aimed at any one or more of the following:

- helping people understand the size and/or impact of adverse environmental events
- helping people understand their personal impact on the environment
- helping people minimise their contribution to environmental issues

Must be in topic domain!

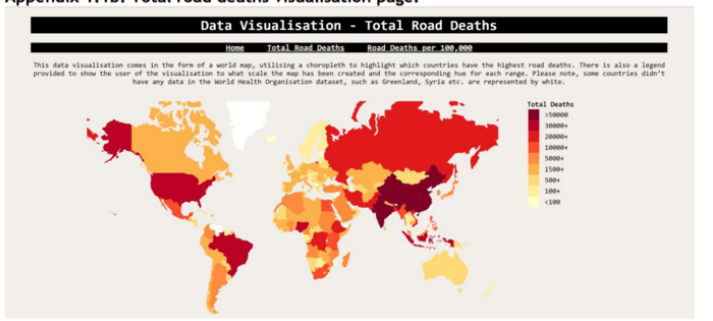
SWINBURNE UNIVERSITY OF TECHNOLOGY

PROCESS BOOK (25%)

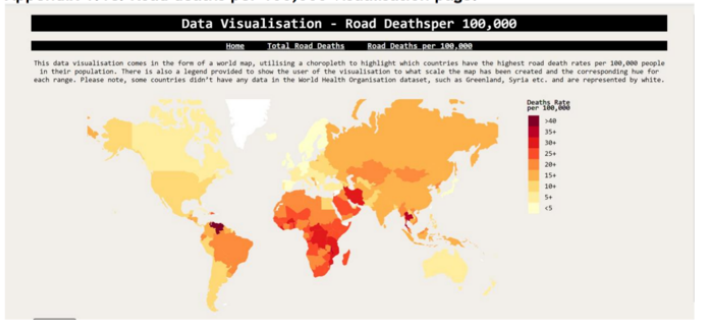
- Design 1: Context and requirements (3)
- Design 2: Data Elements (2)
- Design 3: Process (8)
- Design 4: Alignment (2)
- Validation (4)
- Written Expression (2)
- Presentation (2)
- Integrity (2)

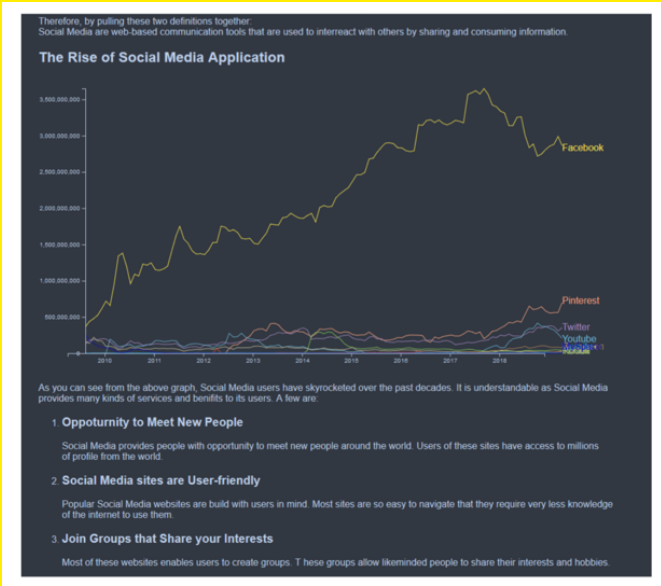
Appendix 1.1a: Index page:

Appendix 1.1b: Total road deaths visualisation page:



Appendix 1.1c: Road deaths per 100,000 visualisation page:





TITLE PAGE

Unit Name

Assessment Name

Title of Project

Student Name and ID

Due date

Link to website

Data Visualisation

COS30045

Sustainable Fishing around Australia

[*Student name and ID*]
[*Due date]*

https://mercury.swin.edu.au/cos30045/

TITLE PAGE

Unit Name

Assessment Name

Title of Project

Student Name and ID

Due date

Link to website

COS30045: Data Visualisation

Process Book:

ANALYSING MENTAL HEALTH AND DISORDERS

[*Student name and ID*]
[*Due date*]
[*website link*]

TITLE PAGE

Unit Name

Assessment Name

Title of Project

Student Name and ID

Due date

Link to website

CHILD MORTALITY PROJECT PROCESS BOOK

Website link- https:/

[*Student name and ID*]
[*Due date*]

TABLE OF CONTENTS

- Use heading styles and auto generate table of contents (Word: Insert/Index and tables)
- Number sections
- Choose a style that is easy to read
- Use page numbers

Contents

1. Introduction 2
   1.1 Background and Motivation 2
   1.2 Project Objectives 2
   1.3 Project Schedule 2
2. Data 3
   2.1 Data Source (Proposal) 3
   2.1 Data Source (Final) 3
   2.2 Data Processing (Proposal) 3
   2.2 Data Processing (Progress) 4
   2.2 Data Processing (Final) 4
3. Requirements 5
   3.1 Must-Have Features (Proposal) 5
   3.1 Must-Have Features (Final) 5
   3.2 Optional Features (Proposal) 5
   3.2 Optional Features (Final) 5
4. Visualisation Design (Proposal) 6
4. Visualisation Design (Progress) 7
4. Visualisation Design (Final) 8
5. Conclusion 9
References 9
Appendix 9

- Use Heading styles

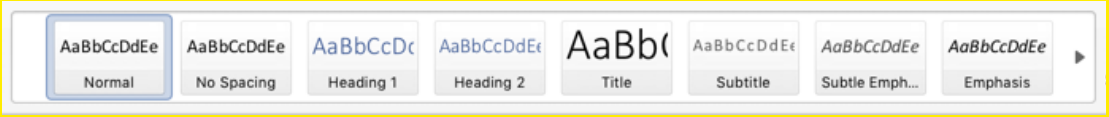


TABLE OF CONTENTS

- Use heading styles and auto generate table of contents
- Number sections
- Choose a style that is easy to read

- Not so easy to read
- Section numbers missing

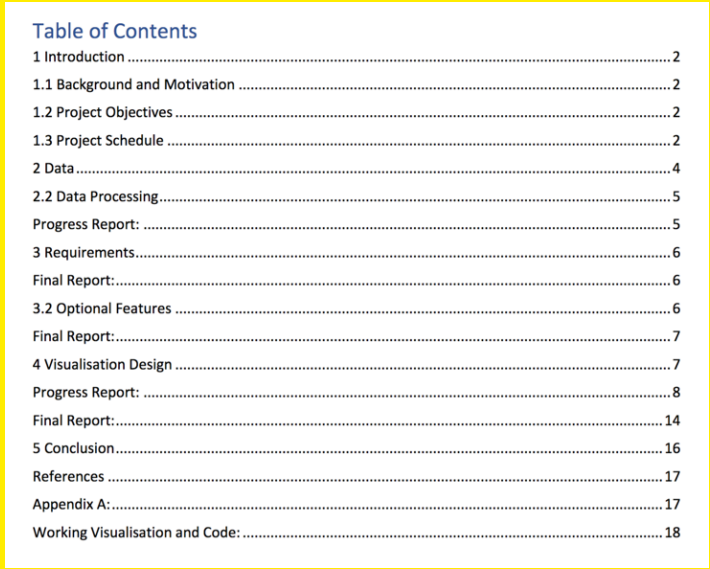


Table of Contents

1 Introduction 2
1.1 Background and Motivation 2
1.2 Project Objectives 2
1.3 Project Schedule 2
2 Data 4
2.2 Data Processing 5
Progress Report: 5
3 Requirements 6
Final Report: 6
3.2 Optional Features 6
Final Report: 7
4 Visualisation Design 7
Progress Report: 8
Final Report: 14
5 Conclusion 16
References 17
Appendix A: 17
Working Visualisation and Code: 18

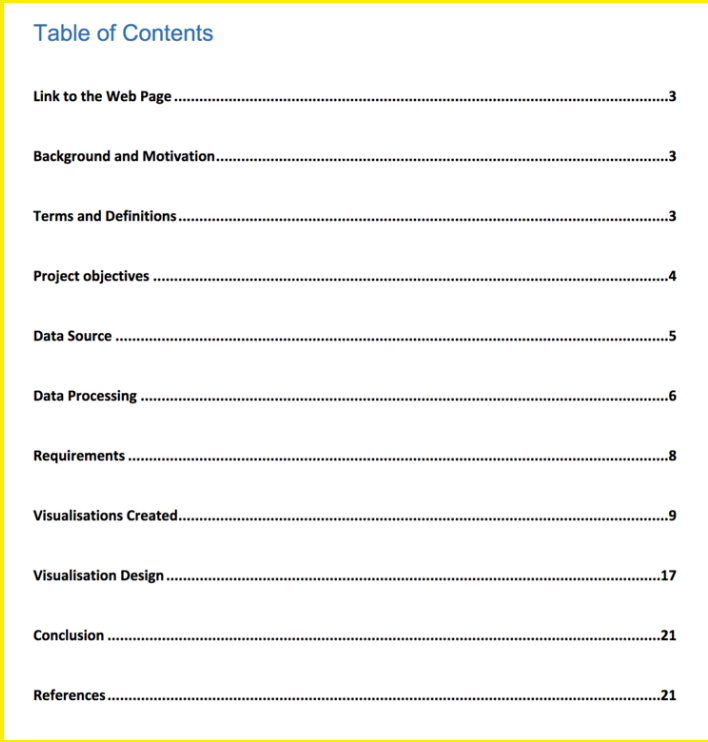


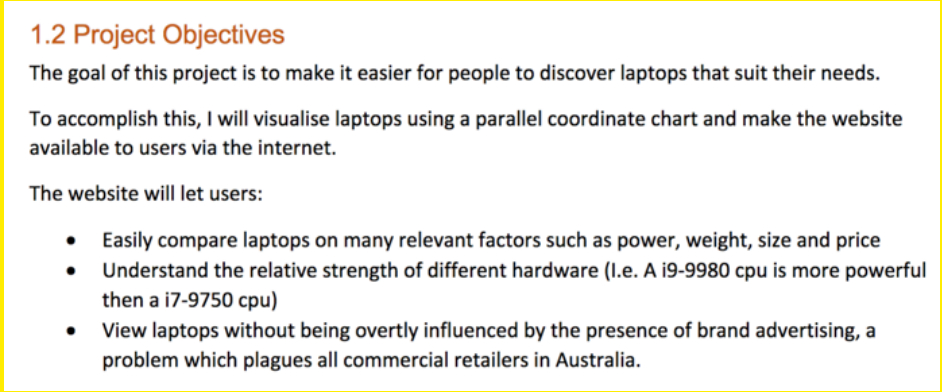
Table of Contents

Link to the Web Page 3
Background and Motivation 3
Terms and Definitions 3
Project objectives 4
Data Source 5
Data Processing 6
Requirements 8
Visualisations Created 9
Visualisation Design 17
Conclusion 21
References 21

1 INTRODUCTION

- Background and motivation
- Visualisation Purpose

- Give the reader some background to the topic domain (i.e., what will they need to know to understand your visualisation and why it is importance)



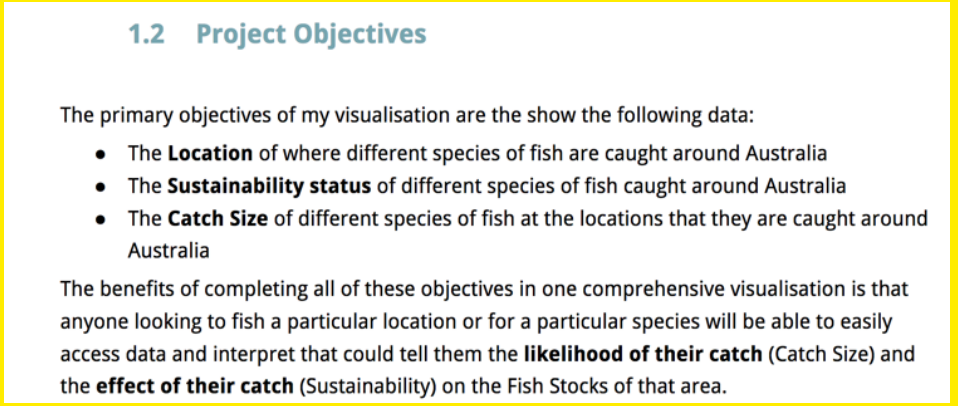
1.2 Project Objectives

The goal of this project is to make it easier for people to discover laptops that suit their needs.

To accomplish this, I will visualise laptops using a parallel coordinate chart and make the website available to users via the internet.

The website will let users:

- Easily compare laptops on many relevant factors such as power, weight, size and price
- Understand the relative strength of different hardware (I.e. A i9-9980 cpu is more powerful then a i7-9750 cpu)
- View laptops without being overtly influenced by the presence of brand advertising, a problem which plagues all commercial retailers in Australia.



1.2 Project Objectives

The primary objectives of my visualisation are the show the following data:

- The **Location** of where different species of fish are caught around Australia
- The **Sustainability status** of different species of fish caught around Australia
- The **Catch Size** of different species of fish at the locations that they are caught around Australia

The benefits of completing all of these objectives in one comprehensive visualisation is that anyone looking to fish a particular location or for a particular species will be able to easily access data and interpret that could tell them the **likelihood of their catch** (Catch Size) and the **effect of their catch** (Sustainability) on the Fish Stocks of that area.

This visualization could be helpful to **Fisheries** or **people who like to fish as a hobby** to help them answer the following questions:

- **What is Australia's current situation when it comes to fishing sustainability?**
- **What parts of Australia are more sustainable to fish?**
- **Which species of fish are more sustainable to catch?**

- What questions will your Vis answer for the user?
  - Questions, Data and Visualisation design must all match!
  - Help your reader see the link - don't make them figure it out for themselves!

2 DATA

- Source
- Data Processing
- Data types and encoding options

- Clearly explain the characteristics of the data
- Clearly show the steps in processing/transforming the data

3. Using functions and relationships to form a single table with relevant fields.

- I used PowerPivot for Excel to manipulate the data into one relevant table using relation functions. I retrieved the year from the ACCIDENTDATE field, the number of people killed from the NO_PERSONS_KILLED field, and the number of people injured by adding NO_PERSONS_INJ_2 and NO_PERSON_INJ_3 fields from the Accident dataset.

[Year] =YEAR(RELATED(ACCIDENT[ACCIDENTDATE]))

	ACCIDENT...	Year	LGA_NAME	Region Name	Postcode No	Num_Killed	Num_Injured
1	T20060000010	2006	DANDENONG	METROPOLITAN...	3175	0	1
2	T20060000018	2006	CASEY	METROPOLITAN...	3804	0	1
3	T20060000022	2006	MORNINGTO...	METROPOLITAN...	3939	0	1
4	T20060000023	2006	DANDENONG	METROPOLITAN...	3173	0	1

Figure 2: Usage of relation functions to summarise and simplify data into new calculated fields.

4. Using Excel PowerPivot to create calculated fields and summarise the data into years.

- Using PivotTables, I created three tables containing the relevant data for the visualisation. The first table contains the count of accidents in each Local Government Area (LGA) from 2006 to 2019. The second table contains the sum of the number of injuries in accidents, and the third table contains the sum of the amount killed in accidents. Using the row and column filters, I removed any blank data cells.

Count of ACCIDENT_NO	Column Labels														
Row Labels	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Grand Total
(FALLS CREEK)	1		4	1		1	2	3	4	1	1	2	4	1	25
(FRENCH ISLAND)				1								1			2
(LAKE MOUNTAIN)		1	1	1	1	3	1	2	1	2	4	1	3		21
(MOUNT BAW BAW)			2	1	1	1	1	1	1	1	2				11
(MOUNT BULLER)	1	4	3		2	1	1	2		4	2	1	2		23

Sum of Num_Injured	Column Labels														
Row Labels	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Grand Total
(FALLS CREEK)	1		6	1		1	2	5	13	9	1	2	4	1	46
(FRENCH ISLAND)				4								0			4
(LAKE MOUNTAIN)		1	1	1	1	3	1	3	1	2	4	1	3		22
(MOUNT BAW BAW)			2	1	0	1	1	1	1	1	6				14
(MOUNT BULLER)	1	5	5		3	1	1	2		4	2	1	1		26

Sum of Num_Killed	Column Labels														
Row Labels	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Grand Total
(FALLS CREEK)	0		0	0		0	0	1	0	0	0	0	0		1
(FRENCH ISLAND)				0								1			1
(LAKE MOUNTAIN)		0	0	0	0	0	0	0	0	0	0	0	0		0
(MOUNT BAW BAW)			0	0	1	0	0	0	0	0	0				1
(MOUNT BULLER)	0	0	0		0	0	0	0		0	0	0	1		1

Figure 3: The three calculated tables that contain the count of each attribute by LGA and Year.

5. Correcting LGA names to correspond to the VIC LGA JSON file.

- As the LGA names in the dataset do not all correspond to the names used in the VIC LGA JSON file, I had to change them to match, otherwise, the data would not load properly. I created a spreadsheet where I could manipulate the LGA names. I then corrected the names for any LGAs that did not match the JSON code (highlighted) and then used a PROPER function to convert the text case to match the LGAs in the code.

=PROPER(B2)

	A	B	C
1	Original LGA	LGA Name Correction	LGA Case Correction
2	(FALLS CREEK)	Falls Creek Alpine Resort (Unincorporated)	Falls Creek Alpine Resort (Unincorporated)
3	(FRENCH ISLAND)	French-Elizabeth-Sandstone Islands (Unincorporated)	French-Elizabeth-Sandstone Islands (Unincorporated)
4	(LAKE MOUNTAIN)	Lake Mountain Alpine Resort (Unincorporated)	Lake Mountain Alpine Resort (Unincorporated)
5	(MOUNT BAW BAW)	Mount Baw Baw Alpine Resort (Unincorporated)	Mount Baw Baw Alpine Resort (Unincorporated)
6	(MOUNT BULLER)	Mount Buller Alpine Resort (Unincorporated)	Mount Buller Alpine Resort (Unincorporated)
7	(MOUNT HOTHAM)	Mount Hotham Alpine Resort (Unincorporated)	Mount Hotham Alpine Resort (Unincorporated)
8	(MOUNT STIRLING)	Mount Stirling Alpine Resort (Unincorporated)	Mount Stirling Alpine Resort (Unincorporated)
9	ALPINE	ALPINE	Alpine

Figure 4: LGA Name Correction to match JSON file.

6. Prepared tables for CSV extraction and filled in blank cells.

- Before extracting the data into a CSV file, I created 3 spreadsheets and copied the data and corrected LGA names into each. Blank cells were filled with zeros to prevent errors which were done by performing by using the Find & Select function (Go to special -> blank -> Function: =0 -> Ctrl + Enter). The years were changed to have a 'y' at the beginning so it can be referred to in JavaScript.

LGA	y2006	y2007	y2008	y2009	y2010	y2011	y2012	y2013	y2014	y2015	y2016	y2017	y2018	y2019
Falls Creek Alpine Resort (Unincorporated)	1	0	4	1	0	1	2	3	4	1	1	2	4	1
French-Elizabeth-Sandstone Islands (Unincorporated)	0	0	0	1	0	0	0	0	0	0	1	0	0	0
Lake Mountain Alpine Resort (Unincorporated)	0	1	1	1	1	3	1	2	1	2	4	1	3	0
Mount Baw Baw Alpine Resort (Unincorporated)	0	0	2	1	1	1	1	1	1	1	2	0	0	0
Mount Buller Alpine Resort (Unincorporated)	1	4	3	0	2	1	1	2	0	4	2	1	2	0

Figure 5: Example of one of the final spreadsheets in the Excel file after final processing.

7. Exporting the final manipulated data tables as CSV files for the three different data variables.

- After each final Excel spreadsheet table was finalised with its data processing, they were each exported as a CSV file. The data is now in ordinal format post the data processing.

Figure 6: The final three CSV datasets used for the Choropleth visualisation. (zoom-in for larger image)

8. Creating an additional dataset for bar chart summary statistics.

- Using the grand totals from the PowerPivot tables in step 4, I was able to combine the tallies of the total accidents, injuries and deaths of each year and into a dataset for the bar chart visualisations. This was hardcoded as arrays in the bar chart code later.

	Measure	y2006	y2007	y2008	y2009	y2010	y2011	y2012	y2013	y2014	y2015	y2016	y2017	y2018	y2019
2	Accidents	13517	13566	14059	13760	13571	13994	13900	14040	14491	14662	14581	12455	11273	1043
3	Injuries	17650	17745	18165	17721	17649	17817	17499	17460	18010	18111	18120	15693	14029	1355
4	Deaths	342	339	308	296	290	290	284	245	254	257	293	262	199	24

Figure 7: Fourth Dataset used for the bar chart visualisations displaying total accident, injuries and deaths each year.

Note:

- Figure captions! All figures and tables must be number and captioned

2 DATA

- Source
- Data Processing
- Data types and encoding options

- Clearly explain the characteristics of the data
- Clearly show the steps in processing/transforming the data

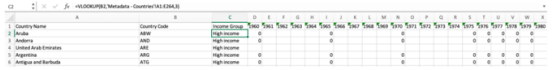
Data Processing

This dataset will only require some basic transformation to enable visualisations to correctly display the data, this will involve basic excel procedures and the use of simple functions such as average. In order to answer some of the questions identified under project objectives, a few calculated fields will need to be made, such as the percentage of mobile subscriptions of the total population, this will utilise excel functions and transformations between two datasets.

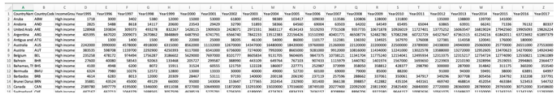
Data Transformation

The first step in preparing the data for visualisations was to remove the unnecessary data attributes from the excel workbook, the Indicator Name and Indicator Code was removed since it has no relevance in our application.

The next step was to add in an extra column for the income group of each of the countries, this new attribute is Categorical (Text). A VLOOKUP excel formula was used to retrieve the associated 'Income Group' value from the related 'metadata-indicator' sheet using the 'Country Code' as the lookup value.



The next step was to make a copy of the dataset, while only copying the Data from 1995 to 2017. In order for the year attribute names to be used in JavaScript, they had to be converted from a number to text. This was done by changing the first value (1995) to (Year1995) and then filling the rest of the columns using excels automatic fill function which continued to (Year2017). The resulting dataset is seen below and was exported as a csv file, for use with the D3 library for the Map Visualisation.



Data Source

The data used for this visualisation will consist of mobile cellular subscription data from 1960 to 2017 sourced from the world bank organisation. This data is in tabular form and consists of the following attributes;

(Dataset: World Bank, 'Mobile Cellular Subscriptions')

Attribute	Description	Type
Country Name	This field contains the name of the country for the subscription data.	Categorical (Text)
Country Code	This is a unique 3 letter identifier of each country and is referenced in a separate dataset, with data regarding the Region and Income group.	Categorical (Text)
Indicator Name	This field contains the name of the indicator from world bank.	Categorical (Text)
Indicator Code	This is a unique code that represents the indicator in a separate dataset, containing the full description.	Categorical (Text)
Year[s]	These attributes range from 1960 to 2017 and contain the subscription number for each country.	Ordinal (Number)

Within the same dataset, a separate sheet named 'metadata – countries' is present and contains the following information;

Country	This field contains the name of the country for the subscription data.	Categorical (Text)
Region	This field contains the geographical region of the country. E.g. North America or South Asia	Categorical (Text)
Income Group	This field categorises the country based on overall income, there are 4 groups: Low Income, Lower Middle Income, Upper Middle Income & High Income.	Categorical (Text)
Special Notes	This field contains a variety of extra details/notes on the source of the data and what is included/excluded	Text

- Show the reader you are confident in handling the data

3 VISUALISATION DESIGN

- Demonstrate the design process
  - low fidelity prototyping (sketches)
  - iteration
  - use of design guidelines

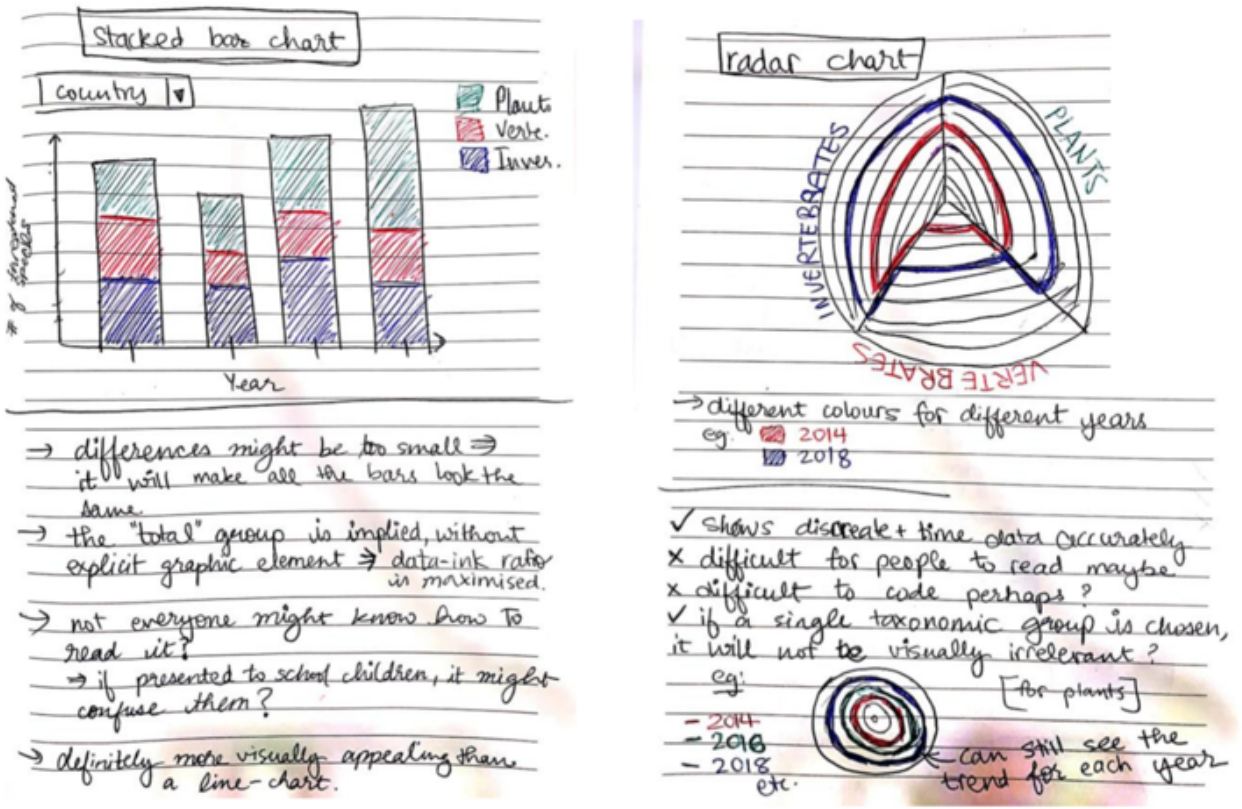


Figure 1: Initial designs of a stacked bar chart (left) and a radar chart (right), that were considered for the visualisation.

3 VISUALISATION DESIGN

- Demonstrate the design process
  - low fidelity prototyping (sketches)
  - iteration
  - use of design guidelines

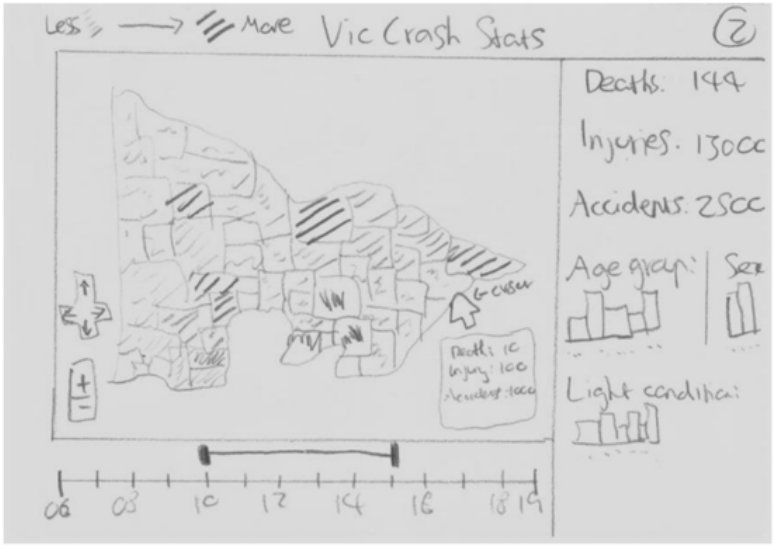


Figure 9: Visualisation Design 2

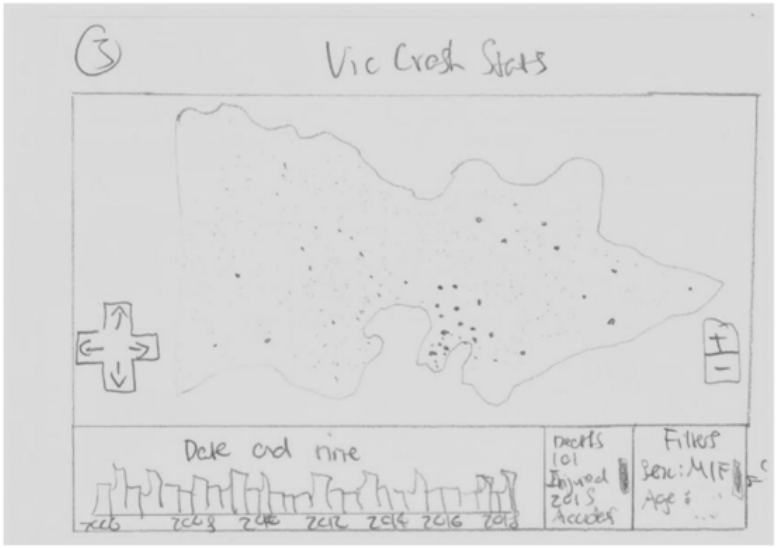


Figure 10: Visualisation Design 3

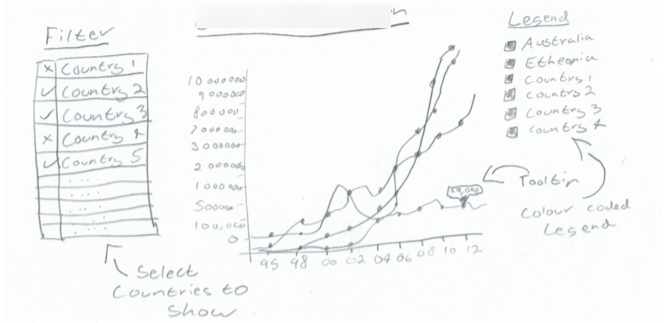
3 VISUALISATION DESIGN

- Demonstrate the design process
  - low fidelity prototyping (sketches)
  - iteration
  - use of design guidelines

Visualisation Design

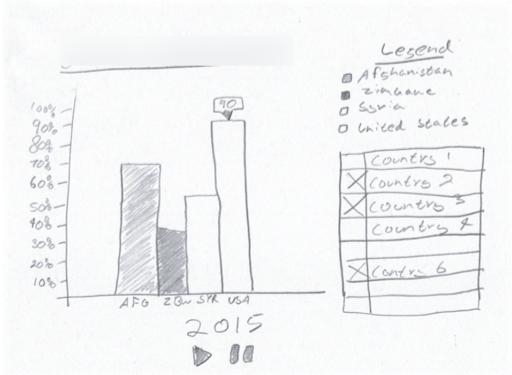
Proposal

Multi-Line Chart

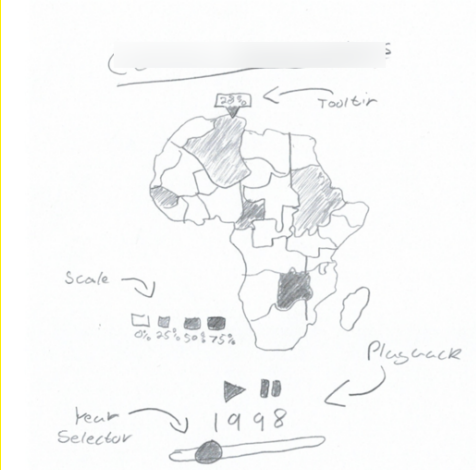


This line chart is an effective method of showing continuous data to show the changes in mobile phone subscription rates over time across multiple countries. The cellular subscriptions are treated as quantitative data and encoded through the position of data points on the line/scale. The line chart should include data points with tooltips for user interactivity, allowing users to reveal the exact data value for that point. In addition, the user should be able to filter out countries using a table/selection tool, there should also be a colour coded legend to identify the different line series.

Bar Chart



Map Choropleth



This map visualisation would show the [redacted] percentage for the different states in each country by altering the colour saturation (Darker colours indicate higher number). This visualisation should also have the ability to scrub through the years, displaying data over time using the slider as well as the automatic playback with pause functionality. Data tooltips should also be integrated to enable the user to see the exact value represented by the colour. A scale should be included to give the user context for the colour saturation.

3 VISUALISATION DESIGN

- Demonstrate the design process
  - low fidelity prototyping (sketches)
  - iteration
  - use of design guidelines

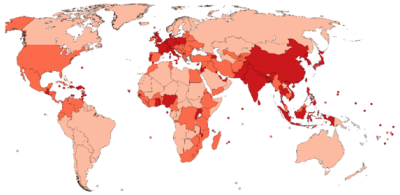


Figure 8: Design for total unemployment rate visualization using heat map

The design for total unemployment rate visualization using bubble map is displayed in figure 9. Similar to the heat map design, this bubble map allows users to easily determine countries which are neighbors to make comparisons. The unemployment rate is encoded by circle size so the effectiveness of displaying unemployment rate variation is acceptable.

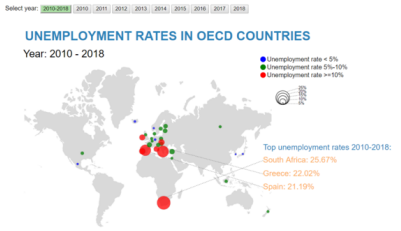


Figure 9: Total unemployment rate visualization using bubble map

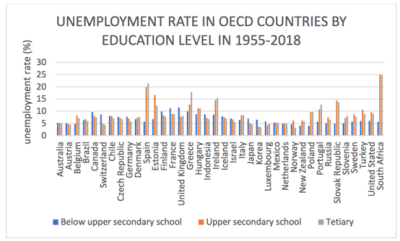


Figure 10: Unemployment rate by gender/educational level line chart design

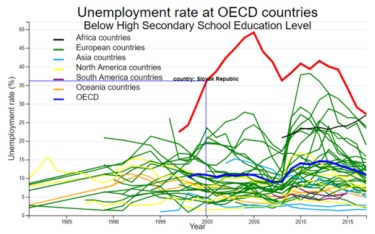


Figure 11: Unemployment rate by gender/educational grouped line chart design

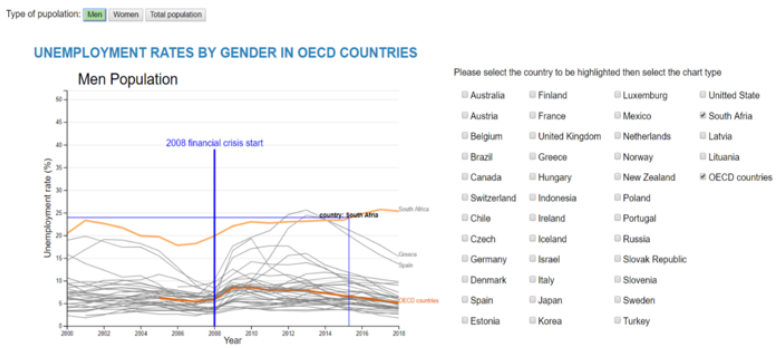


Figure 12: Unemployment rate by gender line chart design

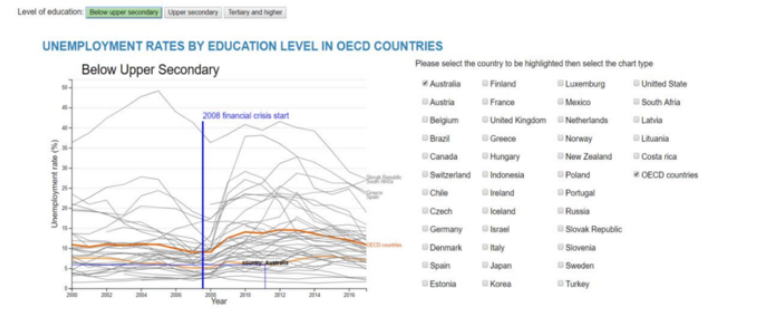


Figure 13: Unemployment rate by educational level line chart design

3 VISUALISATION DESIGN

- Demonstrate the design process
  - low fidelity prototyping (sketches)
  - iteration
  - use of design guidelines

4.2 Updated Design

As it is decided that all the data for each country ranged from 1990 to 2015 and the average alcohol consumption are derived, the first bar graph will show each country with the associated average alcohol consumption over the year range. The figure below shows the updated design based on the initial design.

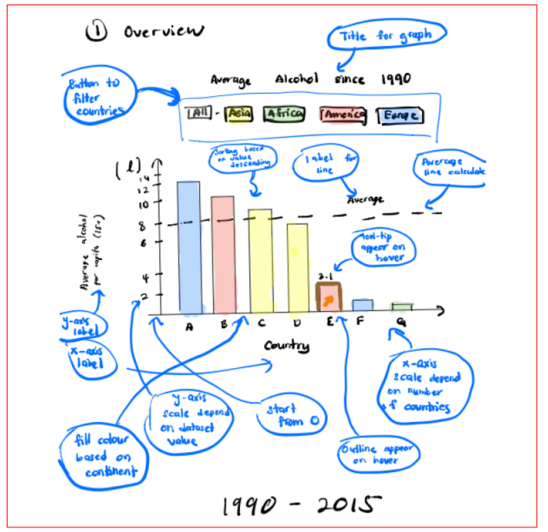


Figure 5 – Bar chart sketch

The second graph is the area chart which replaces the line graph. The change of graph from a line graph to an area chart is due to area chart can show more differences and the differences can be seen clearly.

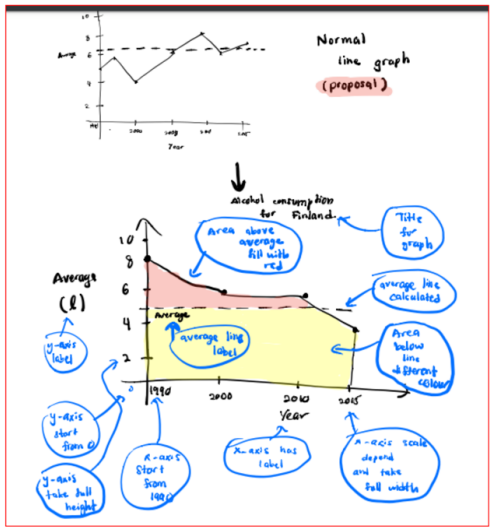


Figure 6 – Area Chart sketch

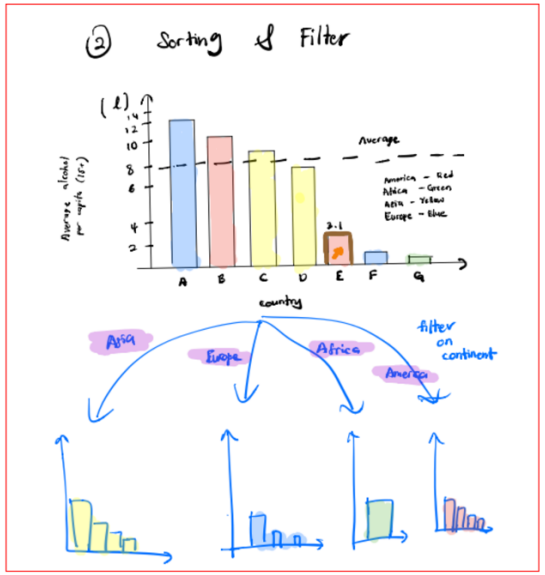


Figure 7 – Filtering feature sketch

3 VISUALISATION DESIGN

- Demonstrate the design process
  - low fidelity prototyping (sketches)
  - iteration
  - use of design guidelines

Design discussion and justification

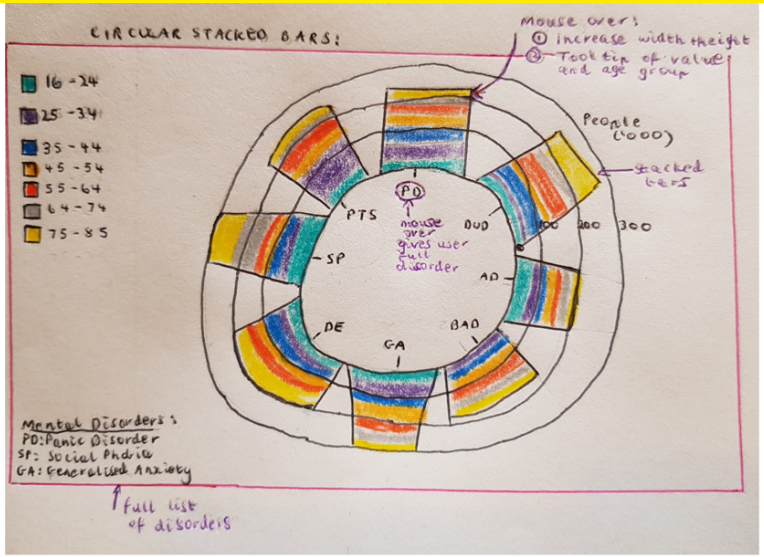


Figure 8 : Design Sketch 3

Design Sketch 3 utilises a circular stacked bar chart. This chart is unique in style however lacks interactivity for the user. The legend will be displayed on the top left corner and will be colour coordinated based on colours from Color Brewer. Ideally, there will be a sort button available for the user to press which will start from the highest/ lowest value and work its way around clockwise. However doing this may be quite difficult in d3. Since there will be little room in the inner circle, the disorders will be abbreviated and the full list will be displayed on the bottom left of the screen.

A feature that would be advantageous we be including not just the typical mouse over for each bar, increasing the size, changing the colour and displaying the value, but to have a mouse over that will display the full disorder when the user highlights the abbreviated one.

Similarly to Sketch 1 and 2 there would be transitions on page load. The lengths/ sizes of the bars help encode the data as the bigger the bar, and the closer it is to another circle, the bigger the value is.

ADVANTAGES	DISADVANTAGES
Values appear on hover, colour changes and bar becomes larger	Hard to compare values of each age group
Explores all the different types of disorders from the dataset	Doesn't explore other characteristics such as gender, smoking status etc.
Colours are used to encode the different age groups for each bar	No accompanying text to explain the graph further.
Easy to see **total** people suffering for each mental disorders	Overwhelming to read, no sort function
Unique and different to any other graph	Coding this d3 may be difficult

Table 4: Sketch 3 Pros and Cons

3 VISUALISATION DESIGN

- Demonstrate the design process
  - low fidelity prototyping (sketches)
  - iteration
  - use of design guidelines

Design guidelines

4.3.1 Colour Hue Combination Used

As this bar chart is used to show relationship between categorical data which in this case would be the country, having one colour hue for each country would not be neat. However, hue can be used here to represent the continents as this can make user able to differentiate the countries based on continents. The colour hue combination here is a distinctive colour selected from colorbrewer2.org. There are four continents that are classified in this visualisation, hence there would be four hues that will represent each one of the continents.

Figure 9 – Colour Code for the Hue Used

These hue colours are chosen because this combination is colour blind friendly. It means that even for those with colour blind would be able to distinguish these hues so that it does not mislead them. This is tested by using the ColorBlinding extension in google chrome where the bar chart with these hue combinations are tested with Protanopia.

Final

The final design includes a table which has an alternating highlighting to show off each laptop. Smaller tweaks have also been made to the heading and navigation bars in order to keep consistent themes.

Because price is one of the most important fields when comparing laptops, we encode the price of the laptop in the hue of the line. Gold represents expensive laptops while aqua is used to represent cheaper laptops.

As each field is brushed or filtered, the table is also filtered to only include laptops that are within the brushed area. This provides a consistent representation between the chart and the table which helps the user connect the data in the chart with the data in the table.

The brush is drawn with the same white as the y-axis to imply that it is interacting with it. We also made the brush transparent to ensure it doesn't obscure any information, while still clearly indicating the selected area.

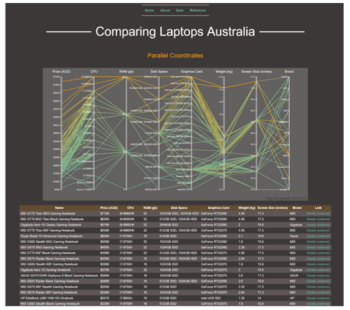
The cursor also changes to a cross when hovering over the y-axis to indicate that the element is interactable. A cross was chosen as it makes it easier to draw precise rectangles.

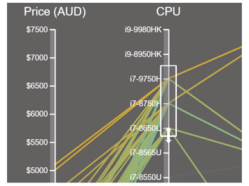
The brush can also be moved along the y-axis. This is indicated to the user by displaying a "dragging" cursor when the user hovers over an existing brush.

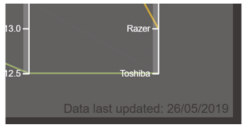
We also ensure that unselected laptops are faintly visible in the background to show that they still exist and can be selected. This can also aid users to slightly widen their search if there is a laptop close to their current criteria that is not selected.

The last updated date is displayed faintly in the bottom right of the graph since we're working with data that changes daily and will rapidly become out of date. This ensures that users do not rely on the data if it is out of date.

The table also highlights when a row is hovered to help users correlate fields of the same row.







Razer Blade 15 Advanced Gaming Notebook
MSI GS65 Stealth 9SG Gaming Notebook
MSI GS75 9SG Gaming Notebook
MSI GT75 8SF Black Gaming Notebook

Progress

The initial design contained a basic parallel coordinate chart with no axis or colour placed on a background image of a city. The colour scheme was chosen at random without much though for consistent colours.

As I developed the chart, I started to play with the hue of the lines to help the user follow the lines throughout the chart.

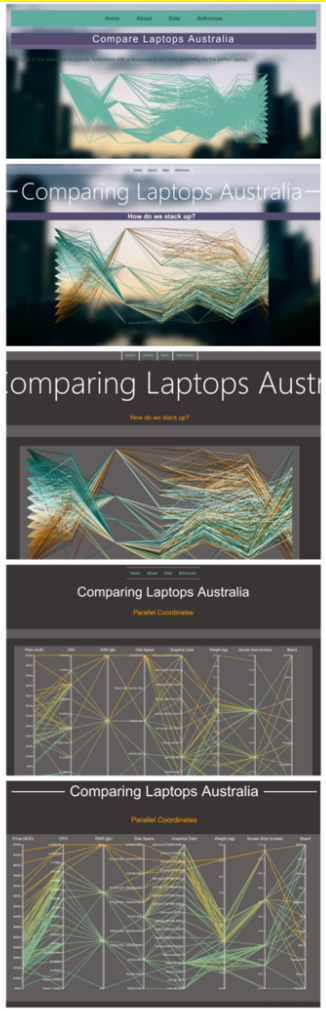
As I started to critique each element in the site, trying to decide if it needed to exist, I opted to remove the background image and transition the site to a more modern look.

The allowed the visualisation to become the centrepiece without any additional visual noise compromising it's impact.

I also restricted myself to a palette of four colours for all elements of the website other than the parallel coordinate chart.

This gave the website and visualisation a harmonious vibe and helped with readability. I tried many spacing options to get the proportions just right. The also included testing the colour and layout on different machines and browsers to ensure the chart remained consistent.

Another significant improvement was the introduction of the y-axis. This massively improved the readability of the visualisation, making it clear what each point in the chart represented.

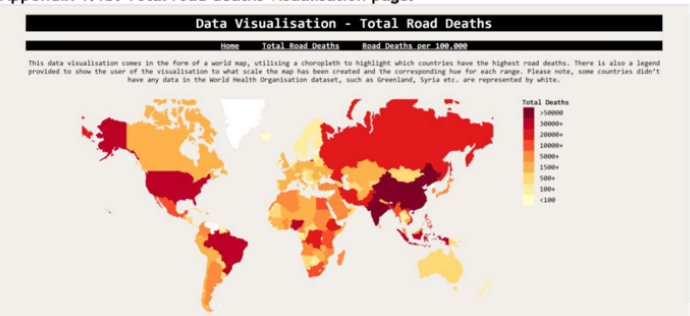


WEBSITE (20%)

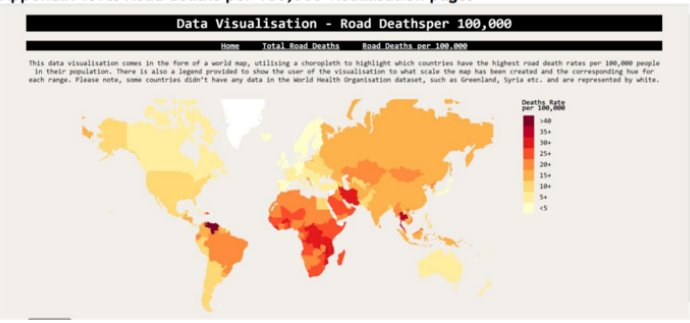
- Implementation 1: Coding Practice (5)
- Implementation 2: Complexity 1(5)
- Implementation 3: Interactivity and Vis design (5)
- Implementation 4: Story telling and aesthetic design (5)

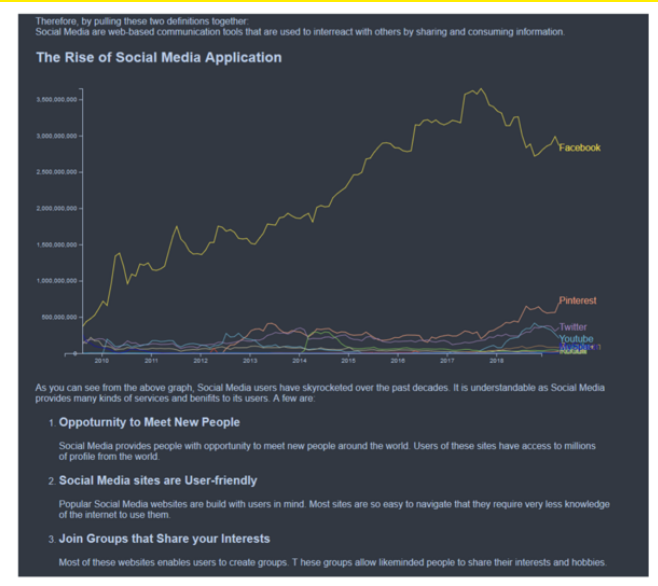
Appendix 1.1a: Index page:

Appendix 1.1b: Total road deaths visualisation page:



Appendix 1.1c: Road deaths per 100,000 visualisation page:





REFLECTION (20%)

- Data Vis Concepts (7)
- Programming concepts (7)
- Team work and Peer Assessment Survey (4)
- Written communication (2)

- 1,5000 to 3,000 words

REFLECTION

2.1.1 Data Binding

This is the technique that I use for my project to connect the dataset into the graphical elements in the visualization. For the graphs or charts to be able to display the right graphical information, binding the data needs to be done so that the actual data is linked to the chart to be updated. For example, in my final project I have used data from UN Migration which I have filtered to get necessary data and retrieved it using D3.

Figure 7: Code for Retrieving CSV Data using D3

This dataset is then bind into the graph which is drawn using svg to specify the characteristics of the data. To append the svg into the container in the DOM, select function is used to specify which container will be displaying the svg. Also, I have have appended g (group) which is used to group the following graph. Once these are set, I draw the graph by entering a new element to the group and use data function to link the previous retrieved data and append which element I want the data to be drawn. Here in the example below, I use path as I want to draw a stacked line chart that requires area as the mark for the items. Following to this, I have updated the existing visual elements to reflect changes in the dataset, which is done by applying attributes, transition and many more.

Figure 8: Append Svg in Body Element

REFLECTION

I have also challenged myself displaying the tooltip container to be adjusting to the current mouse position. This is done by using mouse event and use its x and y position to retrieve the exact position of the mouse. However, as I wanted to make the tooltip to be above the mouse, I have configured its x and y values so that it will be appearing in an appropriate position in the graph.

Figure 13: Positioning Tooltip Using Mouse Event

Besides applying hover features, I have also applied sorting data using buttons as options for users to choose. In this application, I have tried to link these buttons to trigger sorting data ascendant or descendant using d3's sort function combined with either ascending or descending function. This is implemented in my migration project as to show the lowest to highest or vice versa of the destination countries migration ranking.

Figure 14: Sort Dataset Ascending

Figure 15: Sort Dataset Descending

REFLECTION

Figure 4. First data visualisation of the group project

During my group project, understanding the SVG tag was instrumental in creating various visual elements and interactions. For example, I utilised SVG shapes to represent geographical regions on a choropleth map, with each shape encoding different sub region. The ability to incorporate the SVG element with CSS and JavaScript allowed for interactive feature that can update the data being presented in the data visualisation. In the case of my project, the year slider mechanism serve as a good example of the combination of Javascript to handle user interaction and changed the value bind to the visualisation. Furthermore, SVG's path element enabled the creation of line connection graphs, showcasing migration trends between different sub-regions. The use of SVG allowed for smooth animations, highlighting the connections and enhancing the user's understanding of the data.

REFLECTION

seamless connection between the underlying data and the visual representation. The example of my understanding in this concept can be seen in my group project code for choroplast map, since the geoJSON dataset used to draw the world map is made of coordinates that draw the shape of each subregion, path element can be used because it can draw line based on list of screen coordinates. The code below shows that the D3 can select all path elements in the SVG and then bind data of geoJSON coordinate on that path element to draw the shape of each sub region.

Figure 5. Code for selection and data binding in my project

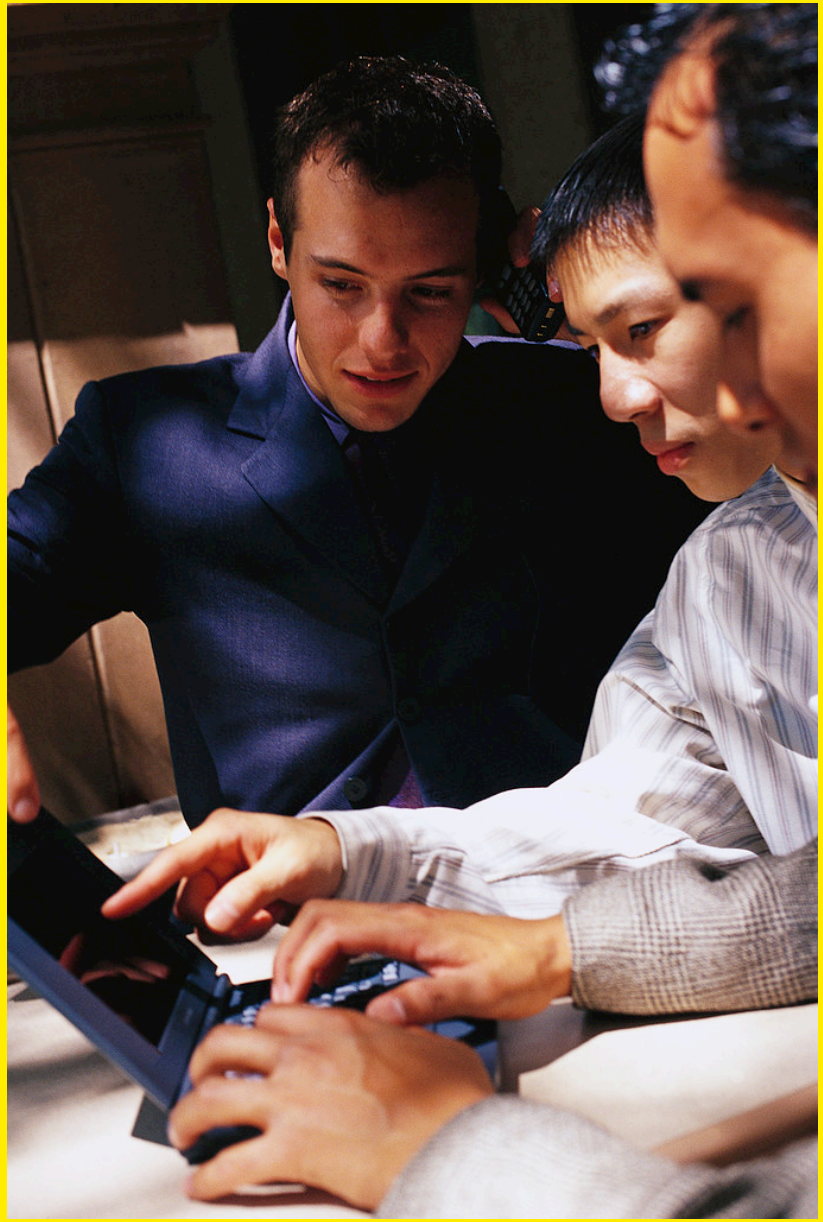
REFLECTION

Based on our project, we used HTML, CSS, JavaScript d3 and JavaScript coding. We use JSON and CSV files to save our data and load it into our JavaScript file. We learned these in our unit COS30045 and used this knowledge to improve the presentation of our project. And the beginning and during the project, we created a process book to analyse every step and decision we made and record everything inside.

Visualisation, using JavaScript d3 to code, is the most important knowledge point of the whole semester. Learning different visualisation charts and elements from the unit and using JavaScript d3 to display different charts, we are able to use JavaScript d3 to create and display our own visualisation to express the data we want to present accurately.

PEER ASSESSMENT SURVEY

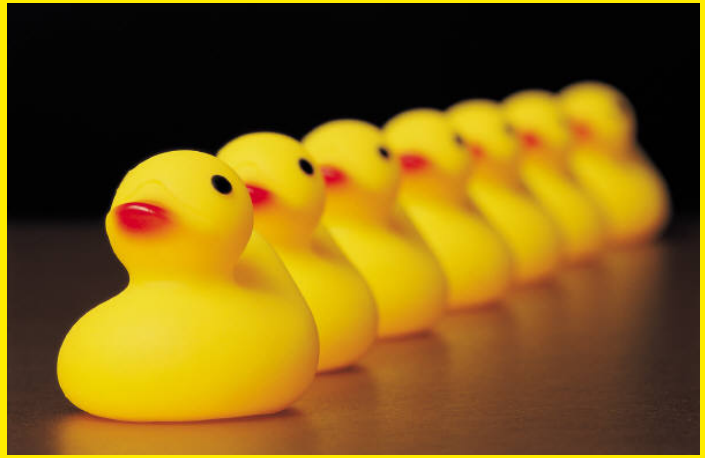
- Convenor will review Peer Assessments
- If issues, student/s will be emailed about attending an interview
- Interview to be held online - assess knowledge of and engagement in project



DATA VIS PROJECT

FINAL ADVICE

Good Luck!



Keep an eye on the marking criteria

- Time spent on documentation vs time spent fixing bugs

Put some effort into presentation of document:

- make it readable
  - 1.3 line spacing
  - use nice font (not Times New Roman)
- use heading styles defined from word processor template (do not use underline as a heading style!)
- use captions on all figures and tables
- use bullet points sparingly (i.e., only for lists, other wise use paragraphs)