**VISUALISATION PROCESS BOOK,2019**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **UNIT DETAILS** | | | | | | | | | | |
| Unit name | | Data Visualisation | | | | | | Class day/time | Friday/8.30 | Office use only |
| Unit code | | COS30045 | | | Assignment no. | | 2 | Due date | 31/5 |  |
| Name of lecturer/teacher | | | | Karola von Baggo | | | | | |  |
| Tutor/marker’s name | | | Humphrey Obie | | | | | | | Faculty or school date stamp |
| **STUDENT(S)** | | | | | | | | | | |
| Family Name(s) | | | | | | Given Name(s) | | | | Student ID Number(s) |
| (1) | To | | | | | Wilson Wee Sen | | | | 100076027 |
| (2) |  | | | | |  | | | |  |
| (3) |  | | | | |  | | | |  |
| (4) |  | | | | |  | | | |  |
| (5) |  | | | | |  | | | |  |
| (6) |  | | | | |  | | | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **DECLARATION AND STATEMENT OF AUTHORSHIP** | | | | | | |
| 1. I/we have not impersonated, or allowed myself/ourselves to be impersonated by any person for the purposes of this assessment. 2. This assessment is my/our original work and no part of it has been copied from any other source except where due acknowledgement is made. 3. No part of this assessment has been written for me/us by any other person except where such collaboration has been authorised by the lecturer/teacher concerned. 4. I/we have not previously submitted this work for this or any other course/unit. 5. I/we give permission for my/our assessment response to be reproduced, communicated, compared and archived for plagiarism detection, benchmarking or educational purposes.   I/we understand that:   1. Plagiarism is the presentation of the work, idea or creation of another person as though it is your own. It is a form of cheating and is a very serious academic offence that may lead to exclusion from the University. Plagiarised material can be drawn from, and presented in, written, graphic and visual form, including electronic data and oral presentations. Plagiarism occurs when the origin of the material used is not appropriately cited. | | | | |
| **Student signature/s**  I/we declare that I/we have read and understood the declaration and statement of authorship. | | | | |
| (1) | To | (4) |  | |
| (2) |  | (5) |  | |
| (3) |  | (6) |  | |

COS30045 Data Visualisation

Wilson To Wee Sen

100076027

Visualisation Process Book Report

# Introduction

## Background and Motivation

The topic that I have chosen for this project is road traffic injuries in Australia. Road traffic accidents are one of the world’s largest public health and injury prevention problems. Australia has seen over 190,000 fatalities since an accurate car record commenced in 1925. The reason that for choosing this topic is to gain more understanding about road traffic injuries. This project would be beneficial for all citizens who are living in Australia.

## Project Objective

* How many deaths are caused by road traffic accident in Australia?
* What are the types of road user which cause road traffic deaths?
* What is the proportion of road traffic deaths by age range.

## Project Schedule

Week 1: Understand the concept of data visualisation

Week 2: Learning about D3 and data visualisation outline

Week 3: Learning about D3 Axis and Scales

Week 4: Make a project proposal

Week 5: Obtain data from reliable website

Week 6: Sketching design for the project

Week 7: Project Progress report, including prototype phrase and simple web application

Week 8: Project Progress report, done prototype and add-in some features.

Week 9: user testing

Week 10: Optimized the data visualisation from the feedback that collected from the users

Week 11: Data Visualisation presentation.

# Data

## Data Source

I obtain the data from the reports that published from Australia Government, department of infrastructure, regional development and cities <https://bitre.gov.au/publications/ongoing/road_deaths_australia_annual_summaries.aspx>. The website contain many datasets which are needed in the project such as Road traffic deaths and Road traffic deaths by type of road user.

Furthermore, I will also take data set from other report that published from National Road Safety Strategy <https://roadsafety.gov.au/performance/road-deaths-age-group.aspx>. This website contain road deaths by age group in Australia which is needed in this project.

The data source provided by these official websites are in Excel file type with 56 tables of data sets. The attributes in data set are mostly region, year, population, and type of road user, genders, and number of deaths.

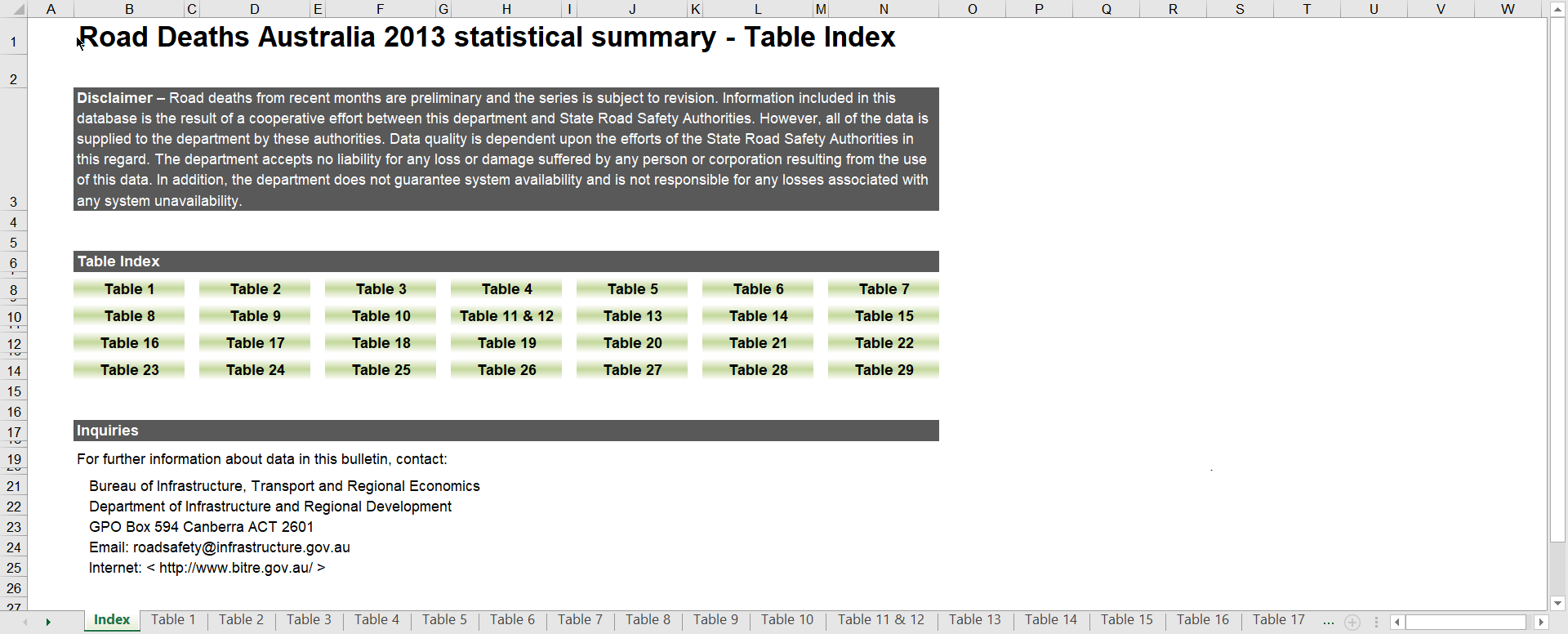


figure 1: Data source from Australia Government website

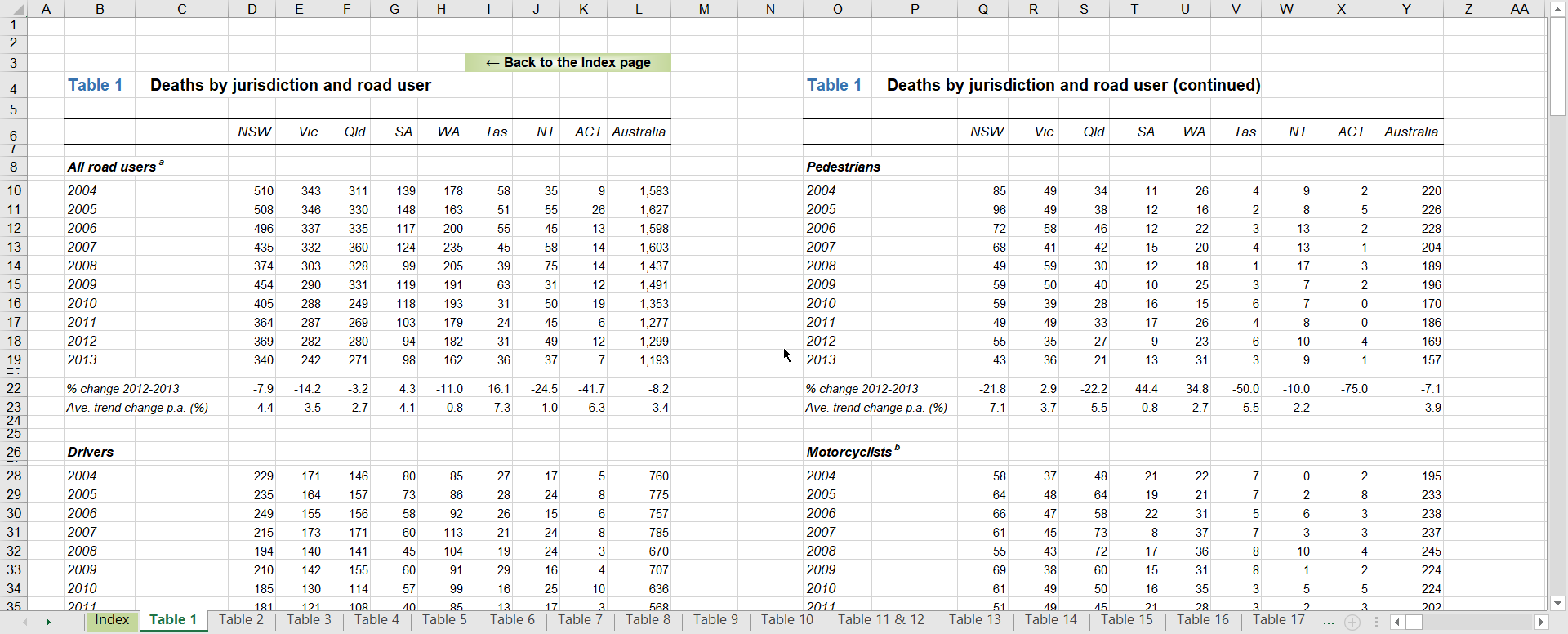


figure 2: One of the table from data source

## Data Processing

### **Proposal:**

Since the data source is already cleaned and legit, I do not plan to do substantial data cleanup. However, I do taking the table sets that are necessary for the project. And also I would change some of the variable names like replace the space by dash-line and convert the file to csv so it can be read by d3.

### **Progress Report:**

For data processing, I have gathered the data from number of injuries from year 2004 to 2013 of each type of road users. The data attributes that I have taken from the data source are type of users, years, age groups, number of road traffic injuries, genders, and number of injuries in states.

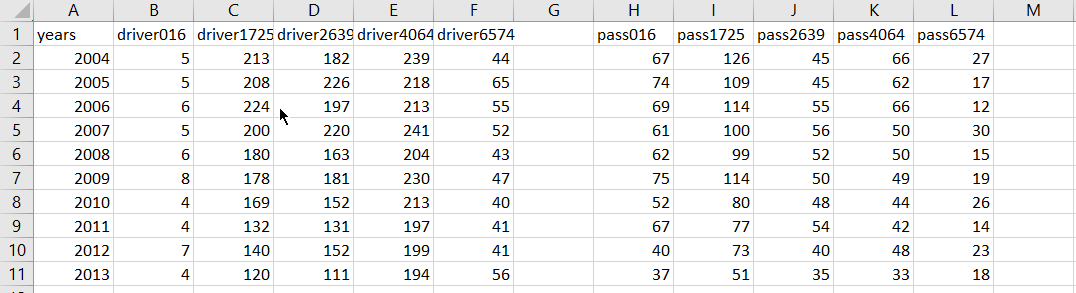


figure 3: example of data processing progress

The figure above shows the data of deaths by road users and age group. I have combined type users and age group in order to make the data visualisation more comparable. It shows the year from 2004 to 2013 and number of deaths by road users and age groups. There are 5 type of road users which are drivers, passengers, pedestrians, motorcyclists, and pedal cyclists. And also, there are 5 age groups have been collected from the reliable source which are 0-16 year old, 17-25 year old, 26-39 year old, 40-64 year old, and 65-74 year old.

### **Final:**

At the end, I have created 3 csv files in order to fulfilled 3 visualisation in this project.

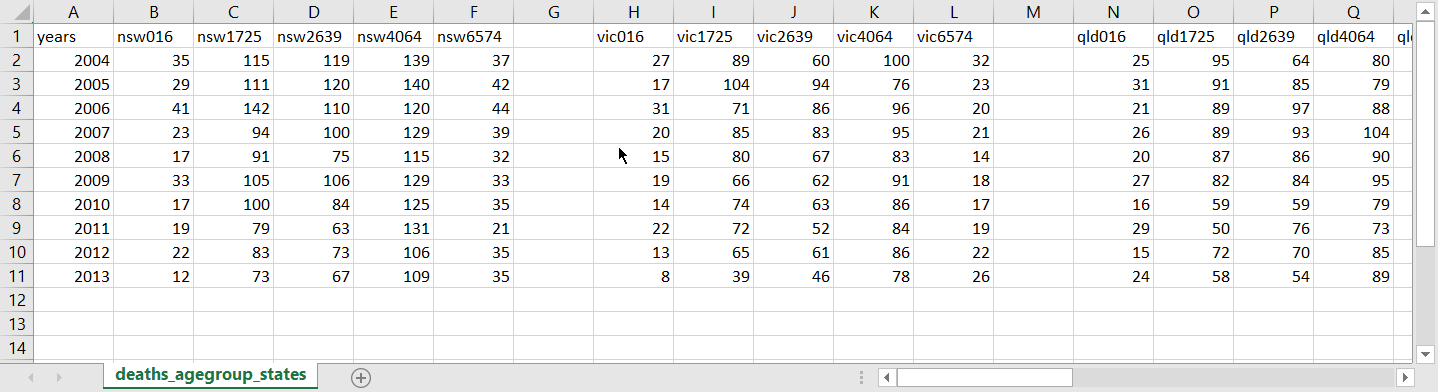


figure 4:data of deaths by states and age groups

The figure above shows the data of deaths by states and age groups. I have combined states and age groups so the visualisation will look more comparable. It shows the number of road traffic injuries from year 2004 to 2013.

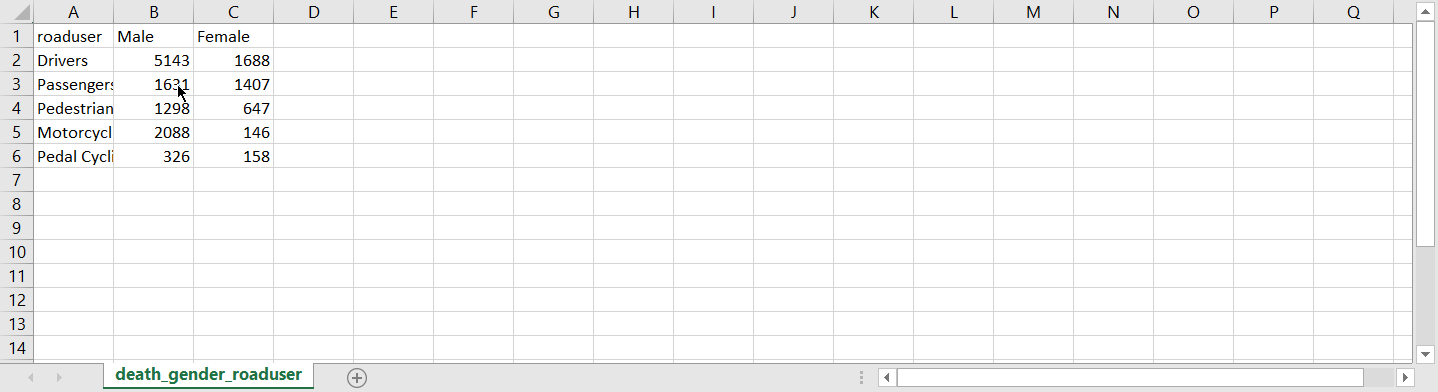


Figure 5: data of deaths by genders and road users

The figure above shows the data of deaths by road users and genders. I have combined the data of deaths by genders and road users. This is looks simple because I have done calculation from multiple tables set of data.

# Requirements

## Must-Have Features

### **Proposal:**

I have considered a hover feature in this project. For example, a data will be shown whenever the mouse move to any bar in the bar chart. And also, I would also consider a button that can switch data. For example if there are two button 2017 and 2018, when user click on 2017 button, it will show 2017 data visualisation. Furthermore, I also plan to implement transition when changing the data.

### **Final:**

I have fulfilled all the must have requirements in this project. As promise, I have implemented hover feature which data value will be shown whenever the mouse move to any bar in the bar chart. And also, I have implemented radio button for each type of road users so the it has interaction between each type of data. Furthermore, the lines and bars transition also implemented while change the data.

## Optional Features

### **Proposal:**

A bootstrap implementation would be the optional feature for this project. It would be nice to have because it can capture the interest of the audience and makes an good impression. And also, I plan to implement interaction between each visualisation. Like if reader click on any data in first chart, the second chart change the data as well. It connects each data set and it could catch reader attraction.

### **Final:**

For optional features, I were not able to implement bootstrap for the website due to running out of time. However, I used CSS to replace bootstrap. I made a navigation bar to link 3 visualisations by CSS and HTML. Due to I could not find the data sets which are relatable, I did not implement interaction between each visualisations.

# Visualisation Design

### **Proposal:**

For this project, I plan to do 2 visualisation which are line chart and bar chart. I plan to use multiple lines as mark for number of road traffic injuries and colors as visual channels to represent different age groups. As for bar chart, I plan to use genders to represent the number of injuries in each categories of road users. There will be 2 button to switch the data between male and female.

**Progress report:**

The design still remain unchanged. I have created a simple prototype to make sure the value from csv file can be read and display. The prototype is multiple line graph which shows the number of deaths by road users and age group. Each line shows the each type of road users and there has a group of ratio button to change the type of road users. I have also implemented labels for y-axis and x-axis.

For next improvement, I plan to do line transition when changing categories of road users. Furthermore, I also plan to do legend area to tell users the meaning of the lines’ colour.

### **Final:**

1. **Deaths by road users and age groups**

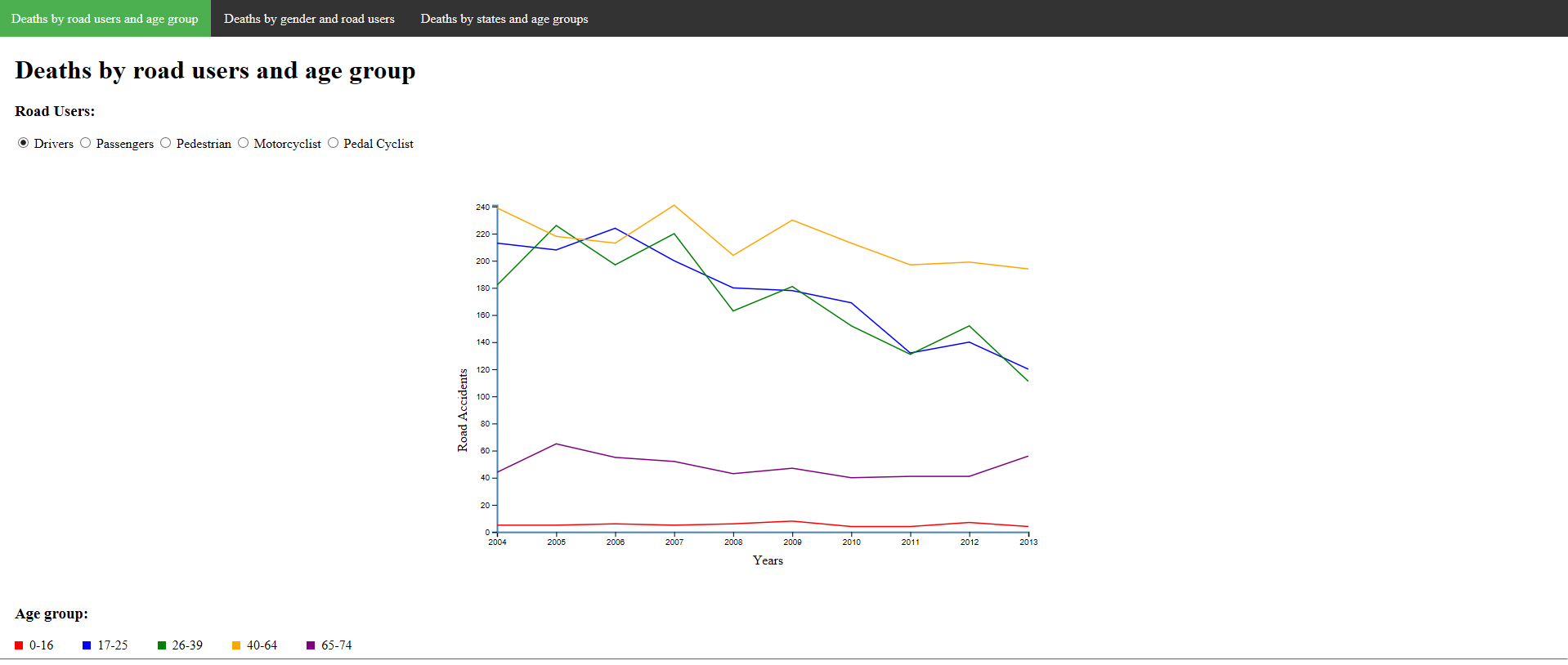


figure 6: visualisation 1--deaths by road users and age group

This is the final design of my visualisation. This visualisation shows the deaths by road users and age group. There are 5 groups of road users which are drivers, passengers, pedestrians, motorcyclist, and pedal cyclist. I made these road users into radio button group so reader can change the data by clicking the radio buttons. In this line chart, there are 5 lines in the chart with different colors. Each color stands for different age group. Red stands for age 0-16, blue is 17-25, green is 26-39, and yellow is 40-64, and purple is 65-74.

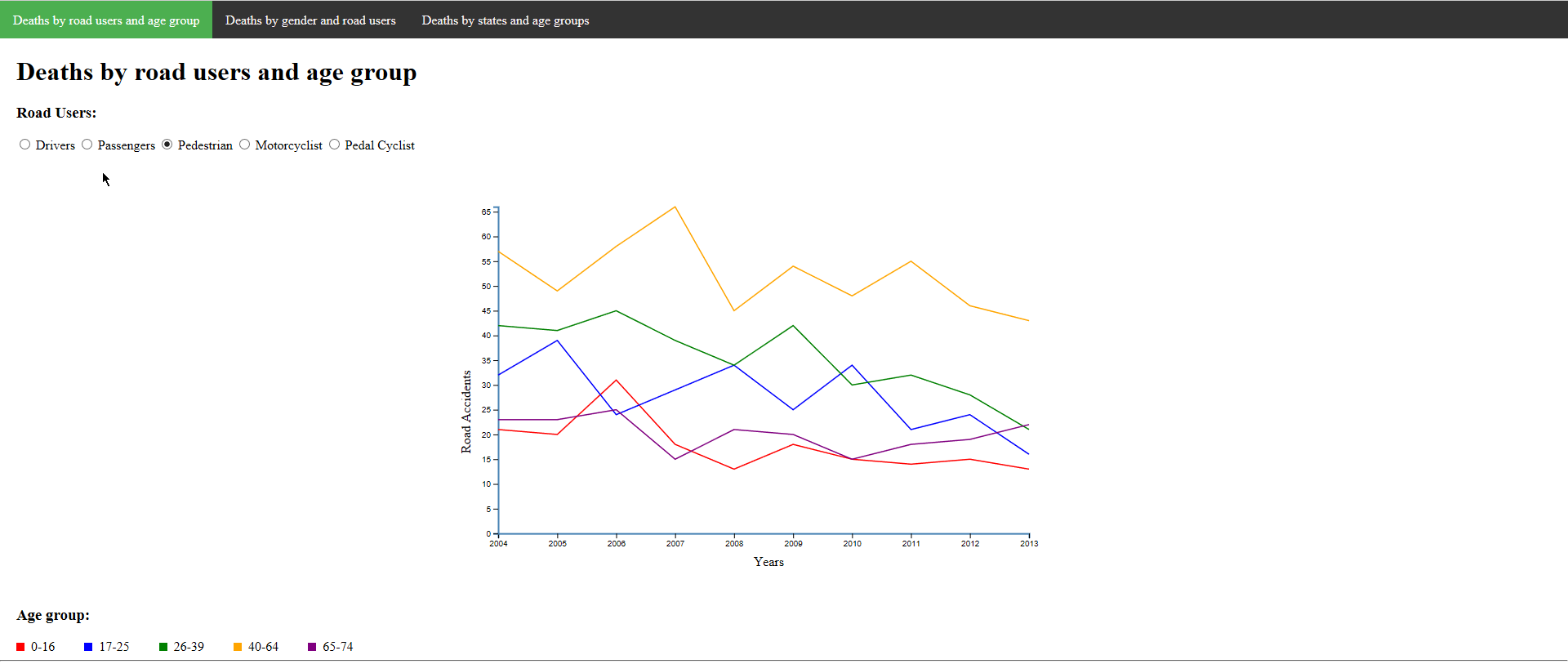


figure 7:visualisation 1--deaths by road users and age group

When clicking on any radio button, not only the lines change, the y-axis as well. The reason why I change the y-axis value when clicking other radio button is because the gap of number of injuries between drivers and pedal cyclist are too big. If I do not change the y-axis, the readers would not be able to compare the value in pedal cyclist category.

1. **Deaths by genders and road users**

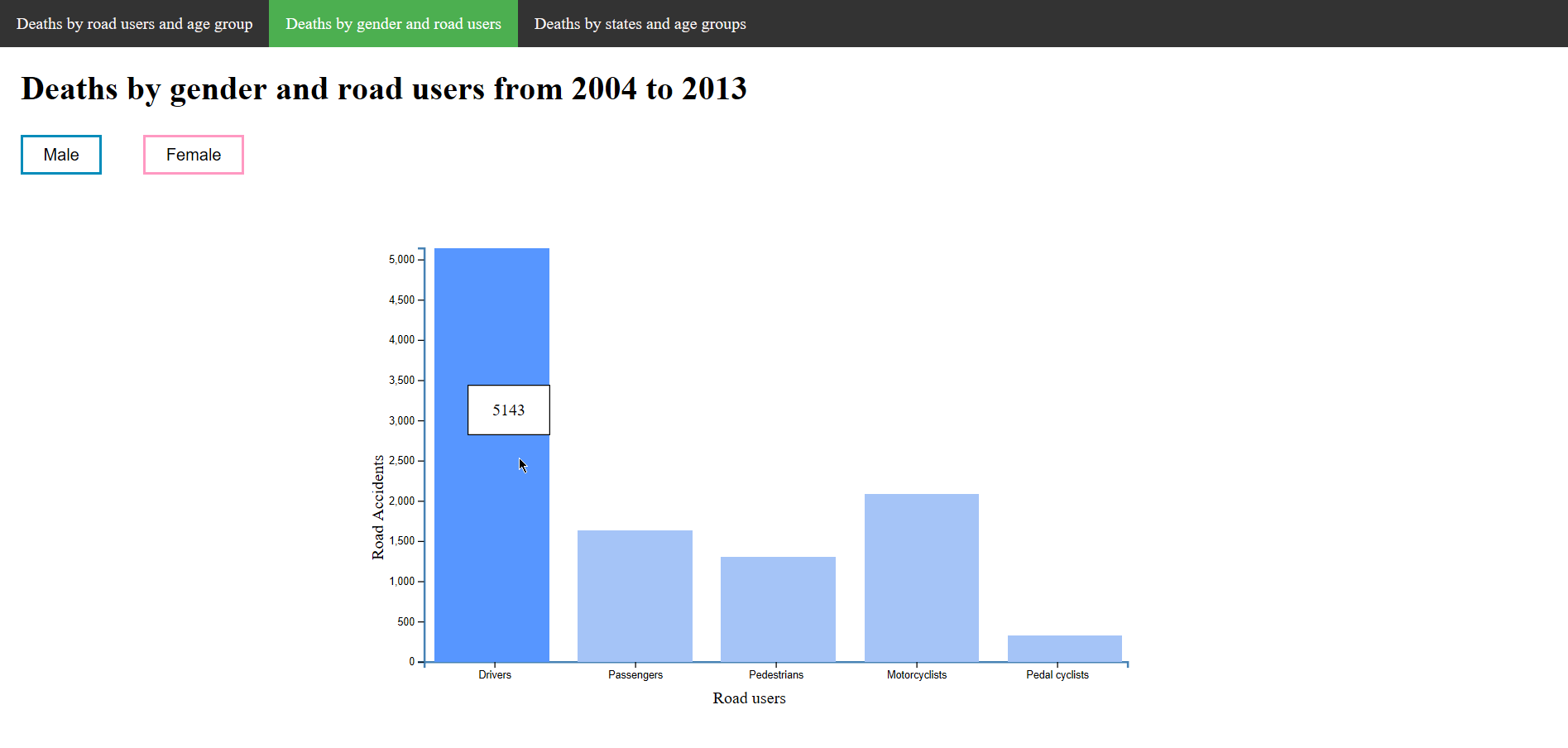


figure 8: visualisation 2—Deaths by gender and road users

This is the final design of my bar chart. This bar chart shows the deaths by genders and road users from 2004 to 2013. I gather all the data from 2004 to 2013 of each type of road users and show it in bar chart. As promised feature, the exact number of injuries will be shown when the mouse move over the bar. Not only number, the color of the bar will change deeper as well in order to create color contrast.

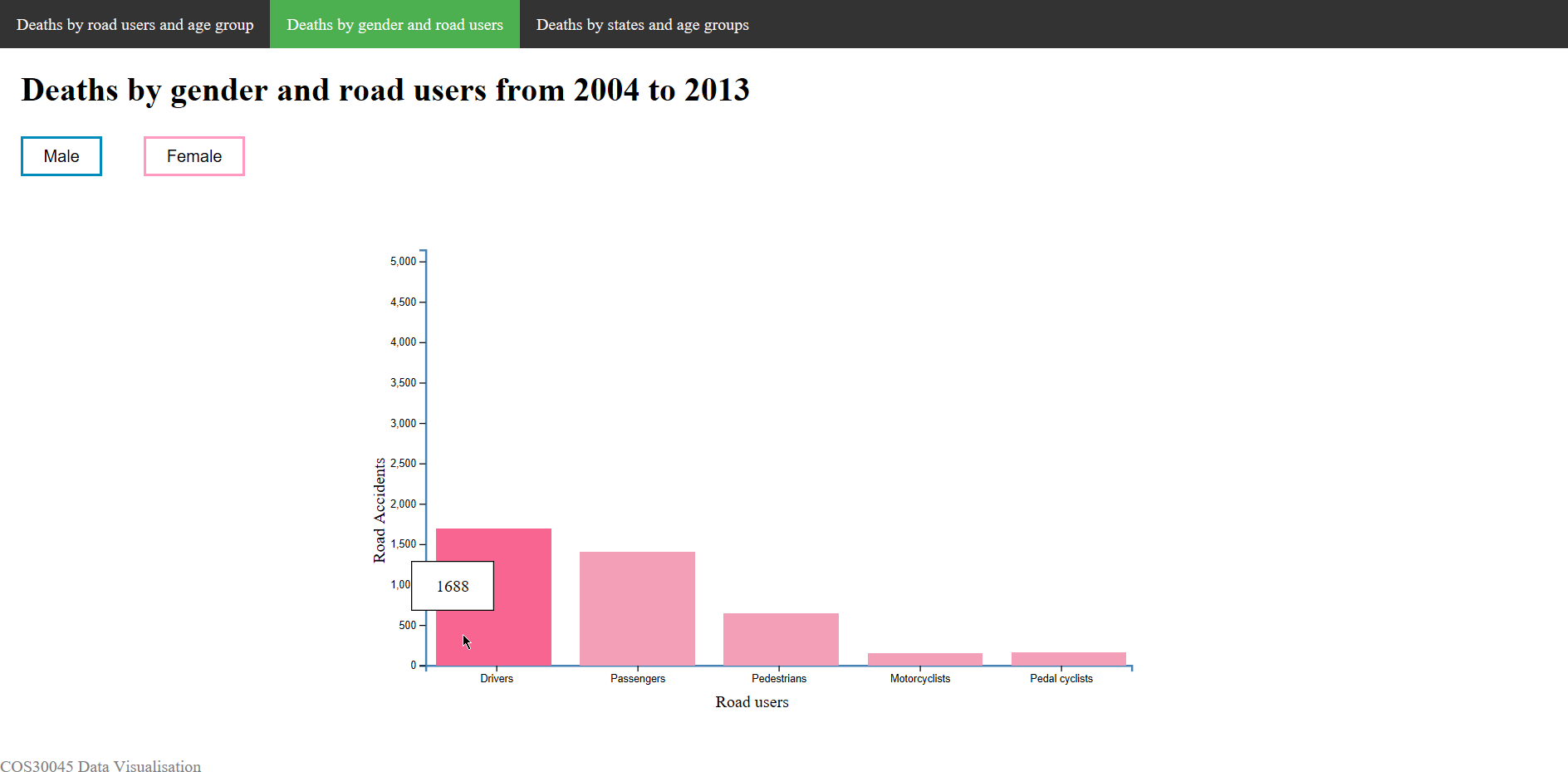


figure 9: visualisation 2—Deaths by gender and road users

This bar chart shows the death by female and road users from 2004 to 2013. When reader click on the female button, the bars and colors will change to same color with the button. So the reader can tell which type of data are they watching.

1. **Deaths by states and age groups**

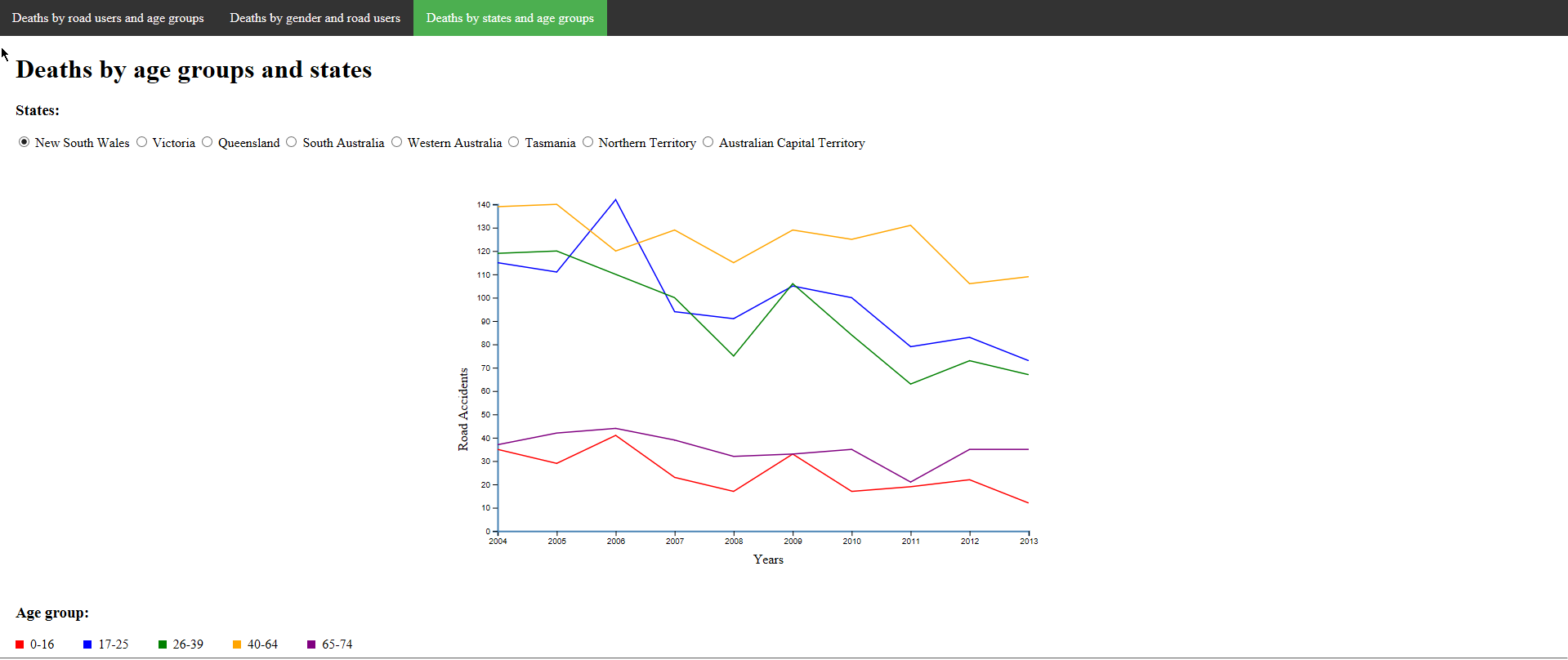


figure 10: visualisation3--deaths by age group and states

The design is exactly the same with the first one. 5 lines with different colors in the chart and each color represent different age groups. There are 5 radio button group which are New South Wales, Victoria, Queensland, South Australia, Western Australia, Tasmania, Northern Territory, Australian Capital Territory.

# Conclusion

This project provides me a great opportunity to learn about visualize data by using D3 and how to analysis data in real world. It also allow me to understand design concepts while comes to design the line chart and bar chart. Furthermore,