Mercury % github link

Table of Contents

1 Introduction

According to Australian Bureau of Statistics (2021), around 30% of Australia’s population comprises immigrants. Among these, the United Kingdom ranked as the top country of origin, followed by India and China. Migration has been a norm for humans, as people always move from one area to another area due to various reasons, such as better living environment, work opportunities or education purposes. In Australia, migration has played a crucial role in the country’s diversity, as there are millions of people from around the globe staying in Australia. Moreover, Thapa (2016) has stated that migrants have made a significant contribution to Australia’s economy, by adding around $330 billion to the GDP.

1.1 Background

The target audience for this visualisation includes but is not limited to policymakers, researchers, students and the general public. As the visualisation of migration data can provide insights, different categories of users are able to use it to carry out various tasks. For instance, policymakers may utilize the visualisations to gain a better understanding of complex data, which can be used to inform and update policies. Other than that, researchers could use the findings to analyse the pattern and trends of migration whereas students can utilize them to gain a better understanding of migration environment in Australia. In short, the visualisation of migration data is crucial as it provides valuable insights to various type of users, which can be utilized to inform policy decisions, analyse pattern and trends, gain better understanding of the migration environment and many more.

1.2 Visualisation purpose

The main purpose of the visualisation is to answer the following questions:

1. How many immigrants are coming to Australia from different countries, and how has this changed over time?
2. What are the different types of visas that migrants are using to come to Australia?
3. Is there a correlation between the number of international students and the number of migrants in Australia?

The visualisation of the first question can be used to understand migration patterns. With clear presentation of data, it allows users to gain a better understanding of migration patterns and trends in Australia. It could provide insights to policymakers and help them in making informed decisions about immigration policies.

1.3 Project Schedule

(To be Finalized)

|  |  |  |
| --- | --- | --- |
| Week # | Schedule | Task to be submitted |
| Week 3 | Getting started (Setting up communication channel, GitHub etc…) |  |
| Week 4 | Getting feedback from tutor and finalizing dataset. | Project Standup 1 |
| Week 5 | Find dataset and clean data |  |
| Week 6 | Starting on design (sketch) and process book. |  |
| Week 7 | Development of website and data visualization |  |
| Week 8 | Development of website and data visualization |  |
| Week 9 | Development of website and data visualization |  |
| Week 10 | Getting feedback from tutor and finalizing the website design | Project Standup 2 |
| Week 11 | Getting feedback from tutor and finalizing the website design | Project Standup 3 |
| Week 12 | Getting feedback from tutor and finalizing the website | Project Standup 4 |
| Week 13 |  | Project Process Book, Peer Assessment, Project Reflection, Final website |

2 Data

2.1 Data Source

Source 1: <https://www.abs.gov.au/statistics/people/population/overseas-migration/latest-release>  (Net overseas migration by country of birth, state/territory - financial years, 2004-05 to 2021-22)

A screenshot of a computer

Description automatically generated with medium confidence

Source 2: <https://www.abs.gov.au/statistics/people/population/overseas-migration/latest-release>  (Overseas migrant arrivals and departures by visa groupings, state/territory - financial years, 2004-05 to 2021-22)

A screenshot of a computer

Description automatically generated with medium confidence

Source 3: <https://www.education.gov.au/international-education-data-and-research/international-student-monthly-summary-and-data-tables#toc-international-student-data-full-year-data-based-on-data-finalised-in-december-2022-> (International Student Data – full year data (based on data finalised in December 2022)A screenshot of a computer

Description automatically generated with medium confidence

All of the datasets above are xlsx files, which contains the table of data required in building the visualization. All of the data are ratio data which contains the number of immigrants arriving in Australia from the year 2004 – 2022, the type of visa that immigrants arrived on from the year 2004 – 2022 and also the number of international student enrolment from the year 2002 – 2022. Some of the data in the dataset might need some cleaning as the dataset downloaded covers a wide range, for example in dataset 1 there are 9 tables in total in the xlsx file and therefore we only choose the one that we need (table 1.1). Other table or data that are not needed in building the dataset will be cleaned out so to make the csv file easier to access and read. A picture containing text, screenshot, font, number

Description automatically generated

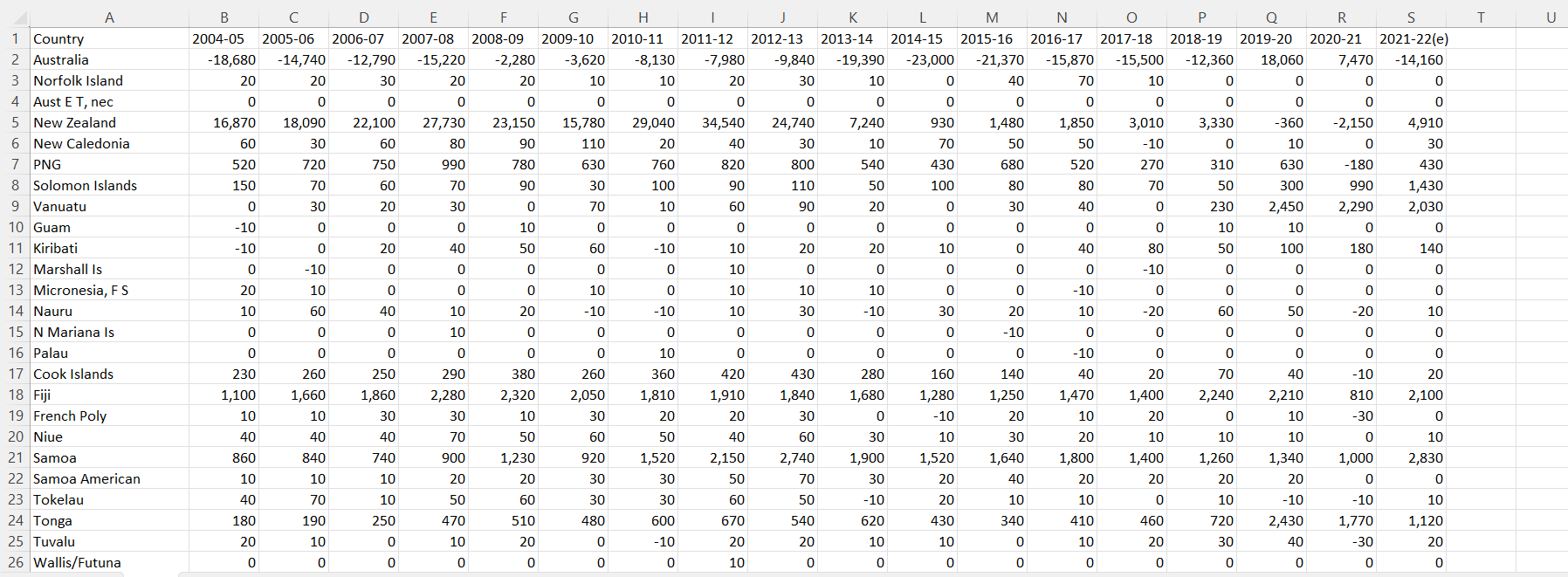
2.2 Data Processing

Dataset 1: Before Cleaning

A screenshot of a computer

Description automatically generated with medium confidence

After Cleaning

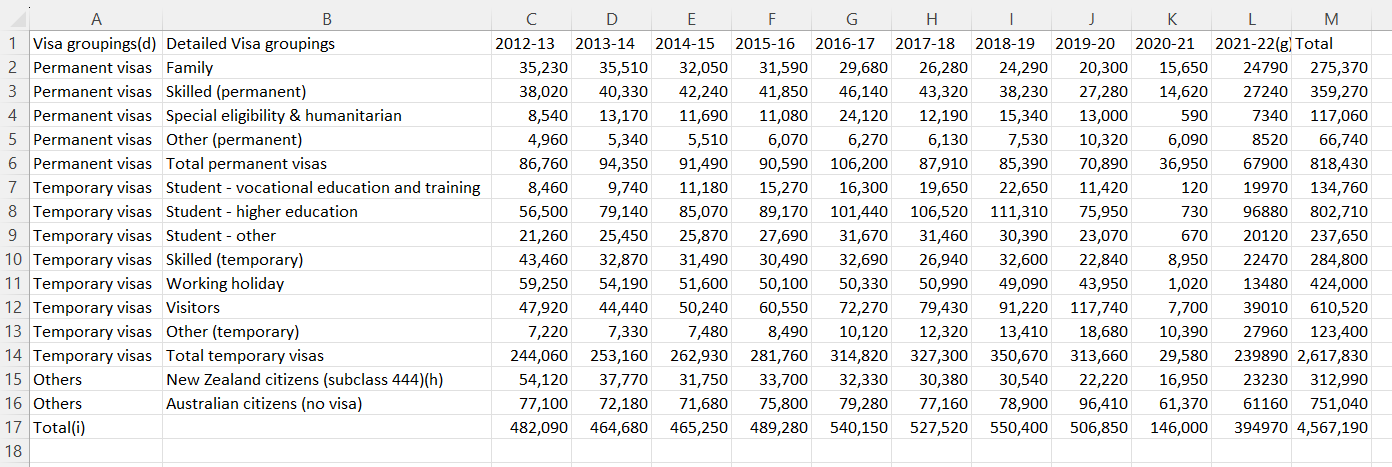


Manual cleanup of data was chosen as the method for cleaning up the dataset as we only required to remove the excessive data and contents that were in the dataset such as title and such and export it to a csv file to be further read in the code. No further alteration was made to the dataset such as the format of integer stored or the Country name as those will be processed inside the code. The data is then saved as csv format called “data1.csv”.

Dataset 2: Before CleaningA screenshot of a computer

Description automatically generated with medium confidence

After Cleaning



Manual cleanup of data was chosen as the method for cleaning up the dataset as we only required to remove the excessive data and contents that were in the dataset such as title and such and export it to a csv file to be further read in the code. Grouping of the visa was added and excessive year such as data from 2004-2011 was remove as the data was too old to be used for analysis and building the visualisation.

A variable called total is also added at the last column and is derived from the sum of visa type of immigrants in the past 10 years to be used in building the visualization. The calculation of the derived variable is calculated using the =SUM() method inside Microsoft Excel. The data is then saved as csv format called “data2.csv”.

Dataset 3: Before CleaningA screenshot of a computer

Description automatically generated with medium confidence

After CleaningA picture containing text, font, line, number

Description automatically generated

Manual cleanup of data was chosen as the method for cleaning up the dataset as we only required to remove the excessive data and contents that were in the dataset such as title and such and export it to a csv file to be further read in the code. Filters that was needed is being filtered to get the data that we required such as filtering all nationality of international student, in every state of Australia and so on. Excessive data such as Commencement data of international student are also removed from the dataset as we only required the use of Enrollment data from the dataset.

A variable called Average is also added at the last row and is derived from the average of international student throughout the whole year for each year from 2002 – 2022 to be used in building the visualization. The calculation of the derived variable is calculated using the =AVERAGE() method inside Microsoft Excel. The data is then saved as csv format called “data3.csv”.

3 Requirements

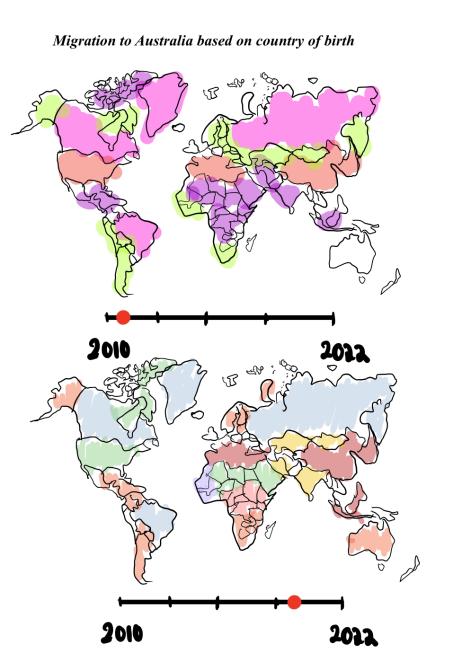
3.1 Must-Have Features   
These are features without which you would consider your project to be a fail-  
ure. Were you able to deliver all the promised features? If not, explain why.

3.2 Optional Features   
Those features which you consider would be nice to have, but not critical.   
Were you able to deliver any of these extra features?

4 Visualisation Design   
How will you display your data? Provide some general ideas that you have for the vi-  
sualisation design. Include sketches of your design. Include at lease 2-3 alternative   
ideas for your visualisation. Describe and justify your choice of visual encoding and   
idioms. Show the evolution of your design. How has it progressed? Justify the visual-  
isation idioms you have chosen to represent your data.   
Description (including screen shots) and explanation of final design.   
[NOTE 1: You are encouraged to provide your own structure to this section (i.e., section headings etc).   
NOTE 2: You MUST show evidence of iterative design (i.e., sketches of alternative and preliminary designs). ]   
Include screenshots of final design.

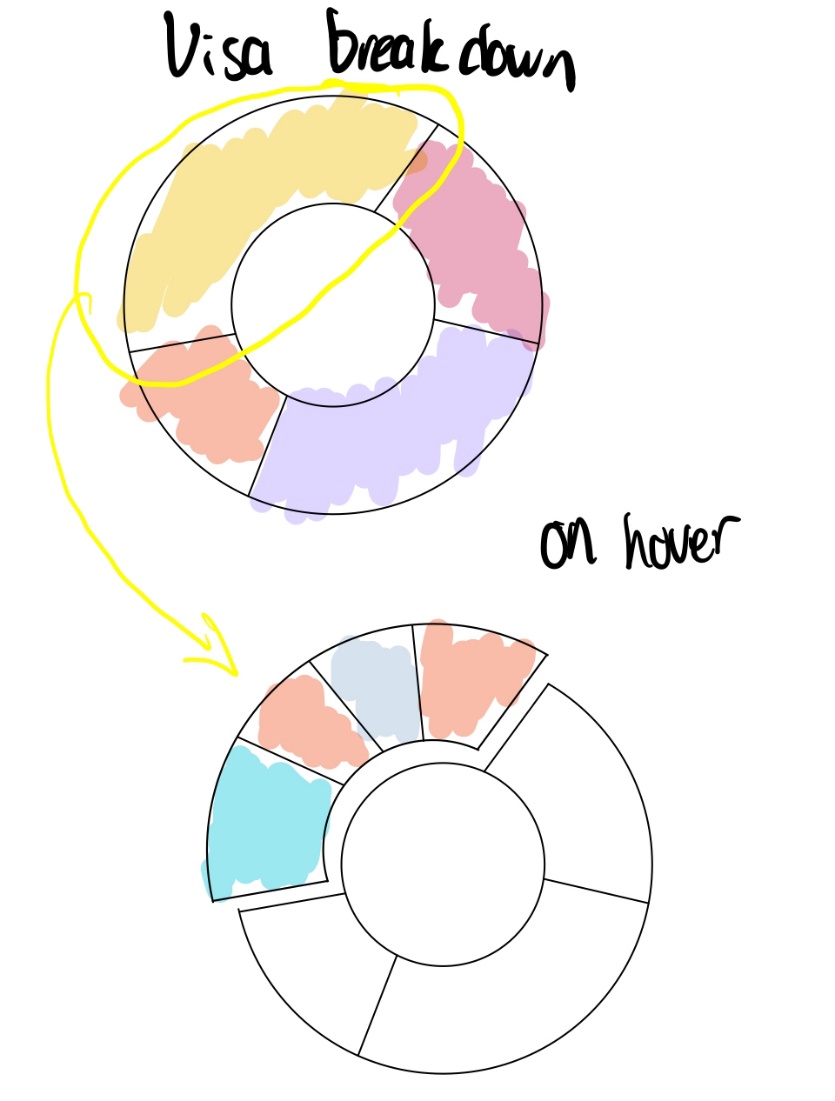
Visualisation 1:



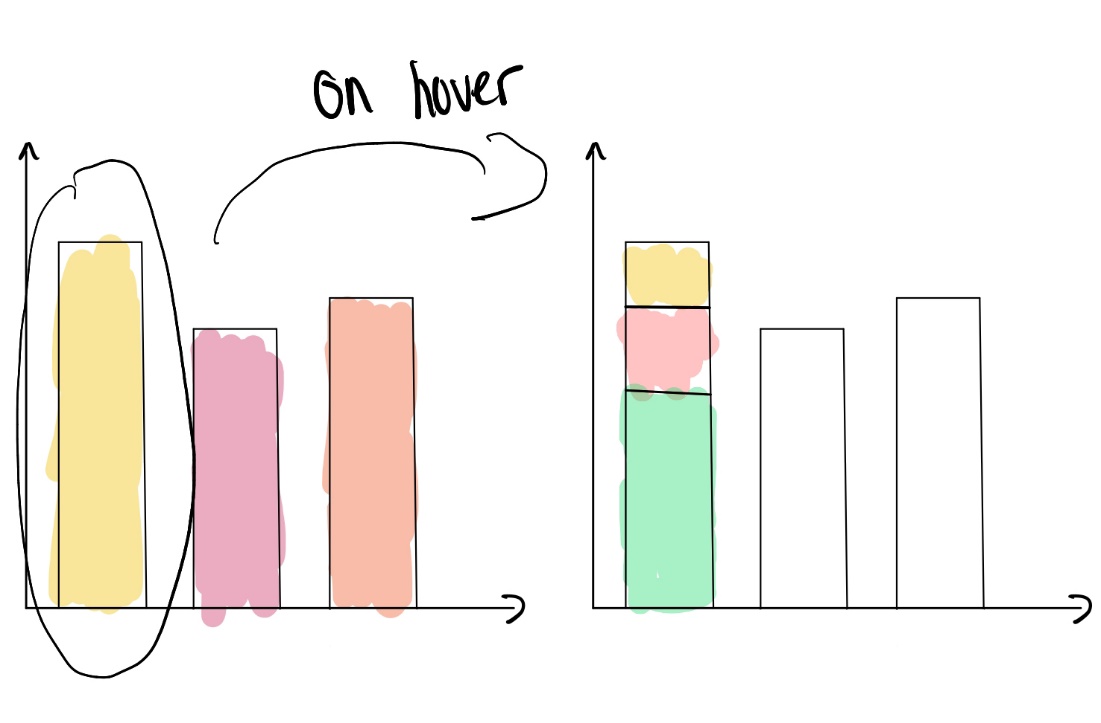


Visualisation 2:

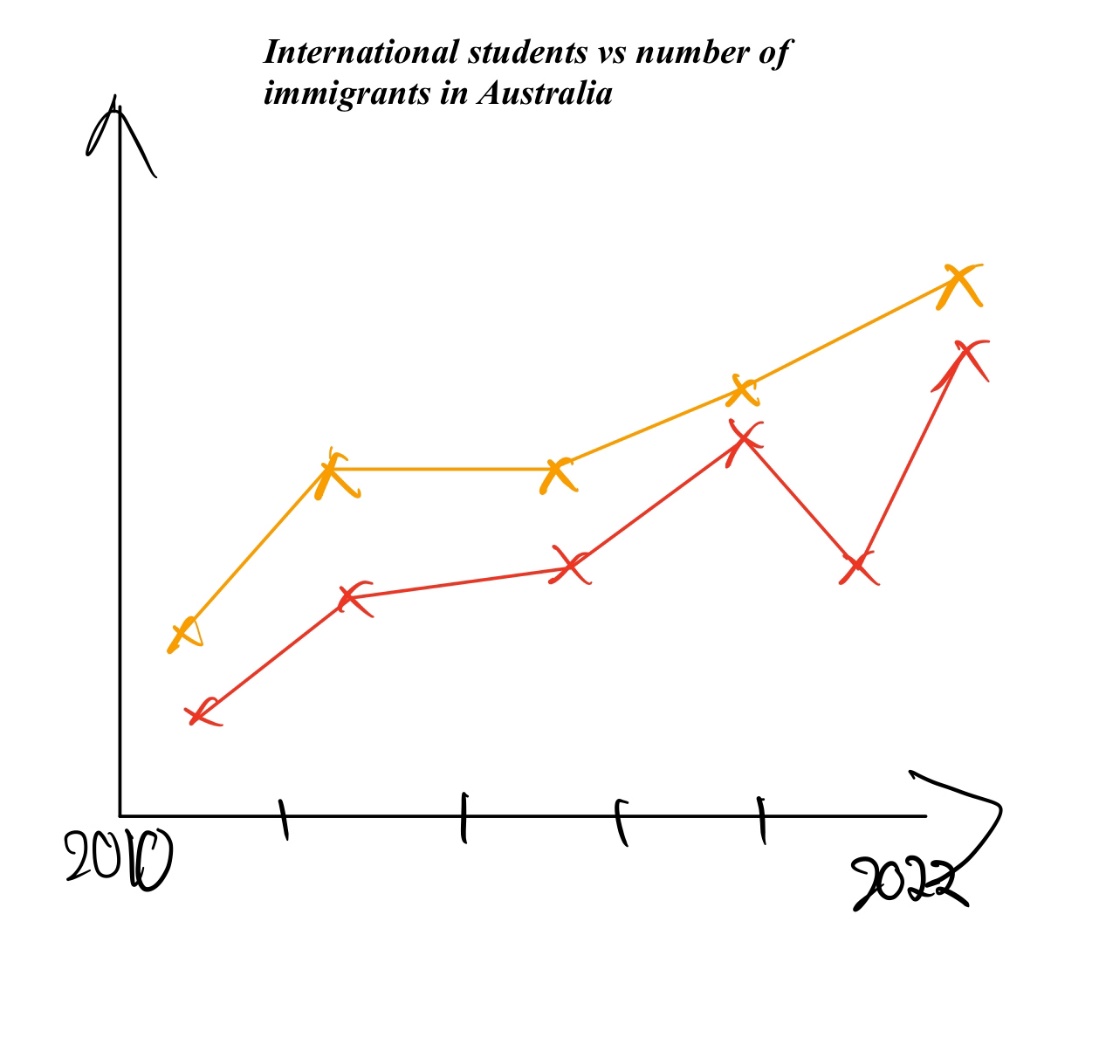
option1



Option 2



Visualization 3



5 Validation [optional - Bonus Points]   
Test your visualisation with users and report the results.

6 Conclusion   
Provide a summary of the project and what you learnt from doing it.   
References   
References consulted (blogs, books, academic papers, discussion/help forums - for   
both design and programming)