SWINBURNE UNIVERSITY OF TECHNOLOGY

SCHOOL OF SCIENCE, COMPUTING

AND ENGINEERING TECHNOLOGIES

===***===

COS30045 – DATA VISUALISATION

PROCESS BOOK

AUSTRALIAN MIGRATION TRENDS AND PATTERNS

Gia Hung Tran (103509199)

Thao Ngan Ngo (104055130)

Tutorial: Thursday 12:30 PM – 2:30 PM

Link to website: Migration (checkiejan.github.io) | mercury.swin.edu.au/cos30045/s103509199/project/

Word count: 9106

Executive Summary

The goal of this process book is to chronicle our group's progress in developing our data visualisation project "Australian Migration Trends and Patterns". It outlines our systematic approach to displaying data-driven graphics in a web browser, documenting each step from our initial selection of project questions through the selection and cleaning of the suitable dataset to the formulation of graph requirements.

The process book describes our though testing method, where we experimented with numerous graphic representations on the same dataset before arriving on our final design. Furthermore, it records our initiative to conduct usability testing, where participants were asked to execute a series of tasks using our visualisations. We identified the most common problems that our users experienced and investigate the possible causes of these problems.

Throughout this project, we discovered that the design of visualisations can greatly vary even with a single dataset and question, each with its own set of benefits and drawbacks. This process book serves not just as a record of our experiences, but as a reflective pierce, offering insights into our design decisions and illuminating the fascinating learning journey we undertook.

Table of Contents

Table of Figures

1. Introduction

1.1 Background and Motivation

Australia is the world's six-largest landmass and has had a long history of migration, which has significantly influenced the development of its society and economy (Geoscience Australia, 2023). Although indigenous people had lived on the continent for over sixty thousand years, in the 18th century, Europeans entered Australia and started colonising, which resulted in considerable population shifts. Following then, waves of immigration from other countries have added to Australia's cultural variety (About Australia, n.d).

This project's goal is to investigate migration to and from Australia through the use of various visualisations. We are focusing on gathering and displaying data on the total number of long-term arrivals and departures in the past two decades, conducting further research on the countries which rank in the top 10 by number of citizens who have migrated to Australia, and analysing the make-up and movements of different categories of visa holders between Australian states.

By visualising this data, we intend to make the information easily accessible and understandable to our target audience. Through clear and visually appealing representations, we aim to highlight the trends and changes in migration flows, allowing users to better understand the patterns and factors influencing migration to Australia.

Our target audience is people who are interested in moving to Australia for reasons such as searching for work, seeking better economic opportunities, or a suitable environment for studying abroad.

1.2 Project Objectives

We aim to use our visualisations to provide insights into the following questions:

- What patterns can be discerned in the migration and immigration trends in Australia over the past two decades?
- Which ten countries predominantly contribute to the migrant population in Australia, and how have those trends evolved over time?
- What are the main categories of visa holders moving to each state in Australia, and how do these compare across states?

Our visualisations will provide a clear and intuitive way to answer these questions, allowing users to quickly understand and interpret the data and identify key insights effectively.

1.3 Project Schedule

To ensure the completion of the visualisations and the Process Book within the given deadline, our group established the following schedule:

- Weeks 1 through 6:

    - Collect the necessary data on migration in Australia (including arrivals, departures, countries of origin, and temporary visa holders.)
    - Start brainstorming ideas for visualisations while sketching the design concepts.
- Week 7:
    - Determine the most appropriate visualisation types for each dataset, considering the target audience and the information to be conveyed.
    - Select the effective graph types, charts, and maps that will enhance comprehension and engagement.
- Week 8:
    - Begin processing the datasets and developing visualisations, while still continuing to research other possible sources of data.
- Week 9:
    - Implement the chosen designs, incorporating the cleaned and pre-processed data, and start writing the Process Book.
    - Ensure the visualisations are interactive so that users may explore and manipulate the data based on their interests.
- Week 10:
    - Compile the visualisations, accompanying explanations, and analysis into a cohesive and well-structured Process Book.
- Week 11:
    - Continue programming the visualisations for the project and writing the Process Book.
- Week 12 to the due date:
    - Finish writing the Process Book, ensuring clarity, coherence, and accuracy of the content, checking for grammar, spelling, and formatting errors, and complete the visualisations.

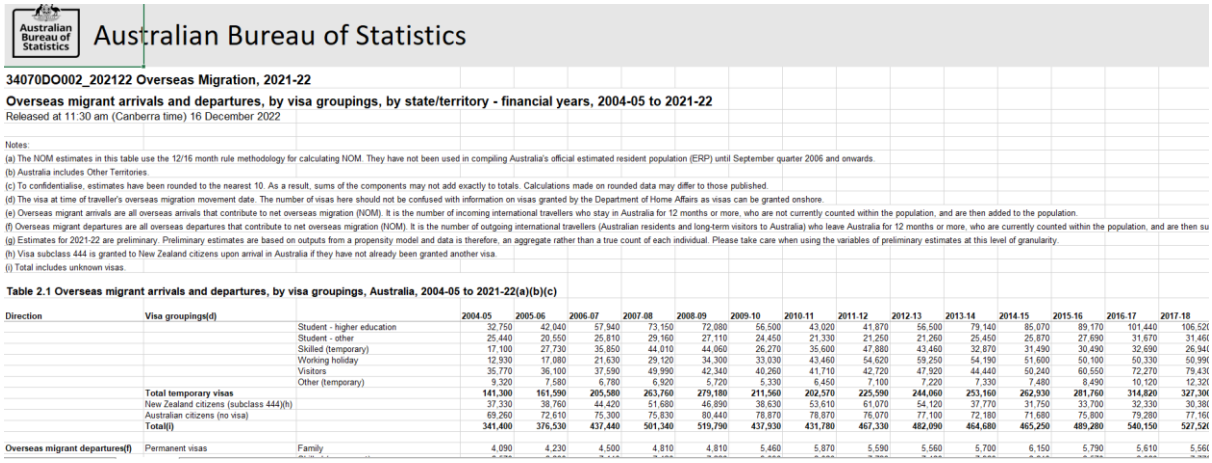
2. Datasets

2.1 Arrival and Departure Trends in Australia

2.1.1 Data Source

To ensure the reliability of our data sources, our focus was on government entities and large organisations. For gathering essential information on Australian migration, we relied on the Australian Bureau of Statistics (ABS). The ABS provides comprehensive data in the form of an Excel file, encompassing details about the number of people who arrived and departed each year, categorised according to their respective visa types. The file comprises multiple sheets, covering individual states as well as the entire country.

Each sheet presents a table, featuring columns that represent the record years, visa groupings, and directions of migration. In total, there are nine sheets for Australia, each corresponding to a specific state. The "direction" attribute denotes whether the migrant arrived or departed, while the "visa grouping" attribute classifies individuals into various categories, such as permanent and temporary visa groups. Both "direction" and "visa grouping" attributes are considered nominal data, whereas the "year" attribute represents ratio data, indicating the number of people involved.



Australian Bureau of Statistics

34070DO002_202122 Overseas Migration, 2021-22

Overseas migrant arrivals and departures, by visa groupings, by state/territory - financial years, 2004-05 to 2021-22

Released at 11:30 am (Canberra time) 16 December 2022

Notes:
(a) The NOM estimates in this table use the 12/16 month rule methodology for calculating NOM. They have not been used in compiling Australia's official estimated resident population (ERP) until September quarter 2006 and onwards.
(b) Australia includes Other Territories.
(c) To confidentialise, estimates have been rounded to the nearest 10. As a result, sums of the components may not add exactly to totals. Calculations made on rounded data may differ to those published.
(d) The visa at time of traveller's overseas migration movement date. The number of visas here should not be confused with information on visas granted by the Department of Home Affairs as visas can be granted onshore.
(e) Overseas migrant arrivals are all overseas arrivals that contribute to net overseas migration (NOM). It is the number of incoming international travellers who stay in Australia for 12 months or more, who are not currently counted within the population, and are then added to the population.
(f) Overseas migrant departures are all overseas departures that contribute to net overseas migration (NOM). It is the number of outgoing international travellers (Australian residents and long-term visitors to Australia) who leave Australia for 12 months or more, who are currently counted within the population, and are then su
(g) Estimates for 2021-22 are preliminary. Preliminary estimates are based on outputs from a propensity model and data is therefore, an aggregate rather than a true count of each individual. Please take care when using the variables of preliminary estimates at this level of granularity.
(h) Visa subclass 444 is granted to New Zealand citizens upon arrival in Australia if they have not already been granted another visa.
(i) Total includes unknown visas.

Table 2.1 Overseas migrant arrivals and departures, by visa groupings, Australia, 2004-05 to 2021-22(a)(b)(c)

Direction	Visa groupings(d)		2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18
		Student - higher education	32,750	42,040	57,940	73,150	72,080	56,500	43,020	41,870	56,500	79,140	85,070	89,170	101,440	106,52
		Student - other	25,440	20,550	25,810	29,160	27,110	24,450	21,330	21,250	21,260	25,450	25,870	27,690	31,670	31,46
		Skilled (temporary)	17,100	27,730	35,850	44,010	44,060	26,270	35,600	47,880	43,460	32,870	31,490	30,490	32,690	26,94
		Working holiday	12,930	17,080	21,630	29,120	34,300	33,030	43,460	54,620	59,250	54,190	51,600	50,100	50,330	50,99
		Visitors	35,770	36,100	37,590	49,990	42,340	40,260	41,710	42,720	47,920	44,440	50,240	60,550	72,270	79,43
		Other (temporary)	9,320	7,580	6,780	6,920	5,720	5,330	6,450	7,100	7,220	7,330	7,480	8,490	10,120	12,32
	Total temporary visas		**141,300**	**161,590**	**205,580**	**263,760**	**279,180**	**211,560**	**202,570**	**225,590**	**244,060**	**253,160**	**262,930**	**281,760**	**314,820**	**327,30**
	New Zealand citizens (subclass 444)(h)		37,330	38,760	44,420	51,680	46,890	38,630	53,610	61,070	54,120	37,770	31,750	33,700	32,330	30,38
	Australian citizens (no visa)		69,260	72,610	75,300	75,830	80,440	78,870	78,870	76,070	77,100	72,180	71,680	75,800	79,280	77,16
	Total(i)		**341,400**	**376,530**	**437,440**	**501,340**	**519,790**	**437,930**	**431,780**	**467,330**	**482,090**	**464,680**	**465,250**	**489,280**	**540,150**	**527,52**
Overseas migrant departures(f)	Permanent visas	Family	4,090	4,230	4,500	4,810	4,810	5,460	5,870	5,590	5,560	5,700	6,150	5,790	5,610	5,56

Figure 1. Raw Dataset of Australian Overseas Migrant Arrivals and Departures

Additionally, for visualising the map of Australia, we have the JSON file available at https://raw.githubusercontent.com/rowanhogan/australian-states/master/states.geojson. This file contains the coordinates for the various states of Australia. It is formatted according to the standard GeoJSON format, making it convenient to use without requiring any adjustments.

2.1.2 Data Processing

The dataset we initially obtained was not in a suitable format for visualisation purposes, as it contained unorganised rows and unnecessary columns. To address this, we utilised the Pandas library in conjunction with Jupyter Notebook to process the data. The code responsible for data cleaning can be found at Data-Processing/process.

Our first step involved examining each row to identify any null values and subsequently dropping them to ensure the consistency of the dataset. Next, we selected the specific attributes we wished to visualise. This included collecting all the year columns and corresponding rows for total arrivals and departures. Additionally, we gathered the same data segmented by state, resulting in separate CSV files for arrivals and departures.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
State	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
New South Wales	125180	132750	150600	170960	169020	144210	141980	140860	152630	155530	160910	170890	198300	191450	194270	172430	70530	145820
Victoria	81230	89860	108600	123450	132640	109130	104900	112280	119250	122250	128170	140220	159630	161170	168310	150640	27140	120480
Queensland	65390	73070	86490	99670	103110	84890	84250	93210	92800	83800	78980	83170	87820	85440	91660	85590	23440	58580
South Australia	16620	20550	23490	25160	27260	24820	20160	22980	22580	23620	23280	23410	23820	24100	26650	27740	6020	20940
Western Australia	39980	47170	53830	66150	69710	57920	64930	77980	74380	60030	53290	51540	48640	44190	48080	50400	12970	30560
Tasmania	3170	3400	3440	3900	4290	4010	3470	3780	3920	4190	4210	4210	4700	5110	5880	6510	1610	5150
Northern Territory	4430	3910	4560	4710	5540	4670	4810	6750	7840	6260	6900	6370	6900	5020	5550	4120	1450	4910
Australian Capital Territory	5410	5830	6420	7350	8190	8270	7270	9490	8690	8990	9520	9470	10330	10970	9950	9380	2810	8450
Australia	341410	376540	437430	501350	519760	437920	431770	467330	482090	464670	465260	489280	540140	527450	550350	506810	145970	394890

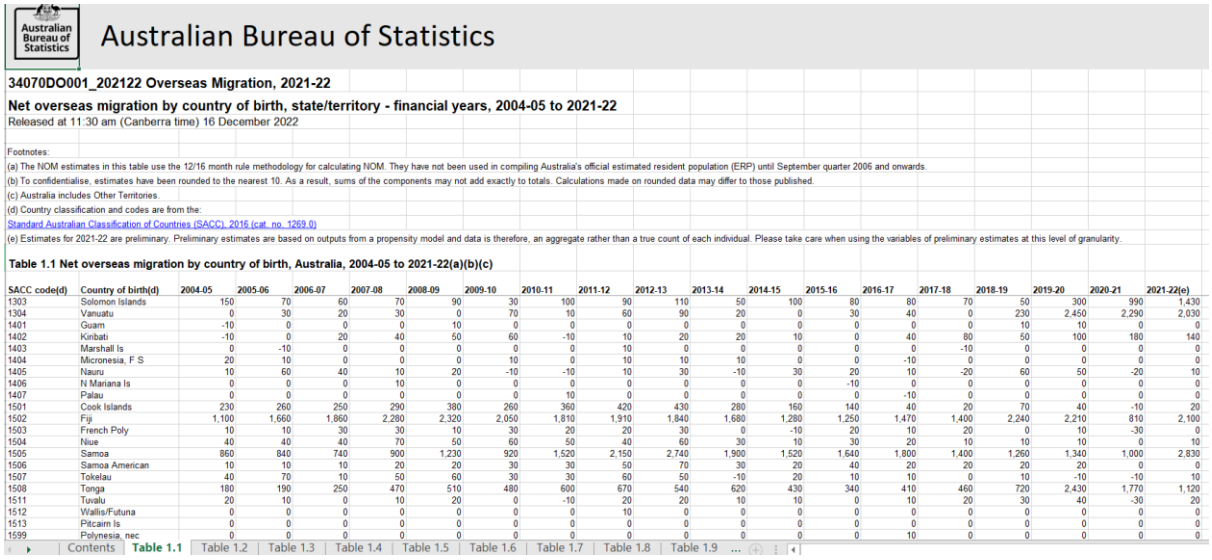
Figure 2. Data After Processed

The final dataset comprises the "state" attribute, representing nominal data for both individual states and Australia as a whole. The remaining year columns represent ratio data, indicating the number of people who arrived or departed, depending on the respective file.

2.2 Top 10 Source Countries for Migration to Australia by State

2.2.1 Data Source

The Australian Bureau of Statistics (ABS) has made available a comprehensive dataset on overseas migration categorised by country of birth. The dataset can be found at this link(https://www.abs.gov.au/statistics/people/population/overseas-migration/latest-release#net-overseas-migration). This dataset offers valuable insights into migration patterns to and from Australia. It is presented in the form of an Excel file, with individual sheets corresponding to each state. The dataset records data from 2004 to 2021, with each year column representing the ratio data of net overseas migration—the net gain or loss of population through international migration (abs.gov.au, n.d). Furthermore, the attributes "SACC" and "country of birth" represent nominal data, providing information on each specific country involved in the migration flows.



Australian Bureau of Statistics

34070DO001_202122 Overseas Migration, 2021-22

Net overseas migration by country of birth, state/territory - financial years, 2004-05 to 2021-22

Released at 11:30 am (Canberra time) 16 December 2022

Footnotes:

(a) The NOM estimates in this table use the 12/16 month rule methodology for calculating NOM. They have not been used in compiling Australia's official estimated resident population (ERP) until September quarter 2006 and onwards.

(b) To confidentialise, estimates have been rounded to the nearest 10. As a result, sums of the components may not add exactly to totals. Calculations made on rounded data may differ to those published.

(c) Australia includes Other Territories.

(d) Country classification and codes are from the:

Standard Australian Classification of Countries (SACC), 2016 (cat. no. 1269.0)

(e) Estimates for 2021-22 are preliminary. Preliminary estimates are based on outputs from a propensity model and data is therefore, an aggregate rather than a true count of each individual. Please take care when using the variables of preliminary estimates at this level of granularity.

Table 1.1 Net overseas migration by country of birth, Australia, 2004-05 to 2021-22(a)(b)(c)

SACC code(d)	Country of birth(d)	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22(e)
1303	Solomon Islands	150	70	60	70	90	30	100	90	110	50	100	80	80	70	50	300	990	1,430
1304	Vanuatu	0	30	20	30	0	70	10	60	90	20	0	30	40	0	230	2,450	2,290	2,030
1401	Guam	-10	0	0	0	0	10	0	0	0	0	0	0	0	0	0	10	10	0	0
1402	Kiribati	-10	0	20	40	50	60	-10	10	20	20	10	0	40	80	50	100	180	140
1403	Marshall Is	0	-10	0	0	0	0	0	10	0	0	0	0	0	-10	0	0	0	0
1404	Micronesia, F S	20	10	0	0	0	10	0	10	10	10	0	0	-10	0	0	0	0	0
1405	Nauru	10	60	40	10	20	-10	-10	10	30	-10	30	20	10	-20	60	50	-20	10
1406	N Mariana Is	0	0	0	10	0	0	0	0	0	0	0	-10	0	0	0	0	0	0
1407	Palau	0	0	0	0	0	0	10	0	0	0	0	0	-10	0	0	0	0	0
1501	Cook Islands	230	260	250	290	380	260	360	420	430	280	160	140	40	20	70	40	-10	20
1502	Fiji	1,100	1,660	1,860	2,280	2,320	2,050	1,810	1,910	1,840	1,680	1,280	1,250	1,470	1,400	2,240	2,210	810	2,100
1503	French Poly	10	10	30	30	10	30	20	20	30	0	-10	20	10	20	0	10	-30	0
1504	Niue	40	40	40	70	50	60	50	40	60	30	10	30	20	10	10	10	0	10
1505	Samoa	860	840	740	900	1,230	920	1,520	2,150	2,740	1,900	1,520	1,640	1,800	1,400	1,260	1,340	1,000	2,830
1506	Samoa American	10	10	10	20	20	30	30	50	70	30	20	40	20	20	20	20	0	0
1507	Tokelau	40	70	10	50	60	30	30	60	50	-10	20	10	10	0	10	-10	-10	10
1508	Tonga	180	190	250	470	510	480	600	670	540	620	430	340	410	460	720	2,430	1,770	1,120
1511	Tuvalu	20	10	0	10	20	0	-10	20	20	10	10	0	10	20	30	40	-30	20
1512	Wallis/Futuna	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0
1513	Pitcairn Is	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1599	Polynesia, nec	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0

Contents | Table 1.1 | Table 1.2 | Table 1.3 | Table 1.4 | Table 1.5 | Table 1.6 | Table 1.7 | Table 1.8 | Table 1.9 ...

Figure 3. Raw Dataset of Net Overseas Migration by Country of Birth

2.2.2 Data Processing

The data presented in Figure 3 is intended for a Sankey diagram, but the raw dataset itself cannot be effectively utilised for this purpose. Although the dataset is clean, free of any null values, the inclusion of some derived variables is necessary. Our focus was on visualising the top ten countries migrating to Australia during three distinct periods (i.e., 2004-2009, 2010-2015, 2016-2021). To determine these top countries, we aggregated the values across all years within each period, sorted them in descending order, then obtained the final list. Utilising this list, we calculated the corresponding values for each country within each state and exported the results to separate CSV files. The script responsible for deriving this data can be found at Data-Processing/sankey.

A	B	C
source	target	value
Vietnam	Western Australia	2580
Vietnam	Victoria	14270
Vietnam	Tasmania	150
Vietnam	South Australia	2050
Vietnam	Queensland	3910
Vietnam	Northern Territory	200
Vietnam	New South Wales	11410
Vietnam	Australian Capital Territory	500
UK	Western Australia	44170
UK	Victoria	27340
UK	Tasmania	1100
UK	South Australia	14660
UK	Queensland	36670
UK	Northern Territory	880
UK	New South Wales	39900
UK	Australian Capital Territory	1470
Sri Lanka	Western Australia	2480
Sri Lanka	Victoria	16460
Sri Lanka	Tasmania	130
Sri Lanka	South Australia	1320
Sri Lanka	Queensland	2850
Sri Lanka	Northern Territory	110
Sri Lanka	New South Wales	6030
Sri Lanka	Australian Capital Territory	490

Figure 4. Top 10 Australian Migration Based on Country of Birth

After processing, we obtained three CSV files, each corresponding to one of the three periods. The data was further filtered to ensure there were no negative values. The attributes "source" and "target" represent nominal data, indicating the country from which individuals migrated and the state to which they migrated, respectively. The “value” attribute represents ratio data, indicating the net oversea migration of the respective period.

2.3 Reasons Given for Migrating to Australia by Visa Type

2.3.1 Data Source

The visualisation employs data drawn from the Australian Bureau of Statistics (ABS), specifically from the same file referenced in Figure 1. However, data was derived differently. This document provides such a rich exploration into migration trends, segmented according to visa types. It consists of individual sheets, each correlating to a specific state. Notably, the data presents two contrasting migration directions: arrivals and departures. While our first visualisation focused purely on total numbers, this current one delves deeper, unpacking component elements to explore migration patterns more intricately. Moreover, in this visualisation, we place our focus solely on arrival patterns.

2.3.2 Data Processing

To prepare required data for our visualisation, we extracted a variety of visa groupings associated with migration to Australia. In order to achieve that, we used Pandas library within the Jupyter Notebook environment to efficiently manipulate data. Initially, we selected rows wherein the visa type was either temporary or permanent, and subsequently categorised them based on the visa's purpose. For instance, within the 'education' category, there are diverse visa types such as higher education and vocational education. To streamline the dataset and enhance the visual clarity of the visualisation, these were combined into a single group. Similar process was applied to skilled visa types and other categories. The script responsible for processing the data can be found at Data-Processing/Line. While the original datasheet also includes data for returning Australian citizens, we chose to exclude this in favour of focusing on other visa types. The script, developed within the Jupyter Notebook, contains comprehensive comments to assist users in navigating the process.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
type	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Family	350	400	440	410	510	490	500	580	540	550	520	440	450	460	370	330	250	440
Special eligibility & humanitarian	220	190	90	170	190	120	130	70	50	150	120	190	290	190	210	170	10	100
Education	1150	1280	1610	1930	2360	2470	1980	2270	2440	3060	3320	3410	3970	4360	3330	2170	70	3720
Working holiday	80	100	120	170	190	230	250	400	370	310	270	260	280	350	260	190	10	70
Visitors	390	400	480	570	580	580	600	620	640	600	670	700	800	1120	1350	2190	120	800
Skilled	690	830	870	1130	1310	1200	900	1960	1330	1460	1690	1170	1250	1050	1070	750	280	700
Other	210	180	150	180	160	160	210	230	300	320	350	330	410	550	510	460	250	690

Figure 5. Oversea Migrant arrival To Australia by visa grouping

After processing, eight CSV files were generated, each corresponding to a state and to Australia as a whole. As depicted in Figure 5, these files serve as representative samples. The "type" attribute is nominal data denoting visa groupings. The subsequent year columns represent ratio data, signifying the count of individuals who migrated to Australia.

3. Requirements

3.1 Must-Have Features

To better determine the success metrics of the project, we defined some objectives in advance to lay the foundation to our work. A thorough understanding of the problem space, target audience, and the characteristics of data is a key factor to ensure a successful data visualisation (The Interaction Design Foundation, 2023). These were based on project requirements and some design guidelines to improve the readability of visualisation:

- There will be a dynamic, interactive choropleth map combining with a bar chart to elucidate migration patterns to Australia, delineated by states and over a specific timeline. Hence, this visualisation should allow users to discern the variability in migrant numbers across different states and through time. By simply hovering over components on the map, users should be provided with detailed information corresponding to the selected element. This feature not only adds depth to the data visualisation but also facilitates a more intuitive and engaging exploration of the migration data.

- There will be a chart to describe the top ten countries contributing to migration to Australia with the emphasis on the popularity of regions that are most frequently chosen as departure points. This chart should highlight the proportionate differences in migration volumes between each country, giving a clear perspective on the most popular origins of migrants.
- There will be a chart to explore the motivations behind people's decision to migrate to Australia. This data will be derived from grouping visa purposes, providing a comprehensive breakdown of the underlying reasons for migration. The chart will distinctively visualise the differential proportions of each motive, enabling a clear understanding of the popularity of various reasons.

To adhere to the best practices and standards of data visualisation, the following guidelines must be observed for all charts:

- Title: Each chart must feature a prominent title at its top, properly capitalised, and bolded, to succinctly communicate the core subject of the graph.
- Axis: Appropriate charts, such as bar and line charts, should have axes clearly labelled. The axes ought to be suitably scaled with a reasonable number of tick marks to provide comprehensive details for the chart's interpretation.
- Legend: Every chart should incorporate a legend explaining the representation of categorical data, often illustrated through colour hues. Additionally, the legend should clarify how gradations in colour luminance are employed to signify variations in data volume.
- Colour: The application of colour should conform to accessibility guidelines, carefully chosen to assist colour-blind readers in comprehending the visual data.
- Chart Description: Lastly, each chart should include a description detailing its content and providing clear guidance on the interpretation of complex visualisations. This will aid users in gaining a more thorough understanding of the data presented.

3.2 Optional Features

To further augment users' comprehension of the charts, we have identified a set of optional features. While these features could enrich the data visualisation experience, their omission, in case of time constraints, would not significantly compromise the quality of the data representation.

- Visualisation for arrivals and departures trends
  - The graph should include an automatic timeline progression, eliminating the need for user manual sliding.
  - Features such as zooming, and panning should be integrated to enhance user interaction with the chart.
  - Bars corresponding to the current year should be highlighted in the chart as the timeline progresses.

- When a specific state is selected, its corresponding data should be displayed in the bar chart.

- Visualisation for top ten countries migrating to Australia
  - The visualisation should allow users to select and view data from different time periods.
  - On selecting a particular source or destination, the graph should filter and display only the relevant node information.
- Visualisation for motivations of people migrating to Australia
  - A special tooltip should be used to sort different categories based on descending values when hovered over, enhancing the chart's readability.
  - A feature should be integrated to highlight the year corresponding to the line being hovered over, enabling users to easily track the associated data.

4. Visualisation Design

4.1 Arrival and Departure Trends in Australia

4.1.1 Initial Design 1: Line chart

Given the nature of the ratio data, which includes the number of arrivals and departures, we sought a visualisation method capable of effectively portraying value changes over time by utilising a magnitude channel. Consequently, we initially considered line marks paired with the position channel, leading us to design a line chart as our primary visualisation that presented the total number of individuals migrating to Australia from 2004 to 2022. The 'Year' attribute, being interval data, is encoded with a position channel and placed on the x-axis. Simultaneously, the 'value' attribute is represented via a position on a common scale channel, denoted precisely on the y-axis.

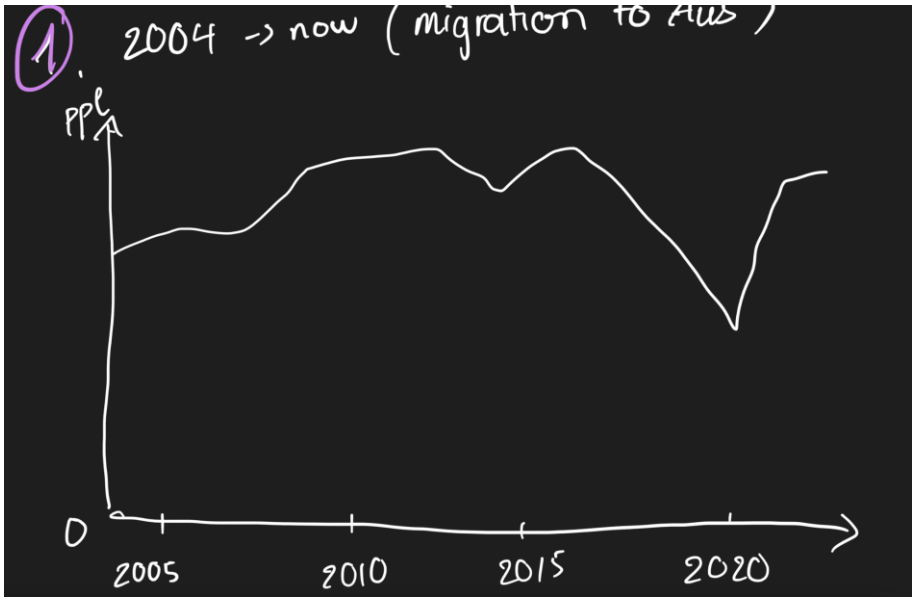


Figure 6. Line Chart for The Trend of People Moving to Australia for Each Year

This graph has some advantages of representing continuous data over an extended period, making it particularly suited to our dataset, which spans from 2004 to 2022. It

efficiently highlights trends and patterns, facilitating the identification of rises or declines in numbers. While its main objective is to underscore changes over several years, it may appear somewhat simplistic. This is mainly due to its lack of distinct visual elements that could potentially attract the audience and help them interpret the chart better. Thus, while it delivers on data accuracy, it may not be as engaging due to the absence of striking visual features.

4.1.2 Initial Design 2: Line chart

Interactivity significantly enhances a reader's experience and aids in comprehending complex datasets (Munzner et al., 2015). Reflecting on this, we have enriched our visualisation design by incorporating additional features to our line chart. Instead of only displaying the total number of migrants moving to Australia, the line charts can now illustrate the fluctuations in arrivals and departures across individual states in Australia. Utilising a drop-down menu, we can separately display the line chart for arrival and departure specifically within each state. Moreover, an interactive mouse hover feature has been implemented that illuminates the precise value of data points corresponding to each year on the line to provide a more engaging and informative user experience.

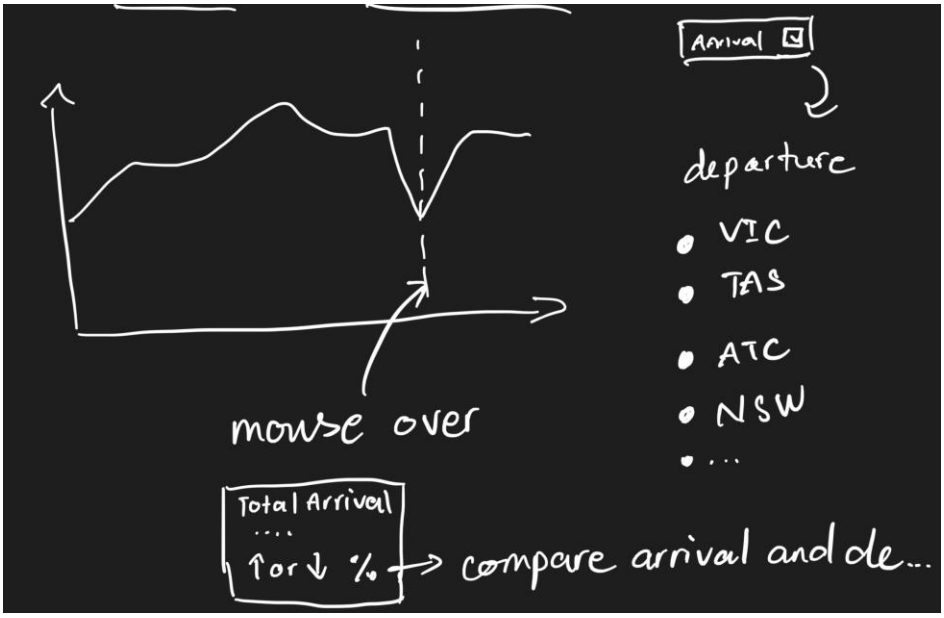


Figure 7. Line Chart for the Trend of the Arrivals and Departures in Each State in Australia

This graph clearly illustrates the fluctuation of migration and immigration numbers across different years and states like VIC, ATC, NSW, etc. Additionally, it has interactive elements that help users access more detailed information.

On the other hand, we decided against using this graphic because we considered that it presented an overwhelming amount of data. Additionally, it did not emphasise the contrast between arrivals and departures adequately, which was the primary focus of this chart. If we were to include the data for each state in the same graph, it would result in a congested visual space, making it challenging to follow the trends. Such a cluttered visualisation might reduce users' comprehension and diminish the effectiveness of the visual representation.

4.1.3 Initial Design 3: Choropleth Map with Bar Charts

Reflecting on our previous iterations, we recognised the need for a more effective way to allow readers to discern the variations in values among states while still highlighting temporal trends for these regions. Consequently, we chose a Choropleth map combined with bar charts for our third design iteration.

In this visualisation, area marks and colour saturation channels are employed to portray the ratio data of the number of arrivals or departures, enabling clear differentiation among states. Alongside the Choropleth map, a bar chart is utilised to present data about arrivals and departures for a specific state or Australia as a whole over a given time period.

Having the bar chart parallel with the Choropleth map mitigates the clutter often associated with standalone line charts, making it more comprehensible. The design also incorporates mouse-over interactivity, providing on-demand details such as annual totals of arrivals and departures and specific values for each year, which significantly enhances the user's interpretation and engagement.

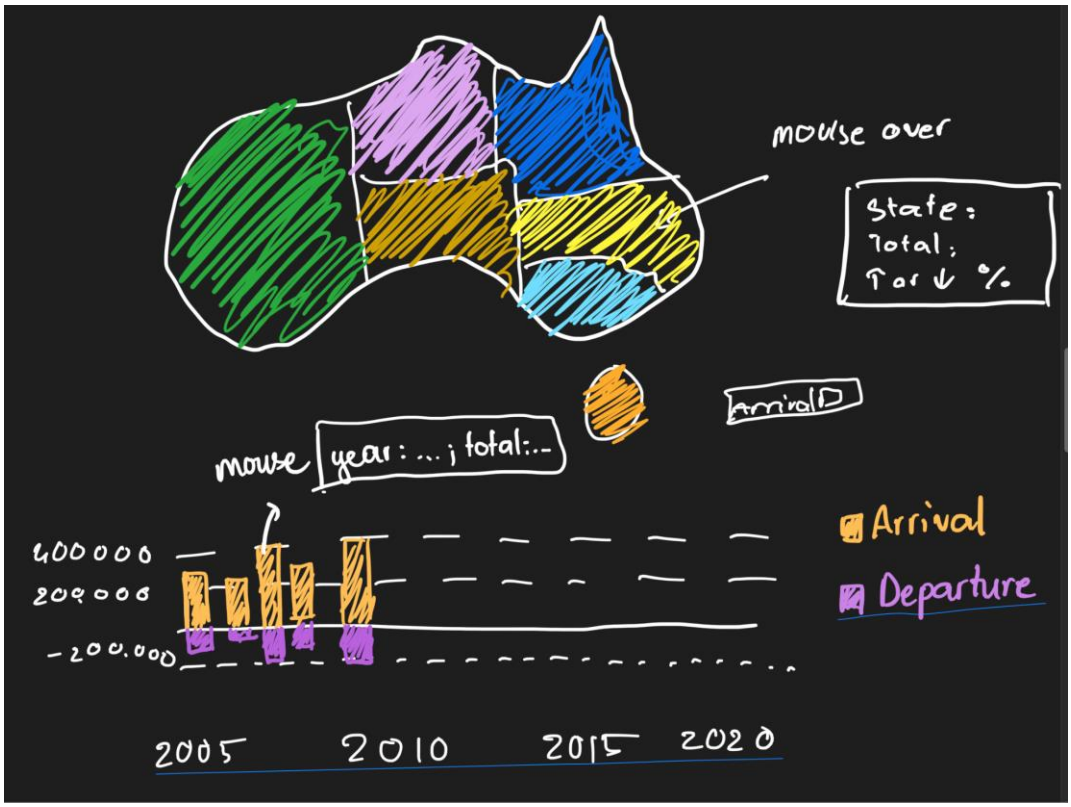


Figure 8. Choropleth Map and Bar Chart for the Trends of Arrivals and Departures in Each State in Australia

This visualisation offers the advantage of presenting the states in their true geographical positions while concurrently showing the data of arrivals or departures. This dual functionality significantly improves the readability of the chart. By employing a choropleth approach, viewers can gain a better understanding of the discrepancies in migration trends among different states. Additionally, the bar chart integrates both

arrivals and departures within the same visual space, presenting the departures upside down. This design enables users to easily draw comparisons between these two categories, which further enriches their understanding of the data.

4.1.4 Final Design

a. Graph Overview

To implement visualisation on the website, we decided to incorporate the design that integrates the Choropleth map and bar chart. We believe this design contains the optimal features for effectively communicating the data to our readers. It aids in imparting insightful information about the ongoing trends of immigration and migration across Australia to facilitate a deeper understanding of the subject matter.

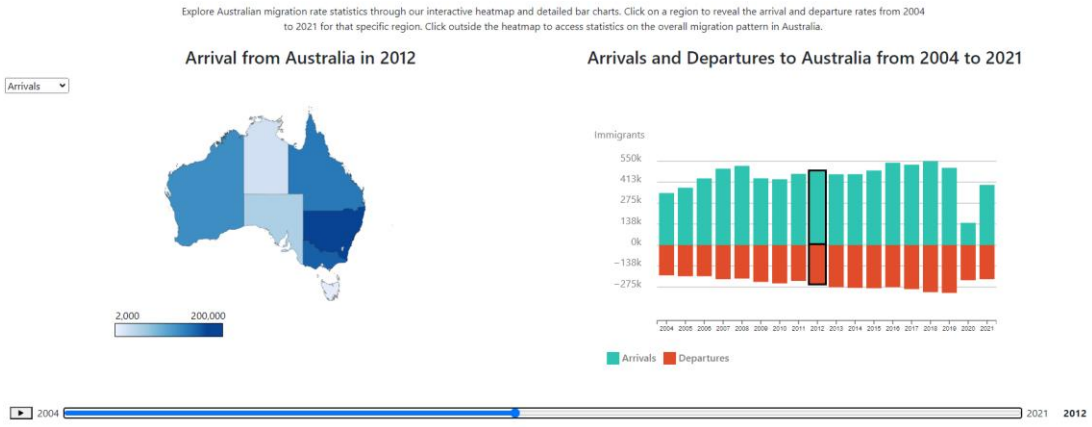


Figure 9. Website Implementation of Visualisation for Arrivals and Departures Trends across Australia

b. Graph Description

The "state" attribute in Figure 2, representing nominal data, is effectively visualised using area marks in the form of a geographical map. Since there are two separate charts, each chart has its own title. By default, the bar chart displays data for the entire Australia, but users can select data for an individual state by clicking on the corresponding region in the choropleth map.

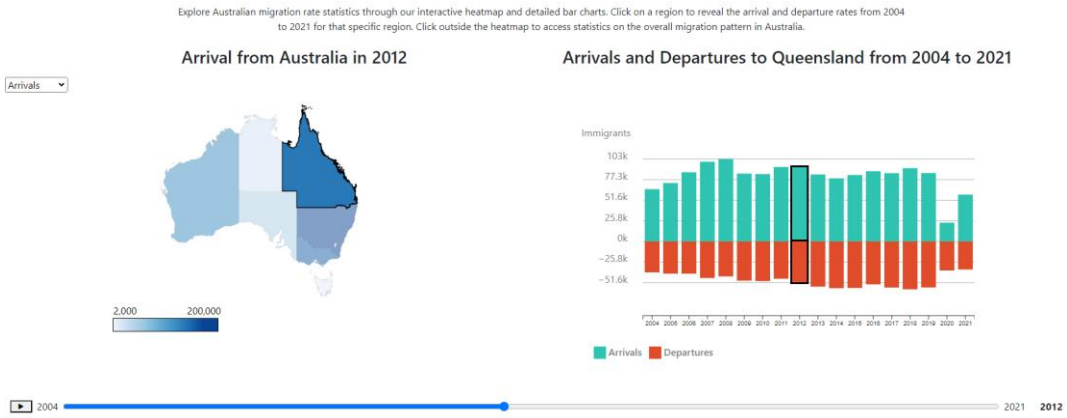


Figure 10. Visualisation Showing Data for Queensland

Upon clicking a state in the choropleth map, the title of the bar chart dynamically changes to reflect the selected state. As depicted in Figure 10, the chosen state is highlighted while the others are blurred, making it more prominent. This view manipulation technique effectively avoids overwhelming visual clutter and directs readers' attention to the key information (Munzner et al., 2015), which ensures a clear understanding of the bar chart's content. It helps prevent any potential misinterpretation by readers.

To handle the complexity of the data, incorporating changes over time is a popular and flexible approach in visualisation design (Munzner et al., 2015). Consequently, a timeline feature is provided, allowing users to slide and observe the fluctuations in the choropleth map between each year. The accompanying bar chart intuitively displays the numerical values over the selected period of time, aiding in the perception of trends. The timeline can be played automatically with the play button, making it easier for readers to perceive changes over time. Furthermore, as the timeline runs, the corresponding bar for each year is highlighted, allowing users to track and observe the differences more effectively. We used a pop-out effect to better gain attentional focus from readers.

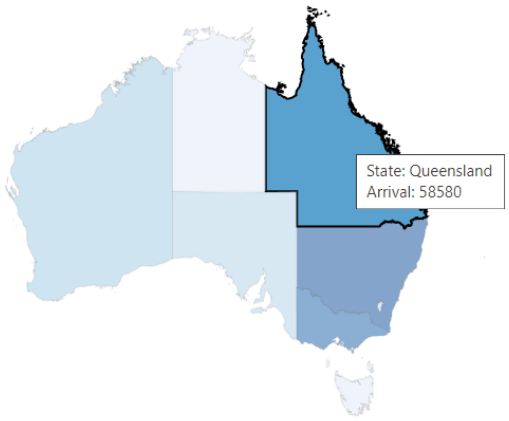


Figure 11. Tooltip when Hovering Over Choropleth map.

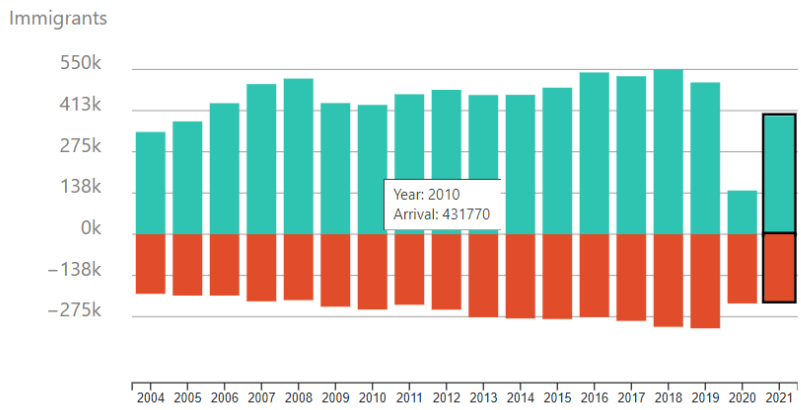


Figure 12. Tooltip when Hovering Over the Bar Chart.

When hovering over a state or each bar in the chart, a tooltip appears, presenting users with the corresponding numeric information. This feature enables users to easily access specific details on demand, enhancing the interactive experience. The map can also be zoomed and dragged so users can conveniently see the region they want.

Furthermore, to represent the volume of arrivals or departures, colour saturation is used, which is suitable for encoding ratio data. Additionally, hue colours are utilised in the bar chart to differentiate between arrivals and departures. In order to ensure accessibility, we conducted a thorough examination of different colour sets to determine the ones that can be easily distinguished by individuals with colour blindness. As a result, we chose the red and green combination. The effectiveness of this colour set for colour-blind individuals can be observed in Figure 13, demonstrating its ability to aid their distinction. To validate the colours, we employed an online tool for colour testing.

Figure 13. Red and Green Colour Set Perceived by Colour-blind People

Figure 14. Legend for the Choropleth Map.

Each chart in the visualisation comes with its own legend to explain the colour scheme used. The bar chart's legend indicates how colour hues are applied to distinguish between arrivals and departures, whereas the Choropleth features a linear-gradient colour bar to explain the use of colour saturation in representing the volume of numbers. When displaying low-to-high numerical values, as in the Choropleth map, the use of sequential colour palettes proves especially effective (Dougherty & Ilyankou, 2021).

The colour scheme for the Choropleth map is linearly scaled according to colour wavelength to mitigate any distortion in scaling. When considering the magnitude channel for the Choropleth map, saturation was preferred over luminance for a few key reasons. Firstly, saturation aligns well with the size channel utilised by the map. Secondly, saturation is perceived more strongly by viewers compared to luminance, which contributes to better comprehension of the data (Munzner et al., 2015).

While our visualisation tries to maintain a high data-ink ratio, thereby it maximises the clarity and efficiency of the data presented, which may initially seem quite complicated due to the density of information. It requires viewers to spend some time familiarising themselves with the various aspects of the graph, including the legends and colour-coding scheme. The interplay between the Choropleth map and the bar charts may also require some acclimation. However, once the initial learning curve is overcome, the visualisation offers a rich and detailed insight into Australian migration patterns over a period of time.

4.2 Top 10 Source Countries for Migration to Australia by State

4.2.1 Initial Design 1: Bar chart

Considering that the "value" attribute signifies ratio data, we felt a visualisation method aptly suited for displaying quantitative attributes was necessary. Consequently, our first draft of this visualisation comprises a horizontal bar chart. Ten countries are represented along the y-axis, which effectively communicates the nominal attribute. The total count of individuals migrating from each country to Australia is represented via the length of the corresponding bars. A noteworthy feature of this bar chart is its sortability in either ascending or descending order, making it easier for users to make comparisons. Furthermore, colour hues are applied to individual bars to distinguish each country visually. There would be a slider for the timeline for users to choose the year they want to display.

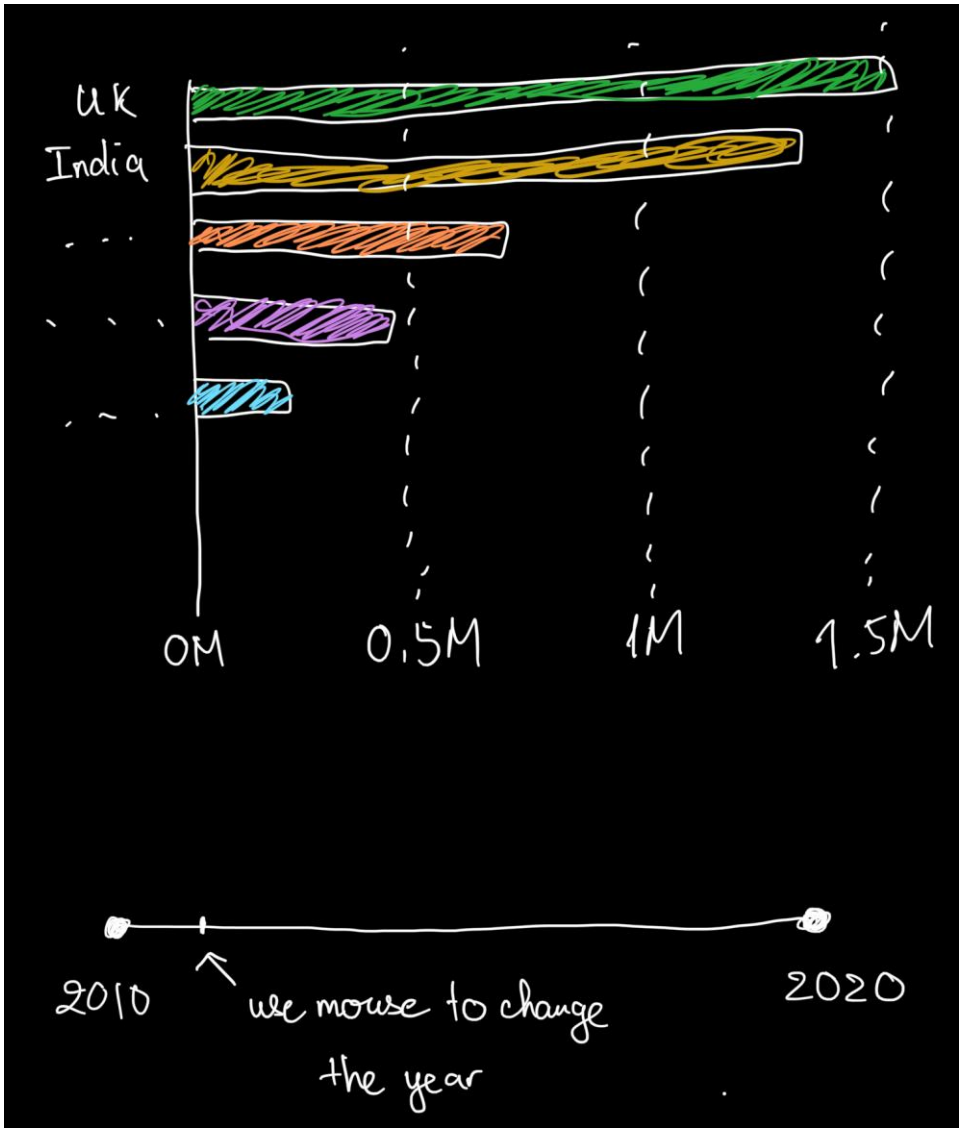


Figure 15. Bar Chart for Comparing the Sum of People Moving to Australia from the Top 10 Countries in The World

The benefit of this graph is that it can facilitate easy comparison between the total number of individuals migrating to Australia from different countries. Furthermore, the horizontal bar chart is more space-efficient compared to other chart types. We can accommodate a larger number of categories without overcrowding the visualisation, making them suitable for displaying comprehensive datasets.

While this design brings about numerous benefits, it still has some shortcomings. Chief among them is its limited information depth, as the diagram primarily communicates the yearly total of migrants to Australia without delving into other critical aspects of migration trends.

Additionally, the necessity of a drop-down menu to select data from different states introduces another challenge. This mechanism, although beneficial for managing a substantial dataset, can potentially hinder the user's holistic understanding. Specifically,

it constrains the user's ability to simultaneously compare migration data between different states.

For instance, our target audience, including parents seeking the most culturally compatible environment for their children, might struggle with this visualisation. They could find it challenging to get a comprehensive view of the migration flow, as they can only observe the data of one state at a time. This limitation not only obstructs a holistic understanding of Australia-wide migration patterns but also inhibits the depth of user interpretation. Consequently, it might hamper their ability to make well-informed decisions based on the visualisation.

4.2.2 Initial Design 2: Sankey diagram

To enhance the comprehension of the previous design, we looked for ideas from D3 Gallery, and we found a Sankey diagram that perfectly fits our expectation. This visualisation uses colour hues to differentiate between various nodes, which are either sources or destinations. Further, the Sankey diagram leverages line marks and area channels to represent the volume and direction of data flow, providing a comprehensive picture. Links are drawn from each country to each Australian state, offering a clear image of migration routes. Moreover, to promote user interaction, hovering over a node illuminates all associated linkages. Similarly, hovering over a link will pop up a tooltip with the total number of people involved in that particular flow.

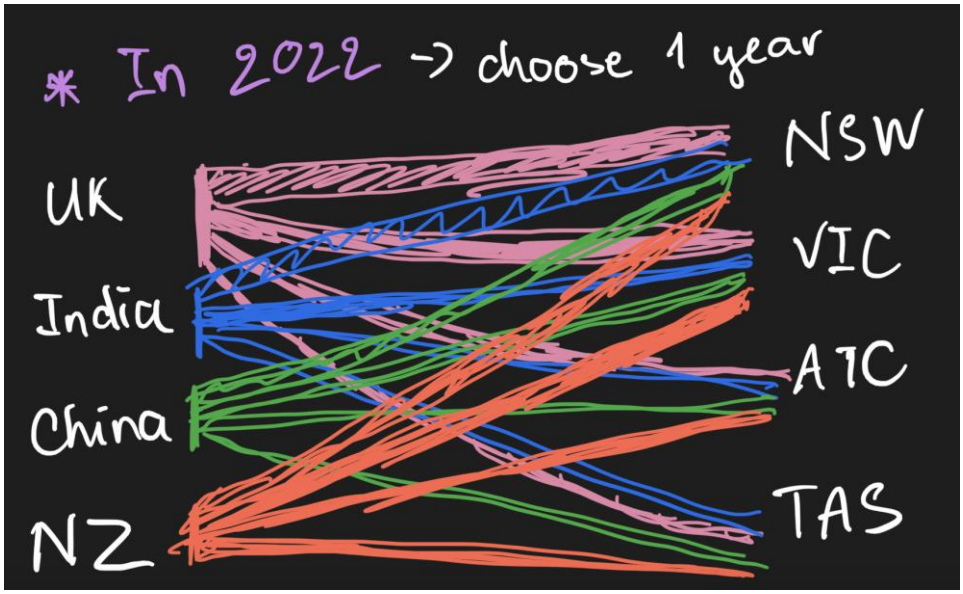


Figure 16. Sankey Diagram of The Sum of People Moving to Each State in Australia

This graph is excellent in illustrating the flow of migrated people between different countries. It effectively demonstrates the migration paths, allowing audiences to understand the relationships and interactions between various components. Moreover, the width of the flow lines in a Sankey diagram is proportional to the quantity being represented. This approach of representation provides a visual grasp of the relative magnitudes of different components, thereby allowing quick comparisons and insightful observations.

However, Sankey diagrams have their downsides. Primarily, they require comprehensive and flawless datasets, ensuring every node connection is without negative or absent values. In addition, Sankey diagrams may demand significant visual space, impacting their spatial efficiency and potentially making the chart appear extremely large.

4.2.3 Final Design

a. Graph Overview

For the final design, we deemed Sankey diagram to be the most suitable. The reason why we chose to go implement that for our website is driven by its striking capacity to deliver critical insights from the dataset. This design visually portrays the flow of migration to all Australian states concurrently, making the experience comprehensive and cohesive for our users. What makes this design even more compelling is the proportional representation of data, which brings clarity to the comparative magnitude of migration across states. By displaying the relative size of the data visually, we make the patterns more intuitive, which aids users in understanding the nuances of migration flows.

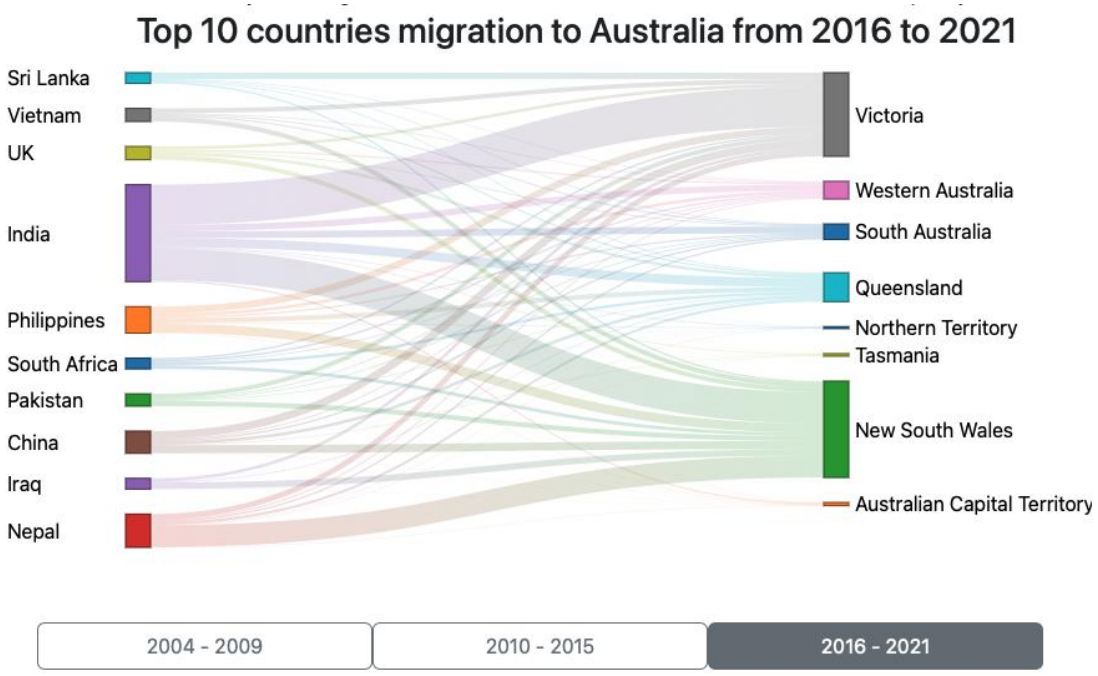


Figure 17. Sankey Diagram for Comparing the Total Number of People Moving to each state in Australia from the Top 10 countries in the world

b. Graph description

This visualisation represents the migration flow from top ten countries to all states across Australia. Each node in the graph signifies either a source (the origin countries) or a destination (the states), with their nominal attributes distinctly represented through a colour hue channel. Due to high data-ink ratio of the graph, the link between each node

employs a linear-gradient colouring system, starting with the colour of the source node and transitioning to the colour of the destination node. This not only enhances the readability of the graph but also prevents clutter, thereby fostering a user-friendly and visually appealing representation of the data.

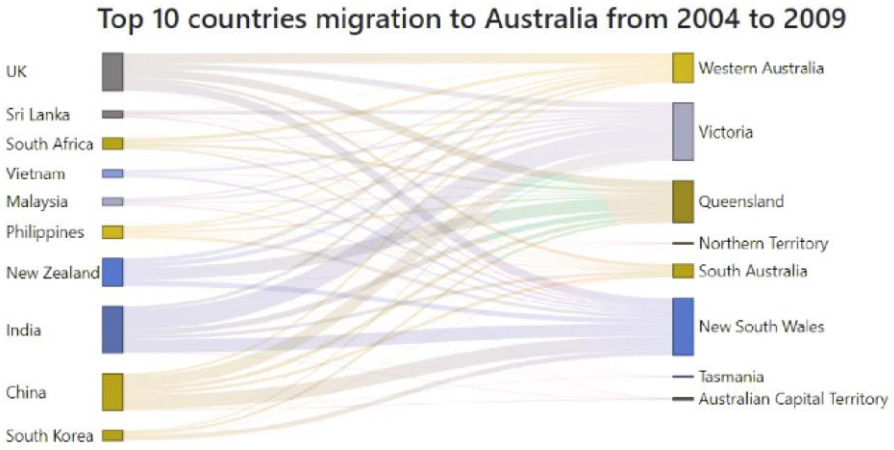


Figure 18. The Flow Chart Perceived by People with Protanopia

The colour hue channel used to distinguish different nodes comes from the D3 library. However, despite its effort to minimise colour repetition between nodes, it still happens due to the library's ten-colour scheme. This limitation, coupled with a large number of nodes, results in insufficient unique colours, potentially causing issues for colour-blind individuals.

To address this, we've applied a linear-gradient colour to the links. This technique uses luminance colours, which helps differentiate nodes for colour-blind individuals. Additionally, we've ensured that two adjacent nodes will not share the same luminance colour, preserving clarity in such cases. The luminance channel provides high-resolution edge information, and to create a design more accommodating to colour blind users, it's advisable not to rely solely on hue colours, as suggested by Munzner et al., 2015. Additionally, we've employed an online tool to verify that our graph is, to a significant extent, comprehensible for colour blind individuals, as demonstrated in Figure 18.

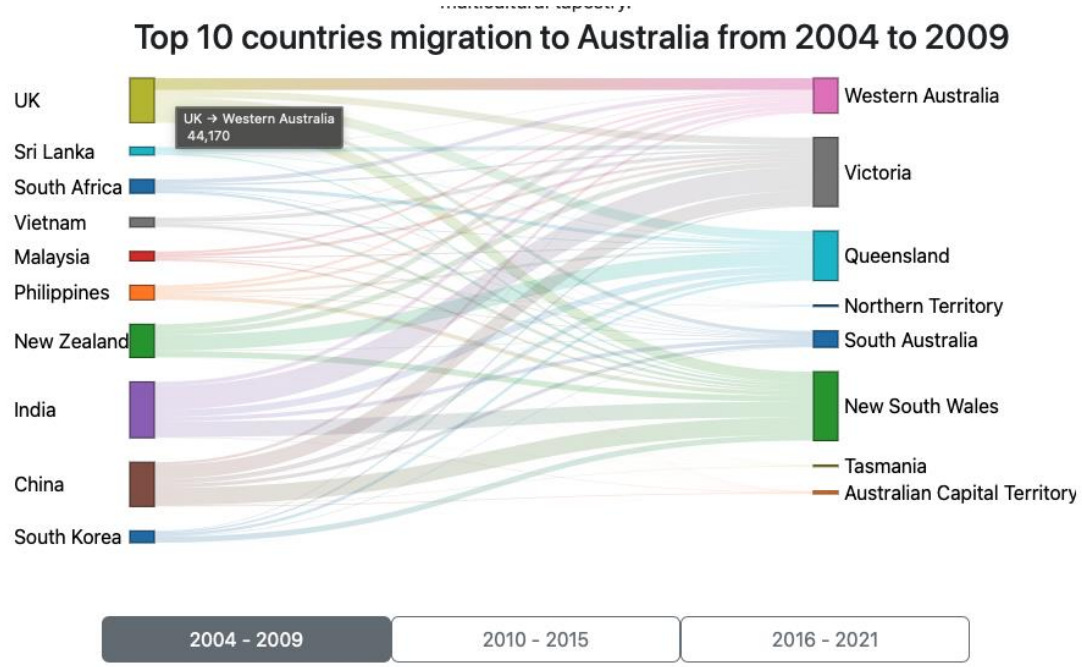


Figure 19. Tooltip when Hovering Over a Node

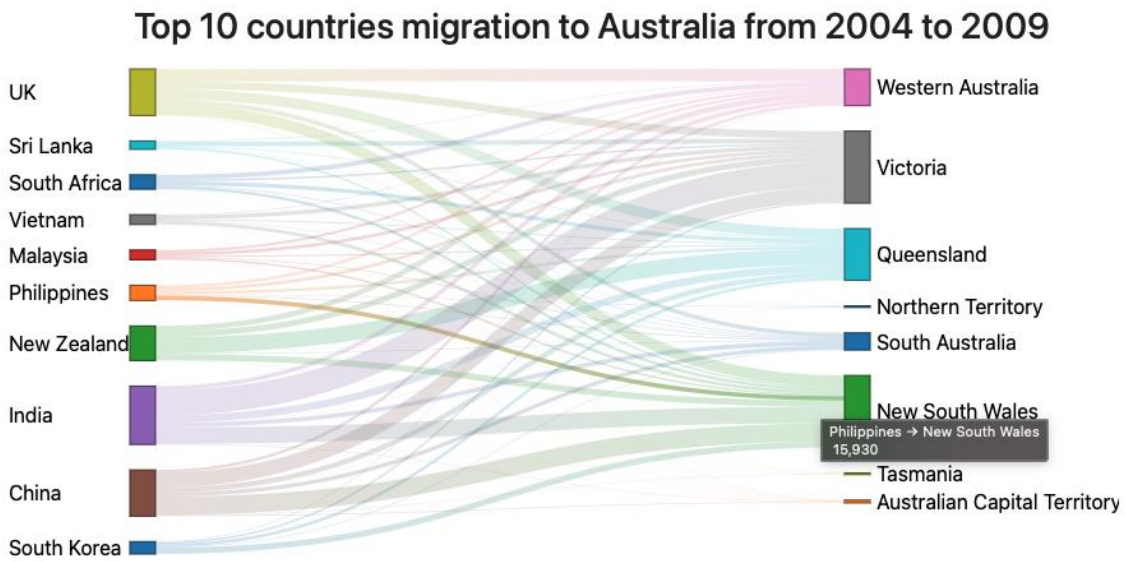


Figure 20. Tooltip when Hovering Over a Flow

To enhance user engagement and interpretation, the graph incorporates an interactive tooltip feature. When a user hovers over a node, the tooltip unveils the exact number of people associated with that node, offering insights into migration volumes. Similarly, hovering over a link brings up a tooltip that discloses detailed information about its originating country and destination, along with the number of individuals migrating between the particular country and the state. By delivering these precise details, this interactive tool significantly bolsters the user's understanding of the specific migration patterns showcased.

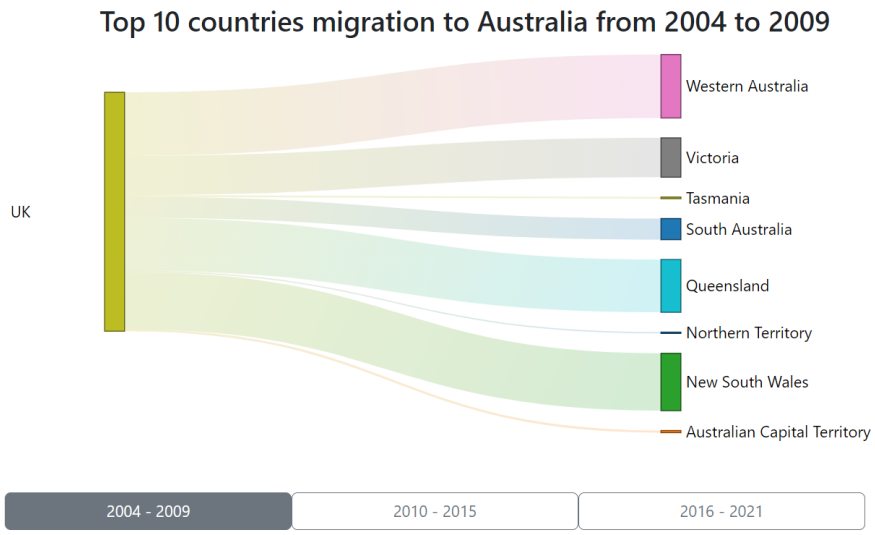


Figure 21. Interaction when Clicking on a source node

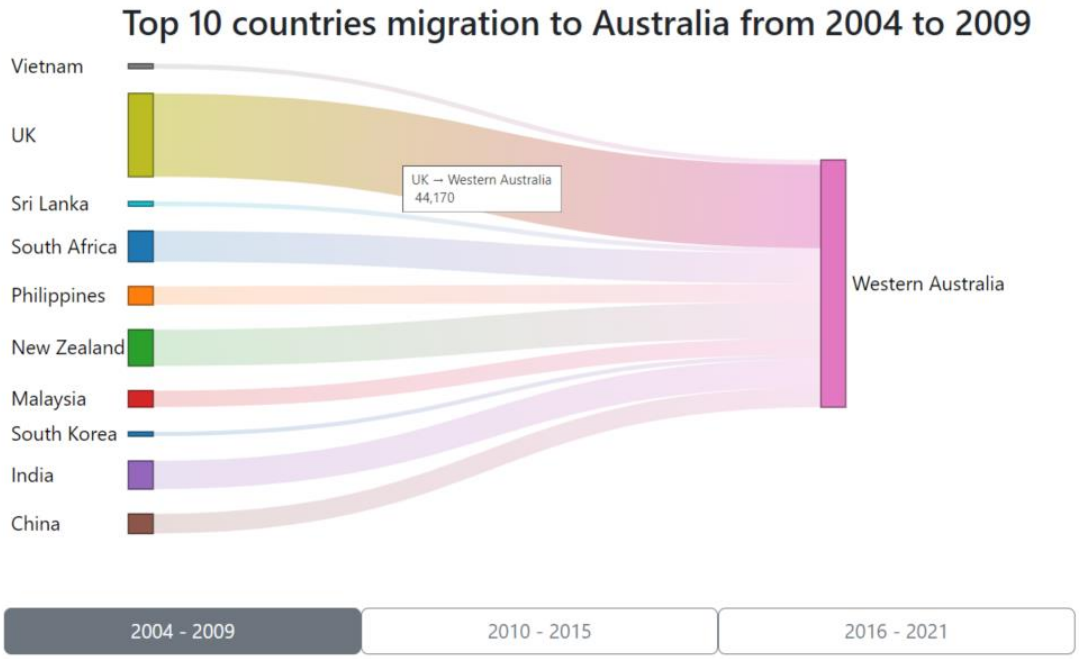


Figure 22. Interaction when clicking on a destination node

As suggested by Munzner et al., 2015, adjusting the view is a powerful tool for handling complex datasets. To address potential clutter due to numerous links, we have incorporated an interactive feature that lets users see only the information relevant to a particular node, including both its source and destination nodes. Upon clicking a node, the Sankey diagram reacts by emphasising the specific country and its connected state links, effectively isolating the migration trend of interest (Figure 21). The exact number of migrants for this flow is also displayed (Figure 22). This focused visualisation streamlines the process of data exploration, allowing users to concentrate on individual migration patterns without distraction from the wider dataset.

Despite the complexity of Sankey diagrams due to their high data-ink ratio, we've incorporated various interactive features to aid users in comprehending the information. These interactions also help in managing the potential clutter within the chart, thereby ensuring an effective and clear visualisation.

4.3 Reasons Given for Migrating to Australia by Visa Type

4.3.1 Initial Design 1: Pie Chart

To represent the ratio data, specifically the number of individuals migrating to Australia, we employed area marks and colour hue channels in our first visualisation iteration – a pie chart. This chart presents various categories of visa holder types as distinct segments. The size of each slice corresponds to the proportion of the data it presents. The larger the percentage, the larger the corresponding slice. Labels are placed outside the pie chart and have a specific colour to indicate the category it represents. Besides, we create a tooltip showing the name of the category and the exact number of that category.

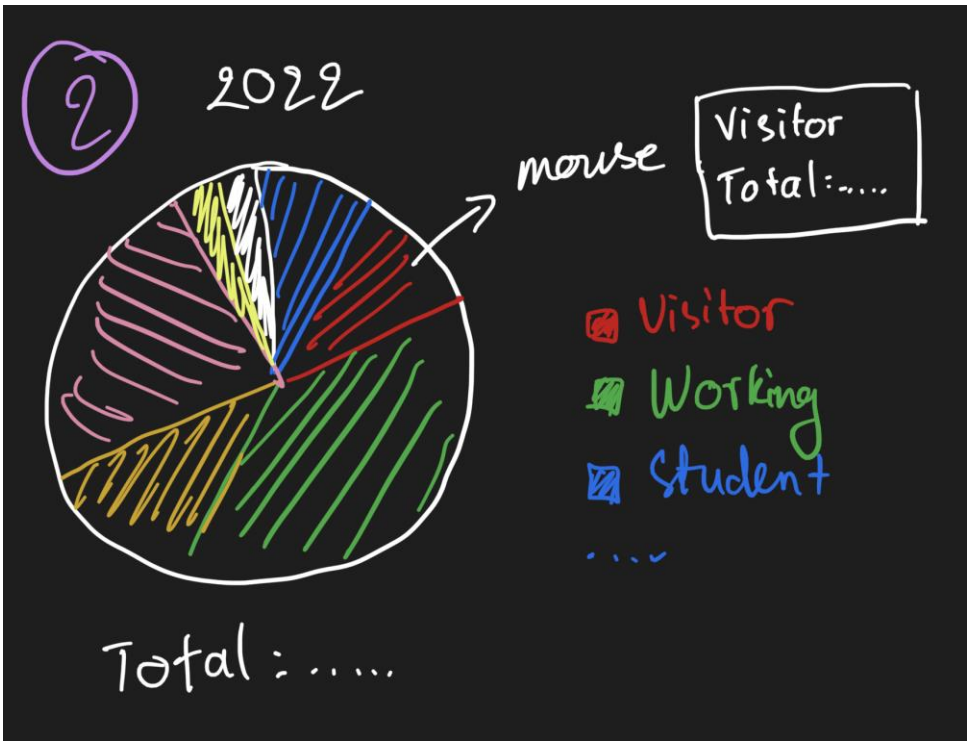


Figure 23. Pie Chart of The Percentage of Visa Holders Moving to Australia in 2022

The advantage of this graph is this visual representation allows viewers to compare the sizes of the slices, enabling them to understand the relative proportions of each type of visa holder.

While the pie chart has its own benefit, it also has some limitations for this data. This type of chart can struggle to effectively demonstrate small values. Small slices may cause visually challenging to perceive, and accurately interpreting their percentage can be difficult, which leads to potential misinterpretation of the data. Furthermore, when the number of categories becomes too large, a pie chart can become cluttered and

visually overwhelming. The small slices can blend together, making it challenging to distinguish between them from the visualisation.

While a pie chart can be a valuable visualisation tool, it might not be the most effective choice when dealing with a large number of categories. Particularly, distinguishing between slices representing nearly equal quantities can be challenging for readers. Therefore, considering this situation, a pie chart may not be the optimal choice for our scenario.

4.3.2 Initial Design 2: Line Charts

Our second design for this data is two-line graphs including two types of visa holders. We decided to encode two different types of data into the line chart. It is commonly used to show the trend or relationship between two continuous variables over a specific period or range. In a line chart, the horizontal axis represents the independent time, and the vertical axis represents the data values of the number of people coming to Australia.

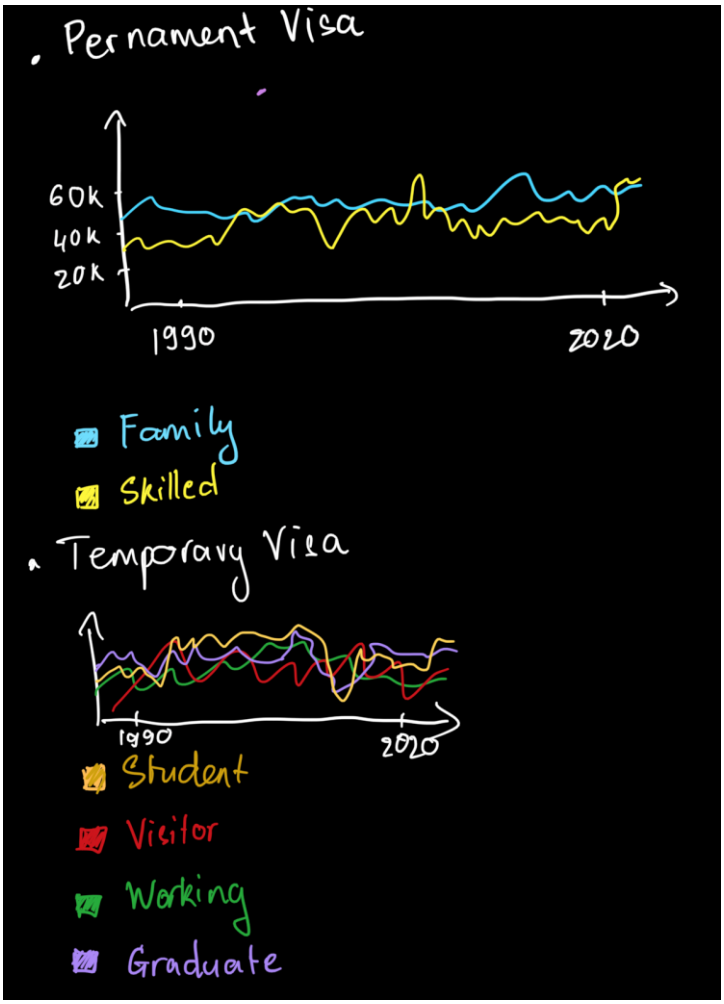


Figure 24. Line Charts of Two Types of Visa Holders

The advantage of these charts is that line charts are excellent for identifying trends over time. By connecting data points with lines, it becomes easier to see the upward and downward trends and overall changes in 20 years.

The only drawback that we could find is when we collected the data, the information we received from permanent visa seemed nothing too special to demonstrate.

4.3.3 Initial Design 3: Line Chart

Our third design for this dataset is only choosing one type of visa holder to let the data clear and easy to highlight the changes over the period.

After careful consideration, we have made the decision to use dots as the form of the points in our visualisations and, specifically construct them for the years given on the x-axis. This design choice ensures that the line's characteristics are readily apparent at first glance, rather than relying solely on mouse interaction to reveal relevant information.

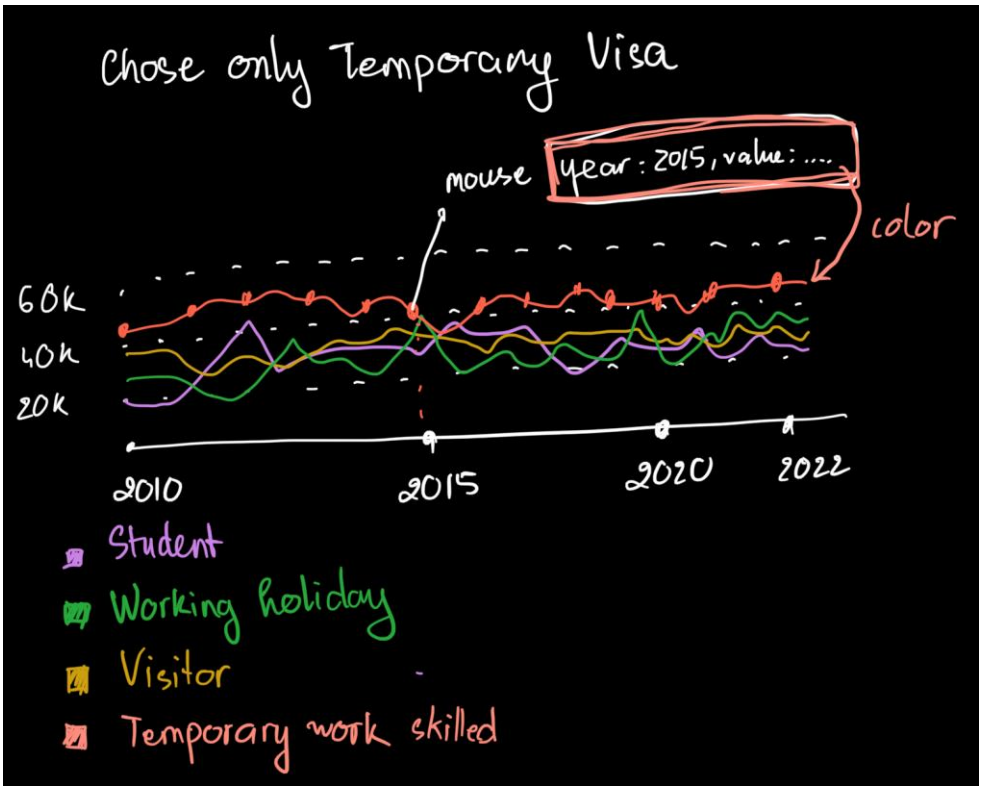


Figure 25. Line Charts of Temporary Visa for The Total Number of People Visiting Australia

The good point of using this type of graph is users can quickly identify the general trend and pattern of the data without needing to rely on additional interaction hover actions. Also, each dot represents a specific value, making it easy to track and compare data across different time periods or categories. This clarity helps users grasp the relationships and fluctuations within the dataset. In that case, audiences can focus on overall trends and identify key data points without being overwhelmed by complex graphical elements.

While using dots as the form of points in a line graph offers advantages, there are also some potential drawbacks to consider. Depending on the density of data points and the size of the chart, using dots can lead to visual clutter. Also, when using dots instead of connecting lines, the graph's smoothness may be compromised.

4.3.4 Final Design

a. Graph Overview

In our website's implementation, we opted for a line chart as it excellently captures the ongoing trends regarding why people are migrating to various Australian states. This line chart not only offers a broad view of the trends but also allows readers to filter for a specific state of interest, thereby enhancing their understanding of popular migration choices. Unlike our previous design that solely focused on temporary arrivals, this updated version merges both temporary and permanent migration data, offering a more comprehensive perspective on migration patterns.

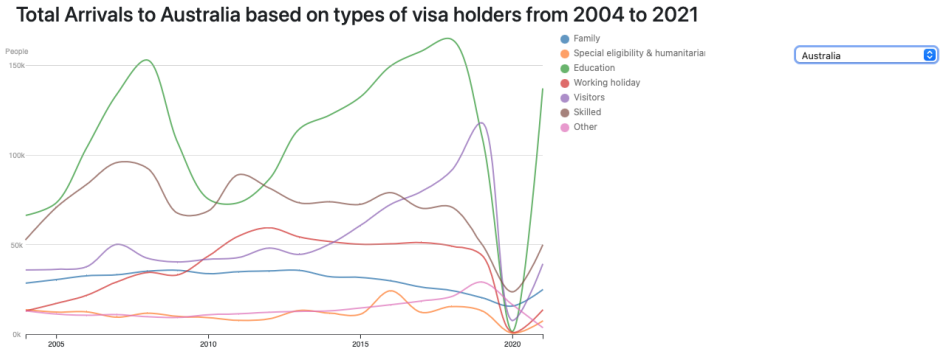


Figure 26. Line Chart of Total Arrivals to Australia Based on Types of Visa Holders from 2004 to 2021

b. Graph Description

This visualisation effectively depicts the total number of arrivals in Australia from 2004 to 2021. The ratio attribute of the data is ingeniously communicated via the position on a common scale channel. The original dataset encompassed more than 13 types of visas, including returning Australian citizens and several visa types serving the same purpose. Hence, to enhance interpretability, we merged visa types based on their respective purposes. Additionally, we intentionally excluded the category of returning Australian citizens, as it doesn't significantly contribute to understanding the reasons for migration to Australia. As reduction is one of the key strategies for handling complex datasets (Munzner et al., 2015), we selectively merged categories while ensuring the integrity of data.

We incorporate a drop-down menu for users to choose to select and view data specific to an individual state or for the entirety of Australia. This interactive feature ensures the accessibility of more data while simultaneously avoiding an overly congested display in the graph.

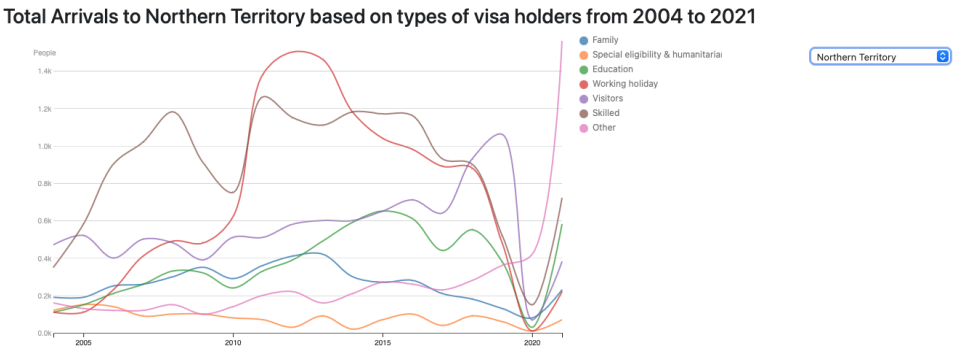


Figure 27. Line Chart of Total Arrivals to Northern Territory State

While our design utilises the colour hue channel to distinguish between lines, given the broad range of categories, the D3 library doesn't offer a colour scheme with eight distinct, colourblind-safe hues. This might make it challenging for individuals with colour vision deficiencies to distinguish between the lines just by referencing the legend.

To solve that problem, we employ a special tooltip to enhance the interpretation of the graph. When users hover the mouse over the graph, the program calculates the position of the mouse in relation to the corresponding year. At this point, circle markers appear, showing the position of each line for that specific year, thereby facilitating data comparison within the year. Additionally, the tooltip also presents data from all categories, arranged in descending order. This arrangement allows users, including those with colour vision deficiency, to understand the value of each line, moving from the top to the bottom in the tooltip. This feature ensures that our visualisation remains accessible to all users.

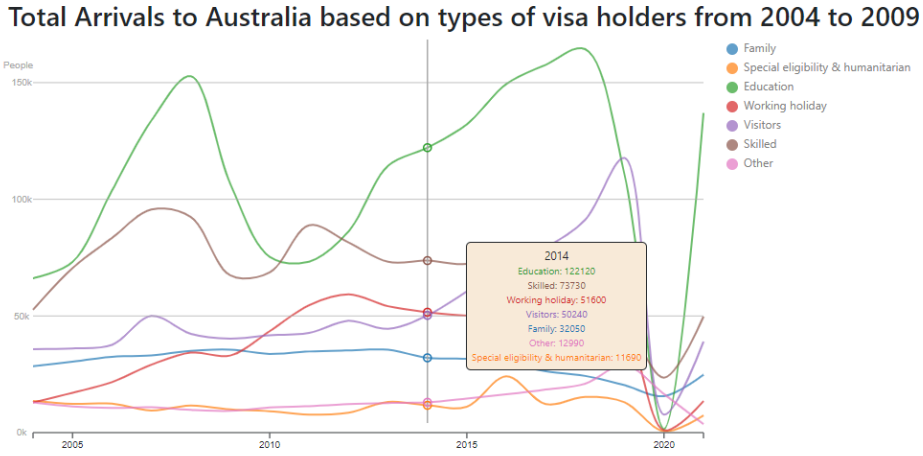


Figure 28. Tooltip for The Line Chart.

5. User Validation

5.1 Methodology

We performed a user evaluation survey by using a completed website (https://checkiejan.github.io/COS30045-Data-Visualization-Project/). Users engaged with our project by first reviewing and accepting an informed consent form.

Participants in our study were given a series of tasks designed to evaluate the usability and interactivity of the three visualisations. These tasks required responding to scenario-based questions utilising information from the visualisations, followed by assessing the task difficulty on a scale ranging from "Very Easy" (1) to "Very Hard" (5). Our aim was to make the visualisation as user-friendly as possible. Therefore, we strove to design it in such a way that the level of difficulty perceived by users would be as low as feasible.

Our survey participants were chosen via convenience sampling, with selection criteria including English fluency, familiarity with graphs and technological devices, and a willingness to participate to the study.

The whole questionnaire, including the informed consent form, can be found at https://forms.gle/icVmzHnuRgjF2hFj8.

Our participant pool consisted of eighteen individuals, all falling within the 18 to 34 age bracket. The majority, at 72.2%, were versed in programming, with a significant portion of them having specific experience in the field of web development.

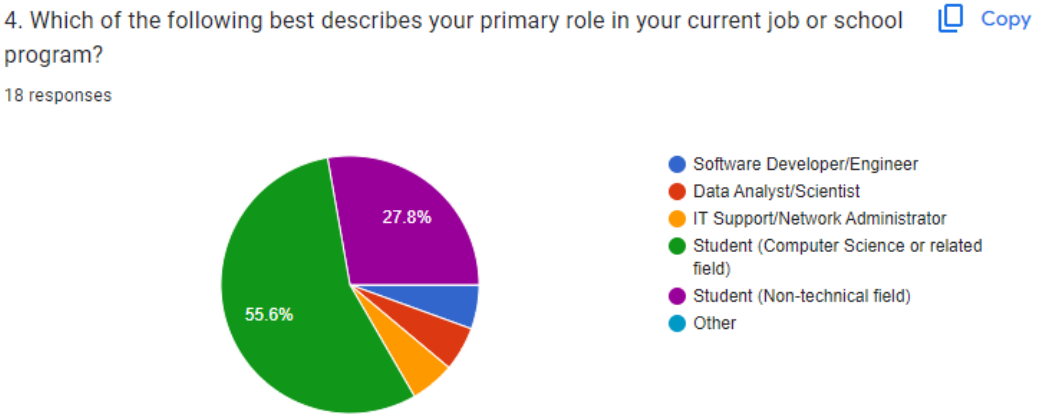


Figure 29. Participants Background Overview

5.2 Arrival and Departure Trends in Australia

5.2.1 Survey Responses

For the first visualisation, we asked three questions that required participants to look for information from the graph. These questions were designed to determine how well users could familiarise themselves with a range of functionalities of the graph, such as the interactive elements of hovering, sorting, and locating specific information.

Question 1 was intended to test the efficacy of the colour usage in the Choropleth function. Participants were expected to identify the answer either by observing the intensity of the colour or by using the hover functionality to reveal the tooltip. This task proved to be quite successful, with 83.3% of participants providing the correct answers. Furthermore, 77.8% found the task ranging from very easy to medium in terms of difficulty.

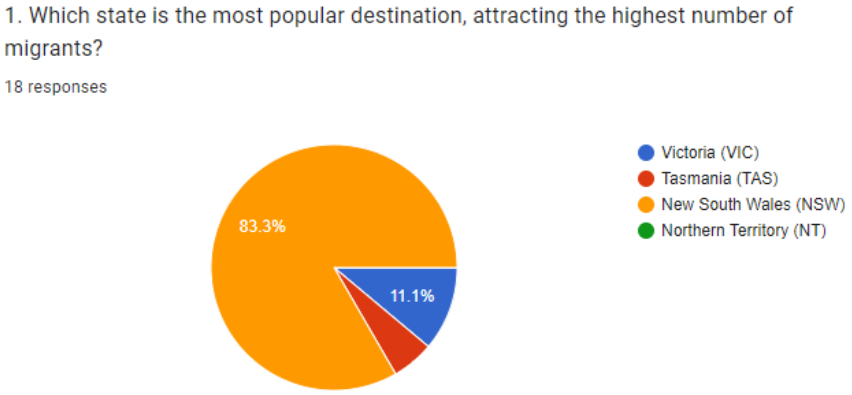


Figure 30. Result of Question 1

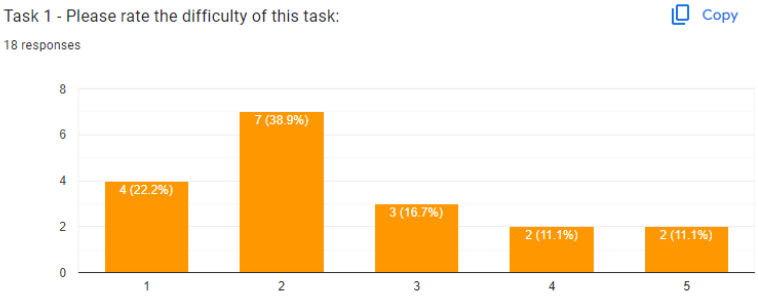


Figure 31. Difficulty Rating of Question 1

For question 2, our goal was to determine how effectively users could interpret the bar chart. The results were not as promising as those for question 1, potentially due to the fact that users had to handle two charts simultaneously, which required some time to get accustomed to them. This assumption is further reassured by the fact that only 66.7% of participants rated this task in terms of difficulty as ranging from very easy to medium.

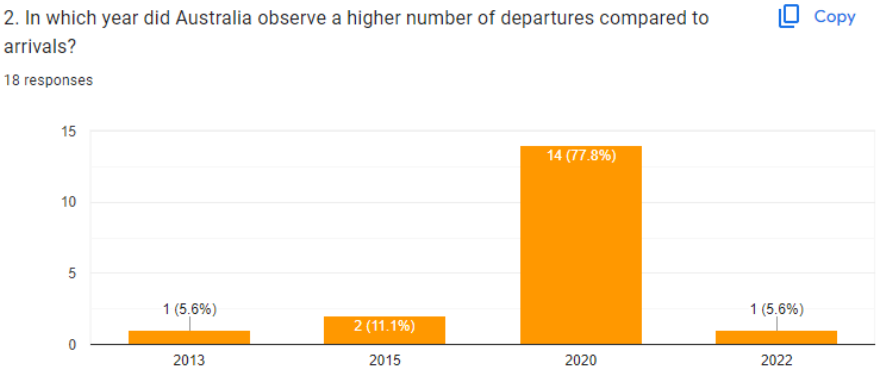


Figure 32. Result of Question 2

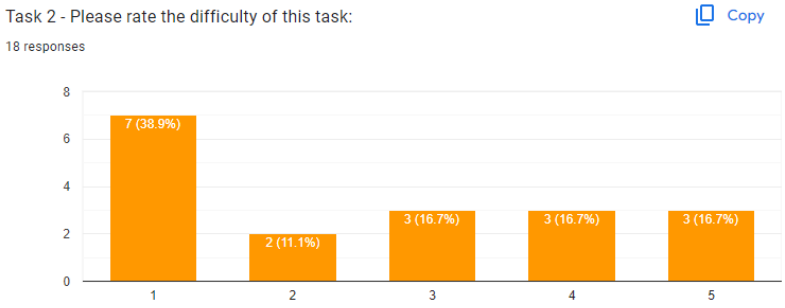


Figure 33. Difficulty Rating of Question 2

For question 3, we sought to determine how effectively participants could combine information from the two charts and leverage the timeline functionality to answer accurately. The outcomes were rather promising, with 83.3% providing correct answers and 72.3% rating the difficulty level of the task between very easy and medium.

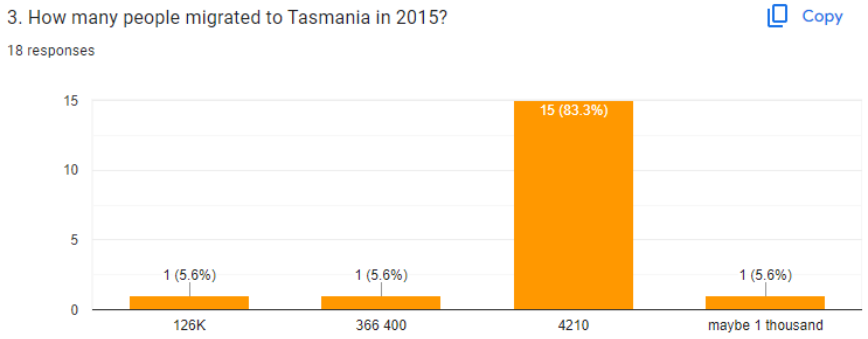


Figure 34. Result of Question 3

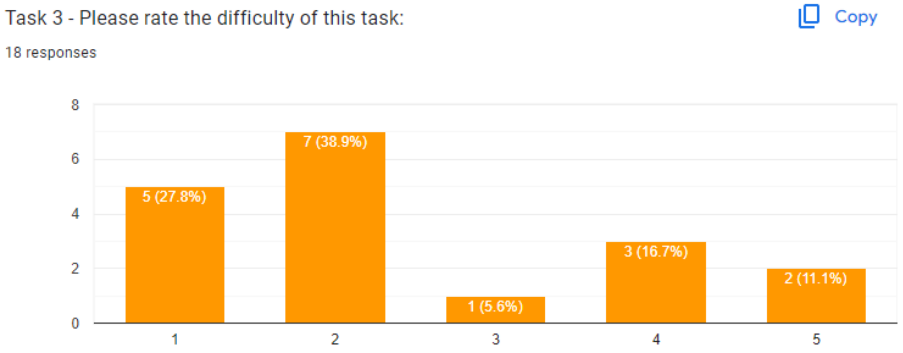


Figure 35. Difficulty Rating of Question 3

Overall, the three questions demonstrated an average accuracy rate of 81.5%. Furthermore, 72.3% of participants found the usage of the first visualisation to range from very easy to medium in terms of difficulty.

5.2.2 Recommendations

Based on recommendation from participants, we found that some users considered that some hover effects were not very smooth, and it took a significant amount of time for

the tooltips to appear. Some users thought that the colour was too bright and there should be a dark mode.

5.3 Top 10 Source Countries for Migration to Australia by State

5.3.1 Survey Responses

In order to evaluate the efficacy of the second visualisation, we posed three questions to our participants, requiring them to extract information directly from the graph. The focus of these questions was to assess how easily users could adapt to and engage with the functionalities of the Sankey diagram - an unconventional yet potent visualisation tool. Key interactions such as hovering, as well as identifying specific details from nodes or links were assessed. Given being unfamiliar with the Sankey diagram, a learning curve was expected, and participants' adaptability was part of the assessment.

Question 4 was meant to test ability of users to accurately interpret the proportionate width of the nodes in the Sankey diagram. Considering the unfamiliarity of many participants with this type of diagram, the results were encouraging with 66.7% providing correct answers. Additionally, a majority (83.3%) found the task to range from very easy to medium in difficulty.

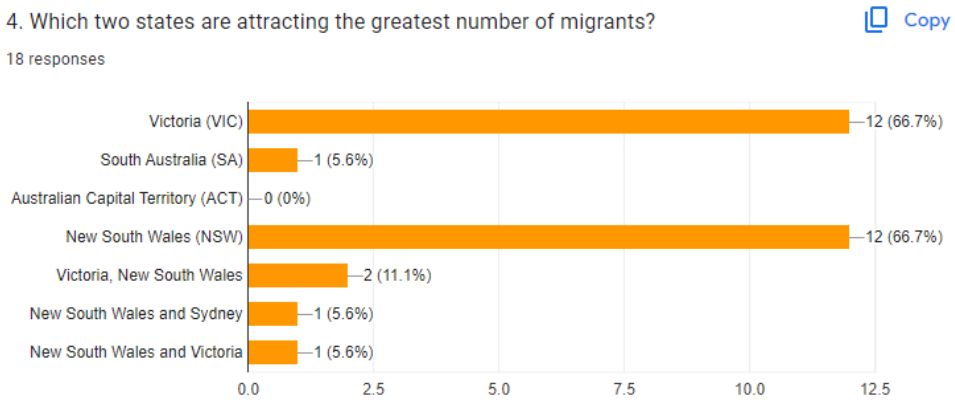


Figure 36. Result of Question 4

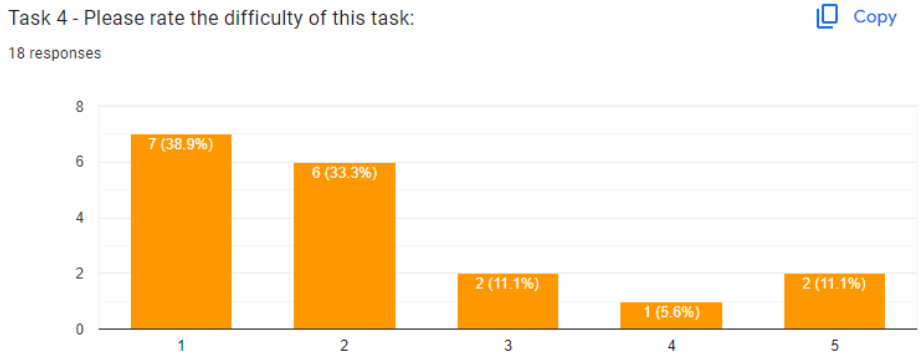


Figure 37. Difficulty Rating of Question 4

For the question 5, we aimed to test how effectively users could use hovering interactivity of the chart to get specific information. The result was quite good when up

to 72.2% participant answering correctly and 72.2% rating the difficulty level of the task between very easy and medium.

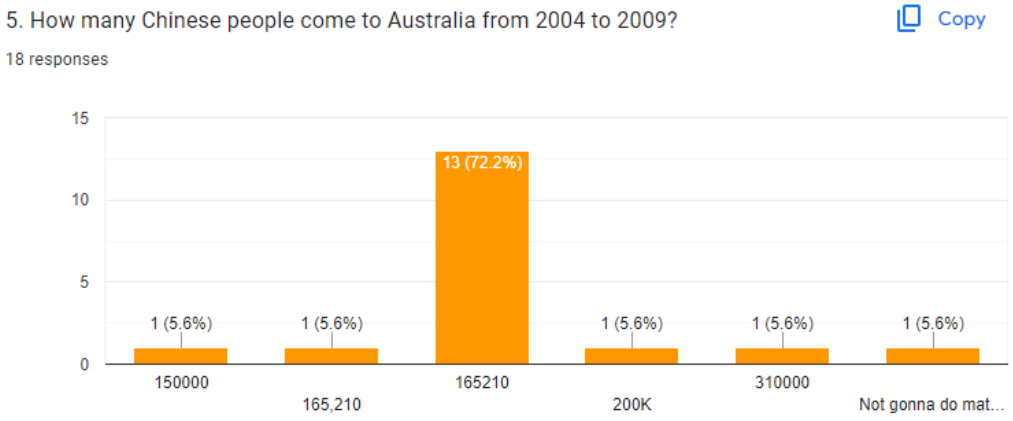


Figure 38. Result of Question 5

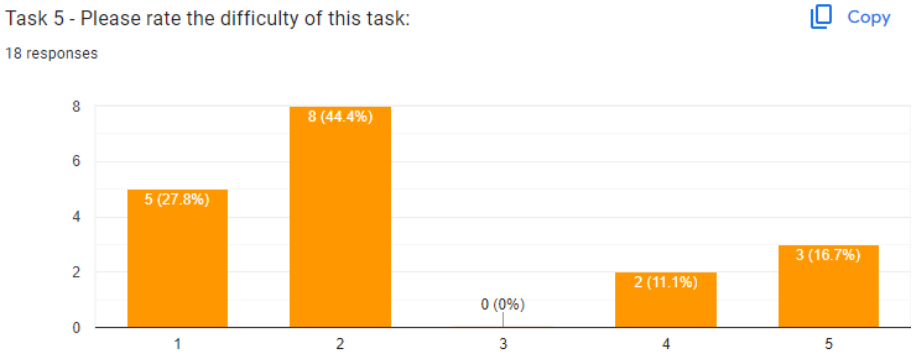


Figure 39. Difficulty Rating of Question 5

Question 6 was designed to evaluate whether users could successfully engage with the interactive feature of clicking a node to isolate its corresponding information. The outcomes were rather promising, with 83.3% providing correct answers and 83.3% rating the difficulty level of the task between very easy and medium.

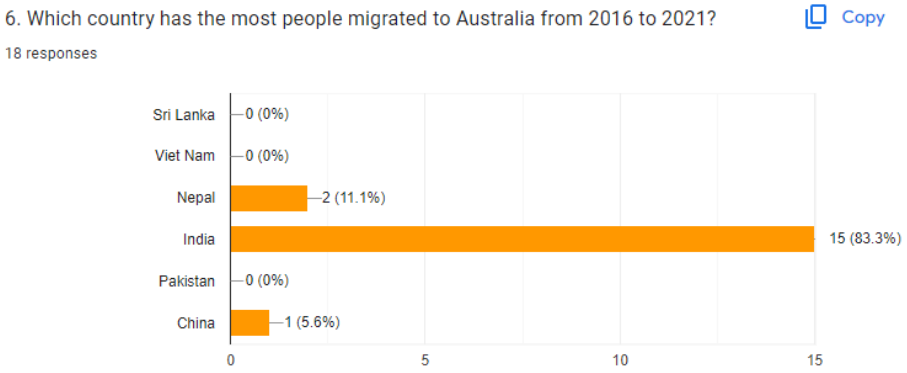


Figure 40. Result of Question 6

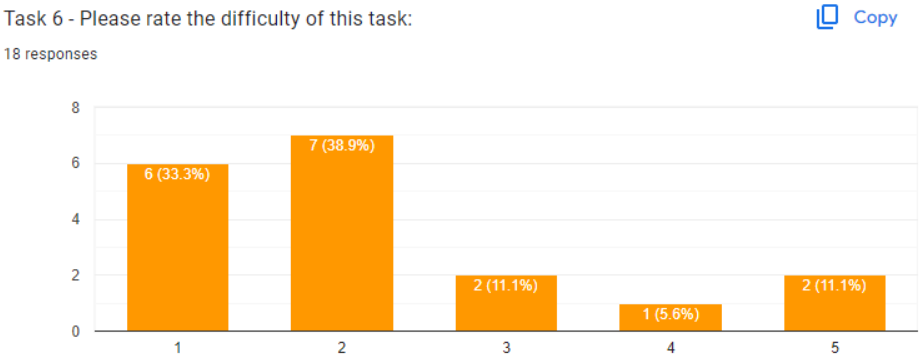


Figure 41. Difficulty Rating of Question 6

5.3.2 Recommendations

Many users expressed their excitement towards the visualisation, as they rarely had chance to use this type of diagram before. Nevertheless, a portion of the participants observed some areas of potential improvement. They mentioned that the graph appeared somewhat cluttered, which could potentially hinder data interpretation. Additionally, they pointed out the colour scheme employed in the visualisation, which, in certain areas, made it difficult to discern details. Furthermore, they noted that some links were so small that they could barely be seen.

5.4 Reasons Given for Migrating to Australia by Visa Type

5.4.1 Survey Responses

To assess the effectiveness of the final visualisation, we presented three questions to our participants, each demanding direct data extraction from the graph. Much like the other evaluations for the previous visualisations, the aim of these questions was to determine the ability of user to quickly acclimate to, and effectively interact with the functionalities of the graph, such as hovering, sorting, and looking for specific information.

Question 7 was intended to test ability of users to accurately comprehend data from the line chart using the tooltip feature. The results were less promising than anticipated, with only 66.7% of the participants responding correctly. Furthermore, only 72.2% of participants categorised the difficulty level of the task between very easy and medium. This may suggest that the visual design of the line chart and its interactive features could be challenging to comprehend for some users.

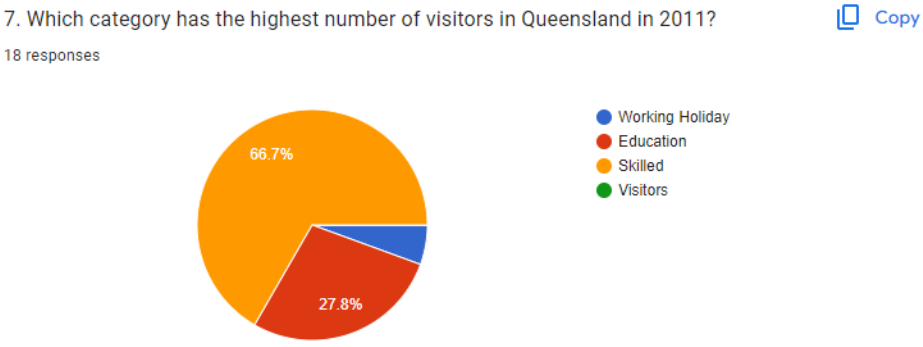


Figure 42. Result of Question 7

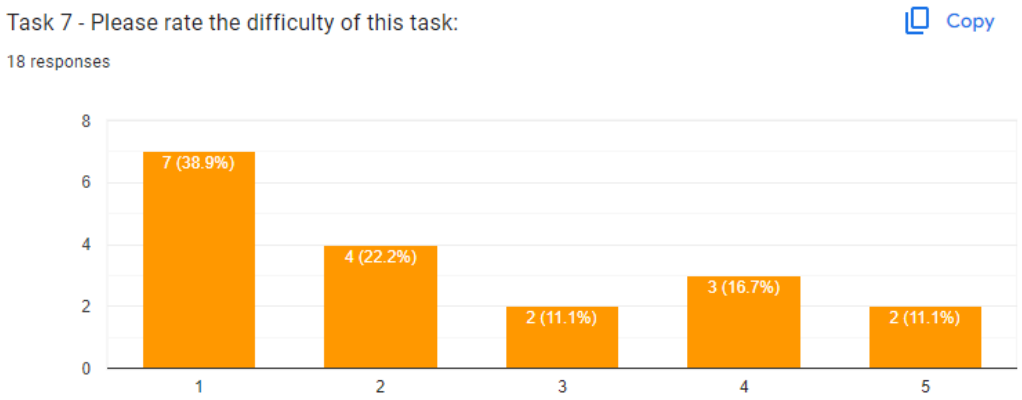


Figure 43. Difficulty Rating of Question 7

Question 8 asked participants to look for a specific detail, so we can test how our visualisation can deliver detail on-demand. The result was quite similar to Question 7 when 66.7% answering correctly, and 77.7% finding the difficulty level of the task between very easy and medium.

Figure 44. Result of Question 8

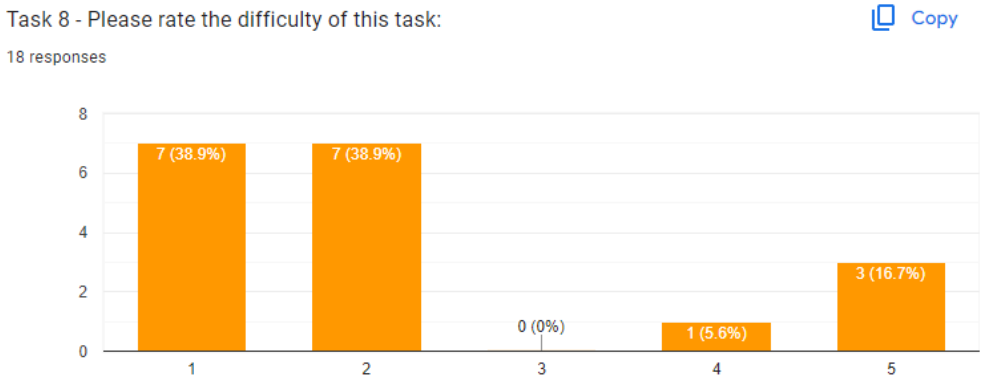


Figure 45. Difficulty Rating of Question 8

Question 9 served the same purpose as Question 8, but Our aim was to test the ability of users to differentiate subtle variations within the data. The results were notably impressive, with 94.4% of participants providing correct answers, and 77.7% of

respondents categorizing difficulty level of the task as ranging from very easy to medium. This suggests that despite the subtle variations in the data, users were highly successful in accurately interpreting the information.

Figure 46. Result of Question 9

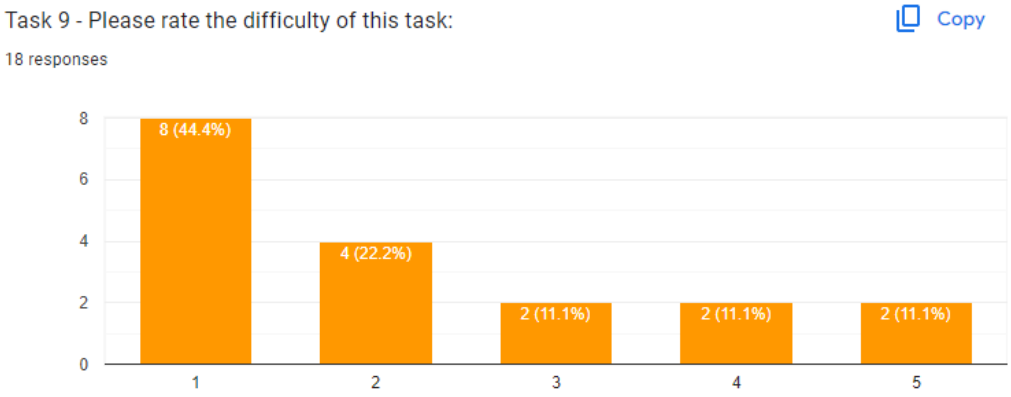


Figure 47. Difficulty Rating of Question 9

5.4.2 Recommendations

Compared to the other visualisations, this one generally saw a lower level of data interpretation accuracy among participants. Furthermore, some users suggested improving the readability of the tooltip content by modifying the tooltip colour background.

6. Conclusion

In conclusion, this Process Book has shown the method of our group for developing visuals for the project "Australian Migration Trends and Patterns ", with a focus on the project's development from ideas to data to designs and finally interactive graphs. Through the use of interactive maps, bar charts, and other visual tools, we were able to effectively present and analyse migration data, shedding light on various trends and patterns. Furthermore, it also demonstrates the information we have gained over the course of this semester.

By explicitly identifying our target audience and the specific problems we aim to address through our visualisations, we have realised a significant improvement in our graph

designs. Instead of creating generic visualisations that attempt to cater to all purposes, we have developed a more focused approach that directly answers specific questions and solves particular challenges.

Throughout the design process, we have been astounded by the multitude of graph options available to visualise the same data, each with its own set of benefits and drawbacks. The application of critique techniques learned during the course has played a crucial role in narrowing down our design and eventually combining various graphs to create a final design that incorporates the strengths of the previous iterations while mitigating their weaknesses.

In summary, our experience has highlighted the value of explicitly identifying the target audience and problem statements when designing visualisations. Utilising this experience, we have successfully achieved our objectives of gathering, visualising, and analysing data on long-term arrivals and departure migration patterns in Australia, the countries contributing most significantly to Australia's immigration intake by state, and the make-up and movements of different categories of visa holders between states. Moreover, the skills we have developed in data processing, graphic design, and programming will contribute to our professional growth, equipping us with the tools necessary to excel in our future endeavours.

7. References

Overseas migration, 2021-22 financial year (no date) Australian Bureau of Statistics. Available at: https://www.abs.gov.au/statistics/people/population/overseas-migration/latest-release#net-overseas-migration *(Accessed: 04 June 2023).*

Munzner, T. and Maguire, E. (2015) *Visualization Analysis & Design*. Boca Raton: CRC Press Taylor & Francis Group.

The Interaction Design Foundation (2023) *How to design an information visualization*. Available at: https://www.interaction-design.org/literature/article/how-to-design-an-information-visualization (Accessed: 04 June 2023).

About Australia (no date) *Australian Government Department of Foreign Affairs and Trade*. Available at: https://www.dfat.gov.au/about-australia (Accessed: 04 June 2023).

Dougherty, J. and Ilyankou, I. (2021) *Hands-on Data Visualization: Interactive storytelling from spreadsheets to code*. lugar: O'Reilly Media.

Geoscience Australia (2023) *Australia's size compared*, *Geoscience Australia*. Available at: https://www.ga.gov.au/scientific-topics/national-location-information/dimensions/australias-size-compared#:~:text=Australia%20is%20the%20planet%27s%20sixth,is%20the%20world%27s%20largest%20island. (Accessed: 04 June 2023).

Nichols, D. (no date) *Colouring for colourblindness*, David Nichols. Available at: https://davidmathlogic.com/colorblind/#%232FC4B2-%23E14D2A (Accessed: 04 June 2023).

Coblis - color blindness simulator (no date) *Colblindor*. Available at: https://www.color-blindness.com/coblis-color-blindness-simulator/ (Accessed: 04 June 2023).