

Group: TU8_G07

STUDENT NAME ID:

AZMAIN(103525182)

SAMMO(103512692)

TUTOR: RAIHANEH AGHAIE

TUTORIAL: TUESDAY 5.30-7.30

SEMESTER 1, 2023

WORD COUNT: 6000

LINK TO MERCURY HOSTED WEBSITE:

mercury.swin.edu.au/cos30045/s103525182/assign_2/final/index.html



A DEEP DIVE INTO MIGRATION PATTERNS OF AUSTRALIA

Table of Contents

1 Introduction	02
1.1 Background and Motivation	02
1.2 Visualization Purpose	02
1.3 Project Schedule	07
2 Data	08
2.1 Data Source	08
2.2 Data Processing	15
3 Requirements	20
3.1 Must-Have Features	20
3.2 Optional Features	21
4 Visualization Design	22
5 Validation	32
6 Conclusion	33
7 References	33


1.1 Information and Background

The intended audience for this visualization would be users who want to critically analyze the relationship between the net migration in the seven states of Australia with the overseas migration in capital cities. Furthermore, another visualization delves deep into the grouping of overseas migrants establishing the premise of this visualization's influence. The main motivation behind this project is to provide an array of visualizations which intends to showcase Australia as a land of opportunities and thus summarizes the inflow of people in the country and migrating people to different regions within Australia. All in all, this visualization shows the flow of information in three layers (Net **interstate** migration -> **Capital City** overseas migrants gain -> **International** migration) . Each year many people from overseas come to Australia. The distribution of these migrants across the country, however, is uneven, with some states and capital cities drawing more immigrants than others. Recognizing the importance of understanding these migration patterns, we undertook this project to visualize these trends. Our aim was to capture the number of migrants in each state, their propensity to settle in capital cities, and their countries of origin. The visualization creates a simple understanding and develops relationships comprising dataset and information flow. Anyone with an interest in migration and especially overseas skilled migration would want to analyze the combined visualization from this website. The beauty of this combined visualization is, it offers information in three layers of interconnected information for an overseas migrant. The importance of this data is immense as it offers key insight of Australia's planning and development policies, Resource allocations and Human resource development.

1.2 Visualization Purpose

Sankey and donut chart visualization (Net interstate migration analysis)

For this visualization, we have decided to focus on the net interstate migration dataset. The greater purpose of the visualization is to showcase the relationship of migrating people from different **states** against natural increase, overseas effect on internal migration and internal migration with capital **cities**. This creates the first layer of visual information (State à capital cities) on our combined visualization. Moreover, the visualization creates a simple



understanding of this complex concept through **data manipulation** and develops relationships through **visualizations** comprising information flow. In the Australian Bureau of statistics website, Overseas, internal and natural increase migration data is used in correlation with net interstate migration data to show the comparison in numbers of migrating people. This visualization shows the same compare and contrast visually. So, the intended audience can get answers for a lot of questions from this visualization.

Choropleth Map and Bar Chart Visualization (Net Overseas Migration Analysis)

Our visualization incorporates crucial data from the year 2022, bringing to light the preferences of immigrants when selecting their new Australian homes. This visualization aims to shed light on the dynamics of foreign migration throughout the various states of Australia. Every year, a large number of immigrants from a wide variety of nations come to Australia because it serves as a magnet for world residents. By showing both the migrants' preferences for their home states and their selections for the capital cities, our visualization aims to clarify this complicated picture.

The visualization tries to reveal the degree to which each state draws immigrants from outside, providing insight into how this inflow of immigrants is dispersed across Australia's vast geographic area. The focus simultaneously shifts to the capital cities, revealing the tendency of immigrants to live in major urban centers. In order to highlight patterns of urban preference among the migrant population, the proportion of immigrants choosing to live in the capital compared to the total number of immigrants entering each state is carefully portrayed.

Additionally, our investigation delves deeply into the ancestry of these foreign residents. We offer a thorough bar chart showing the top 18 nations that make up the majority of Australian immigrants. This map provides a clear image of the multicultural mix for each chosen state, enabling us to understand the distinct make-up of the immigrant population.

Our visualization, which combines these many components, acts as an exploration tool meant to spark the reader's interest and encourage them to learn more about Australia's intricate migratory patterns. It allows for a thorough comparison and provides an all-encompassing glimpse of the complex nature of immigration from outside to Australia.


Questions

Some key questions that the visualization provides to the intended audience are as follows:

- How does net interstate migration compare to natural increase, overseas migration, and internal migration with capital cities in each state?
- Are there any states where net interstate migration is significantly higher or lower compared to other migration factors?
- How does net interstate migration contribute to the population growth or decline in each state?
- Are there any patterns or trends in the relationship between net interstate migration and other migration factors across different states?
- Which capital cities attract the most migrants from other states?
- Which states experience the highest net interstate migration?
- Are there any states where net interstate migration has a significant impact on the overall population change?
- Are there any states where net interstate migration and natural increase, overseas migration, or internal migration with capital cities show contrasting patterns?
- How does net interstate migration contribute to the population distribution across different states and capital cities?


Some key Questions from the Cheropleth map visualization is given below:

- 1.Which Australian state saw the most influx of international migrants in 2022?
- 2.How do the migration trends change between the different Australian states?
- 3.What percentage of immigrants made their homes in each state's capital city? Do they favor urban life in any discernible trends?

- 
4. Is there a relationship between a state's population or economic standing and the amount of immigrants it draws?
 5. Which nations are the main sources of immigrants in Australia, and how are they distributed throughout the states?
 6. How do the cultural backgrounds of immigrants differ amongst Australian states?
 7. Are there certain Australian states that are more popular with immigrants from a particular nation than others?
 8. How do the demographics of the Australian states, particularly the capital cities, change as a result of the immigrant population?

Benefits

The analysis of migration data yields a huge amount of information which needs to be summarized in order for the general audience to perceive. The visualization accomplishes this task wonderfully. The main benefit of the visualization lies in this regard. The visualization presents complex data and information flow in a simplified and easy-to-understand format. This makes it easier for users to analyze and understand the relationship between net migration in the 7 states of Australia, and overseas migration in capital cities. The visualization also allows users to identify patterns and trends in the data that may not be apparent from raw numbers. This can provide insights into Australia's planning and development policies, resource allocations, and human resource development. The visualization can aid decision-making processes related to migration by presenting data in an informative and interactive way. It can also provide valuable information to policymakers, researchers, and stakeholders who are interested in the relationship between migration and its impact on Australia.



This visualization design skilfully combines clarity and complexity to create a visually stunning tableau that conveys a compelling tale of international migration in Australia from a complicated matrix of data.

The design is really powerful yet being straightforward. It elegantly and simply communicates complex migratory patterns, making it accessible to a wide range of audiences regardless of their demographic expertise. The audience may quickly identify patterns and anomalies by comparing states and making comparisons between them.

But what really distinguishes this visualization is the breadth of the information. An extra degree of sophistication is added by the inclusion of data from the top 18 countries of origin, which reveals the rich tapestry of multicultural influences that shape the demographic landscape of Australia. The viewer's understanding of the immigrant population's vast diversity and cultural plurality is enhanced by this nuanced vision.

A panoramic picture of the immigrant landscape is also captured by this design, encompassing both the enormous geographical expanses of the states and the vibrant, bustling energy of the capital cities. The audience may track the movement of immigrants not just across state boundaries but also across urban and rural differences inside those borders thanks to the holistic approach's wide-ranging insights.

This design's ability to encourage exploration is one of its most appealing features. It gently prods the audience to ask more probing questions, arousing interest and fostering discussion. The audience is drawn into a lively investigation of the facts by this catalyzing impact and the possibilities for involvement.

This architecture is a potent tool for data-driven decision-making, as well. Its findings may be used by academics, social scientists, and policymakers to develop sensible legislation, customize urban design, and allocate resources.

In essence, this visualization design is an intricate dance of data and aesthetics - a choreography that enlightens, engages, and inspires.

1.3 Project Schedule

Week 1:

Understand project requirements and objectives.

Week 2:

Draft project plan and timeline.

Week 3:

Gain basic understanding of chart visualization.

Week 4:

Research net interstate migration dataset and visualization techniques.

Week 5:

Exploring Sankey, Choropleth diagram and gathering resources.

Week 6:

- Explore various design options for the donut chart.
- Gather data for donut charts from the Australian Bureau of Statistics.

Week 7:

- Explore Sankey, Choropleth diagram visualization and its implementation.
- Develop the basic structure and components.
- Process data for the donut, bar charts visualization.

Week 8:

- Experiment with different layouts and node arrangements.
- Review and refine donut, radial chart visualization.

- Ensure design consistency.

Week 9:

- Enhance user experience with tooltips and labels.
- Integrate donut chart with the Sankey diagram.
- Integrate Bar charts chart with the Choropleth diagram.
- Iterate on design and functionality based on feedback.

Week 10:

- Conduct thorough testing and debugging.
- Refine user interface and interactions.
- Enhance with polylines and tooltips for donuts.

Week 11:

- Conduct usability testing for feedback.
- Test functionality and interactivity.

Week 12:

- Finalize project report.
- Submit deliverables by deadline.
- Reflect on the project's strengths, lessons learned.

2.1 Data source

The data used in the four visualizations are obtained from the Australian Bureau of Statistics. The data collection has been accomplished in distinct methods for the different models of visualizations created on the website.

Sankey diagram:

For the Sankey diagram, certain criteria were to be considered before loading the data in the diagram. From the literature (), Sankey diagram focuses on flow of information that visually represents the flow of data, energy, or material through a system. It uses arrows of varying thickness to depict the quantity or volume of the flow. A Sankey diagram has few key components that constitute the visual story telling of it. The key components are:

Nodes: Identify the different entities or categories that participate in the flow. These can represent various stages, processes, or components in your system. Each node is represented by a rectangle or a box in the diagram.

Links: Define the connections or flows between the nodes. These links indicate the movement or transfer of data, energy, or material from one node to another.

Flow values: Assign values or quantities to the flows between the nodes. This data represents the magnitude or amount of the flow and determines the width of the links in the Sankey diagram. It could be numerical values or percentages.

Direction: Specify the direction of the flow between the nodes. This information helps determine the layout and arrangement of the diagram.

Labels: Provide labels or names for the nodes and the links. This helps identify and describe the entities and flows in the diagram, making it more understandable for viewers.

Source of data for Sankey diagram: (Australian bureau of statistics) [Migration, Australia, 2019-20 financial year | Australian Bureau of Statistics \(abs.gov.au\)](#)

The key findings section from the (dataset type: **table**) is used to load into the links. There are three different kinds of Sankey diagrams for visual representation. This website uses the flow Sankey diagram style. Nodes act as the “Source and target” for the links to originate and terminate. These links showcase the flow of net migration from one state to another. The thickness of the links corresponds to the quantity or volume being transferred. This is a wonderful feature of Sankey which allows the user to develop a relationship between numbers of the data with the perceived understanding. Similarly, the nodes of the Sankey diagram have adjusted height values according to the data. These node heights basically allow the user to understand that the states (represented by the nodes) had a higher or lower number of people migrating. The data set includes

information of two attributes. The "State" attribute indicates the geographic locations. The data set provides the numerical values representing the net overseas migration for each state. "Net overseas migration (Persons)" attribute reflects the quantity of individuals involved in the net overseas migration for each respective city. In terms of data type this tabular form describes the dataset.

Table 1 "Dataset: Australian Bureau of statistics, Net instate migration."

Attribute	Description	Application in visualization
States	The data for the state attribute is categorical in nature. Each value represents a distinct category or label, identifying a specific state.	Nodes
Net overseas migration (Persons)	The data for the net overseas migration attribute is quantitative and represents a ratio or count of persons.	Links

There are two more fields from the table that have not been used in the Sankey flow diagram (namely the arrivals and departure fields). The flow of information provides better perceived information to the viewer if the net migration is showcased in the Sankey flow diagram (Data processing section has more information on the calculation of net migration for certain states). The inclusion of arrivals and departures could not support the flow

model of the Sankey and disrupt the visual understanding of the viewer. A cluttered visualization would not be apt to answer the **outlined questions**.

Donut chart:

Source of data for Donut diagram: (Australian bureau of statistics) : [Regional population, 2021-22 financial year | Australian Bureau of Statistics \(abs.gov.au\)](#)

Data set type: Table.

Attributes:

Attributes	City	Natural Increase	Internal Migration	Overseas Migration
Description	This attribute represents the geographic locations.	The numerical values representing the population change due to natural increase (births minus deaths) for each city.	The numerical values representing the population change due to migration within the country (movement between different regions or cities) for each city.	The numerical values representing the population change due to migration from other countries (international migration) for each city.

Data types:

Attribute	Data type	Description
City	Categorical	Each value represents a distinct category or label, identifying a specific city.
Natural Increase, Internal Migration, and Overseas Migration	Quantitative	They are ratio/quantitative data types, as they involve numerical values that can be measured and compared.

No exclusions from this table. All the data fields have been utilized to convey detailed information on the migration flow for answering the **outlined questions**.

Choropleth Map:

To make a choropleth map at first I have collected the json file to draw the australia map in the screen. And I have got the Australia json file from the github which was contributed by Gerado Furtado. And the link for the Aust.json is:

https://gist.github.com/GerardoFurtado/02aa65e5522104cb692e?short_path=7003276

```

1  {"type":"FeatureCollection","features":[{"type":"Feature","properties":{"STATE_CODE":"1","STATE_NAME":"New South Wales"},"geom
2  {"type":"Feature","properties":{"STATE_CODE":"2","STATE_NAME":"Victoria"},"geometry":{"type":"MultiPolygon","coordinates":[[[1
3  {"type":"Feature","properties":{"STATE_CODE":"3","STATE_NAME":"Queensland"},"geometry":{"type":"MultiPolygon","coordinates":[[[
4  {"type":"Feature","properties":{"STATE_CODE":"4","STATE_NAME":"South Australia"},"geometry":{"type":"MultiPolygon","coordinates
5  {"type":"Feature","properties":{"STATE_CODE":"5","STATE_NAME":"Western Australia"},"geometry":{"type":"MultiPolygon","coordinat
6  {"type":"Feature","properties":{"STATE_CODE":"6","STATE_NAME":"Tasmania"},"geometry":{"type":"MultiPolygon","coordinates":[[[1
7  {"type":"Feature","properties":{"STATE_CODE":"7","STATE_NAME":"Northern Territory"},"geometry":{"type":"MultiPolygon","coordinat
8  {"type":"Feature","properties":{"STATE_CODE":"8","STATE_NAME":"ACT"},"geometry":{"type":"Polygon","coordinates":[[[148,80484547
9  {"type":"Feature","properties":{"STATE_CODE":"9","STATE_NAME":""},"geometry":{"type":"MultiPolygon","coordinates":[[[150,59397

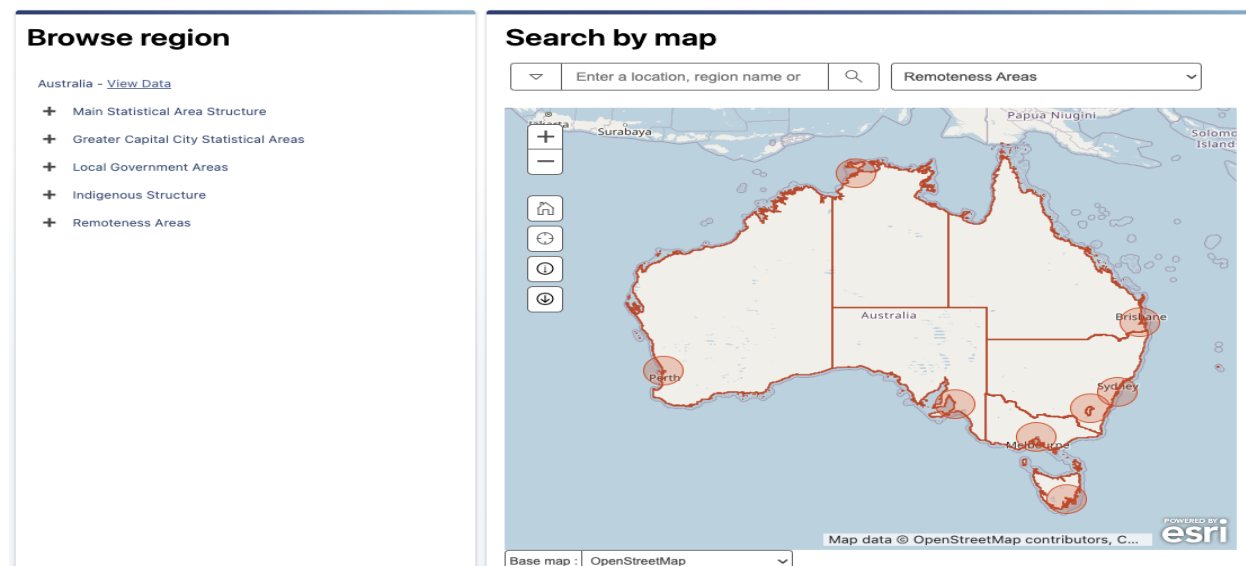
```

And then to make a choropleth map the most important thing is the migration data of the states. Which I collected from the Australian Bureau of Statistics. And the link is:

<https://dbr.abs.gov.au/>

Data by region

For information about the statistics and geographies view [Data by region methodology](#).



Well the data collection took a long time for me, because from the each region I have to manually collect the data and make a separate csv file input the data. That's how I got the regional data.

After that as I was going to do the bubble map on top of the choropleth map to show the capital of each states and the migration of the state, I had to collect the Latitude and the longitude of the cities so that I can be make a bubble in the exact location. So I took help

from a website called: simple maps. And the link of the website is:

<https://simplemaps.com/data/au-cities>



Australia Cities Database

Below is a list of 323 prominent cities in Australia. Each row includes a city's latitude, a subset of all 13,842 places in Australia (and only *some* of the fields) that you'll find in subset for free under an [MIT license](#). You're free to use the data below for personal or downloaded in [.csv](#), [.json](#), and [.xlsx](#) formats.

Notable Cities: The capital of Australia is *Canberra*, which is in the territory of Australian Capital Territory, with an urban population of 4,840,600.

Download Data in Spreadsheet (323 cities) ▾


Get all cities in Australia (13,842 cities)

Loc

city	lat	lng	country	iso2	admin_name
Sydney	-33.8678	151.2100	Australia	AU	New South Wales
Melbourne	-37.8142	144.9631	Australia	AU	Victoria
Brisbane	-27.4678	153.0281	Australia	AU	Queensland
Perth	-31.9559	115.8606	Australia	AU	Western Australia
Adelaide	-34.9275	138.6000	Australia	AU	South Australia
Gold Coast	-28.0167	153.4000	Australia	AU	Queensland
Cranbourne	-38.0996	145.2834	Australia	AU	Victoria
Canberra	-35.2931	149.1269	Australia	AU	Australian Capital
Central Coast	-33.3000	151.2000	Australia	AU	New South Wales
Wollongong	-34.4331	150.8831	Australia	AU	New South Wales
Hobart	-42.8806	147.3250	Australia	AU	Tasmania
Geelong	-38.1500	144.3500	Australia	AU	Victoria
Townsville	-19.2500	146.8167	Australia	AU	Queensland
Ipswich	-27.6167	152.7667	Australia	AU	Queensland
Newcastle	-32.9167	151.7500	Australia	AU	New South Wales
Cairns	-16.9200	145.7800	Australia	AU	Queensland
Darwin	-12.4381	130.8411	Australia	AU	Northern Territory

And then to collect the migration report of the city I again head towards the ABS website. And went to the data category by major cities and collected the data and that's how the choropleth bubble map is completed.

For every state when hovering over the states I have shown a Bar chart which is showing the data of the top 18 countries from where people came to the state and live there. I have



also collected the data from the Australian Bureau of Statistics. That's how the bar chart is done.

2.2 Data processing

Sankey diagram:

For the Sankey diagram the link is loaded with the net migration value from the key findings section in the ABU website. But few states had only the arrivals and departures, not the actual net migration value. This calculation was completed in excel.

Steps:

- The key findings section provided most of the values for the Sankey diagram.

Table 2"Example of key findings section used to obtain data."

were at their highest since December 2019.

- In net terms, New South Wales only gained people from Victoria (+880) and lost most to Queensland (-3,800).

Greater Sydney

- There was a net loss of 8,200 people for Greater Sydney in the March 2021 quarter, compared with 9,300 in the previous quarter.
- Sydney had net losses of 5,100 people to the rest of the state and 3,100 interstate.

Victoria

- There was a net loss of 4,900 people from internal migration in the March 2021 quarter, compared with a loss of 6,500 people in the previous quarter and a gain of 590 in the March 2020 quarter.
- Interstate arrivals increased from 18,500 in the previous quarter to 18,900 in the March 2021 quarter, while departures decreased from 25,000 to 23,800.
- March 2021 quarter arrivals were at their highest for a quarter since the December 2019 quarter, and departures were at their lowest since September 2020.

- In net terms, Victoria did not gain people from any state or territory and lost most to Queensland (-2,900).

Greater Melbourne

- There was a net loss of 8,300 people for Greater Melbourne in the March 2021 quarter, compared with 8,500 in the previous quarter.
- Melbourne had net losses of 4,800 people to the rest of the state and 3,400 interstate.

-This dataset section is used to show the displaced count of interstate migrants. The cleaning up of the data was done by omitting the unwanted information from the spreadsheet. Data for Australian Capital and Northern Territories are not included in the Sankey diagram.

- Then creating a new header for a table named as net migration I made a simple subtraction using excel built in functions. **(Note that this section is used to showcase the net migration happening in each of the states, not the flow of migration from one state to another.** The must have section has more information on this feature.) When hovered over the nodes this migration data can be seen at the top with transition.

	A	B	C	D	E	F	G	H
1	Graph 4.1 Interstate migration - arrivals and departures - Australia - year ending June 2020(a)							
2	State	Interstate arrivals	Interstate departures	Net migration				
3	NSW	89,873	110,760	20,887				
4	Vic.	74,954	77,197	2,243				
5	Qld	76,441	101,789	25,348				
6	SA	23,726	25,886	2,160				
7	WA	29,211	31,621	2,410				
8	Tas.	11,749	12,962	1,213				
9								
10								
11								
12	a. Estimates for 2019-20 are preliminary - see Interstate migration in paragraph 9 of the Methodology.							
13								
14	Source: Australian Bureau of Statistics, Migration, Australia 2019-20 financial year							
15								
16								
17								

-Finally, the net values are loaded into the visualization.

Donut diagram:

This section of the data set did not require substantial data clean up because from the ABU website the formatted data was sourced.

Table 3 "Formatted data table for donut chart"

Graph Table

Download

Components of population change by capital city

	Natural Increase	Internal Migration	Overseas Migration
Brisbane	16,328	28,111	14,717
Melbourne	27,011	-26,212	54,239
Sydney	34,177	-51,738	54,886
Perth	13,426	9,139	9,681
Adelaide	4,613	-64	11,513
Canberra	3,252	-3,318	3,200
Hobart	801	-798	1,643
Darwin	1,543	-2,289	1,527

The arcs of the donut chart are loaded to occupy the three formatted data variables.

Data Processing for the Choropleth Map:

The json file for the map was already prepared and processed. So I didn't have to change anything for the json file. The screenshot of the json file is:

```
1 [{"type": "FeatureCollection", "features": [{"type": "Feature", "properties": {"STATE_CODE": "1", "STATE_NAME": "New South Wales"}, "geometry": {"type": "MultiPolygon", "coordinates": [[[["159.070808704", -31.518288762], [159.084682656, -31.543376649000002], [159.098790528, -31.5586035365], [159.090078912, -31.580085557], [159.07853824, -31.592290649000002], [159.068456, -31.595036992500003], [159.068912, -31.5779779945], [159.074980128, -31.5705556835], [159.076747808, -31.542503708], [159.057703968, -31.520500844], [159.070808704, -31.518288762]]], [[["150.701604288, -35.122861729], [150.661052992, -35.117823957], [150.657864128, -35.105357954], [150.64039296, -35.1034419645], [150.639977152, -35.1132890335], [150.621445088, -35.108865998], [150.611643008, -35.1001479655], [150.594382176, -35.0996854285], [150.5873832, -35.088590479000004], [150.563691008, -35.093661033000004], [150.564723136, -35.1054970185], [150.556159072, -35.118028974], [150.57397584, -35.12253402], [150.586460992, -35.120234969500004], [150.596947072, -35.130944971], [150.592867936, -35.143937003], [150.604695712, -35.157900415], [150.601698016, -35.168886454500004], [150.591248992, -35.1724724945], [150.593970976, -35.1850184915], [150.572751008, -35.193035004], [150.539507008, -35.216613994], [150.530678016, -35.232444], [150.541479008, -35.245496009], [150.514506016, -35.267739151], [150.501200992, -35.2723790065], [150.479656, -35.289831], [150.472027008, -35.316660995], [150.476044, -35.346913009], [150.484914144, -35.3511535975], [150.486786016, -35.365900003], [150.466726016, -35.376267995], [150.448190016, -35.3942309955], [150.445442368, -35.407571993], [150.42001408, -35.424163052000004], [150.391113024, -35.468995064], [150.39571104, -35.4799570355], [150.389432256, -35.488337036000004], [150.392456, -35.504443006500004], [150.401128992, -35.5120600035], [150.407670016, -35.5296159935], [150.391644, -35.537441993], [150.375959008, -35.5719009995], [150.364722016, -35.570780991], [150.338540992, -35.6053850005], [150.32256, -35.613879997], [150.329806016, -35.6302929935], [150.314383008, -35.635299001], [150.299379008, -35.647214], [150.297391008, -35.662983992], [150.302886016, -35.667934], [150.300572, -35.683164994500004], [150.276795008, -35.7111270045], [150.277700992, -35.7193450005], [150.262124, -35.721131009000004], [150.246834016, -35.709789991], [150.228811008, -35.702560006], [150.198299008, -35.701532997], [150.179810048, -35.708734973], [150.19854, -35.719091994500005], [150.199226016, -35.729590004500004], [150.216780992, -35.7457260005], [150.211872, -35.75999], [150.234500992, -35.7753570065], [150.236731008, -35.7869879935], [150.225128992, -35.828104003], [150.186699008, -35.8331160045], [150.178083008, -35.8438830045], [150.178700992, -35.853539005500004], [150.156288, -35.876817000500004], [150.150243008, -35.8911389975], [150.157802016, -35.9087420065], [150.160958016, -35.933253008], [150.155196992, -35.947550992000004], [150.160124, -35.9547360035], [150.151067008, -35.9578600035], [150.159812, -35.9968430025], [150.160812992, -36.013432008], [150.149238016, -36.0288100035], [150.141352992, -36.0493549935], [150.141463008, -36.063427000000004], [150.13256, -36.082401007], [150.137539008, -36.097025997], [150.126071008, -36.129105996], [150.124968, -36.158294001], [150.135455008, -36.170687003000005], [150.131882016, -36.182134007500004], [150.131732,
```

After that I had to collect the data for all the states first and then input them into the csv file. After that I collected the land and lon of all the major cities in all the states and after that collected the data for the city migration. In the csv file we can see the data is for 2 years. Initially was thinking about doing the choropleth for 2 different years but after that, due to the lack of data and to collaborate with my teammate I have dropped the plan of doing for 2 different years and went for 1 single year. The screenshot of the csv file is given below:

state	city	lat	lon	citym_20	citym_16	statem_20	statem_16
New South V	Sydney	-33.8678	151.21	50083	80976	58300	104478
Victoria	Melbourne	-37.8142	144.9631	56083	76498	61300	90009
Queensland	Brisbane	-27.4678	153.0281	16169	31656	28600	35199
Western Aus	Perth	-31.9559	115.8606	22214	12998	25000	13384
South Austra	Adelaide	-34.9275	138.6	13665	11243	14900	11668
ACT	Canberra	-35.2931	149.1269	2368	3598	2400	3964
Tasmania	Hobart	-42.8806	147.325	2113	1998	3700	2029
Northern Ter	Darwin	-12.4381	130.8411	132	1729	200	1748

At last I have collected the data for the top 18 countries and the data was prepared to use for what I was trying to do. I just removed some of the extra commas and extra text which was unnecessary. The screenshot of the dataset is given below:

State	NSW	VIC	QLD	SA	WA	TAS	NT	ACT
China	256100	176600	51600	26800	30000	3300	1400	11900
India	153800	182800	53100	29000	53400	2100	4200	10900
Philippines	94100	56000	43100	13300	33400	1800	7000	4000
Vietnam	92900	89900	21800	15700	17600	500	1200	3600
Italy	56000	79100	14800	20200	22000	1000	600	2100
South Africa	47700	30300	44700	7200	45600	1700	1100	2200
Malaysia	34900	55300	16400	8400	32300	1600	800	2600
Scotland	33100	30300	25200	13300	30400	2600	900	2000
Sri Lanka	32600	63200	11100	4200	8900	400	1000	3100
Germany	33700	30000	23100	11200	12800	2400	1200	2500
Greece	36000	57000	4000	10200	2800	600	1800	1200
South Korea	58000	16900	21000	4100	7900	700	500	2500
USA	36900	24300	20900	5200	11500	2000	1700	3300
Hong Kong	47300	24400	12200	4000	7000	500	400	1800
Lebanon	68400	19900	1700	1800	1300	100	100	500
Ireland	28200	18200	13300	3700	21600	800	1400	900
Indonesia	34100	19300	7900	2300	12500	400	1300	1400
Netherlands	18900	22200	15800	7200	10400	2500	500	1100



[Overseas Migration, 2021-22 financial year | Australian Bureau of Statistics \(abs.gov.au\)](#)

Data source of sankey and choropleth

[Labour Force, Australia, March 2023 | Australian Bureau of Statistics \(abs.gov.au\)](#)

[Overseas Migration, 2021-22 financial year | Australian Bureau of Statistics \(abs.gov.au\)](#)

3.1 Must have visualization features

Sankey Chart:

- **Node Layout:** The Sankey chart displays nodes (representing entities or categories) and links (representing flows or connections) between nodes.
- **Flow Direction:** The chart clearly indicates the flow direction of the links, typically from the source node to the target node.
- **Link Thickness:** The thickness of the links should represent the magnitude or value of the flow being visualized.
- **Node and Links Interactivity:** Users are able to interact with the nodes and links, such as hovering over them to display additional information or clicking on them to trigger actions.
- **Tooltips:** The chart provides options of tooltips, to enhance the visual representation and contextual information.

Donut Chart:

- **Circular Layout:** The Donut chart displays a circular layout arc representing different categorical values.
- **Proportional Sizing:** The size of each arc is proportional to the value or percentage it represents, allowing for easy comparison between categories.
- **Colors and Legend:** The chart uses distinct colors for each category.
- **Labels:** Each arc is labeled with the corresponding value to provide clear identification.


Choropleth and Bar Chart:

1. **Geographical Accuracy:** Due to the nature of the data, precise geographic representation is crucial. To ensure the reliability of the immigration statistics being visualized, each state and its capital city should be depicted appropriately.
2. **Clear Indicators:** To aid comprehension, changes in color depth should clearly reflect each state's immigration inflow. Similar to this, the number of migrants in each capital city should be appropriately represented by the size of the bubbles on the bubble map.
3. **State-Specific Visualization:** To enable viewers to obtain state-specific data, interactive features should be included. The visitor should be given a thorough overview of immigration from other nations by just hovering over a state.
4. **Top Country Bar Chart:** A bar chart with the top 18 immigration nations is necessary to depict the varied origins of immigrants. Depending on the state chosen, this visualization should alter dynamically to represent the distinct characteristics of each state's migrant population.
5. **Seamless Integration:** The choropleth map, bubble map, and bar chart should all work together as one cohesive visualization. In order to produce a coherent, captivating, and educational visual tale on immigration in Australia, each element should enhance the others.

3.2 Optional visualization features

(Sankey and donut charts)

In terms of features, the combined visualization maximizes all the available interactivity the charts can offer. One feature I personally wanted to add was making the color of the charts color blind friendly.



Colorblind-friendly palettes aim to provide colors that are more easily distinguishable for individuals with various types of color vision deficiencies. D3 provides a specific color scheme called **d3.schemeTableau10** that is designed to be colorblind-friendly. It consists of 10 colors that are carefully selected to be distinguishable even for people with common forms of color blindness. I have successfully added this to my chart color palette with an extra button for the user interactivity.


(Choropleth Map)

We have taken it a step further in the visual portrayal of immigration from outside to Australia by including information at a national level as well. We have included a second bubble map feature in addition to the main map that shows how many migrants are traveling to each state. The number of migrants who are opting to settle in each state's capital city is explicitly depicted on this bubble map. As a result, it's not only a matter of which state has the most number of migrants; it also depends on how desirable the state's capital city is to them. This additional layer of data helps to provide a more thorough and in-depth picture of migration patterns in Australia.

4 Data visualization design (Sankey and donut)

Decision making

Initially we planned to opt for a stacked bar chart and choropleth combination. But while doing our research for the visualizations we came across the Sankey diagram. This chart particularly caught my attention and I decided to include this in our project. But the main challenge for me was this chart was not directly instructed in our course and I had to self-taught myself about it. My primary source of learning was this literature ([1](#)). The choice was easy, but the implementation was difficult as the Sankey diagram is a complex visualization graph. I have explained thoroughly about this diagram in the (Data section) while discussing the data input for this diagram. So, with this the display of data as flow of



information from two nodes was decided. A lot of general ideas were developed for it while working on the stand up 1 and 2. Ultimately the decision had to be changed for accommodating the dataset available in the Australian bureau of statistics. I believe that this showcases my understanding on data manipulation and encoding skills as the pivot to the dataset available was strenuous, but I pulled it off.

The decision to add donut charts was based on adding more interactivity to the visualization. The addition increased not only the interactivity, but the overall information conveyed increased significantly. This allowed me to showcase a compare and contrast scenario from the visualizations and answer key **outlined questions**. The donut chart was further modified with the addition of polylines. The connection between two diagrams with the hover feature establishes an ease of perceived information for the user. The general idea for the donut charts was to utilize them in a way where the user is fed with more information through user interactivity. I was successful in accomplishing this task, so I did not have to pivot to other decisions. However, the choice of using donut charts to showcase information of the selected dataset was an excellent one. The donut charts perfectly translated the raw numbers into visually processed information just the way I wanted it to.

Design evolution

Sketch 1 is from the time around first stand up where the idea of going for a Sankey diagram was being considered. Initially we thought that we could do a country-to-country data flow about migration. This was one of the **alternate ideas** for our visualization, but our tutor mentored us to have a more meaningful visualization pertaining to a specialized topic. As we gained more insight about this topic we started working on the finalization of the dataset and visualization model.

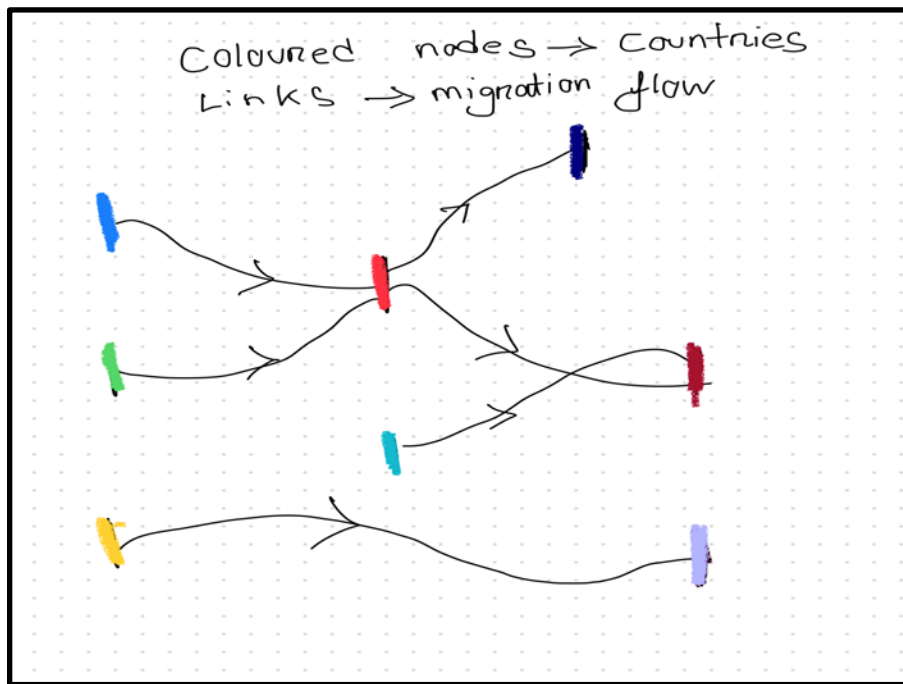


Fig 2 “Preliminary design sketch”

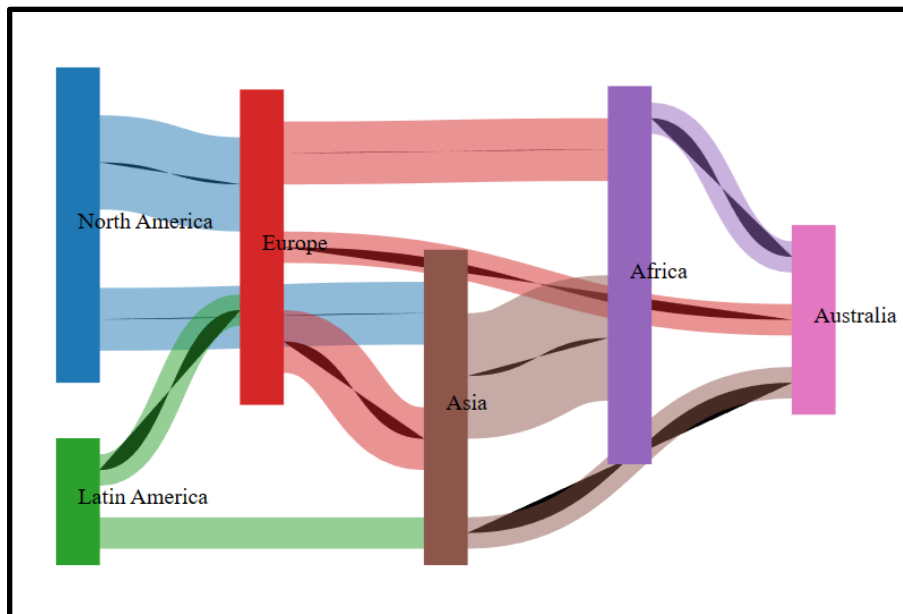



Fig 3 “Preliminary design implementation”

At that time there was no consideration on adding interactivity to the visualization. By week 9, I



finalized on the dataset for the Sankey diagram. One major issue in terms of designing this diagram was the d3 framework had another chart type called chord diagram to showcase the circular relationships in visualizations. Sankey diagram is used extensively to showcase the flow of information between two nodes. From my research I overcame this obstacle by using “the flow style “Sankey diagram. In my visualization the data is shown from one direction only (left to right) and duplicate nodes are created to aid in this process. Sketch 2 showcases the alternate design idea, and it could clutter the diagram extensively as there would be too many links. So, I prioritized clarity and simplicity in this case. Other general ideas were using the color and contrast to aid the visual process, adding interactivity through the links and nodes, making the chart consistent with contextual information. Tooltips on the links, contextual information through “svg” on hovering over nodes added great details to the visualization. All of the general ideas were implemented to the final design. I utilized all the scope of interactivity in this diagram linking to the donut charts.

The donut charts were finalized at the last step of the designing process. The data availability was a huge challenge at this stage. There were multiple data files that I thought would be suitable for the chart, but the design guidelines prohibited me to do so. Furthermore, I emphasized a flow of information style chart combination with these two and when I was adding a lot of contextual information to the donut charts it was very difficult to get the sensible perceived information out of it as a viewer. The (DataRocks, 2021) literature advises the visualization creator to ensure the audience picks up the perfect balance between complex information and efficient visualization.

Final updated design:

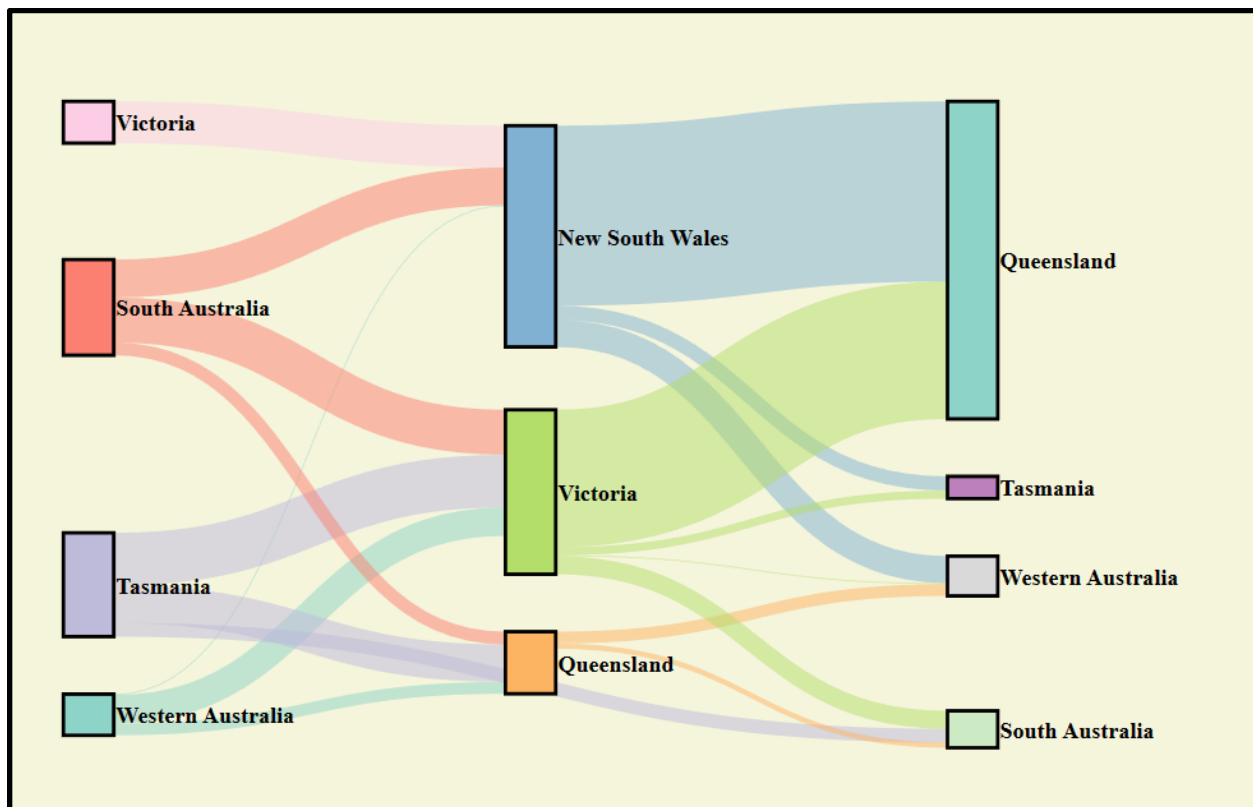


Fig 4 "Final updated design(1)"

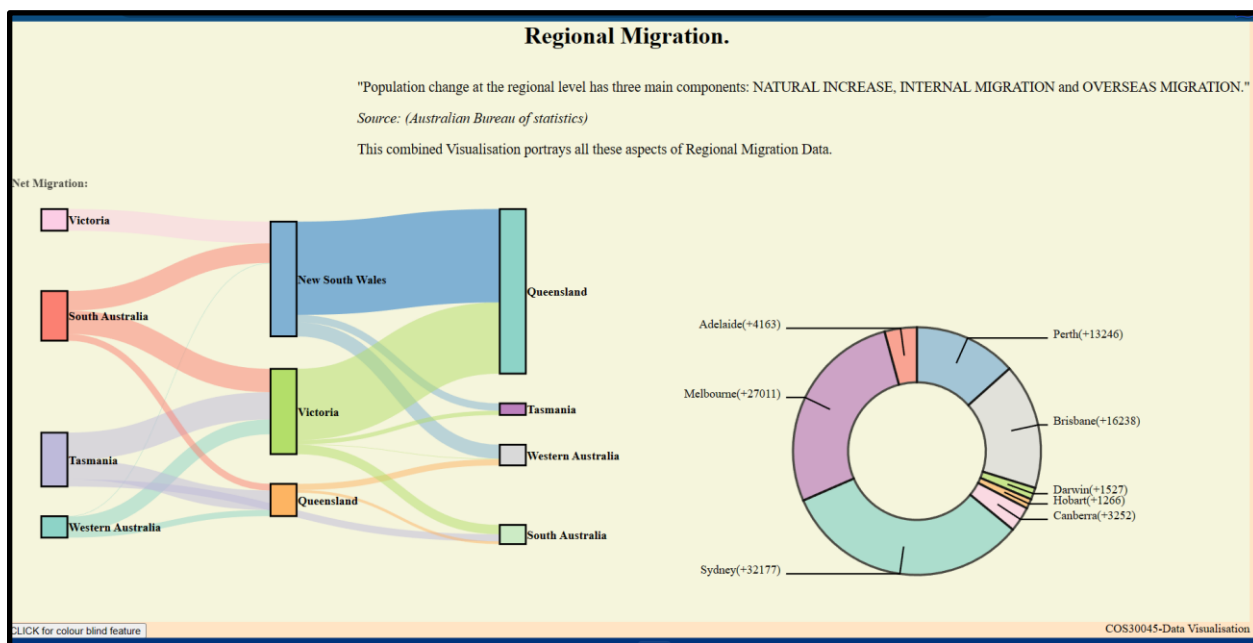


Fig 5" Final updated design(2)"

Net interstate migration analysis → Sankey

Net migration analysis of Interstate, overseas, internal data migration → Donuts.

In the stage of Low-fidelity prototyping, I ensured that the guidelines were a top priority in the visualization.

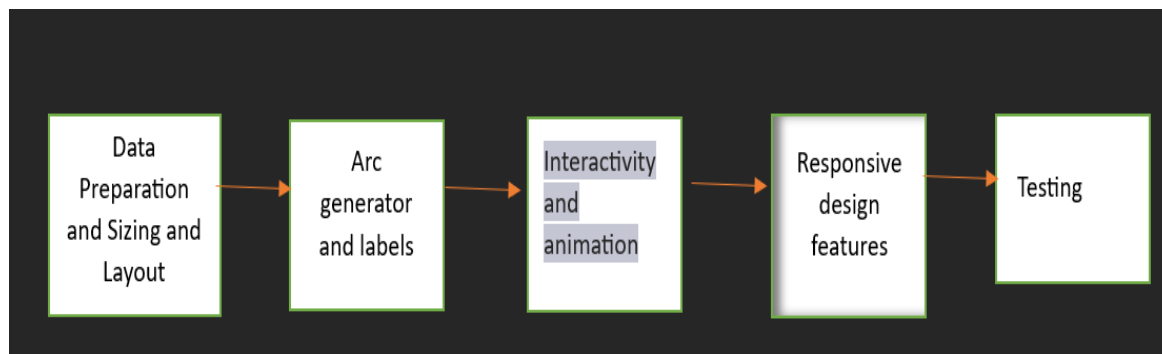
Low-fi prototyping.

Considerations	Implementation
Basic Representation	Conveying the core functionality and structure through tooltips, labels, and polylines(donut)
Rapid iteration capabilities	The Sankey was changed a lot of times when put under this iterative design load, especially the node alignment while loading data was direct use of this consideration guideline.
Risk Reduction	The iterative concept and feedback from the tutor helped reduce the risks of following a faulty visualization guideline.

Greater exploration options


Finally, the low fidelity prototyping allowed a great deal of interactivity exploration in the combo chart model. The nodes and links hovering interactivity were accomplished only because there was stepwise adherence to low fi prototyping.

Basic design guidelines steps (Sankey and donut)



Adding a lot of screenshots and pictures from previous standups

Design justification and discussion



The main justification of this visualization design would be to prove if it can peak the interest of the intended audience. The specific audience here is Researchers, Policy makers, Urban planners, and the general public. Every section of the audience might have different motivations like the Urban planners may see this as a compare and contrast model or the general public may see this as a knowledge pit. So, if the visualization is able to cater to such a broad range of audience it can be undoubtedly said that the overall design process is justified.

The aim of the visualization is to allow the users to develop a compare and contrast analysis in different layers of the information flow. The analysis of net migration whether it be in international level, state, capital cities it allows the user to note the key differences in the data elements.

The flow of information style in the Sankey relates two attributes of data together. One example of this concept is when the arrival and departure information and net migration count was integrated into the Sankey through the links and tooltip above. Moreover, the nodes are designed as a bridge between these two data attributes. The user through simple hovering understands the differences and interprets the changes in trend of the data. Donut chart also offers the same comparative model. Size difference in arcs and included polylines provides key determinants for analysis.

Data visualization design (Choropleth and Bar Chart)

In our quest to create a comprehensive visualization of overseas migration to Australia in 2022, we employed a meticulous understanding of the context and applied astute visualization design principles to convey our findings effectively.

The first layer of our design portrays the volume of migrants settling in each state, allowing an at-a-glance comparison of their relative attractiveness. A deeper dive into the urban-rural divide illuminates the proportion of migrants opting to dwell in capital cities, shedding light on their urban inclinations. In our immersive visualization, we've chosen the canvas of an Australian map to display each state in its unique geographical context. To capture the volume of migrants each state attracts, we've employed the visual appeal of a choropleth

map, turning data into a vivid spectrum of colors. By simply observing the hue of each state, viewers can glean insights into the scale of migration. A deeper color saturation signifies a higher influx of migrants, painting a picture of a state teeming with newcomers. Conversely, lighter shades denote fewer migrants, indicating states where migration was more subdued. This fusion of geography and color in our visualization offers an intuitive understanding of migration trends across Australia, transforming raw data into a chromatic narrative of human movement.

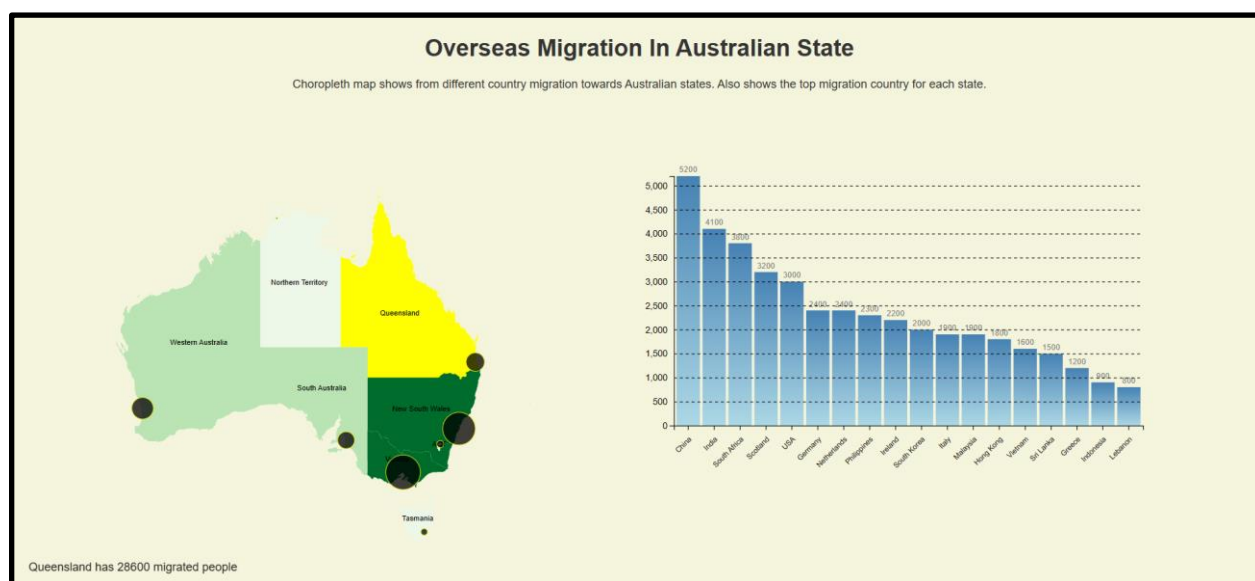



Fig 6” Final updated design of choropleth”

In our compelling second visualization, we've merged the appeal of a choropleth map and a bubble map to dynamically illustrate the geographical spread of overseas migration within Australia. Each bubble, placed meticulously using the longitude and latitude coordinates of each capital, acts as a magnet, signifying the allure of the capital cities among new arrivals to each state. This design captures not just the influx of migrants into each state, but also their propensity to gravitate towards the beating heart of the state - its capital city. The interactivity of the bubbles introduces a novel layer of engagement; on hovering, a tooltip materializes, offering a glimpse into the specific values associated with each capital city. This fusion of maps, bubbles, and tooltips crafts a fascinating visual narrative that is both engaging and informative.



To imbue our visual story with an extra layer of multicultural richness, we've ingeniously woven a detailed bar chart into our design, showcasing the top 18 countries of migrant origin. This striking element reveals the global interconnectedness of migration trends and uncovers distinctive state-by-state patterns, elucidating the diverse cultural influences shaping Australia's vibrant demographic mosaic. Cleverly concealed at first glance, this element morphs into an interactive exploration tool as viewers engage with the map. As they hover over each state, a vibrant cascade of data unravels, rendering a unique narrative of migration for that region in the form of a meticulously ordered bar chart. This interactive design feature not only enhances audience engagement but also invites a deeper understanding of the complex tapestry of immigration that threads its way across the Australian continent.

This layered visualization is carefully tailored towards a diverse audience, catering to the interests of policymakers, researchers, potential immigrants, educators, and anyone intrigued by the dynamics of overseas migration. It equips them with insights and prompts further enquiry, thereby serving its dual purpose of educating and engaging the viewer.

Our design process was iterative and explorative. We generated numerous sketches and considered multiple alternatives before arriving at our final design. Each step was driven by feedback and refined understanding, enabling the evolution of our design from a rudimentary concept to a polished, professional-level visualization.

This visualization design, enriched by our creativity and commitment to excellence, represents an insightful response to the assignment task, enabling a thorough understanding of Australia's 2022 migration landscape. It provides an accessible window into the complex interplay of people, places, and cultures that shape Australia's vibrant immigrant population.

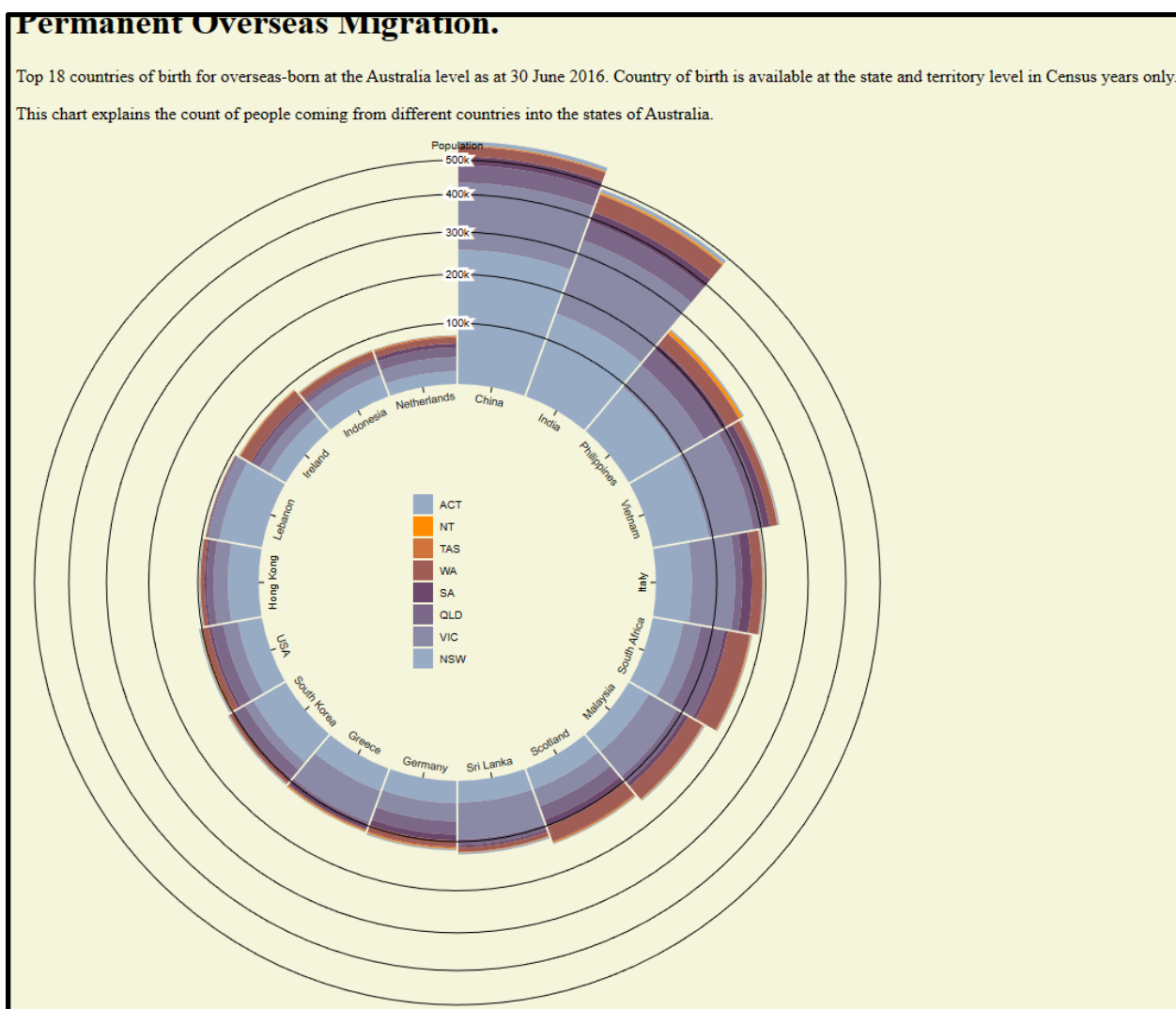


Fig 7” Final updated design of radial chart”

Validation

For validation, the case study of (PRB, 2001) was studied with our combined visualizations. We asked for peer reviews on how the visualizations would benefit the analysis of the case study. Most of the peers suggested that the visualization helped them understand the case study better as they were not required to navigate through raw numerical data. The data accuracy was also judged wonderfully with this validation strategy.

Conclusion

In conclusion, this extensive research successfully translated complicated data into compelling visual tales, offering insightful information on the dynamics of interstate and international migration in Australia. It clarified migratory trends and cultural diversity using a variety of visualization approaches, such as Sankey diagrams, doughnut charts, and choropleth maps and stacked bar charts. Our ability to manipulate data and visualize it was improved through an iterative, feedback-driven process, and viewer engagement was increased through the use of interactive design that was targeted to the audience. The study demonstrated the value of powerful data visualization in comprehending intricate processes by providing a clear image of Australia's migratory landscape in 2022.

References

- Australian Bureau of Statistics. (2021-22). *Regional population*. <https://www.abs.gov.au/statistics/people/population/regional-population/2021-22>.
- DataRocks. (2021). Visualisation Analysis & Design by Tamara Munzner or the What-Why-How of data viz. *Visualisation Analysis & Design by Tamara Munzner or the What-Why-How of data viz*. https://www.datarocks.co.nz/post/data-viz-bookshelf_visualization-analysis-design-tamara-munzner
- PRB. (2001). International Migration Transforms Australia. <https://www.prb.org/resources/international-migration-transforms-australia/#:~:text=Between%201945%20and%202000%2C%20most,%2C%20economy%2C%20and%20settlement%20patterns>.
- Zhu, N. Q. (2017). *Data visualization with D3 4.x cookbook: discover over 65 recipes to help you create breathtaking data visualizations using the latest features of D3*. Packt. https://librarysearch.swinburne.edu.au/discovery/fulldisplay?docid=cdi_askewsholts_vlebooks_9781786469960&context=PC&vid=61SUT_INST:SWIN2&lang=en&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,data%20visualization%20d3&offset=0