

**Data Visualization Project (37%)**  
**FIT5147 - Data Exploration and Visualization**  
**Semester 1, 2024**

**Name: Rayyan Aamir**  
**Student ID: 32065647**

**Tutors: Gavin and Bruno**  
**Applied Session 7**

**Impact of Overseas Migration and Education Levels on**  
**Australia's Labour Force**

<b>1.0 Introduction.....</b>	<b>1</b>
<b>2.0 Design Process.....</b>	<b>2</b>
2.1 Sheet One.....	2
2.2 Sheet Two.....	3
2.3 Sheet Three.....	4
2.4 Sheet Four.....	5
2.5 Sheet Five.....	5
<b>3.0 Implementation.....</b>	<b>8</b>
3.1 Technical Implementation.....	8
3.2 Interactive Narrative Visualization Implementation.....	9
3.3 Using The Implementation.....	12
<b>4.0 Conclusion.....</b>	<b>13</b>
<b>5.0 Bibliography.....</b>	<b>15</b>
5.1 References.....	15
<b>6.0 Appendix.....</b>	<b>17</b>

## **1.0 Introduction**

Migration statistics and education levels are one of the biggest contributors towards changing levels in the labour force within Australia. Different states will have different levels of unemployment, participation, education, and migrations, this brings forth a variety of combinations of gaps in policies or even the population which leads to these numbers. The purpose of this visualization is to provide a general outlook of the biggest statistics that everybody is concerned with. The average person does not care about individual statistics that represent less than even 1% of the population, it is big numbers and statistics that sway an individual in deciding who to cast their vote for in the next election, it is when the population's goals, ambitions, and personal targets are aligned with the person running for the next mayor or prime minister.

With this visualization, the goal is to educate the population and also shed some light on topics and areas that are of concern to relevant policy-makers in their state. Overseas migration is one of the biggest contributors towards increasing high-skilled individuals in various categories of employment, primarily full-time and part-time lines of work. Education levels is another major contributor because the local Australian people also need to retain their jobs and not be pushed out by the people coming from abroad, therefore the market diversifies and grows, requiring individuals of other skills and qualifications that are unique to Australia.

After providing an overview of the issues at hand, it is the task of either the population to bring up their concerns with their local city council and go deeper into the fine-print as to what can be done to resolve these issues or to ignite some fire into policy makers to address these issues in a way that boosts the Australian economy and brings equality within the population through potential pay gaps between part-time and full-time employment. The end goal of this visualization is simply to shed some light on some statistics and figures using some easy to read figures and diagrams, hence making it highly appealing to the general population.

Through the cohesiveness and interactivity implemented within this visualization, it provides the target audience absolute customizability and subjects them to making their own conclusions. It is not possible for all micro-aspects of such a vast topic to be considered within a simple dashboard concept, however, for future projects such an implementation can be considered.

The dashboard aims to encourage the population to call for action and see the data through their own eyes and address their concerns with the right people who are willing to take the step towards bettering work opportunities and providing the right incentives!

## **2.0 Design Process**

Starting off with the planning stage for this project, the design process was conducted through brainstorming ideas and topics and producing dummy visualizations for these ideas, this was done through the Five Design Sheets methodology. This method essentially breaks down the structure of the topic into major categories and follows a clear set of stages and a linear approach towards the problem. This technique was an excellent way to adopt a framework which focused on the comprehensiveness of the exploration of the various ideas and datasets required for this project. Furthermore, it allowed for imagination and innovation to become real where the goal was to achieve these ideas through vigorous coding styles and techniques in order to maximize the end goal, a user-focussed design.

### **2.1 Sheet One**

Beginning with sheet number one, the purpose of the sheet was to have a general understanding of how things must flow throughout this project. Was it going to follow a martini glass styled approach, or just a regular glass type approach? The first step is always to start at the very top, which is why the first sheet focussed primarily on the map of Australia and how things were going to be based on that. Since the project targets the entirety of the country, it was important that the users had access to any and all information that is specific to the state they wish to study, this provides the utmost flexibility to the user, allowing them to come to their own conclusions from the data they see.

The aim here was to not dictate and throw information at the users where they just have a boring read of static visualizations, they must be interactive and allow control over the visualization, this engages the user into the story that we are trying to portray. A map is one of the most consistent ways to start off a story as this brings forth some major geographical relevance and allows the user to view statistics specific to that state and even allow for comparison where the difference in the size of the circle (radius) may determine some underlying issues that the user wishes to research.

Apart from the map which brings a very good opening to the project, the topics of migration and education requires some bold figures and diagrams where the user is extracting maximum information, whilst the project must ensure that the user is also able to understand these visualizations and is not spending significant time simply learning to read some simple statistics, and that the visualization actually makes sense. Beginning with, sheet one also displays the implementation of pie charts, bar and line graphs, word clouds, time series graphs, and heatmaps.

Pie charts were thought to have been useful to portray some proportions but might not display enough information and would require too much filtering for the user when looking at year-wise and state-wise data. Word clouds are only useful when studying in more detail for the topic of employment which this project does not focus on, but it would provide the user some insights into major job titles in full-time and part-time employment.

Heat maps would represent the relationship between various variables against one another, however they are difficult to read with accuracy. The goal here is to display the general statistics whilst going into a decent level of depth without having the user be confused.

## **2.2 Sheet Two**

Sheet two focuses purely on the aspect of migration and unemployment and how these statistics bring a change to the labour force. Whilst education is a big topic to consider, initially it was planned that maybe focussing on one of the ideas was more suitable. Hence, as the focus shifts more towards these variables and factors, using the map which was brainstormed in sheet one, this sheet aims to provide detailed analysis and information on how migration and unemployment are linked.

This time however, the map would be more detailed which would display diagrams within hoverable tooltips, this would enable the user to engage with the project in a more fine and detailed manner which would aim to facilitate a more personalized and meaningful analysis and draw the necessary insights. Furthermore, the implementation of a Sankey diagram serves as a very powerful tool which gives the sense of quantity to the different migration types by each state, because as per the human visual system, humans prefer to see numbers in size and figures.

If complex data is not visualized correctly, it makes it difficult for humans to process information literally and they begin to only see numbers, ordinarily the first 2 digits perhaps, therefore information must be presented in a particular, neat manner which aims to grasp key details and patterns and not bombard the user with complexity. By combining interactive elements with visually compelling representations of migration and unemployment data, Sheet Two aims to assist users to make informed decisions and contribute meaningfully to discussions surrounding labour force participation and relevant policies.

### **2.3 Sheet Three**

In sheet three, it is considered if the focus of the project should be towards employment types between genders and the age groups over time. This would display a distinction between the part-time and full-time lines of work between the different genders and across the various age groups. The intersection of all of these variables would be studied through heatmaps whilst everything is selectable by state, granting the user customizability. This section aims to provide an understanding of the differences in earning potentials and other demographics across the different states.

Here, the utilization of bar graphs layered over line graphs offer a variety of information through a single visualization. The relationship between genders, age groups, and employment types, along with their rates can be plotted, this will allow for easily spotting patterns and disparities within the data. The implementation of color differentiation between the genders and different employment categories is an excellent way of implementing the concepts that have been taught within this semester as it ensures clarity whilst facilitating comparisons for the different demographic segments.

Furthermore, using heatmaps adds a layer of complexity whilst also illustrating the complex relationship between the different age groups and employment types by gender. Here however, controlling the color spread is more difficult to achieve whilst also portraying useful information, but it is crucial for users to be able to optimize readability and ensure that they are extracting the utmost information rather than facing visual complexity.

## **2.4 Sheet Four**

In Sheet Four, the focus shifts towards examining the impact of different education levels on employment types (part-time and full-time) and their effects on income. This section aims to dive deeper into the relationship between education, employment, and income, providing users with valuable insights into the factors influencing earning potentials within the labour force.

The implementation of a linear regression model table allows for a more in-depth analysis of the relationship between education levels, employment types, and income. By providing information on coefficients and p-values for different education levels and employment types across various states, users can gain a deeper understanding of the statistical significance of these factors in determining income levels. This adds a layer of necessary depth to the analysis, allowing users to make informed decisions of the data and identify trends.

Furthermore, an area chart depicting the spread of income complements the regression analysis by offering a visual representation of income distributions across different education levels and employment types. This allows users to observe the overall trends and variations in income levels, providing context for the statistical findings presented in the regression model table. An animated boxplot is also implemented which shows the income spread of a specific state over the years and adds another dimension to the visualization by highlighting trends and fluctuations in income distribution. This visualization allows users to view changes in income disparities over time, providing insights into the evolving nature of the labour market within their selected state.

The choice of these visualizations is crucial in following through on the narrative established in sheet one. By focusing on the relationship between education, employment, and income, sheet four aims to build upon the underlying details that drive disparities in earning potentials within the labour force. Through a combination of analysis and highly informative visualizations, sheet four aims to maximize user engagement, and allow the users to derive insights regarding workforce participation and economic opportunities.

## **2.5 Sheet Five**

In Sheet Five, a dashboard-style framework is utilized which provides an opportunity to combine the key findings and visualizations from the previous sheets into a user-friendly interface which aims to provide a general understanding of the purpose of this project and narrative.

Each tab focuses on a specific aspect of the labour market and how the labour force is affected in relation to it. The tabs would go into details for these major topics: migration statistics, employment statistics, and education impacts, this allows users to explore the data according to their interest and efficiently draw conclusions.

#### Tab 1: Migration Statistics

- 1) Interactive Map: the map of Australia serves as a central visualization, providing users with a geographical perspective of migration patterns. Tooltips offer additional information about each state, while the radius of circles indicates the net change in migration, allowing for easy comparison between states.
- 2) Sankey Diagrams: these diagrams visually represent different migration types and whether they witness a net increase or decrease, offering insights into the flow of migration across states. These diagrams provide a clear depiction of migration trends, complementing the map visualization. The diagrams can be individually filtered for migration types or the different states

#### Tab 2: Employment Statistics

1. Bar Graph with Line Graph: this combination allows users to compare employment trends between genders (male and female) and across employment types (part-time and full-time) for the different states. It provides a general overview of workforce participation and trends over time.
2. Animated Boxplot: the boxplot illustrates the spread in participation rates over the years for all of Australia, offering a unique perspective of the temporal variations in workforce levels. This visualization adds depth to the understanding of employment trends and fluctuations over the years.

#### Tab 3: Education Statistics

1. Regression Table: the table provides an in-depth analysis of the relationship between education levels, employment types, and income, allowing users to assess the statistical significance of these factors. It offers a thorough assessment of the data and aids user understanding.

2. Violin Plot: these plots visually illustrate the spread and frequencies of different income levels across education types, offering insights into income disparities within the labour force. This visualization complements the regression analysis by providing a graphical representation of income distributions.

These visualizations were selected for the final project as they offer a balanced and comprehensive representation of the key topics and insights derived from the data. They provide users with multiple perspectives to explore and analyze the complex dynamics of the labour force in Australia, allowing for informed decision-making and policy making.

### **Munzner's What-Why-How Framework**

This framework guides the process from initial brainstorming to final execution, offering an organized method for creating narrative visualizations. By adhering to this framework, the project guarantees quality, efficacy, and clarity while disseminating the main findings and messages to the intended audience, which consists of legislators and the general public. The project defines the primary subjects of interest—statistics on migration, employment, and educational attainment, as well as the major conclusions and insights that will be illustrated through the visuals in the "what" section. This guarantees that the main story and goals are understood clearly. The project highlights the relevance and significance of the selected themes in the "why" section, emphasizing their role in influencing public opinion and policy choices. The project attempts to successfully engage and inform the audience by providing an explanation of the consequences of the data through visuals.

The project uses a range of visualization strategies and instruments in the "how" section to present the data in an understandable, interesting, and engaging way. The project guarantees that users may effortlessly examine and comprehend the data in any way they choose by carefully choosing visualizations that correspond with the goals of every tab. All things considered, Munzner's framework offers an organized and logical method for creating narrative visualizations, directing the work from inception to completion and guaranteeing that the finished result successfully conveys the desired message to the intended audience.



### **3.0 Implementation**

#### **3.1 Technical Implementation**

In order to produce this project, only Rstudio was used, in which a Shiny app was developed using datasets that had been pre-wrangled. By performing much of the data preprocessing externally, the app had been optimized for smoother processing during runtime, this was to ensure the speed and efficiency of the app and a more comfortable experience for the user.

The challenges faced during implementation include extensive data wrangling and formatting of coordinates and date variables, are common in data-driven applications but were difficult and time consuming to overcome. Especially when dealing with diverse datasets of various data types such as spatial, categorical, tabular, numerical, and ordinal require complex implementation for visualizations. Several libraries were utilized for the purpose of producing this Shiny app:

1. shiny: it was used for creating the interactive web application.
2. shinydashboard: it facilitated the creation of a dashboard-style layout for organizing and presenting the visualizations.
3. dashboardthemes: allowed for the customization of the dashboard's appearance with custom themes.
4. tidyverse: includes several packages such as dplyr and ggplot2 which are used for data filtering and visualization purposes.
5. readxl: allowed for the reading of .xlsx files, which contained the wrangled data used in the app.
6. htmlwidgets: was necessary for rendering interactive HTML widgets, this was used for the Sankey diagrams.
7. highcharter: was used for creating interactive and visually appealing Sankey diagrams.
8. lmtest: this package was used to run the regression model for the implementation of linear regression tests, when comparing education levels and the impact on income.
9. leaflet: this package was used for creating interactive and customizable maps within the Shiny app.
10. plotly: Enables the creation of interactive plots with support for various chart types and customization options.
11. kableExtra: Enhances the presentation of tabular data with additional formatting options.

For the implementation of all the codes, putting it all together required a lot of trial and error and researching from various locations on the internet. Several YouTube videos put together is what helped greatly in getting a dashboard styled layout going in the first place. No such frameworks or source codes were adopted from anyone or anywhere, everything was performed and tested locally and all inspiration was gained from the idea of a dashboard styled project.

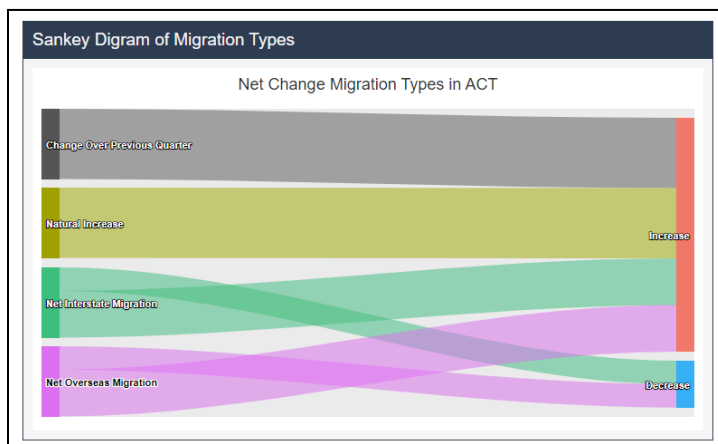
### **3.2 Interactive Narrative Visualization Implementation**

#### **Page 1 - Introduction Page**

Upon entering into the application, the user is greeted with an introduction to the topic and what they should expect of this app, furthermore, there is also a basic how-to-use instruction on the front page. This page is one of four pages on this dashboard and acts as a nice welcome rather than simply jumping into the app without any explanation and leaving the user in the deep-end of the swimming pool and expecting them to figure things out on their own.

#### **Page 2 - Overall**

Page two of four, the first page of the dashboard called "Overall," provides a comprehensive overview of migration statistics using Sankey diagrams and an interactive map. These visualizations offer users a clear understanding of migration statistics and their effects on the labour force across different states. The Sankey diagram is ideal for illustrating migration flows, allowing users to filter data by migration type or a specific state. This interactivity helps users focus on relevant data and get an idea of implications caused by them.

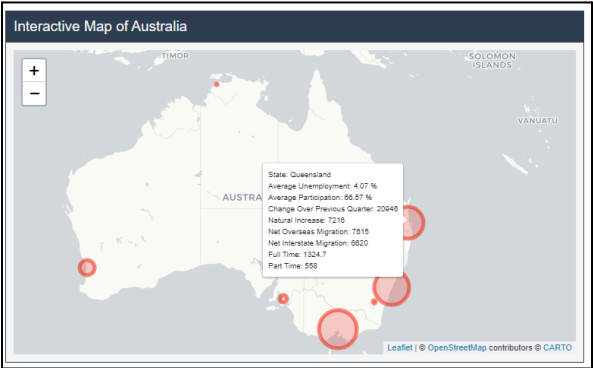


The "Select State or Migration Type" dropdown menu displays a list of options. The "Select Plot:" label is followed by a dropdown menu. The visible options are: "Australian Capital Territory", "Queensland", "South Australia", "Tasmania", "Victoria", "Western Australia", "Change Over Previous Quarter", "Natural Increase", and "Net Overseas Migration".

The interactive map better the user experience by providing detailed state-specific information through tooltips. These tooltips display average migration numbers across the different types, participation and unemployment rates, and average income for part-time and full-time

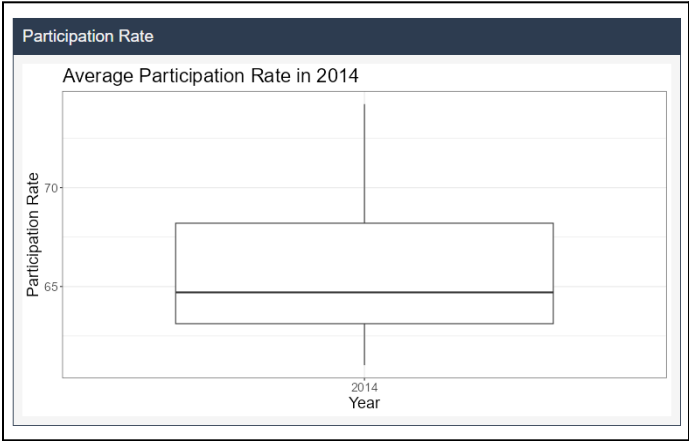
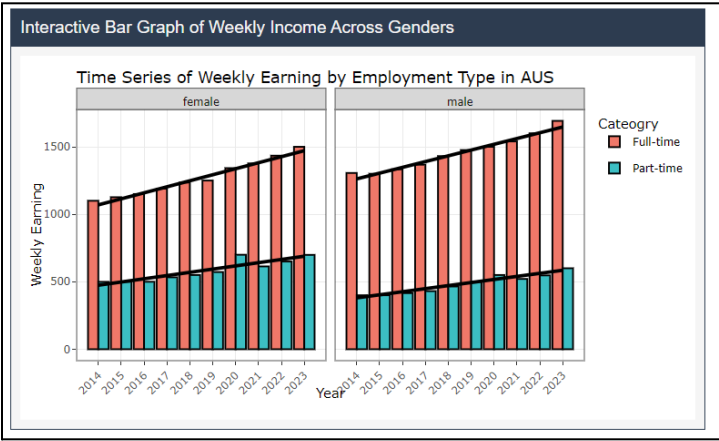
employment. This allows users to quickly compare different regions and identify states that interest them.

Together, these visualizations create a cohesive and interactive experience, enabling users to explore migration and employment data in depth. By starting with a broad overview and offering detailed insights, this page ensures that users can effectively navigate and understand the various statistics displayed through a text paragraph at the bottom of the page which aims to provide a narrative and general idea as to what is being observed through this data.



Page 3 - Employment

Moving onto page three of four, which is the second page in this narrative, it is called “Employment” and it focuses primarily on the weekly earning capabilities of men and women in part-time and full-time employment categories. This is assisted by state-selection buttons which control the concoction of a bar graph and a line graph. The image below is what the outcome of this plot looks like to the user. Furthermore, this page is supported by an animation of a boxplot which shows the participation rates over time in Australia.



The bar graph presents a clear comparison between the weekly earnings of the different genders, differentiated by color, with bars representing the earnings for part-time and full-time employment categories. The line graph layered on top displays the trends over time, adding depth to the comparison by showing how earnings change.

The state-selection buttons allow users to filter the data specific to each state, providing a general view of employment trends and making it easier to identify regional disparities or similarities in earning capabilities.

The animated boxplot has a play/pause button which visualizes participation rates across the years, offering an overall view of how these rates have changed over time. This feature allows users to observe trends and fluctuations in workforce participation, providing insights into broader patterns and labour force dynamics. By animating the boxplot, users can better understand the temporal aspects of employment statistics, such as long-term trends.

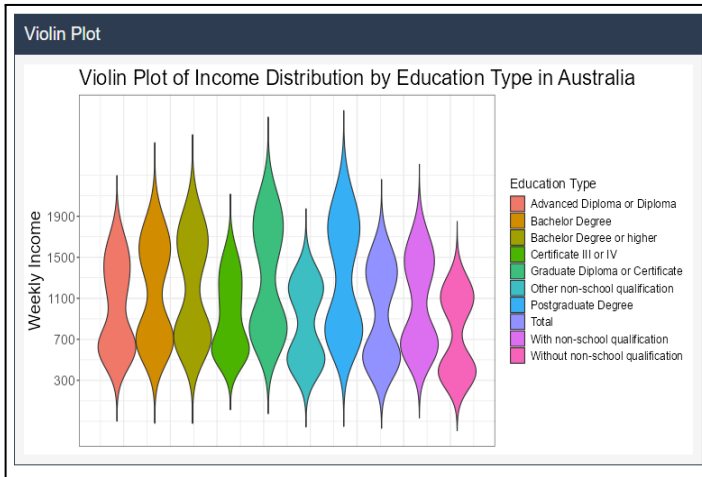
Additionally, the page includes a text based explanation, giving users some analysis for the visualizations and summarizing key findings. This explanation helps users interpret the data more effectively, ensuring that the visualizations are not only informative but also accessible. By combining interactive elements with dynamic visualizations, this page ensures that users can easily navigate and understand complex employment data. The ability to filter by state and animate changes over time enhances user engagement and provides a better understanding of the employment statistics in Australia.

Page 4 - Education

And lastly, page four of four, which is the third and final page of this dashboard, and is called “Education”. It provides an in-depth analysis of the relationship between education levels, employment types, and income. This page features a customizable regression table and a violin plot, both of which offer significant interactivity to enhance user engagement and understanding.

The regression table allows users to explore the relationship between education levels and income, with options to filter by state and employment type (full-time or part-time). Users can select their state of interest using an input selection, and select between full-time and part-time employment using buttons. This customization ensures that users can view relevant data specific to their region and employment category, making the insights more personalized.

Linear Regression Model			
x			
Full Time Linear Regression Model for Australia			
education_type	coefficient	pvalue	state
(Intercept)	1400.860	0.0000000	Australia
Bachelor Degree	196.815	0.0008906	Australia
Bachelor Degree or higher	267.210	0.0000083	Australia
Certificate III or IV	-154.680	0.0086392	Australia
Graduate Diploma or Certificate	401.455	0.0000000	Australia
Other non-school qualification	-199.495	0.0007602	Australia
Postgraduate Degree	410.870	0.0000000	Australia
Total	-38.355	0.5113435	Australia
With non-school qualification	69.425	0.2351406	Australia
Without non-school qualification	-283.005	0.0000025	Australia



The violin plot represents the distribution of average income across different education levels, with filtering options by state. Users can select a state to see how income varies by education level within that specific region. The violin plot's shape and spread provide a clear visualization of income disparities, highlighting differences and trends that may be less apparent in other types of charts. Together, these interactive elements allow users to dive

into the underlying trends of how education impacts income across various states and employment types. By offering interactivity options, the page ensures that users can adjust the data to their specific interests and needs, leading to more meaningful and relevant insights.

At the bottom of the page, a text section provides a general overview of the observed data, helping users interpret the findings from the regression table and violin plot. This contextual information supports users in understanding the broader implications of the data, guiding them through the analysis and enhancing their understanding of the educational impact on income.

### **3.3 Using The Implementation**

To effectively run, view, and use the Data Visualization Project, the following set of instructions must be followed. First, ensure that the user is using Google Chrome to run the app. This is crucial for optimal performance and compatibility. Additionally, set your browser's zoom level to 80%. Due to a software limitation, other zoom levels may cause text overflow and misalignment issues, disrupting the layout and readability.

Before launching the app, make sure all required libraries are installed. And, ensure to be patient with the plots as they may initially appear blank. This is due to the data loading in the background. Depending on your system's performance and the size of the data, this may take a few moments, otherwise the datasets are not big at all and this process should be fairly quick. If a plot appears distorted or does not fit within its designated area, simply re-run the app to rectify the issue. The dashboard is divided into four pages. The first page, "Overall," shows migration statistics using Sankey diagrams and an interactive map. Use state-selection buttons and migration type filters to explore specific data. The second page, "Employment," focuses on

weekly earning capabilities with a bar graph and line graph. Filter by state using buttons and view the boxplot which is animated and shows participation rates over time. The third page, "Education," presents a regression table and a violin plot showing the impact of education levels on income. Adjust the filters of these visualizations by state and employment type as per the user's requirement.

The user may utilize the interactive elements within each page to maximize their experience. In the Sankey diagrams the user may filter by migration type or state, and hover over the links to see detailed data for that specific node. On the interactive map, hover over states to view tooltips with detailed statistics, including migration numbers, participation rates, unemployment rates, and average incomes. For the bar and line graphs, use state-selection buttons to filter the data and observe trends for specific regions, and hover over the plots to see detailed information about the income distribution. The animated boxplot allows you to view changes in participation rates over time with the play/pause button. Lastly, use the input selection and buttons in the regression table and violin plot to customize the data by state and employment type. By following these instructions, you can ensure a smooth and engaging experience with your interactive narrative visualization, allowing you to explore and analyze the data effectively.

#### **4.0 Conclusion**

This project has successfully achieved the goal of providing an interactive and informative visualization of migration statistics, employment data, and education levels across Australia. By using a combination of visualizations that communicate the complex interplay between these factors and their impact on the labour force effectively. The data wrangling required to produce these visualizations has been not only difficult, but time consuming. But, due to the importance of clean and well-organized data in creating meaningful and user-friendly visual representations, this step must be carried out with high care.

One of the most important takeaways learned from this Data Visualization Project is that simplicity does overshadow complexity in data visualization. While it can be tempting to create complex visualizations, the most effective ones are often those that present the most information in the least complicated way and in a straightforward manner that is easy for users to understand. This is one of the aspects of this semester that was valuable in creating visualizations that are more suitable for a general audience, ensuring that the data is accessible and engaging without confusing the user.

Looking back, a more micro-based approach on a specific state, such as Victoria could have enhanced the depth and detail of the analysis, rather than producing visualizations of all of Australia. The very wide scope of this project made it difficult to gather data from various sources, which is why a more focused study might have allowed for a more thorough and in-depth exploration of the issues at hand. Additionally, focusing on a single state could have allowed for the inclusion of more data types, providing even greater insights and more unique visualizations.

Looking forward, future work could benefit from this more focused approach by concentrating on one state, it would be possible to dive deeper into the specific factors influencing migration, employment, and education within that region. Furthermore, incorporating more theoretical frameworks and methodologies learned throughout the semester could enhance the accuracy and complexity of the analysis. This would not only improve the quality of the visualizations but also provide more insights of higher quality for policymakers and the general public.

Overall, this project has been a learning curve, demonstrating the importance of interactive visualizations in demonstrating complex data. It has illustrated the importance of a user focused design and the need for simplicity and clarity in effective data communication.

## **5.0 Bibliography**

### **5.1 References**

- Australian Associated Press (2023, December 9). Anthony Albanese announces plan to reduce immigration levels following Covid influx. *The Guardian*.  
<https://www.theguardian.com/australia-news/2023/dec/09/anthony-albanese-announces-plan-to-reduce-immigration-levels-following-covid-influx>
- Cheng, J., & Xie, Y. (2023). htmlwidgets. R package version  $\geq 1.6$ .  
<https://cran.r-project.org/web/packages/htmlwidgets/index.html>
- Chang, W., Cheng, J., Allaire, J.J., Xie, Y., & McPherson, J. (2023). shiny. R package version  $\geq 1.7$ . <https://cran.r-project.org/web/packages/shiny/index.html>
- Chang, W., Cheng, J., & Bourgon, L. (2023). shinydashboard. R package version  $\geq 0.7$ .  
<https://cran.r-project.org/web/packages/shinydashboard/index.html>
- Data Analytic. (2020, December 9). *[R Beginners guide] - Sankey diagram from your own datasets: Code Included* [Video]. YouTube.  
<https://www.youtube.com/watch?v=k-IN6HBhgq4>
- Gohel, D. (2023). flextable. R package version  $\geq 0.6$ .  
<https://cran.r-project.org/web/packages/flextable/index.html>
- Gohel, D., Aubert, N., & Teucher, A. (2023). highcharter. R package version  $\geq 0.9$ .  
<https://cran.r-project.org/web/packages/highcharter/index.html>
- Hare, J. (2023, July 12). International students and graduates in Australia hit historic high of nearly 1m. *Australian Financial Review*.  
<https://www.afr.com/policy/health-and-education/life-is-about-to-get-tougher-for-the-record-1m-foreign-students-here-20230711-p5dnd4#:~:text=A%20surge%20in%20the%20number,threaten%20the%20university%20sector%20too.>



Hothorn, T., & Zeileis, A. (2023). lmttest. R package version  $\geq 0.9$ .

<https://cran.r-project.org/web/packages/lmttest/index.html>

Pebesma, E., & Bivand, R.S. (2023). leaflet. R package version  $\geq 2.1$ .

<https://cran.r-project.org/web/packages/leaflet/index.html>

Read, M. (2023, April 28). International students drive Australia's migration rebound. *Australian Financial Review*.

<https://www.afr.com/policy/economy/migration-to-hit-a-record-400k-this-year-amid-a-surge-in-students-20230428-p5d3zw>

Sievert, C. (2023). plotly. R package version  $\geq 4.10$ .

<https://cran.r-project.org/web/packages/plotly/index.html>

Walker, K. (2023). openxlsx. R package version  $\geq 4.2$ .

<https://cran.r-project.org/web/packages/openxlsx/index.html>

Wickham, H., Bryan, J., Elberg, A., & Garrett Grolemond. (2023). dashboardthemes. R package version  $\geq 1.0$ . <https://cran.r-project.org/web/packages/dashboardthemes/index.html>

Wickham, H., & Bryan, J. (2023). readxl. R package version  $\geq 1.3$ .



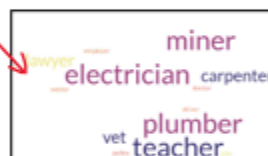

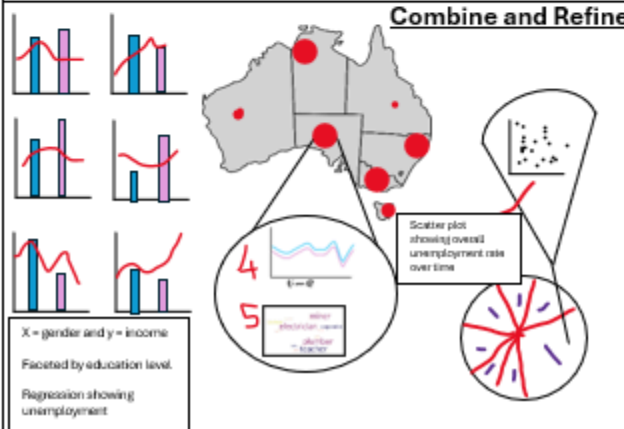
<https://cran.r-project.org/web/packages/readxl/index.html>

Zhu, H. (2023). kableExtra. R package version  $\geq 1.3$ .



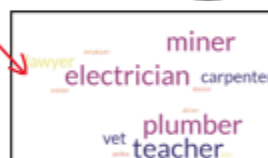

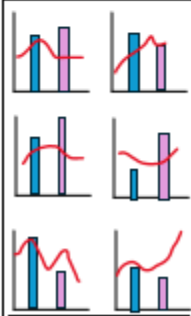
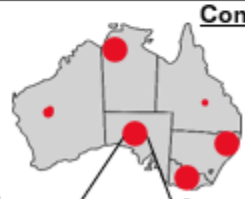

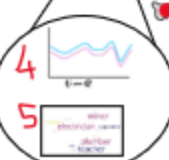
<https://cran.r-project.org/web/packages/kableExtra/index.html>

## 6.0 Appendix

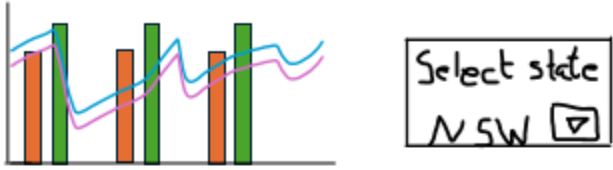
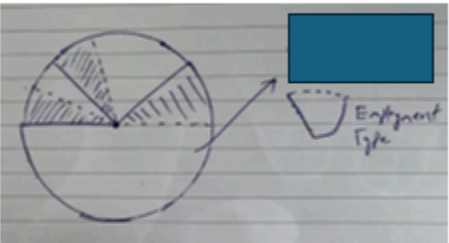

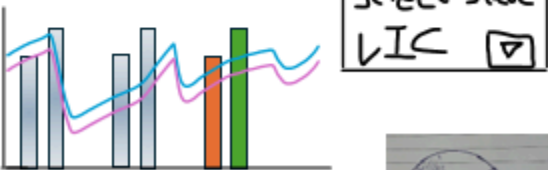

### 1.1 - Five Design Sheets: Sheet 1

<p><b>Name:</b> Data Visualization Project <b>Author:</b> Rayyan Aamir (32065647) <b>Date:</b> 10<sup>th</sup> May 2024 <b>Sheet:</b> Number 1 <b>Task:</b> Ideas for project</p>	<p><b>IDEAS</b></p> <ol style="list-style-type: none"><li>1. dashboard</li><li>2. filtering</li><li>3. categorizing</li><li>4. feature listing</li></ol>	<p><b>Filter</b></p> <p>Idea 1 can be expanded upon to implement some plots within hovering.</p> <p>Idea 2 needs to be further divided to show more information, maybe level of significance of each education level</p> <p>Pie chart potentially providing very limited information rather than useful</p> <p>Idea 4 can be expanded upon such unemployment categories for those with disabilities</p> <p>Word cloud may require bigger dataset looking into individual people's information – census data?</p> <p>Heatmap seems over complicated providing very limited information for trends</p>
<p><b>Ideas Continued:</b></p> <ol style="list-style-type: none"><li>1) Different states, tooltips, hover will show statistics, overall figures</li><li>2) Education levels and income over time</li><li>3) Pie chart for different states and migration levels, purple lines represent the different states and territories.</li><li>4) Unemployment over time, time series perhaps</li><li>5) potential word cloud to study different employment types</li><li>6) Heat map for part time and education level</li></ol>	   	<p><b>Categorize</b></p> <p>Three different categories can be seen here:</p> <ol style="list-style-type: none"><li>1. focusing on migration statistics and overall unemployment</li><li>2. focusing on gender unemployment + disability status</li><li>3. focusing on education level and income in different employment types</li></ol>
 <p>X = gender and y = income Faceted by education level Regression showing unemployment</p>	<p><b>Combine and Refine</b></p> <p>A couple of combinations can be implemented in the map for example.</p> <p>Use of faceted diagrams in R to display significance of different education levels</p> <p>Pie chart to go into greater depth for specific states</p>	<p><b>Question</b></p> <p>The ideas displayed here allow to go into various depths about many different statistics that may or may not be affecting the overall labour force of Australia.</p>

## 2.1 - Five Design Sheet: Sheet 2

<p><b>Name:</b> Data Visualization Project <b>Author:</b> Rayyan Aamir (32065647) <b>Date:</b> 10<sup>th</sup> May 2024 <b>Sheet:</b> Number 1 <b>Task:</b> Ideas for project</p>	<p><b>IDEAS</b></p> <ol style="list-style-type: none"><li>1. dashboard</li><li>2. filtering</li><li>3. categorizing</li><li>4. feature listing</li></ol>	<p><b>Filter</b></p> <p>Idea 1 can be expanded upon to implement some plots within hovering.</p> <p>Idea 2 needs to be further divided to show more information, maybe level of significance of each education level</p> <p>Pie chart potentially providing very limited information rather than useful</p> <p>Idea 4 can be expanded upon such unemployment categories for those with disabilities</p> <p>Word cloud may require bigger dataset looking into individual people's information – census data?</p> <p>Heatmap seems over complicated providing very limited information for trends</p>
<p><b>Ideas Continued:</b></p> <p>1) Different states, tooltips, hover will show statistics, overall figures</p> <p>2) Education levels and income over time</p> <p>3) Pie chart for different states and migration levels, purple lines represent the different states and territories.</p> <p>4) Unemployment over time, time series perhaps</p> <p>5) potential word cloud to study different employment types</p> <p>6) Heat map for part time and education level</p>	   	<p><b>Categorize</b></p> <p>Three different categories can be seen here:</p> <ol style="list-style-type: none"><li>1. focusing on migration statistics and overall unemployment</li><li>2. focusing on gender unemployment + disability status</li><li>3. focusing on education level and income in different employment types</li></ol>
    <p>4</p> <p>5</p> <p>X = gender and y = income Faceted by education level. Regression showing unemployment</p>	<p><b>Combine and Refine</b></p> <p>A couple of combinations can be implemented in the map for example.</p> <p>Use of faceted diagrams in R to display significance of different education levels</p> <p>Pie chart to go into greater depth for specific states</p>	<p><b>Question</b></p> <p>The ideas displayed here allow to go into various depths about many different statistics that may or may not be affecting the overall labour force of Australia.</p>

### 3.1 - Five Design Sheet: Sheet 3

<p><b>Layout</b></p> <p>Bar chart for different employment types, regression line showing gender distribution over time</p>   <p>Heat Map for different age groups</p> 	<p><b>Title:</b> Data Visualization Project</p> <p><b>Author:</b> Rayyan Aamir (32065647)</p> <p><b>Date:</b> 10 May 2024</p> <p><b>Sheet:</b> Number 3</p> <p><b>Task:</b> Employment type by gender, age</p> <p><b>Operations:</b></p> <p>Layout is interactive using <del>Rshiny</del>, all graphs are interactive to select for specific states.</p> <p>Pie chart gets filtered according to the filter above for bar graphs</p> <p>Age group heat map can be filtered by the specific age groups to allow for a more detailed view.</p> <p>May add option to <u>full</u> hide unwanted graphs and running filters in the <u>background</u> nonetheless.</p>
<p><b>Focus:</b></p> <p>Selecting state will highlight the unique bar and make the others translucent</p>  	<p><b>Discussion:</b></p> <p><u>Positive:</u></p> <p>The implementation of a multi-filtering system and the production of duplicate graphs allows for a detailed overview of gender, age group, and employment type of that state. Allowing user for interactivity and control</p> <p><u>Negative:</u></p> <p>Repetition, often confusion of why there are repeated diagrams. User might get confused on the level of complexity in the overlapping graphs and reading the multi-legend wraps.</p>

## 4.1 - Five Design Sheet: Sheet 4

Layout			
Education	Employment Type 1	Employment Type 2	
1	Value	Value	Education - Gender -
2	Value	Value	
3	Value	Value	
4	Value	Value	
5	Value	Value	

**Focus:**

- State
- Disability Status

Education	Employment Type 1	Employment Type 2		
1	Value	Value	Education -	Gender -
2	Value	Value		
3	Value	Value		
4	Value	Value		
5	Value	Value		

**Title:** Data Visualization Project

**Author:** Rayyan Aamir (32065647)

**Date:** 10 May 2024

**Sheet:** Number 4

**Task:** Education level and income in different job types

### Operations:

Allowing for interactivity and studying a regression table for coefficients and intercepts.

Users are able to study area graphs and line graphs for changes of income over time for different genders and those with disabilities.

### Discussion:

#### Positive:

Shows some in-depth analysis of regression techniques displaying how much of a difference each education type sees in its respective employment type

Individual gender proportions in pie chart through density plot, using ~~plotly~~ package perhaps

#### Negatives:

Animation might confusion, might not display the correct information, user may find confusing

#### Solutions:

Perhaps add a filter system and enlarge the boxplots for animation to allow for easier understanding for the user.

## 5.1 - Five Design Sheet: Sheet 5

