Narrative Visualization Project of Australian population tendency

Student name: Yi Chong

Student ID: 30289327

Tutorial Number: 24

Tutor Name: Aniruddha Indurkar

Content

1. Introduction

With the global population increasing, the problem of shortage of resources is becoming more and more obvious. Hence, analyzing and research the population tendency is significant for the public. The definition of narrative visualization is that to explore data visualization by revealing stories within the data which is different from traditional storytelling (Segel & Hee,2010). Implementing an interactive data visualization is a great method to realize this principle.

This project aims by using R(shiny) to implement a data visualization project which can interact with the audiences. It focuses on the public as our intended audience. Through using this project, the user can gather a completed and comprehensive understanding of the population tendency and demographic structure in Australia and predict what is the population trends of Australia in the future.

2. Description of the design process

In this section, the visualization design process will be discussed. For this project, it bases on the methodology of five design sheets to develop the initial design. Robert, Headleand & Ritsos (2016) illustrate that a five-design sheet is a method used to draw some sketches when the user wants to brainstorm their visualization ideas. Sheet 1 is called IDEAS that is used for gathering all brainstorming ideas by discussing the general idea, then to analysis and filter all ideas and move on categorizing and combine. The sheet 2, 3 and 4 is called initial design, which will draw the layout and the focus details from the plot and list the interaction functions or operations about each layout. In the end, it also writes down the advantages and disadvantages of each layout. Lastly, the sheet 5, realization, is utilized to draw a complete layout and focus based on pervious steps.

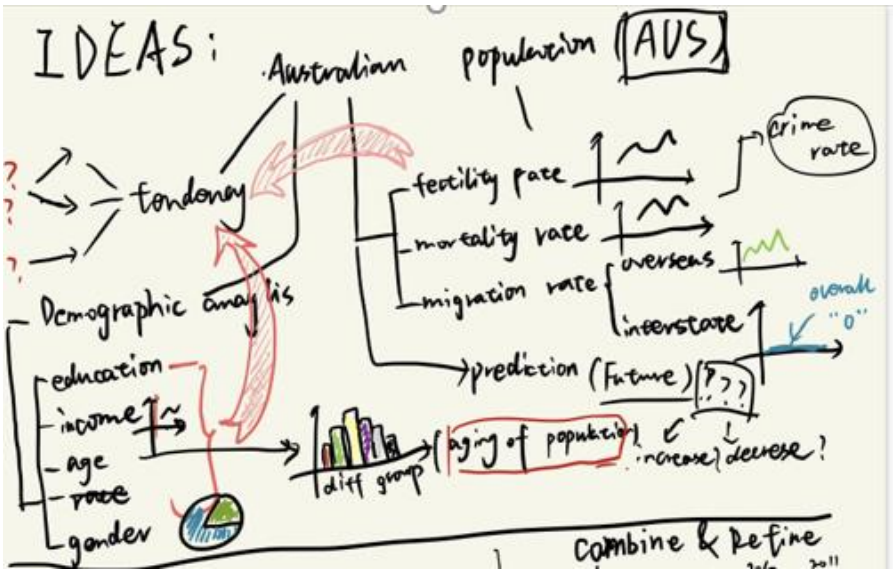


Figure 1. Ideas

In this project, it aims to explore the Australian population from 2000 to 2015. In the sheet 1, it is easily to think out some keywords that relative to the topics and the dataset. For instance, the tendency of the population, demographic analysis and what factors will have influence on the population tendency, etc. For each brainstorming idea, it proposes lots of chart to display the information. In the filter part, based on the dataset and information validity, it removes some insufficient ideas from brainstorming. Then to categorize and combine the remaining thoughts.

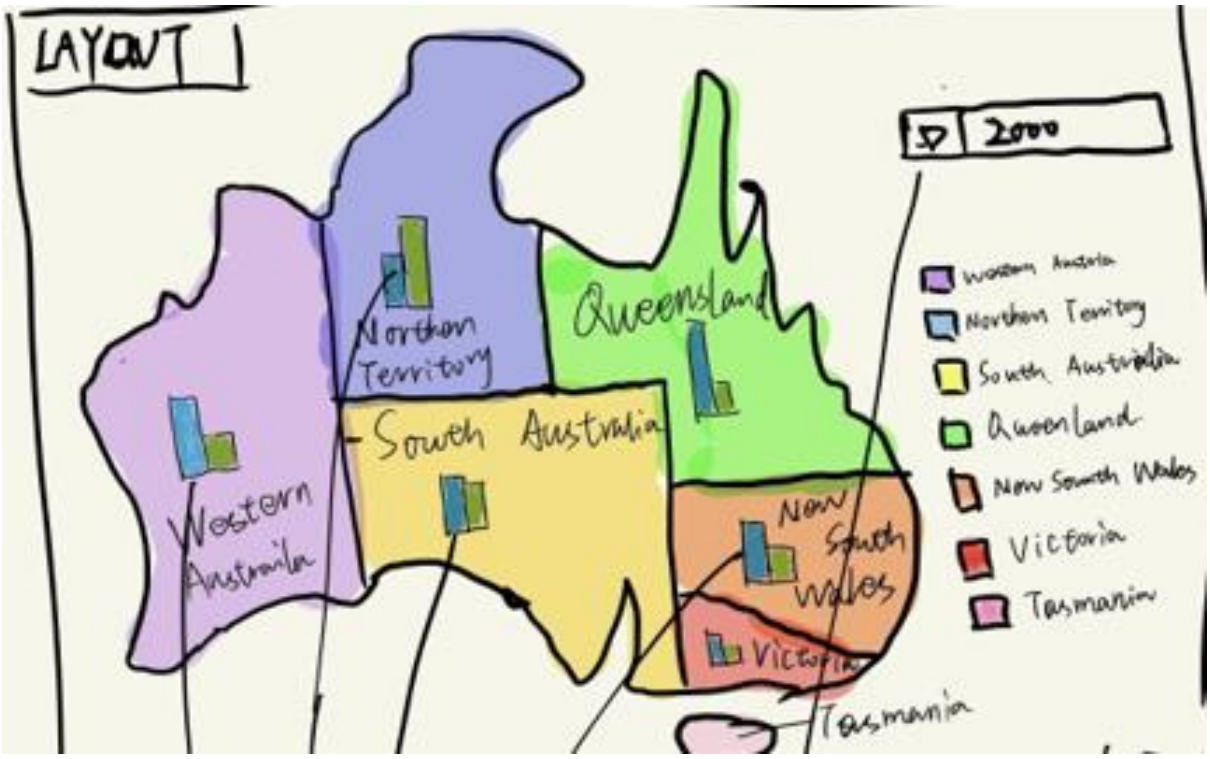


Figure 2. Initial Design 1

In sheet 2, it focuses on the population of each state from 2000 to 2015 by using the combination of choropleth map and proportional symbol to display the information. Each state will be filled in different color and the percentage of the male and female will be drawn on each state. It involves a select bar that allows the user to interact with the chart. It will be using gradual change color to display the changing of the population. For example, the population of the state, which filled in the dark blue, is more than filled by light blue. The advantage of this layout is having a clear and higher interaction. Through observing the color of each state, the changing of the population can easily be known. However, it is hard to know and distinguish the actual changing number.

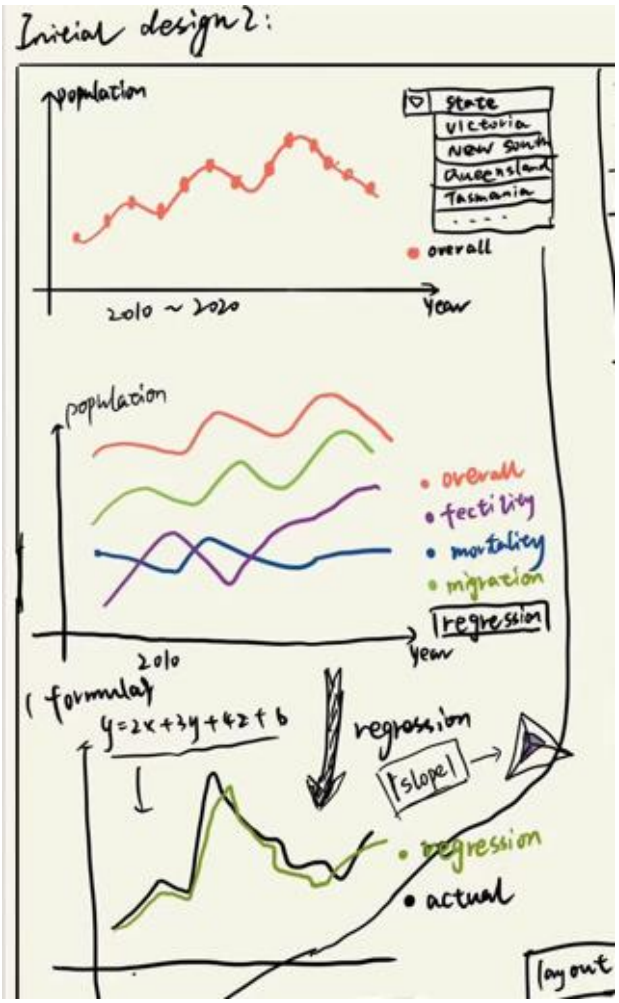


Figure 3 Initial Design 2

Sheet 3 is used to find out what factors will have an effect on the Australian population. Select using a line chart to show the relationships between growth rate, fertility rate, mortality rate and migration rate. Using a filter to select the state then the state's population tendency will be shown in the figure. Each state line will have a unique line type and color that will improve the visual effects. Allow user to select an attribute that they want to find out the relationships between this variable and growth rate. In the end, it will construct the linear relationship between the response variable (growth rate) and explanatory variable (fertility rate etc.). After the regression, the formula of their relationship can be derived out. Based on the consequence, the regression line and actual data line will be drawn. There is a variety of advantages sheet 3. Firstly, through using regression model, it provides some meaningful conclusion for users. Then, it provides extensive options to show state population details. Thirdly, it combines various data in one chart to improve exploration information from the dataset.

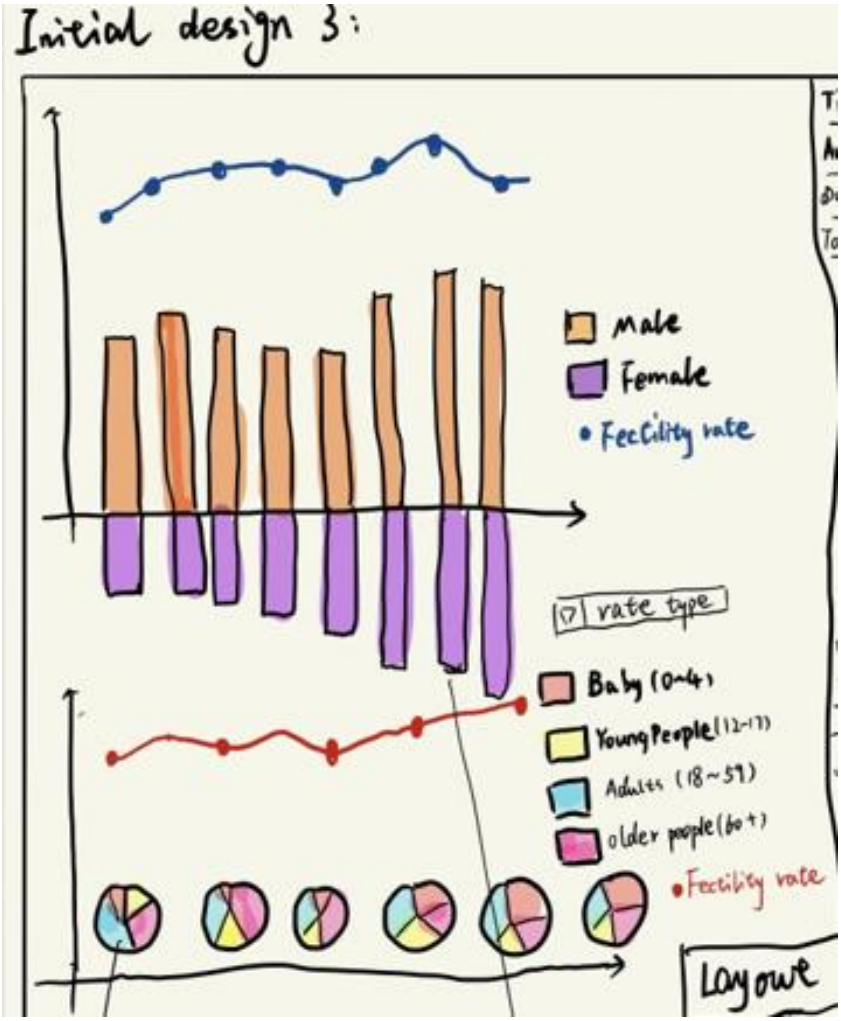


Figure 4. Initial Design 3

Demographic analysis is the analysis focus on the population structure and all the data is composed by education, income, age, race, and gender(32323232). In sheet 4, it will focus on analyzing the demographic structure. It will use two charts to illustrate the changing of the demographic structure. The first chats will connect the fertility, mortality, and growth rate with the percentage of gender. Another chart will show the relationship between the percentage of different age group and those rates. By selecting the rate type, the user can predict the relationships of those variables and also find out the changing tendency of the population structure. The advantage of sheet 4 is combing different charts in one chart that can visually highlight the differences between different sets of data. In addition, combination of population structure changing pie chart with growth, fertility, mortality rate, it can help user to gather extra information from the datasets.

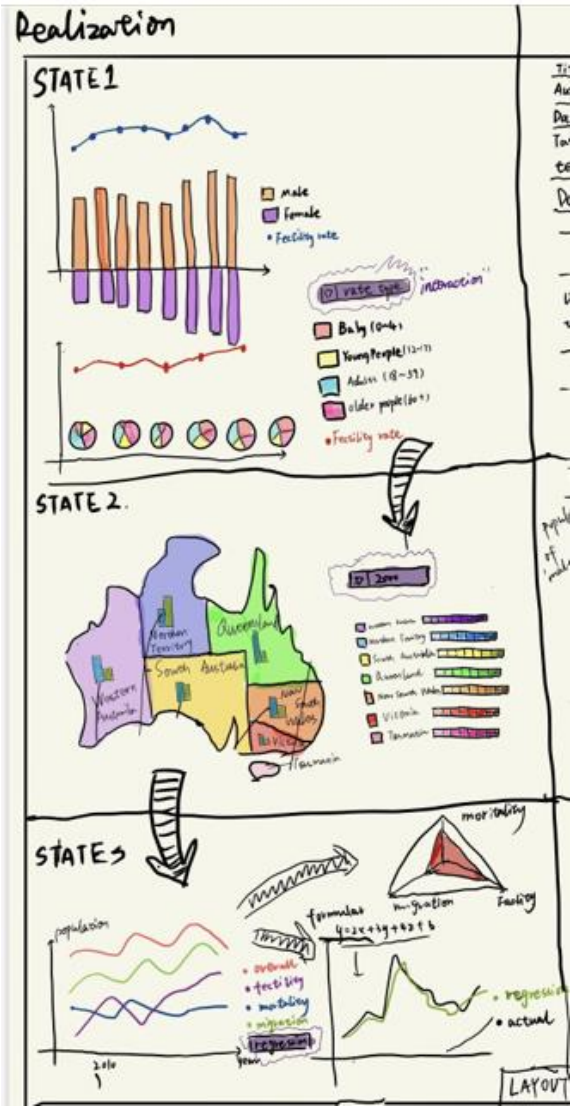


Figure 5. Realization

Finally, in the realization sheet, takes the overall layout from previous steps and remove the repeat information layout. The line chart from sheet 3 is similar to sheet 2. Considering the diversity of the visualization, the line chart about the population tendency of each state will be removed. In that case, using the choropleth map and proportional symbol to show the population of each state. Then it summaries all materials and attentions details in details part.

To reflect on the final design, it provides a variety of interactive interface for the user, which can attract user's curiosity and improve their participation. With clear signposting of messages and narrative, it will give user a good browsing impression. For each layout, the user can gather various meaningful information from that. Through a combination of the different dataset, it can demonstrate extra information for the user. All the information has been showing logically and clearly. Through utilizing different color with each variable,

it will lead to have a positive visual effect.

3. Implementation

After constructing the design process, it comes to the implementation part. In this part, it will illustrate the libraries used in this project and the reasons for the implementation decisions.

It utilizes “shiny” to implement the visualization project which is an open package from RStudio. In addition, it provides a web application framework that can be used to create an interactive visualization web. And it contains some extremely powerful functions. People can utilize it to share their analysis in an interactive manner with the community through building interactive web pages(). Since Shiny provides an automatic reactive binding between inputs and outputs, and it is built using two components, ui.R and server.R. All the UI design and interaction will be put in ui.R part which will lead to the whole project keep structured and it is easier for implementation.

“Shinydashboard” package is also utilized in this project which is usually created with Shiny (RStudio/shinydashboard). It is easy to create attractive dashboards. Through using it, we can build a sidebar to instruct navigation and show the overall content that project contain. And splitting the layout then putting in a separate page that will make the project more clearly and orderly.

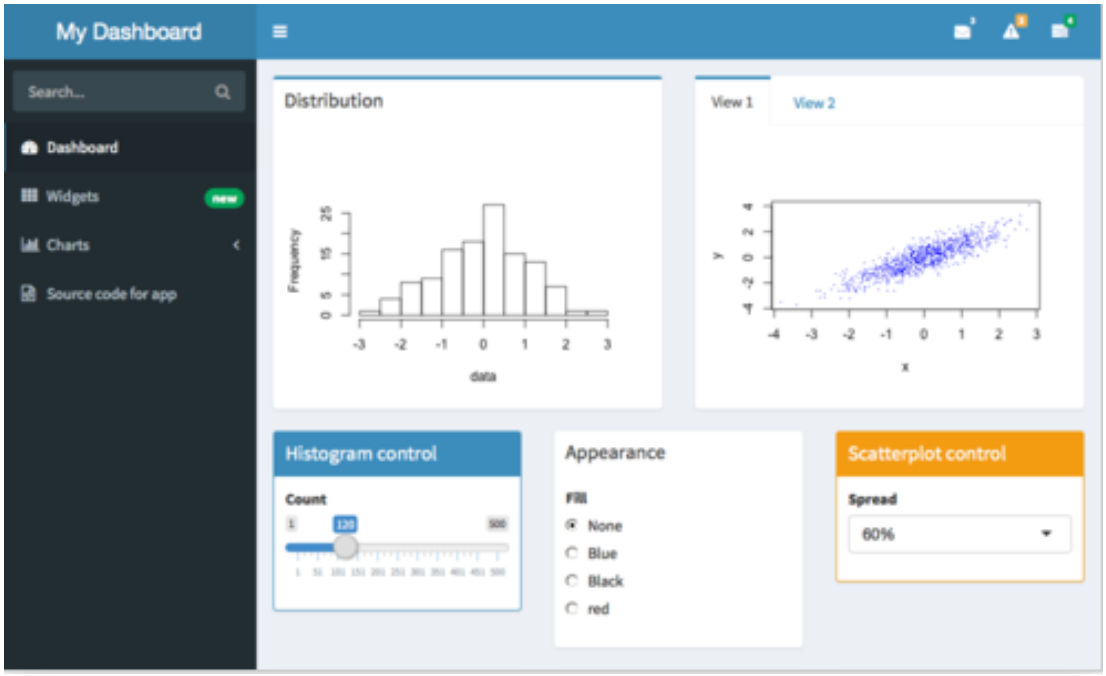


Figure 6. Dashboard

Retrieve from https://rstudio.github.io/shinydashboard/

“leaflet” package is utilized for creating and working with Shiny to post interactive maps to websites(RStudio leaflet). It has many aesthetic features and accepts many types of map objects that make it popular in R community.

By using leaflet to draw a map, it can improve project visualization diversity. In addition, it also improves the user's interaction experience.
tidyr and dplyr package are applied for data wrangling. Data Wrangling is the process of cleaning, gathering, selecting, and transforming data to the desired format which aims to show a good performance of data for further data analysis (Goldstom, 2008). By using some function from those packages, the dataset can be clearly and meaningfully without any worthless data. This step is essential for the further exploration step.

"plotly" package is an R package for creating interactive, publication-quality graphs graphic. It has various function for users to allow user to zoom in, zoom out, download as picture, etc. Comparing with directly mark value in graph and making interactive graph, the latter can help us to builder a clear-looking visualizations and provide viewers with a tool to explore the information based on their interest.

4. User guide

This part will point out the instruction of the narrative visualization project.

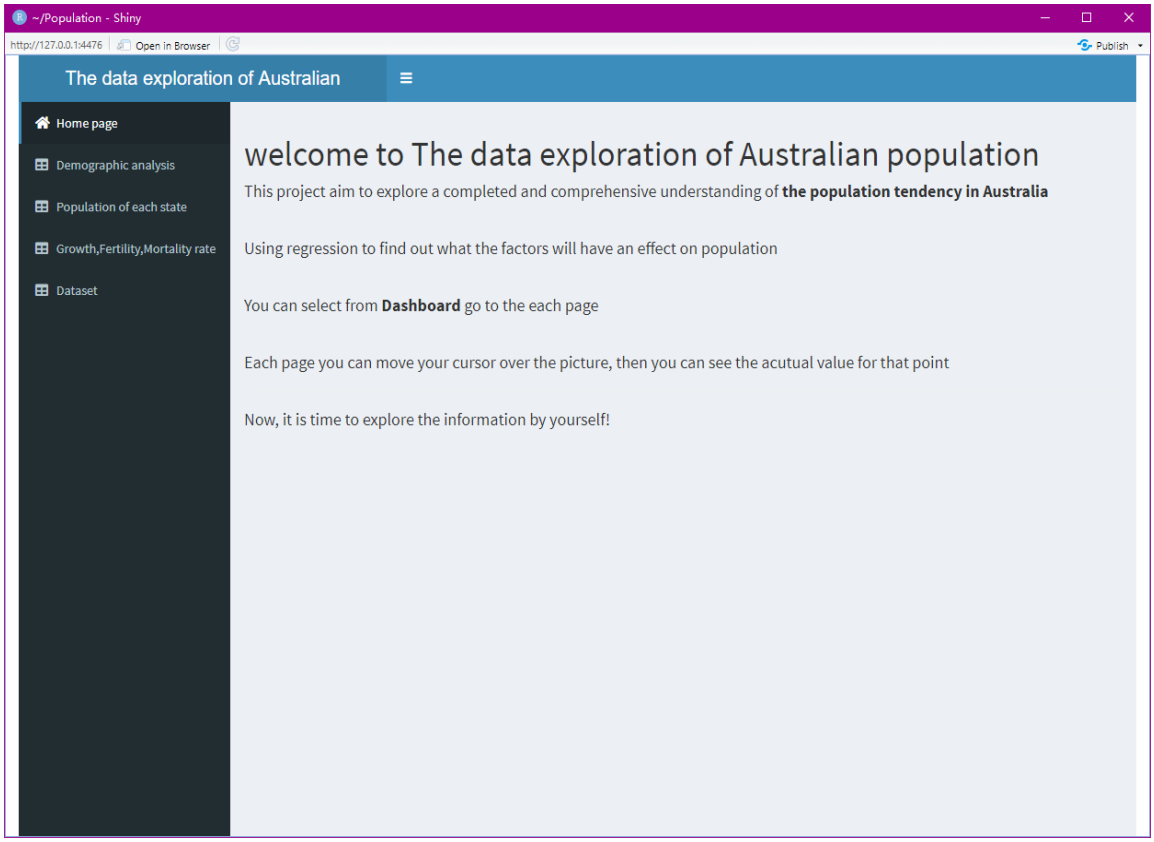


Figure 7. Home page

After the user launches the application, they will see the first page, called the home page, shown above(See Figure 7). This page is used for the introduction page, and it will briefly describe the purpose and function of this project web as well as provide an instruction guideline for users.

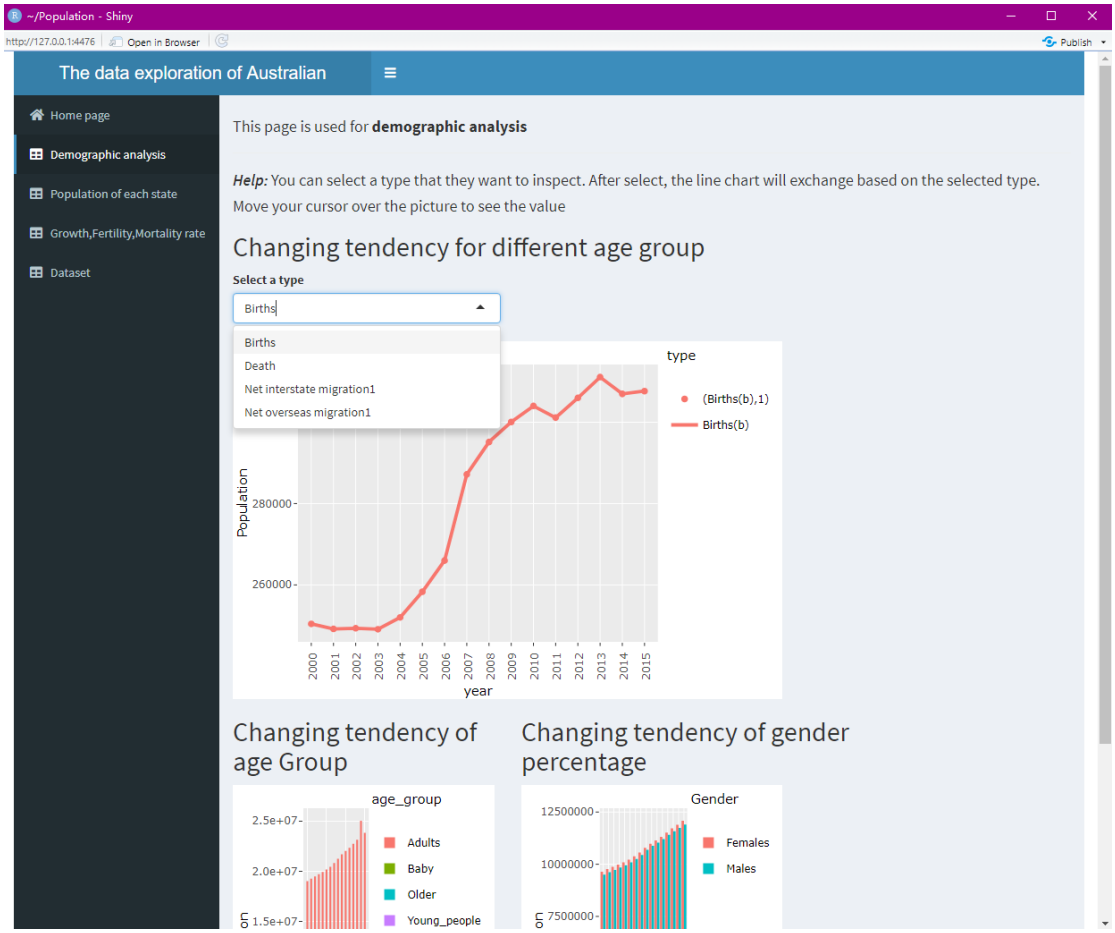


Figure 8. Demographic analysis page(1)

The second page is demographic analysis page. User can select a type that they want to inspect. After select, the line chart will exchange based on the type(See Figure 8).

Figure 9. Demographic analysis page(2)

For each chart, when the user mouse hover the chart, the value of each point will display immediately.

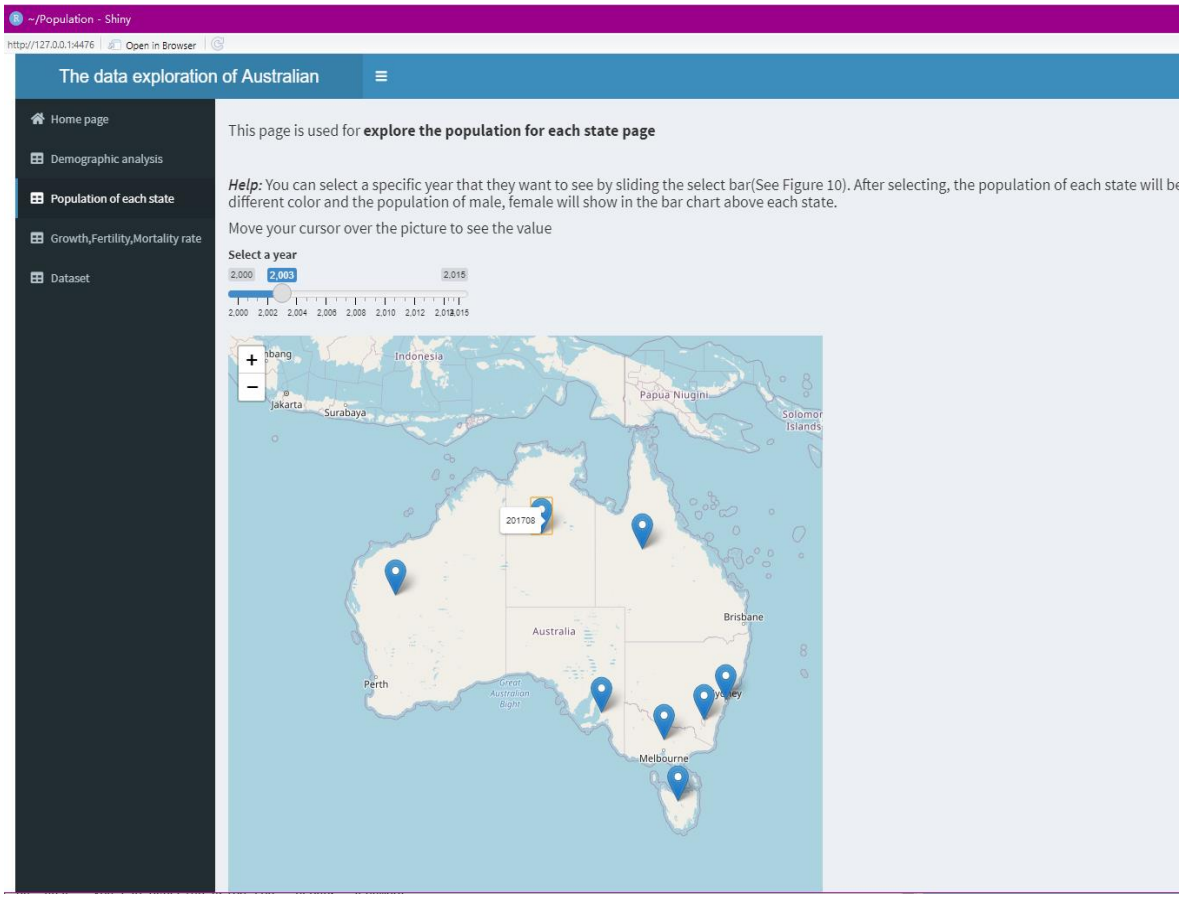


Figure10. Population of each state page

For the population of each state page, user can select a specific year that they want to see by sliding the select bar(See Figure 10). After selecting, the population of each state will be drawn by a different color and the population of male, female will show in the bar chart above each state. When the user moves their cursor over the map and the population of that state will show(See figure 11).

.

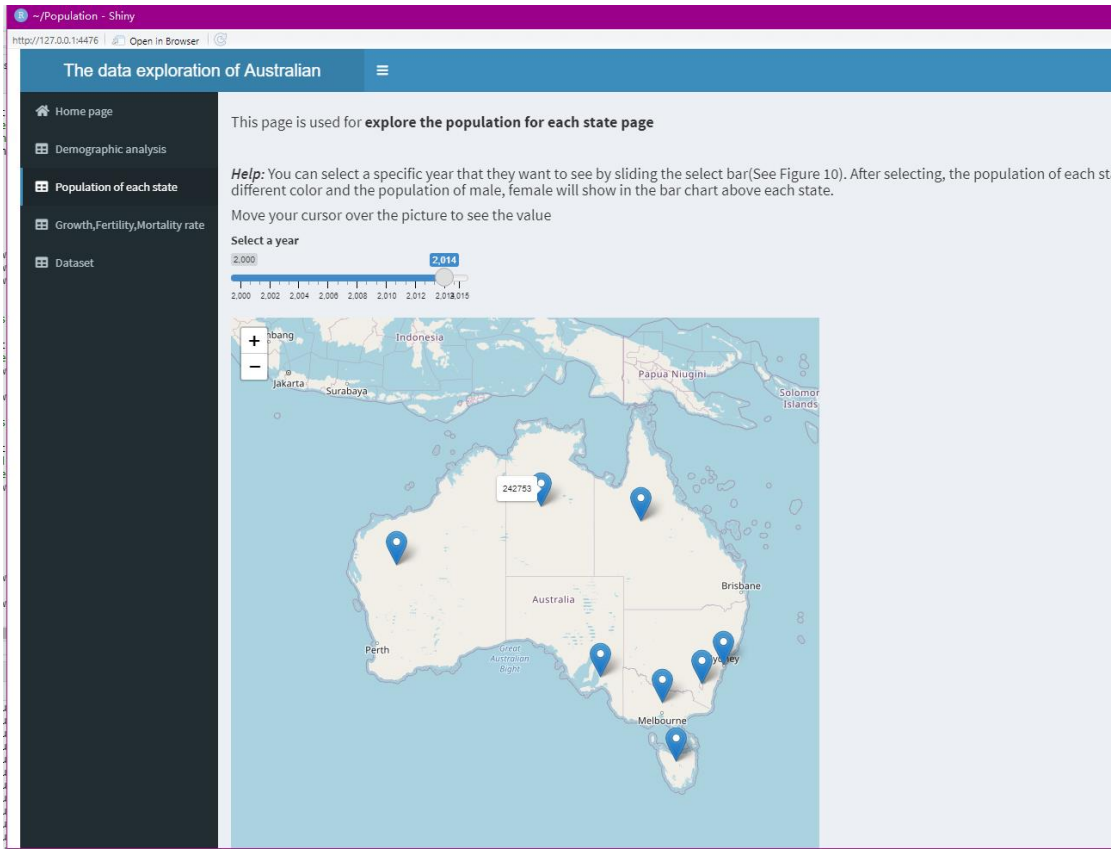


Figure11. Population of each state page

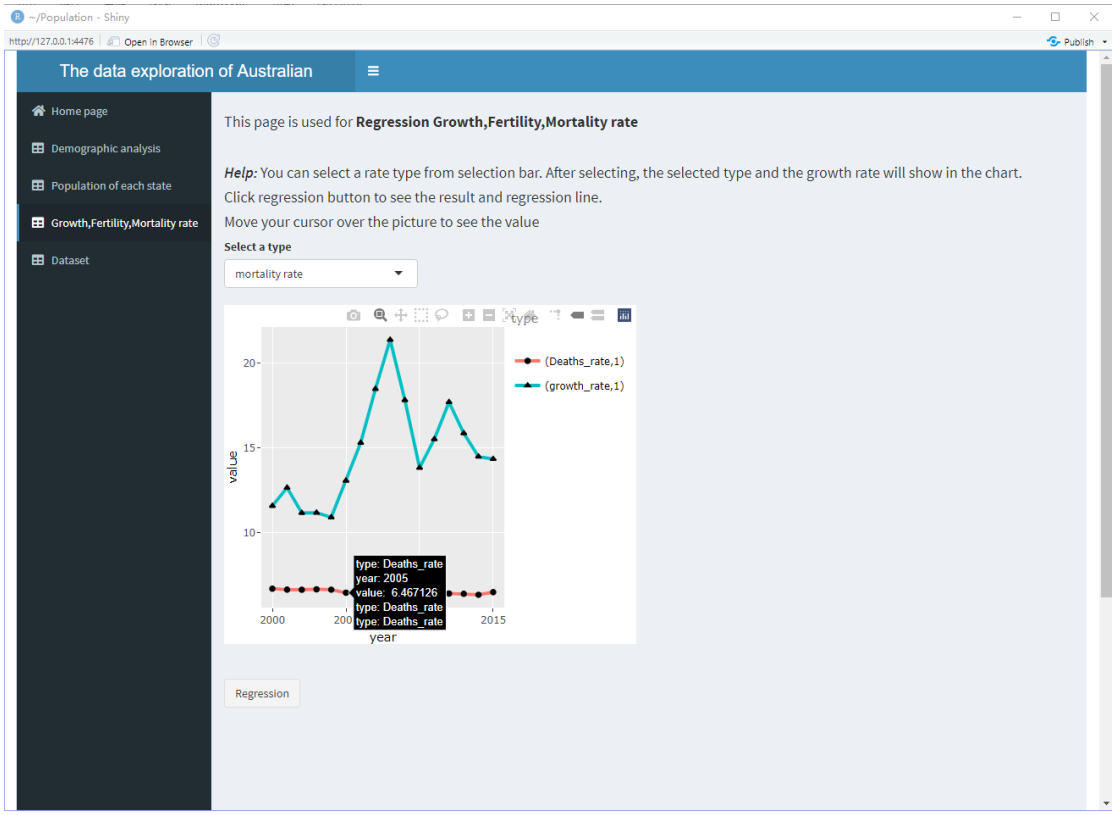


Figure 12. Growth, Fertility, Mortality rate page

In this page, user can select a rate type from selection bar. After selecting, the

selected type and the growth rate will show in the chart. The growth rate will continuously display in the chart that will allow the user can combine it with other rate type. The regression button is utilized for regression process. When the user clicks it, the results of the regression model will show immediately. Based on the value of slope, the regression line and actual data line will show below(See Figure 13).

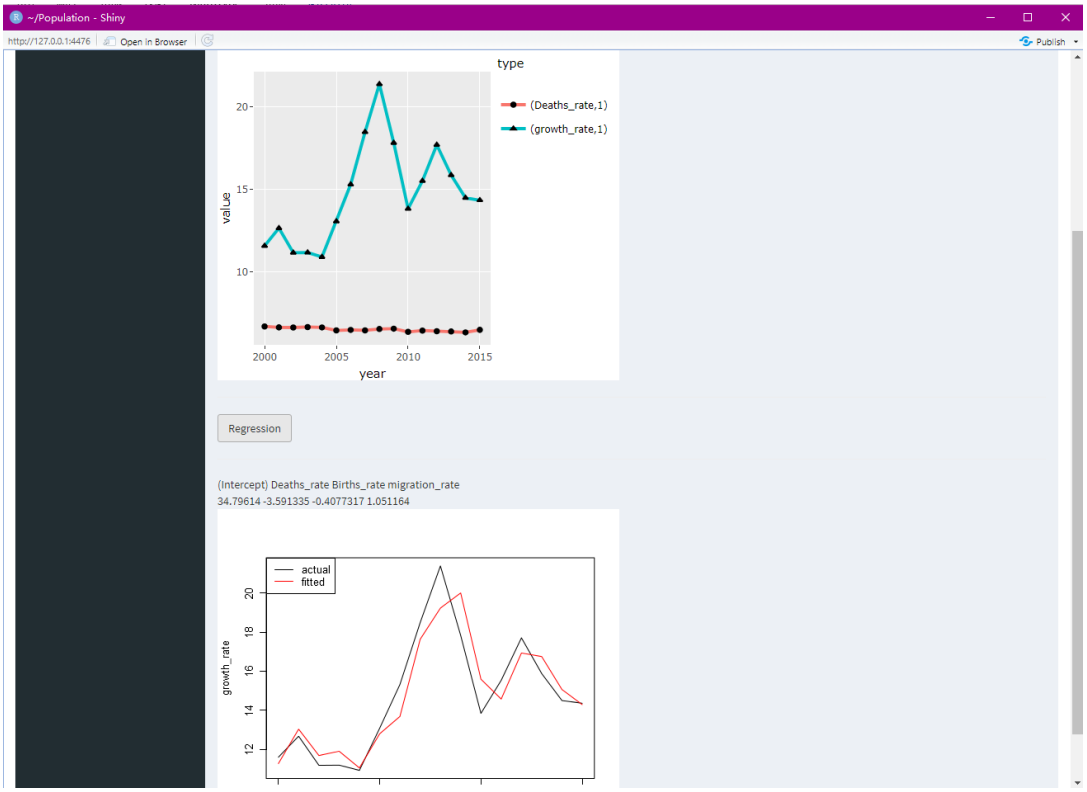


Figure 13. Dataset page

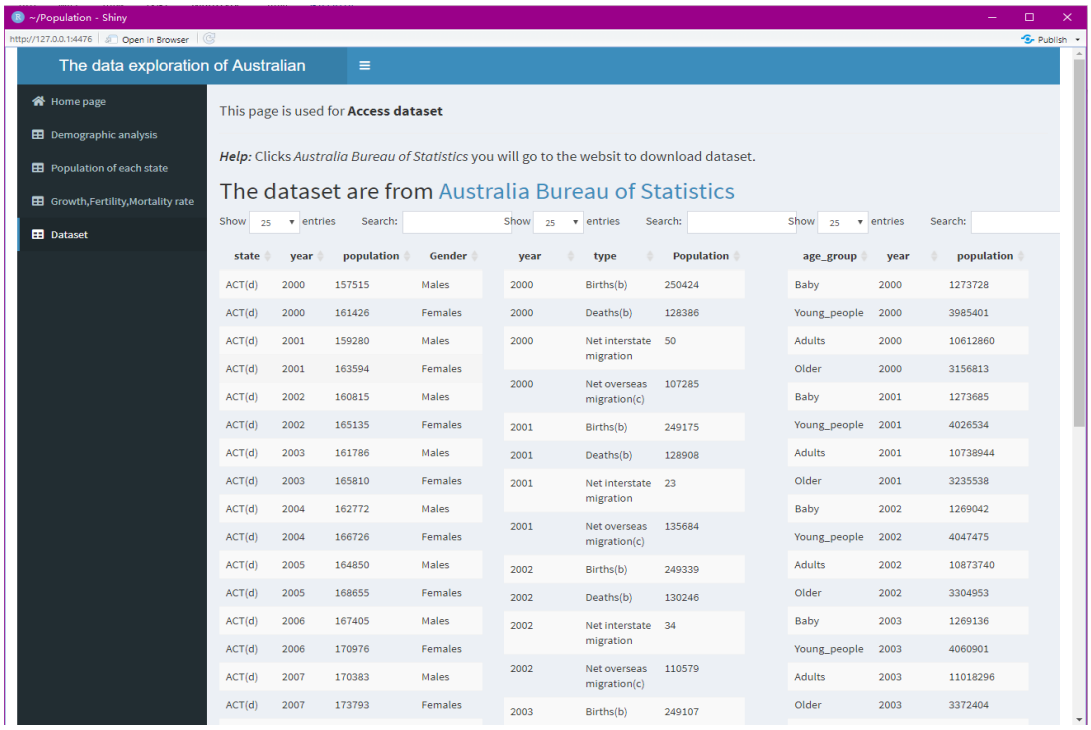


Figure 14. Dataset page

The dataset page will display all the datasets which are used for this project. The reference to data sources also has been provided. When the user clicks “Australia Bureau of Statistics”, it will automatically open the website for the user(See Figure 15). User can select the viewing order as well as search the information that they want to inspect.

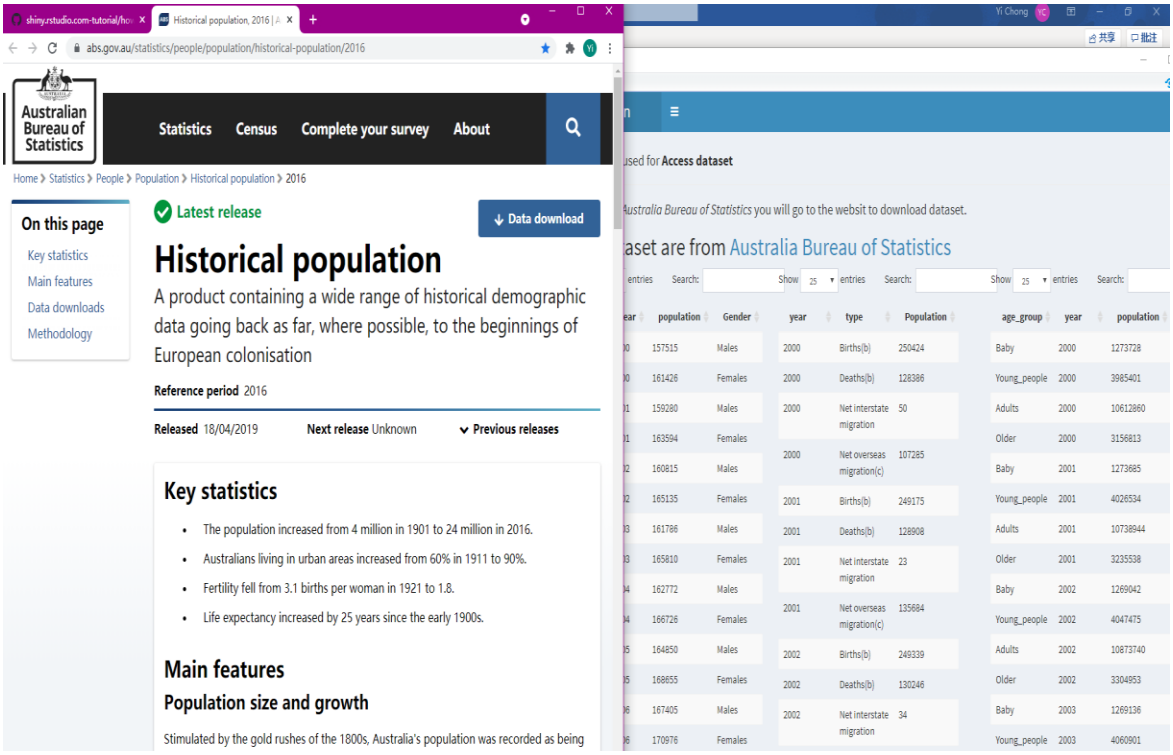


Figure 15. Australia Bureau of Statistics website
Retrieve from https://www.abs.gov.au/statistics/people/population/historical-population/2016

5. Conclusion

To sum up, the interactive data visualization of the project provides an obvious indication of the population tendency and demographic structure in Australia from 2000 to 2015. The overall population of Australia show a continuous growth trend between 2000 and 2015. The total demographic structure trends to the aging of the population which will have a result in future. The Females percentage is always higher than males. The project has been implemented by using R shiny and it achieved the function of interaction. It allows the user to navigate the project and interact with visualization data.

Reflect on this project, I learned about the population trends in Australia and the main factor which has an effect on population changing. Through Implementing an interactive data visualization, I found that the interactive data visualization enables the user to operate and explore the graphical data directly. It can make designer construct clear-looking visualizations and it provides an efficient way that helps

user efficiently and effectively gather information as easier as possible. Hence, making an interactive visualization is important for data statistics. For further exploration, although it has contained a large dataset, it should be covering another type of data except for only tabular data and need to improve that combining different dataset to draw different chart in one picture.

Bibliography

Australian Bureau of Statistics. (2016). Historical population. Retrieve from https://www.abs.gov.au/statistics/people/population/historical-population/2016

Goldston, D. (2008). Data wrangling: collecting and releasing environmental data have stirred up controversy in Washington, says David Goldston, and will continue to do so. *Nature, 455*(7209),15. Retrieve from https://link.gale.com/apps/doc/A188944501/AONE?u=monash&sid=AONE&xid=123f82fb

Narrative Visualization: Telling Stories with Data Edward Segel and Jeffrey Heer

Roberts, J. C., Headleand, C.,& Ritsos, P. D. (2016). *Sketching Designs Using the Five Design-Sheet Methodology*. Retrieve from https://ieeexplore.ieee.org/document/7192707

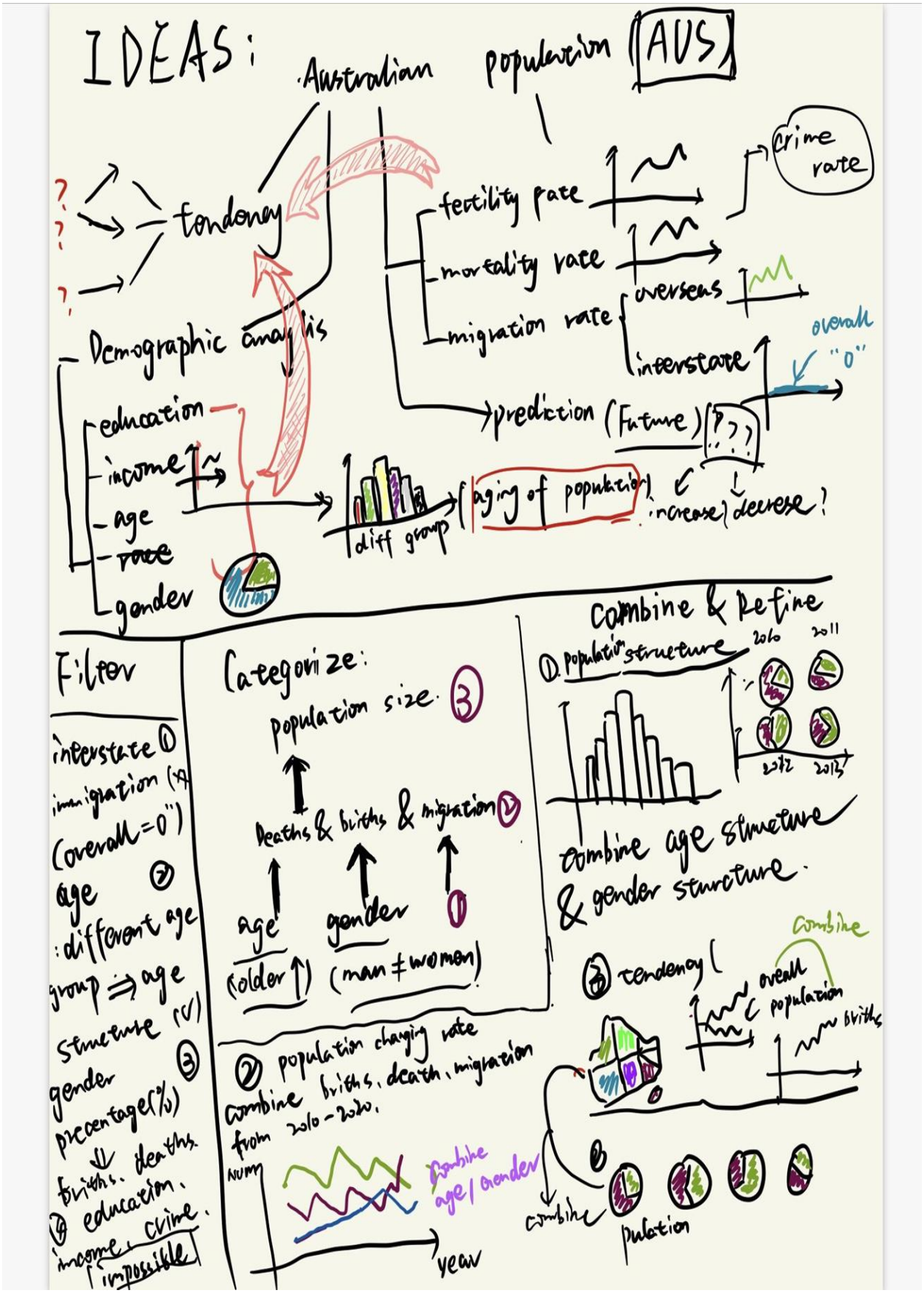
RStudio shinydashboard. (n.d). Retrieve from https://rstudio.github.io/shinydashboard/

RStudio leaflet. (n.d). Retrieve from https://rstudio.github.io/leaflet/

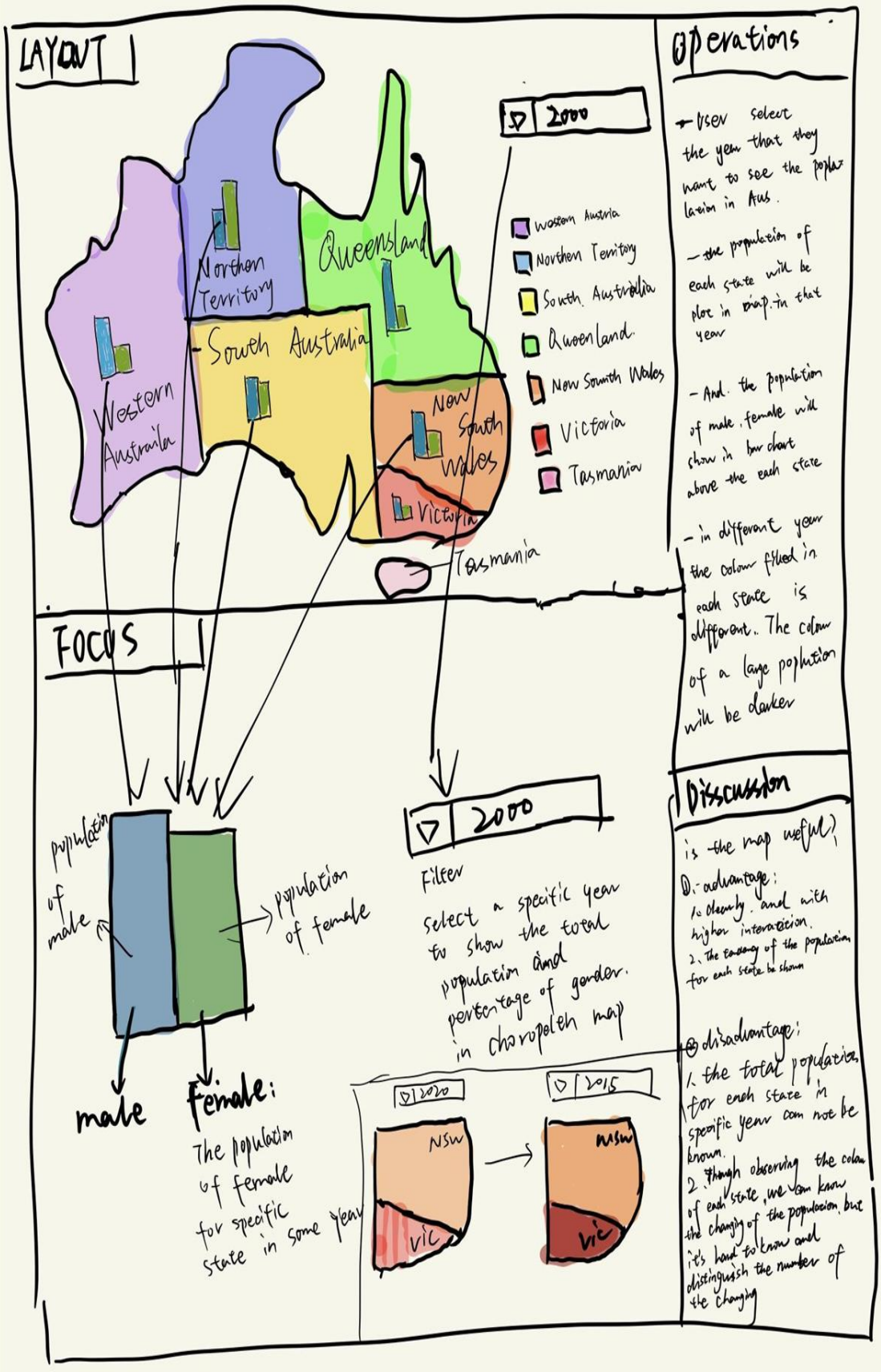
Segel ,E., & Hee,J. (2010). Narrative Visualization: Telling Stories with Data, 16(6),1139-1148. doi: 10.1109/TVCG.2010.179

The Shiny gallery .(n.d). Retrieve from https://shiny.rstudio.com/gallery/

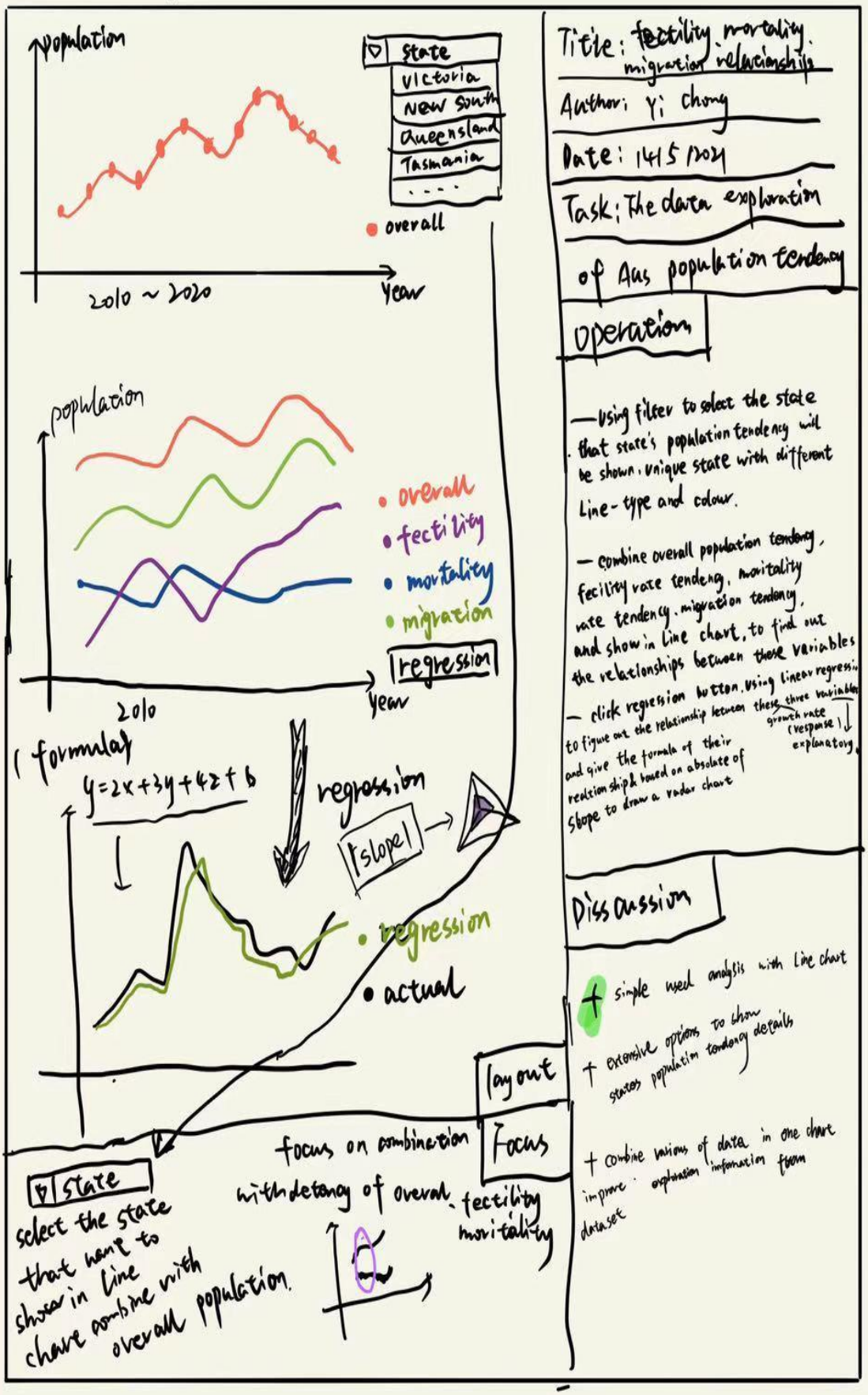
Appendix


IDEAS:
Australian population
AUS
tendency
Demographic analysis
education
income
age
race
gender
fertility rate
mortality rate
migration rate
overseas
interstate
crime rate
prediction (Future)
overall "0"
diff group
aging of population
increase
decrease
Filter
interstate ①
immigration (overall="0")
age ②
: different age group ⇒ age structure (rv)
gender
precentage(%) ③
births, deaths, education, income, crime.
impossible
Categorize:
population size ③
Deaths & births & migration ①
age
gender
older ↑
(man ≠ woman)
② population changing rate
combine births, death, migration
from 2010 - 2020.
num
year
combine age/gender
Combine & Refine
① population structure
2010
2011
2012
2013
combine age structure & gender structure.
③ tendency
combine
overall population
births
combine
population

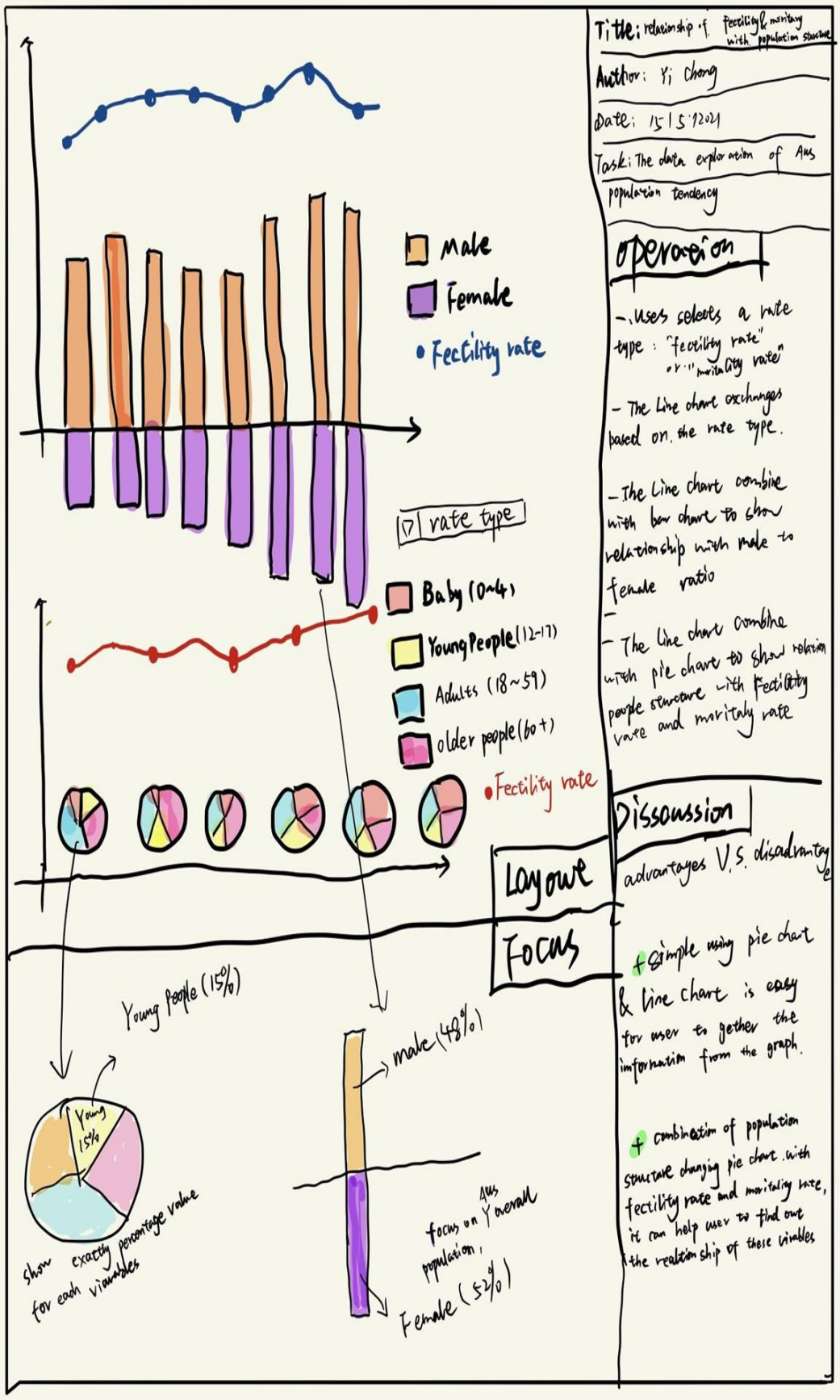
Initial design 1:



Initial design 2:



Initial design 3:



Title: relationship of fectility & maritaly with population structure

Author: Yi Chong

Date: 15/5/2021

Task: The data exploration of Aus population tendency

operation

- uses selects a rate type: "fectility rate" or "maritaly rate"
- The Line chart exchanges based on the rate type.
- The Line chart combine with bar chart to show relationship with male to female ratio
- The line chart combine with pie chart to show relation people structure with fectility rate and maritaly rate

Dissoussion

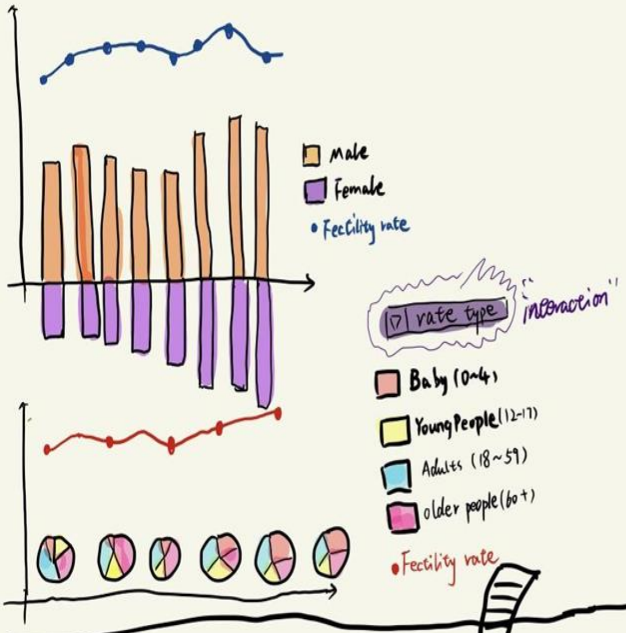
advantages V.S. disadvantage

+ Simple using pie chart & line chart is easy for user to gether the imformation from the graph.

+ combination of population structure changing pie chart with fectility rate and maritaly rate, it can help user to find out the realtionship of these viriables

Realization

STATE 1



Title:

Author: Yi chong

Date: 16/5/2021

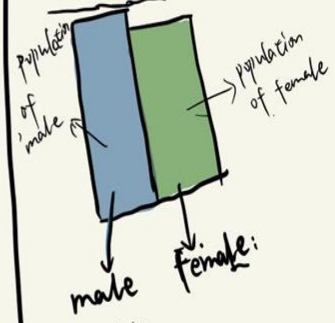
Task: The exploration of AUS population tendency

Details

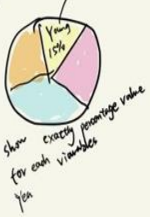
- using Linear regression to predict and find relationship with grow, births, death.
- For different variable, using different line type and line colour, for each year, the value would point on Line chart.
- software: D3 / R(shiny)
- estimate time: 2 days

STATE 2.

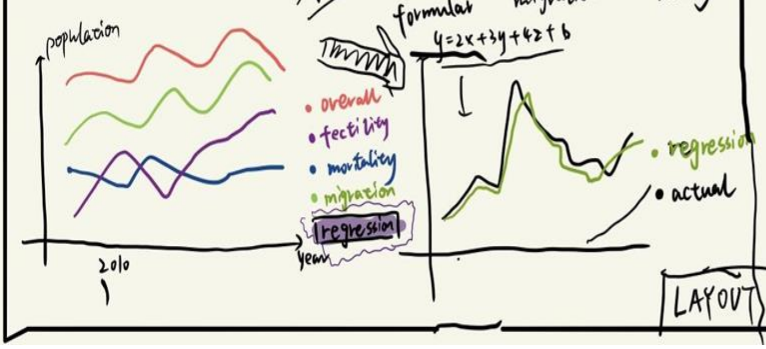
Focus

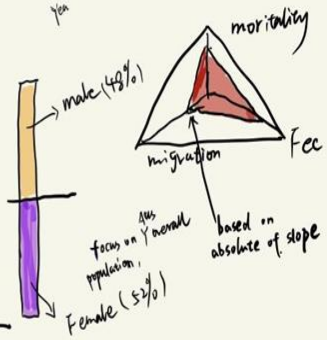


The population of female for specific state in some Young People (15%)



STATES





LAYOUT