# Week 13 Project Submission

Ian Lee

2023-11-22

## Final Submission (Week 13)

### (1) What is the theme of your data story?

The theme of my data story revolves around the assessment of singlehood in Singapore, with a specific focus on two pivotal factors: age and the highest qualification attained. The aim was to assess their impact on the rising trend of singlehood. By dissecting the influence of age and educational attainment, my analysis seeks to unravel the potential reasons behind the increasing prevalence of singlehood across different demographic segments in Singapore. This allows me to gain a better insight into how age and education independently contribute to the evolving landscape of relationships and marital patterns in the country.

### (2) Why is it important to address this question?

Addressing the question of rising singlehood in Singapore is crucial due to its multifaceted implications. Firstly, the alarming increase in singlehood directly impacts the nation's declining fertility rate, a significant concern for the Singaporean government (Ng, 2023). Understanding the factors contributing to this trend helps policymakers come up with effective strategies to address this demographic challenge.

Secondly, the emphasis on educational attainment in Singapore is rising, with more individuals achieving higher qualification levels (Lim, 2021). This shift in educational patterns may be intrinsically linked to the surge in singlehood. Unravelling this correlation is pivotal in comprehending the societal dynamics influenced by education and how it intersects with relationship choices.

### (3) Why do you think the data sources that you have curated can help you answer the question?

The dataset I used spans from the years 1990 to 2022. The extended time frame of the dataset enables us to identify long-term trends. For my project, I focused on the more recent period from 2013 to 2022, enhancing the accuracy and applicability of my analysis.

Firstly, the dataset provides a breakdown across age groups, allowing for an examination of how these demographics contribute to the overall trend. This is crucial in understanding variations in singlehood.

Secondly, the inclusion of the variable "Highest Qualification Attained" in the dataset is important, and the main factor I analysed for this project. By aligning this data with the trends in singlehood, we can observe patterns related to educational achievements and their impact on marital status. This is particularly significant given the societal emphasis on education in Singapore and the hypothesis that higher qualification levels may be linked to the increasing trend of singlehood.

In essence, the inclusion of relevant variables and a broad scope provide a solid foundation for my analysis of how age and educational attainment impact singlehood in the Singaporean context.

## (4) What are the insights from the data and how are they depicted in plots?

The hypothesis I had when I started the project was that Singlehood numbers would increase with the qualification level attained, and decrease with age.

The insights derived from the data and graphs in this project offer a comprehensive understanding of singlehood in Singapore. These insights are visually depicted through my plots, which I hope enhances the clarity of the findings.

Firstly, the overarching trend of rising singlehood is a key observation spanning the years 1990 to 2022. In my datastory tab, we can see this trend, from the first line graph, titled "How does Highest Qualification Attained affect Singlehood? (30-39 Years Old)". It illustrates a gradual increase in the percentage of singles, especially in recent times.

Focusing on the period from 2013 to 2022 provides a more detailed analysis, as seen in the second line graph. This is the line graph where I started incorporating the love-themed colours, which are 5 different shades of pink/red. These colours will be used throughout the project to indicate that the data used were from 2013 to 2022.

Secondaly, analysis based on the highest qualification attained unveils a significant correlation between education and singlehood. This analysis can be found under the "Qualification Attained VS Singlehood" tab. The line graphs vividly showcase the varying percentages of singles across different qualification levels.

I incorporated a Shinyapp visualisation, where the user of the website can filter based on age group and Highest Qualification Attained. This visual representation reinforces the finding that individuals with higher educational qualifications, especially at the university level, exhibit a higher likelihood of being single, as compared to lower educational qualifications, such as at the Secondary school level.

Although the "Below Secondary" category does not follow my hypothesis from 2019 onwards, this could be due to a small sample size, with more people having access to education in Singapore (Lim, 2021). This highlights the evolving landscape of education in Singapore and its potential influence on relationship choices.

Lastly, the age-based analysis demonstrates a clear relationship between age and marital status. The stacked bar plot depicting the percentage of singles across the age groups visually reinforces the trend that as age increases, Singlehood decreases. This analysis communicates the societal expectations and pressures associated with age, influencing marriage decisions.

The Shinyapp visualisation used in this tab shows the contrast between age groups. We can see that there is a lower average singlehood percentage in the 40-49 age group compared to the 30-39 age group. This is also visually captured by the purple-coloured bar plot.

## (5) How did you implement this entire project? Were there any new concepts that you learnt to implement some aspects of it?

To implement this project, I had a systematic approach that involved data collection, cleaning, analysis and visualisation using ggplot2 and Shiny.

Data cleaning for this dataset was relatively simple, as there were no missing values. I removed the explanation of the dataset for my cleaned data. As the original dataset did not have the title "Years", I included it, when transposing my data.

Using ggplot, I was able to create meaningful graphs and plots to analyse my dataset. The Shinyapp platform was also utilised to create interactive visualisations. The use of a consistent love-themed colour scheme enhanced clarity and coherence across different plots. This was also done to make my webpage have a more romantic vibe, since my project topic is related to romance, or maybe... the lack of romance.

While implementing this project, I used my skills in data visualisation and gained proficiency in creating interactive applications using Shinyapp. Although I faced many challenges with my codes along the way,

with the help of my teachers and fellow classmates, I was able to come up with a functional website. With the help of Google, I was also able to incorporate some stuff that was not taught in class.

On my webpage, I have interactive images. When the user clicks on them, they will be brought to a new tab, consisting of the link I inserted. The first instance of this falls under the "Datastory" tab, where I inserted an article from Today. The second instance is a more light-hearted one, where I inserted the Rick Roll image. This falls under the "About Me!" tab and was included to bring a bit of fun to my users as they scroll through my website. I also included a gif, sending a virtual hug to my users, which was meant to make my website more lively. I also found out that I could insert images from my local drive, instead of inserting images from Google. As such, I drew a duck and gave a chat bubble to it. Images with the duck are interactive.

I also learned that I could change the font colour of my titles. I did it to create romantic vibes, just like the love colours in the graphs. The theme used was called "journal" as this uses pink colour for the top of the website, which is consistent with the theme I was going for.

**(Word Count: 1198 words)**

### References

Lim, J. (2021, June 18). 2020 census: More singles in all age groups; slowest decade of population growth since 1970. TODAY. https://www.todayonline.com/singapore/2020-census-more-singles-singapore-all-age-groups-rate-population-growth-slowest-1970

Ng, A. (2023, February 24). Singapore's total fertility rate drops to historic low of 1.05. CNA. https://www.channelnewsasia.com/singapore/singapore-total-fertility-rate-population-births-ageing-parents-children-3301846

## Documentation (Week 12)

### Challenges Faced & How I overcame them

First challenge I faced was tidying my data. I overcame this by reading up on Google and discussing with fellow classmates.

Another challenge I faced was attempting to do a ggplot. I faced some difficulties as I was trying to create a graph based on the year. However, I realised that in my dataset, the years were not titled "years". As such, in my code, I used "2022" directly.

As this is the first time I am designing a website, I am unsure what makes a website look nice. I have been asking design friends for advice as well as searching Google for inspiration.

## Documentation (Week 11)

### (1) List the visualizations that you are going to use in your project (Answer: What are the variables that you are going to plot? How will it answer your larger question?),

Variables to plot: Age Group Sex (Gender) Highest Qualification Attained Relationship Status

### (2) How do you plan to make it interactive? (Answer: features of ggplot2/shiny/markdown do you plan to use to make the story interactive),

I plan to use several graphs to make my website interactive

Dependent Variable: Relationship Status

Independent Variable: Age: I am planning to list down the different age groups. When the user clicks on the icon, the statistic for the number of singles for that age group will appear Sex: I am planning to do a pie chart to display the gender distribution within the dataset, providing insights into gender representation. Highest Qualification Attained: I am planning to use a line graph to show how highest qualification attained may affect relationship status

### (3) What concepts incorporated in your project were taught in the course and which ones were self-learnt? (Answer: Create a table with topics in one column and Weeks in the other to indicate which concept taught in which week is being used. Leave the entry of the Week column empty for self-learnt concepts)

to be answered when I have more progress on the website

## Documentation (Week 10)

### (1) What is the question that you are going to answer? (Answer: One sentence that ends with a question mark that could act like the title of your data story),

Why is Singapore experiencing an increasing number of people who are not married?

### (2) Why is this an important question? (Answer: 3 sentences, each of which has some evidence, e.g., "According to the United Nations..." to justify why the question you have chosen is important),

Singapore's rising singlehood is of significant concern, with one of the world's lowest fertility rates, at 1.05 as of 2022 (CNA,2023), driven partly by delayed marriage and childbirth. This is impacting future population growth and aging demographics.

The surge in singlehood also reflects evolving social and cultural norms, emphasizing individual pursuits and career development, with potential implications on family structures, gender roles and societal values in Singapore.

Lastly, policymakers urgently need to comprehend the reasons behind this trend to develop effective strategies, including family support policies and measures to address the challenges posed by an aging population and a shrinking workforce in Singapore.

### (3) Which rows and columns of the dataset will be used to answer this question? (Answer: Actual names of the variables in the dataset that you plan to use).

I will be using all variables from my dataset. These are datas that show Age Group, Sex And Highest Qualification Attained. They include:

30-39 Years (Total) 30-39 Years (Males) 30-39 Years (Females) 40-49 Years (Total) 40-49 Years (Males) 40-49 Years (Females) Below Secondary Secondary Post-Secondary (Non-Tertiary) Diploma & Professional Qualification University

# Documentation (Week 9)

## (1) What is the topic that you have finalized? (Answer in 1 or 2 sentences)

The topic I chose to analyse is the population of singles in Singapore and the potential causes behind why people are not getting married.

Single: This refers to a person who has never been married.

## (2) What are the data sources that you have curated so far? (Answer 1 or 2 sentences)

https://tablebuilder.singstat.gov.sg/table/TS/M810591

Proportion Of Singles Among Resident Population By Selected Age Group, Sex And Highest Qualification Attained