

# Analysis of Singapore's Economic Statistics, Socioeconomic Backgrounds and Population Demographics.

By Hsu En Yang, Parth Goda, Pham Anh Kim Chau

**Abstract** - In global surveys and media coverages, Singapore is often cited as a safe haven for the wealthy and associated with attention-grabbing headers like “[most expensive city](#)” and “[city with a high number of millionaires](#)”. Although there are elements of truth in these titles, the media noises have drowned out the fact that Singapore is still a humble nation with real problems. With Singapore's publicly available statistical data, we wanted to collect these information, transform the visualisation data and change the way people think about Singapore. We seek to provide a one-stop tool for users of any background to deep dive into the different facets of Singapore, such as income, qualifications, occupations, etc and also use our charts to find correlations between the factors. This new application is deliberate on using choropleth map charts, box plots, bar graphs and pie charts as they are familiar visualisations to the public eye.

**Index terms** - Singapore, Statistics, Economics, Income, Occupation, Housing, Transport, Qualification, Social, Visual Analytics.

---

## 1. Introduction

As the jewel of Southeast Asia, Singapore is a chart topper for many global, accredited rankings. Well known for its prosperity, the media attention for Singapore often focused on the pool of wealth and affluent individuals that own them. When *Crazy Rich Asians* topped the global box office charts, Singapore's status as one of the richest nations solidified and people gradually became blinded to the real problems faced in Singapore.

Coupled with the benefits in this age of information, Singapore's public statistical data such as population census, socioeconomic figures, housing and establishments are now readily available on portals like [data.gov.sg](#) and [singstat.gov.sg](#). However, the abundance of data has become an issue of information-overload for most users. Moreover, the data have been broadly categorised and presented in

over-simplistic visualisations, which poses a hindrance to users trying to understand the data.

The unrealistic portrayal of Singaporeans and lack of comprehensive data to substantiate reality have resulted in an unsettling question for many citizens residing in the country: *why do I not feel rich?*

This passion project seeks to debunk the facade of affluence by unearthing the real problems faced on a daily basis. Our intention is to design an application that is comprehensible, intuitive and visually appealing for the general public to use. The application is deliberate in turning common population data into easy-to-understand visualisations, which serves to cover the entirety of the nation when an attribute is chosen.

This paper will cover the push factors behind our research direction and discuss the proposed solutions for the aforementioned problems. This

is followed by a review of the application architecture and analytics that we have derived from the data points presented in our application. Finally, we will share the future direction of our efforts and close with a summary of this passion project.

## 2. Background

### 2.1. Motivations and Objectives

Our research and developmental efforts were largely motivated by the lack of comprehensible and effective web-enabled visual analytics tool for understanding population statistics in Singapore. While the open-data portal is a good initiative, we feel that the difficulty in navigating through the abundance of data and unappealing formats of data presentation have rendered the platform to be underutilised.

Fuelled by the misrepresentation of Singapore's reality on the media, our team had chanced upon the void between readily available data and its efficacy in conveying accurate information. Specifically, we seek to achieve the following objectives:

- 1) To create a web-based application for users of any background to navigate through and understand the statistical information presented
- 2) Highlight the economic situation, population demographics and socioeconomic backgrounds through interactive, comprehensible and dynamic visualisations
- 3) Debunk media hype by highlighting the realities of Singapore: income inequality and the varying levels of life standard

### 2.2. Data Collection

As mentioned, our data are readily available on the government-backed website: <http://data.gov.sg>. Since the media hype has been centered on Singapore as a haven for lavish lifestyles, we were intentional in picking datasets that have a direct or indirect impact on income. By doing so, we will be able to highlight Singaporeans' *actual problems*, which are mostly income-centric, and portray a more accurate version of this humble nation. In order to make the datasets more relevant to users, our data are categorised by planning areas of Singapore. As a highly connected society, users will be able to draw from their memories and past experiences from visiting the location, and have this sense of imagery surfaced when utilising the tool. After thoroughly sieving through the platform, we have accessed the following datasets:

- a. Highest Qualifications by area
- b. Household size by area
- c. Industry worked by area
- d. Economic Status (active/ inactive/ unemployed) by area
- e. Marital status by area
- f. Mode of transport
- g. Monthly income by area
- h. Occupation by area
- i. Household type by area
- j. Population size by area

These datasets are the backbones to the visualisations available in our application. In order to ensure the comprehensibility of income-related statistics, we will be discussing the steps to transform data in the next section to help remove information tangential to our objectives.

### 2.3. Data Preparation: Trimming and Manipulation

To prevent information overload, we have carefully curated the dataset by removing columns that were too in-depth or had little correlation to income. For example: Household type per area had columns on price per square feet and price sold were taken out.

Next, we begin manipulating the formats of the datasets with the functions available on the various packages in R. For instance, mode of transport/ occupation and industry worked by area were presented in the following form:

no.	Service	Construction	Manufacturing
Yishun	9999	9999	9999
Katong	9999	9999	9999

Table 1: Snippet of transport mode choice

The above table has been transformed to the following output using gather function in dplyr package:

Area	Industry	No.
Yishun	Service	9999
Yishun	Construction	9999
Yishun	Manufacturing	9999

Table 2: Transformed version of Table 1

Another example of a common format presented by the source:

level_2	level_3	level_4	value
Total	Total	Ang Mo Kio	312.2
Total	Female	Ang Mo Kio	235.9
Total	Female	Ang Mo Kio	76.3

Table 3: Snippet of gender by planning area

The reformatting is carried out by removing the aggregated sum in the following:

Area	Gender	Value
Ang Mo Kio	Female	235.9
Ang Mo Kio	Female	76.9

Table 4: Transformed version of Table 3

After summarizing all the datasets, we used dplyr package to find the percentage of different categories under the same planning area. This step is to avoid different scales, as a result of varying number of residents in each planning area, when presented as a visualization.

Each dataset has to be transformed based on the intended visualisation. For instance, data used for choropleth maps have to split across multiple columns. With this, users can choose the type of category distribution they want to see and the map will display the corresponding information from the desired column. On the other hand, datasets prepared for stacked bar charts are aggregated by attributes.

## 2.4. Types of Visualisation

Incorporating different types of visualisation is imperative to highlight the key message for datasets.

- **Choropleth map:** “a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable being displayed on the map, such as population density.”
- **Stacked bar charts:** “to show comparisons between categories of data, but with the ability to break down and compare parts of a whole.”
- **Box plot:** “a graphical method of displaying variation in a set, or multiple sets of data.”
- **Pie chart:** “a type of graph in which a circle is divided into sectors that each represent a proportion of the whole”

## 3. Application

### 3.1. R Shiny Framework

**Packages:** *Readxl*, *shiny*, *sf*, *tmap*, *tidyverse*, *dplyr*, *ggplot2*, *plotly*, *DT*, *stringr*, *RColorBrewer*, *shinythemes*

The Sunny Singapore Application was fully developed in R through utilising the Shiny Framework. With the support of the aforementioned packages, users are able to navigate through the web-based application, which is linked to our prepared data on the backend, in an interactive manner such as filters and tooltips.

### 3.2. Application Description

To appeal to the general public, we invoked a sense of familiarity by using the image of Singapore’s iconic Marina Bay Sands and

immediately contrasted it after the navigation bar with a picture depicting income disparity.

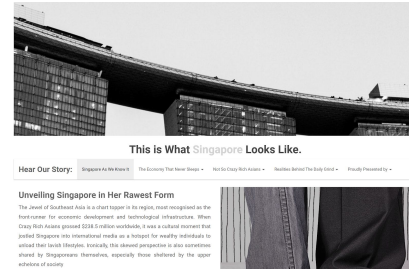


Fig. 1: Homepage

The layout and colour scheme intentionally reflects minimalism, in hope to create the impression that the application is easy to navigate and up to date. During the design, we adopted a more colloquial way of presenting the content, such that users can better interpret the intended messages.



Fig. 2

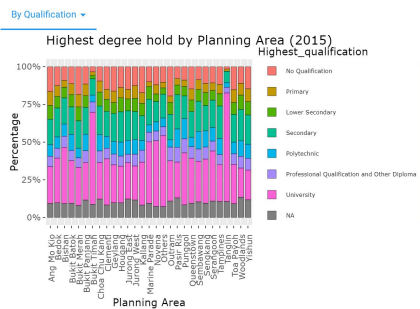
The navigation bar walks users through a series of perspectives in viewing Singapore’s wealth. ‘The Economy That Never Sleeps’ focuses on economic indicators, ‘Not So Crazy Rich Asians’ illustrate the income disparity while ‘Realities Behind The Daily’ underscores the demographics of the citizens. Each sub-tab is supported by Sunny Singapore’s in-house analysis based on the visualisation presented. After navigating through these steps, users should be able to view Singapore objectively.

## 4. Case Studies

Fundamentally, the goal of Sunny Singapore Application is to offer an overview of Singapore’s reality through different lenses. The basis of this approach mainly stemmed from the muted problems of income disparity, which were mainly anecdotal reports covered by

less-mainstream media. With the aid of Sunny Singapore’s visualisations, some attributes demonstrated a distinct and direct relationship with income earned. As such, this section aims to reveal the unnerving divergence between the rich and poor with the cases found in Sunny Singapore.

#### 4.1. The relationship between education and income earned



It is no secret that those with higher qualifications and education tend to find higher paying jobs than those with lower qualifications. Through our application, we can see this empirically as the percentage of people with a university degree staying in Bukit Timah and Tanglin is twice that of people living in other areas. Hence, it comes as no surprise that the highest concentration of people with a monthly income of more than \$15,000 can be found in these 2 places which suggests that there is a significant relationship between education and income earned.

#### 4.2. The relationship between education and income earned

In countries such as America where cars are a necessity due to the vastness of the country, people from all income groups are likely to own a car. However, the opposite is true for Singapore where car prices are exorbitant due to various additional costs and are only owned by the well-to-do citizens. In our application, it can

be seen the predominant modes of transport in most areas are buses and trains. However, the predominant mode of transport in the affluent areas of Bukit Timah and Tanglin is by car, one of the costliest assets one can own in Singapore, thus suggesting that there is a significant relationship between owning a car and earning an above average income.

### 5. Areas of Improvement and Future Work

Currently, our application is limited to three broad categories presenting Singapore. We are eager to explore other factors such as political viewpoints, healthcare situation, psychological well-being, generational trends, etc.

To further improve the authenticity of the application, we hope to create multi-dimensional filters such that users can deep dive into certain attributes (e.g. filter by gender and qualification altogether). Moreover, the interactivity of the application can be enhanced by allowing users to enter their data to generate an insight or simple gamification features.

In our next iteration in enhancing the application features, we hope to incorporate data spanning over multiple years to discern the progression of economic statistics, socioeconomic statuses and population demographics. With historical data, cognizing trends would create the opportunity to strategise prescriptive measures.

### 6. Conclusion

Sunny Singapore is founded on fundamentals of visual analytics, which serve as an important framework to handle big data. Thus, it has been specially designed for users of any background to understand through visualisations the current situation in Singapore.

Through this exercise, we hoped to have debunked the notion that *Singaporeans are generally wealthy*. Beneath these false claims are citizens struggling to make ends meet. The visualisation tools that were adopted should have painted a clear picture of this disparity. Although recognising these problems have begun to gain traction, we are still slow raise this awareness. With the ease of using this application, we hope to reach out to a wider network of users and help people to understand the real problems in Singapore, thus working in harmony to progress as a nation altogether.

## **7. Acknowledgments**

We would like to thank professor Dr. Kam Tin Seong, for his time and valuable guidance throughout the course of this project and IS428 Visual Analytics for Business Intelligence.

## **8. References**

Webb, J. (2017, November 30). Singapore Is Still The Most Expensive City In The World. Retrieved from <https://www.forbes.com/sites/jwebb/2017/11/30/singapore-is-still-the-most-expensive-city-in-the-world/#76fe0d0ac5e9>.

Lim, S. (2018, October 19). There are almost 200,000 millionaires in Singapore – around a thousand with over S\$60 million. Retrieved from <https://www.businessinsider.sg/almost-200000-millionaires-in-singapore-a-thousand-with-over-s60-million/>.