Singapore Society in Numbers

Edited by Shannon Ang

Last updated 20 May 2019

Contents

Preface

This online book is a compilation of resources aimed at advancing quantitative social science in Singapore. It is meant to be a 'living document', so it will be updated as frequently as possible. The main goal is to promote interest, rigour, and transparency in trying to understand Singapore society quantitatively. It does so by:

1. **Providing information on Singapore-relevant datasets** that are currently used to answer research and policy questions (Chapter 1 and Chapter 2). This includes:
    - Descriptions of *publicly available* datasets and how to access them. This overview of the 'data landscape' will be helpful for social scientists to get started with research on Singapore, and prevent wasteful overlap in primary data collection across institutions.
    - A list of *restricted* or *non-publicly available* datasets that could be used to answer important research or policy questions if access was granted. If available, details on the dataset and reasons for data restriction will also be listed. It is hoped that this list will promote greater transparency in data sharing across research teams.
2. **Occasional think pieces by researchers** on best practices and on how to improve quantitative social science in Singapore (Chapter 3).
3. **Maintaining a repository of replicable case studies on Singapore society** (with annotated code, where possible) which can be used for illustrations in any quantitatively oriented college-level class (Chapter 4 onwards). These may be short summaries (blog-length) of published work, or side analyses that may not be appropriate for an academic journal but are useful for Singapore social science nonetheless.

Readers with ideas on how to improve this resource (or who may wish to help me maintain it) may email me at shanang@umich.edu.

Why I started this project

Quantitative research is not (and should not be) the only approach we take to understanding Singapore society, but constant appeals to "big data"[1] or claims of "evidence-based policy"[2] makes it ever more important for members of the public to **critically evaluate the use of numbers** in making arguments or in representations of social phenomena.

Educational institutions have an important role to play in this "data-driven" world. Every year, undergraduates studying the social sciences in our local universities take several courses in research methods to fulfil the requirements of their degrees. Part of this research methods sequence typically involves training in introductory statistics or "quantitative reasoning". Quantitative courses in social science departments differ from those taught in the natural sciences because they are thought to be more applied - the focus is on the use of statistical methods to answer questions about society. Understanding and applying these methods **to the Singapore context** is crucial here - at this point, students learn about (and hopefully are inspired by)

[1] See, for instance, https://www.todayonline.com/singapore/business-big-data-singapore-has-built-cutting-edge

[2] Government agencies such as the Ministry of Social and Family Development often use such a phrase.

the kind of questions they can ask about the very society they live in, given the quantitative tools they are learning.

However, my first exposure to statistics as an undergraduate reading Sociology at NUS[3] was to textbooks containing examples from only Western societies (e.g., Agresti and Finlay, 2009; Treiman, 2009). While the use of these internationally-recognized textbooks may provide some assurance of quality education, sole reliance on foreign material often becomes a missed opportunity to inspire students to build on and improve Singapore social science. Without contextualization[4], abstract statistical concepts (e.g., hypotheses testing, chi-squared tests) seem removed from everyday experience, and impede the ability to take these important concepts beyond the classroom and into public dialogue.

I started this book with the view to use it primarily *as a teaching tool*[5], but it can be used in many other ways. In the long term, I hope that resources in this book will encourage quantitative literacy and research in Singapore by making it easier for interested parties to browse, use, and understand Singapore-relevant data. Social science researchers may use the dataset listings as a springboard for collaboration, or contribute their own interesting case studies for the benefit of the Singapore public. Others (such as journalists, civil servants, or non-profit organizations) may find value in these material as a gateway to quantitative research on Singapore society, and how to think carefully about pertinent issues surrounding such work.

For Singapore social science.

How to contribute

Instructions on how to list a dataset, contribute a case study, or write a think piece for this page.

Acknowledgements

This book is being written through the **bookdown** package (Xie, 2019), which was built on top of R Markdown and **knitr** (Xie, 2015).

Contributors include:

About me

Little write-up about myself

[3](the) National University of Singapore

[4]Notwithstanding the terribly unhelpful stereotype of social science students being "good at writing but bad at numbers".

[5]For instance, the public repository of Singapore-oriented examples and illustrations may be used to supplement courses based on textbooks written by international scholars.

Part I

Datasets for Social Science

Chapter 1

Public Data

List of public data

Chapter 2

Restricted Data

List of restricted data

Part II

Think Pieces

Chapter 3

Thinking about Numbers

Think pieces section

3.1 Think piece 1

Part III

Case Studies

Chapter 4

Out of Proportion

- Contributor: Shannon Ang

Chapter 5

Case study 2

This is another case study

5.1 Witty title for Case 2

- Contributor:
- Dataset:

This is an example of in-line code annotation and output.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Figures can be referenced, e.g., see Figure 5.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 5.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

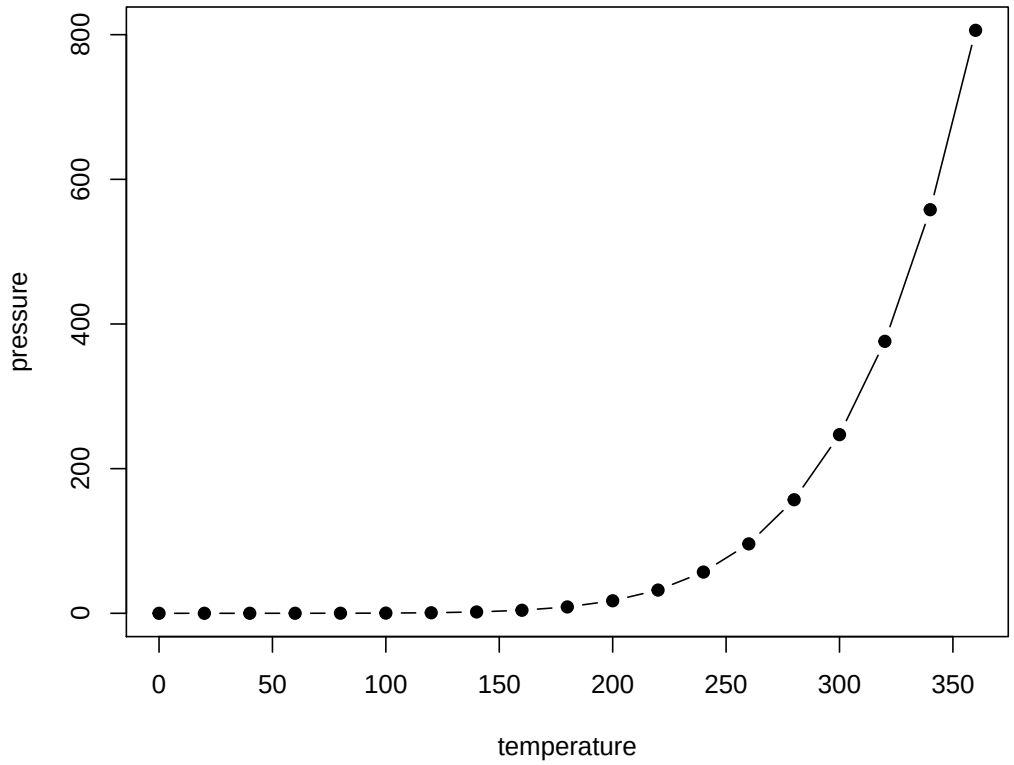

Figure 5.1: Here is a nice figure!

Table 5.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Bibliography

Agresti, A. and Finlay, B. (2009). *Statistical Methods for the Social Sciences*. Pearson Education, 4 edition.

Treiman, D. (2009). *Quantitative Data Analysis: Doing Social Research to Test Ideas*. Jossey-Bass.

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2019). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.10.