A Gender Data Lab Capacity Building Programme in Rwanda

Workshop

Shanghai Jiao Tong University

Shu Fu

Aug 12–14, 2024

Bugesera, Rwanda

Shanghai ⇌ Rwanda

Gender Data Lab Workshop

- National Institute of Statistics of Rwanda
- Gender Monitoring Office
- PARIS21 of OECD
- SJTU (Global Challenge Project)

Workshop Themes

- Reproduction
- Visualization
- Communication

Objectives

- To conduct data analysis in a **reproducible** manner (programming in R)
- To improve the ability of statisticians to convey gender-related content via **visualization**
- To practice **communicating** with Factsheet and Infographic

By the end, hope you can 1) be able to better visualize gender-related data using R packages; 2) reproduce figures with new data in the future

Outline

WHO: Self-Introduction and Ice Breaking

WHY: Why Visualize?

HOW: Workshop Outlines

Self Introduction

Instructor

Dr. Shu FU

- Assistant Professor in School of International and Public Affairs, SJTU
- Ph.D. in Political Science, The University of Chicago
  - American Politics & Empirical Methodology
- My research is published at top journals in political science, such as *Journal of Politics*, *World Politics*, *Political Research Quarterly*, *Presidential Studies Quarterly*
- Passion in teaching: Introduction to American Politics, Causal Inference
  - Awarded the 1st Prize in Young Faculty Teaching Competition, SJTU

My Teaching Philosophy

- As a **facilitator**
  - Teach how to fish, not give a fish
  - Learning the skills of **essence** that participants can use be further applied
  - First time in Africa, I am devoted in working with you. Hope this workshop really useful and make a change.
  - Lecture + Hands-on practices
  - Be engaging with me, with each other, and with facilitators

Facilitators

- Jingjing LI, (Krystal)
- Junliu JIANG, (Willow)
- Sida ZHANG, (Star)
- Han BAI, (Alex)
- Zetong ZHANG, (Zet)
- Tingting XU, (Ella)

- Support participants in solving analytical or coding problems
- Assist participants to produce targeted reports or presentations

Ice Breaking

Why Visualize?

Group Discussion

- Why we need visualization?
- What criteria do you have in mind for a good visualization?

Why visualize?

In short:

- Data are complex.
- Your reader won't be familiar with your data.
- Visualizing your data cuts through the cloud by simplifying the data.
- A picture is worth a thousand words.

A Picture is Worth a Thousand Words

Source: AP Photo by Evan Vucci

Examples

LEGISLATORS	1990		1994		2001		2003		2008		2011		2013		2018		2021		2023	
DEPUTIES	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M
	17	83	14	86	23	77	49	51	56	44	56	44	64	36	61	39	61	39	61	39
SENETORS	NA						34	66	38	62	38	62	38	62	38	62	38	62	38	62

Here is a dataset.

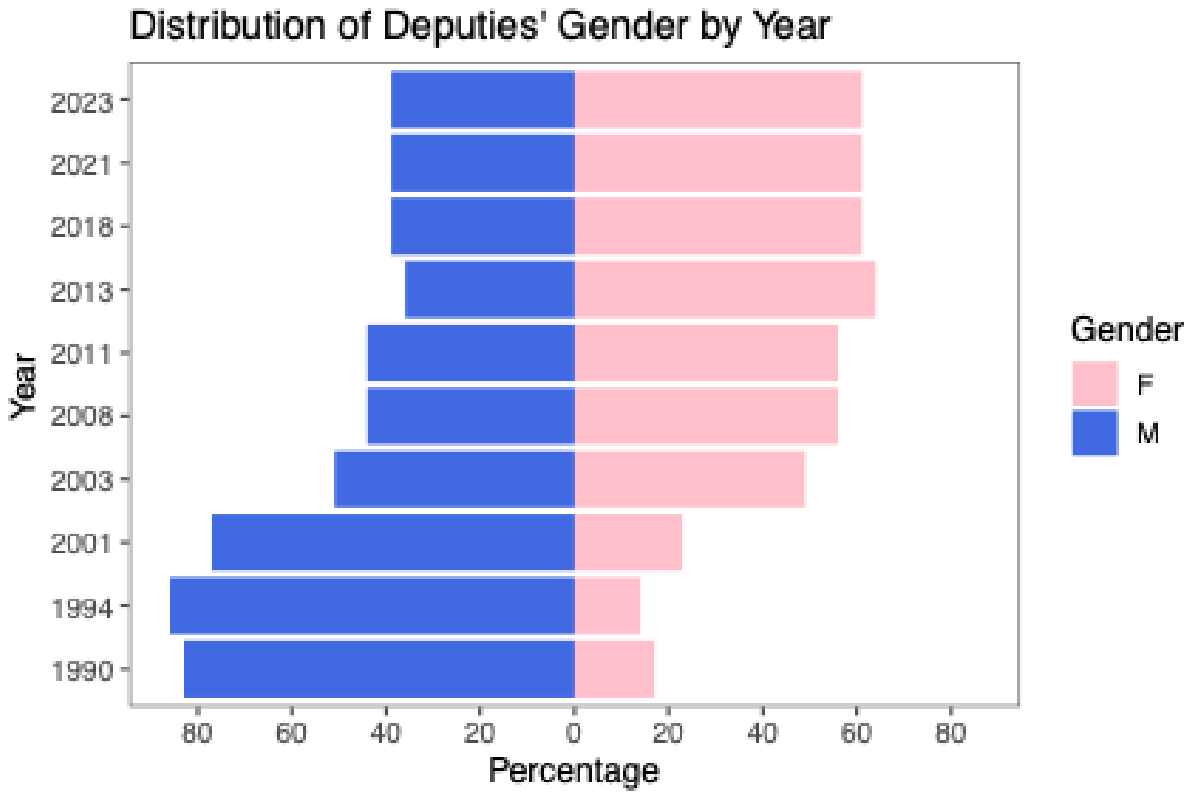
Do you like it?

Anything unclear?

Is it an effective way for communication?

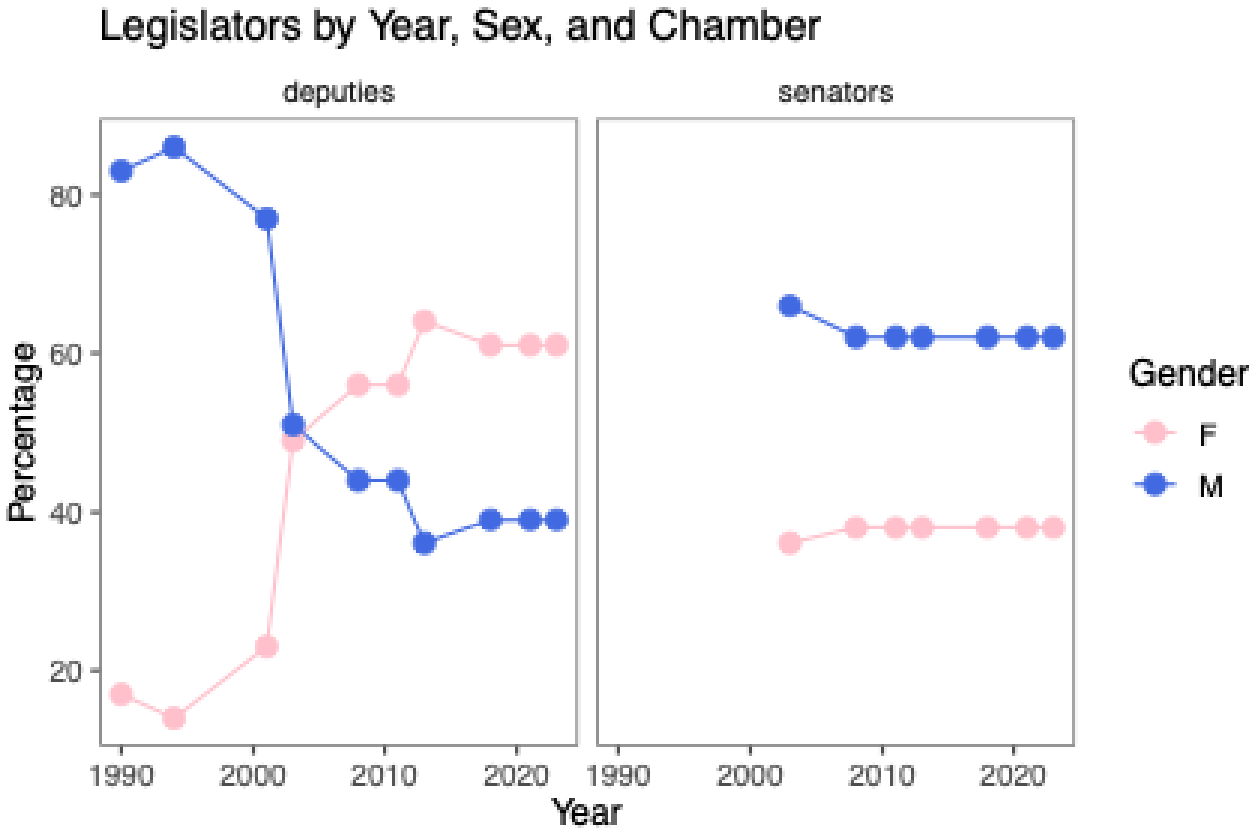
Examples

LEGISLATORS	1990		1994		2001		2003		2008		2011		2013		2018		2021		2023	
DEPUTIES	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M
	17	83	14	86	23	77	49	51	56	44	56	44	64	36	61	39	61	39	61	39
SENETORS	NA						34	66	38	62	38	62	38	62	38	62	38	62	38	62



Examples

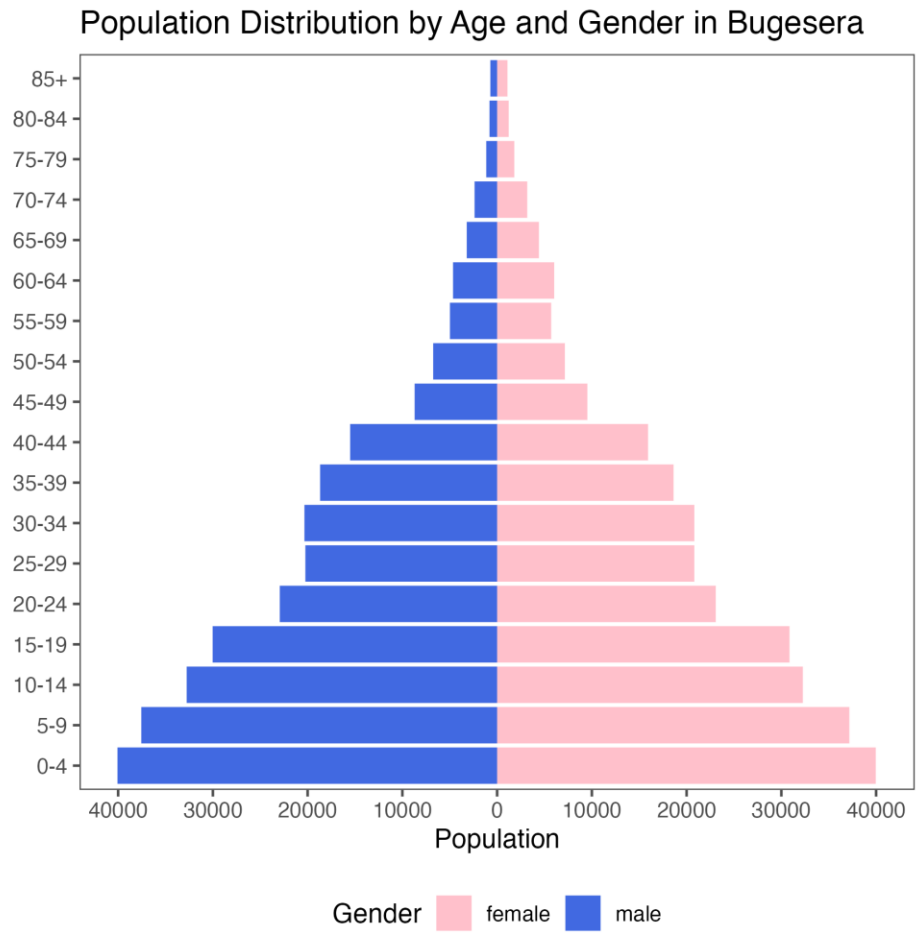
LEGISLATORS	1990		1994		2001		2003		2008		2011		2013		2018		2021		2023	
DEPUTIES	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M
	17	83	14	86	23	77	49	51	56	44	56	44	64	36	61	39	61	39	61	39
SENETORS	NA						34	66	38	62	38	62	38	62	38	62	38	62	38	62



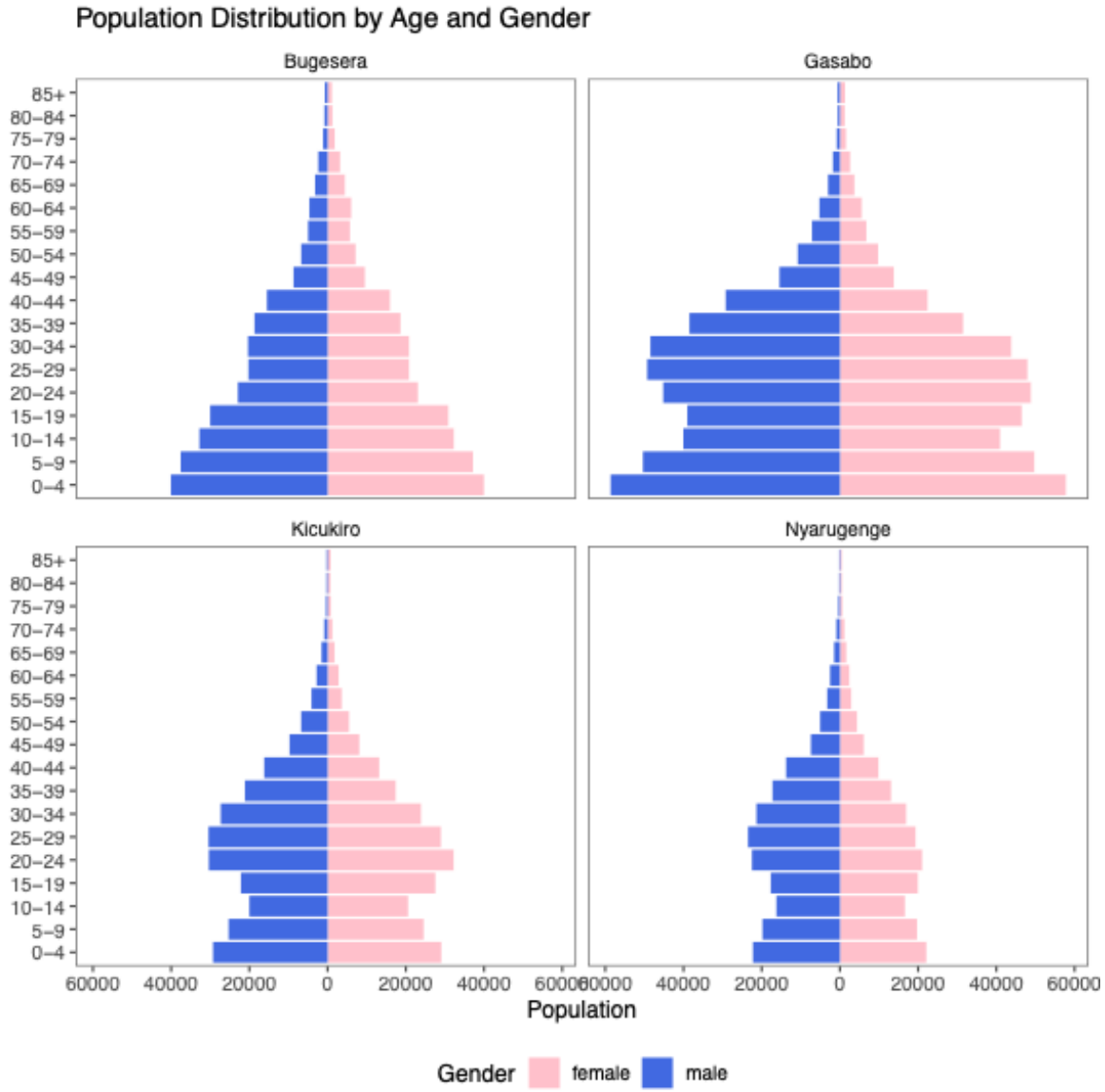
Another Example

District = Bugesera			
	Sex		
	Both Sexes	Male	Female
age group five from 0 to 85			
Total	551,103	271,468	279,635
0-4	80,033	40,069	39,964
5-9	74,730	37,556	37,174
10-14	65,034	32,767	32,267
15-19	60,886	30,026	30,860
20-24	46,017	22,940	23,077
25-29	41,051	20,239	20,812
30-34	41,151	20,337	20,814
35-39	37,304	18,683	18,621
40-44	31,442	15,513	15,929
45-49	18,229	8,705	9,524
50-54	13,899	6,748	7,151
55-59	10,686	4,981	5,705
60-64	10,676	4,653	6,023
65-69	7,616	3,205	4,411
70-74	5,559	2,380	3,179
75-79	2,960	1,147	1,813
80-84	2,024	804	1,220
85+	1,806	715	1,091

Source: Fifth Rwanda Population and Housing Census, 2022 (NISR)



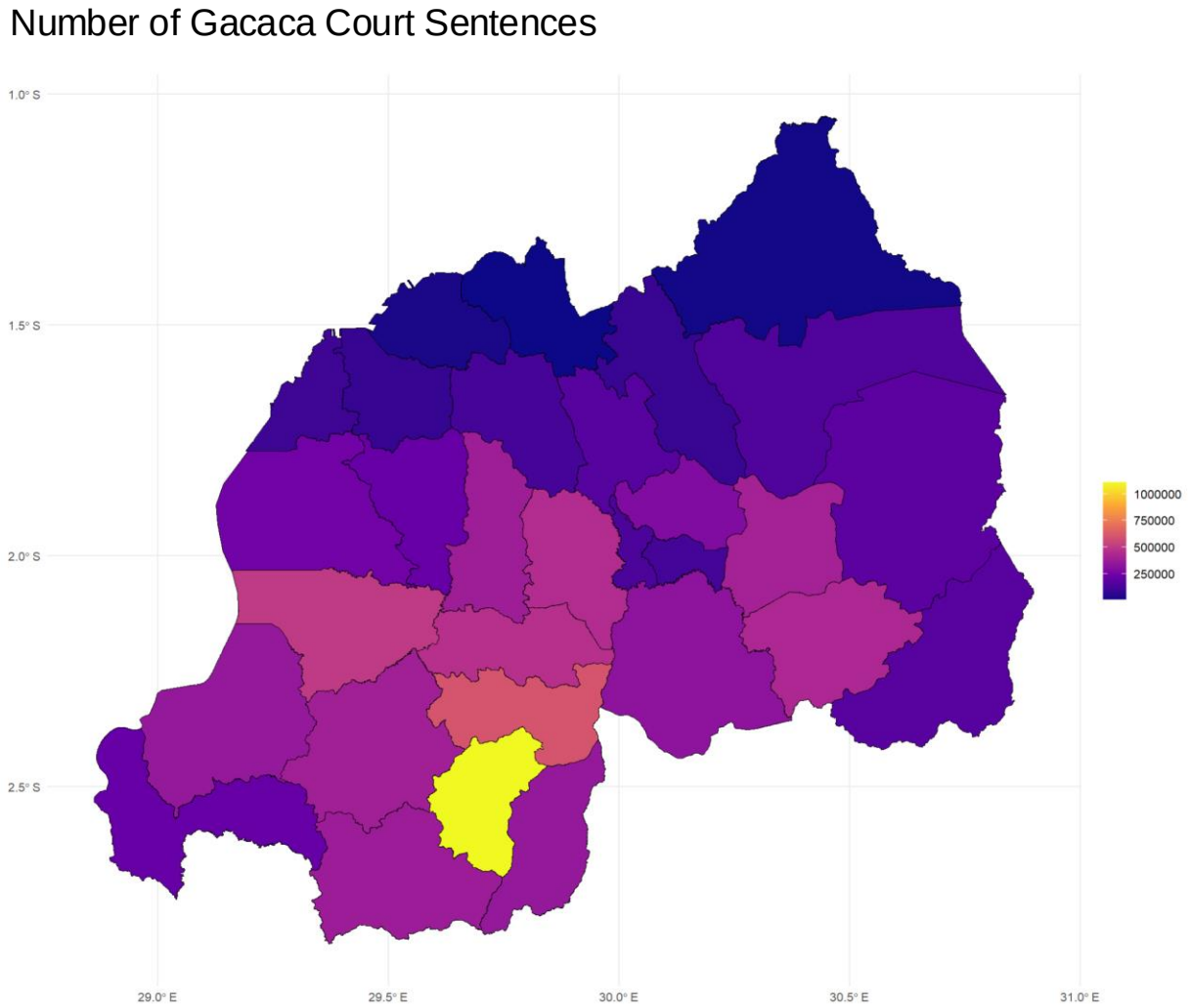
Moreover, it's reproducible!



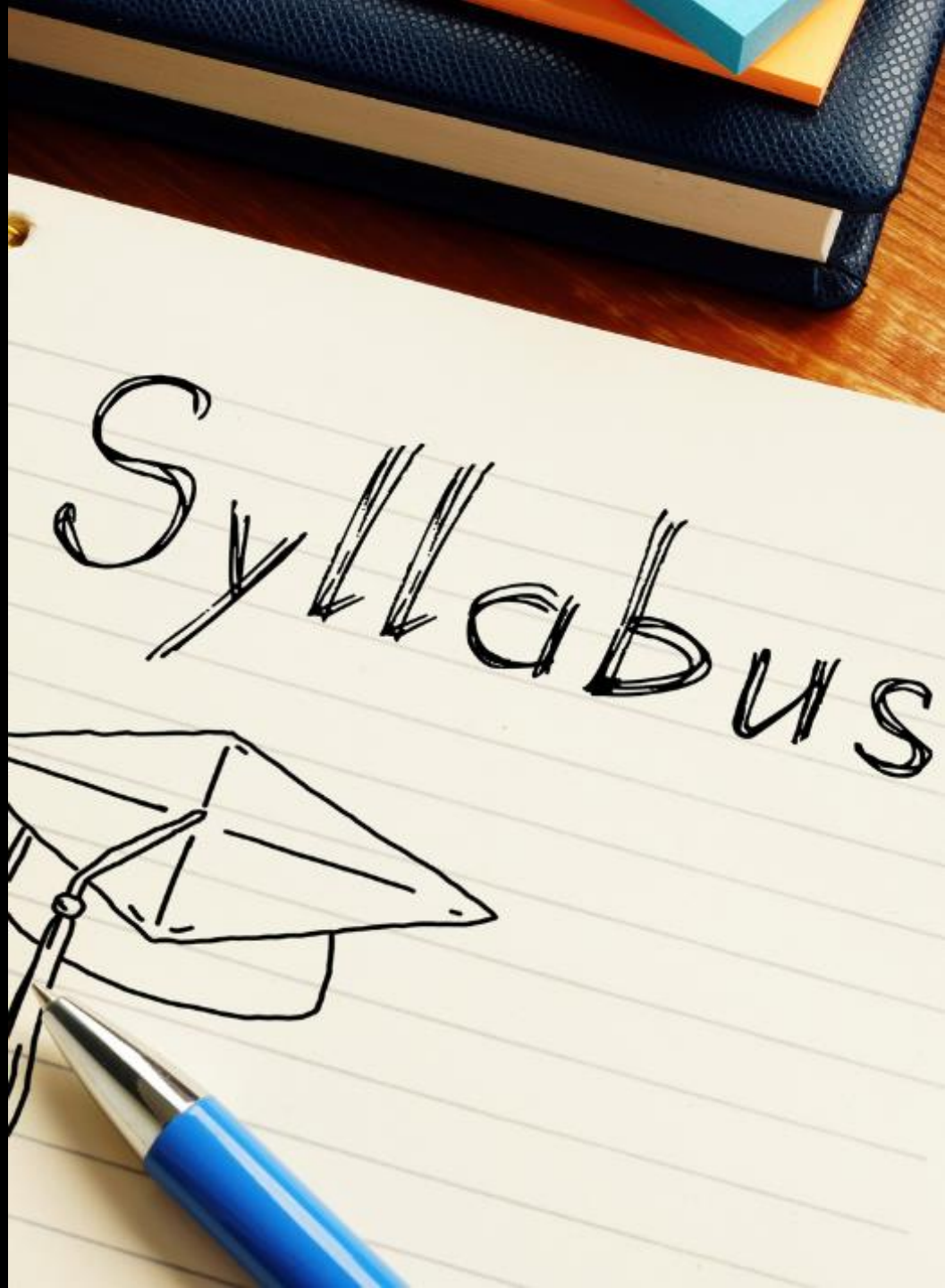
It’s just several lines of code!

```
long_data %>%
  filter(district %in% c("Bugesera", "Nyarugenge", "Gasabo", "Kicukiro")) %>%
  ggplot(aes(x = age_and_district, y = population, fill = gender)) +
  geom_bar(stat = "identity", show.legend = TRUE, ) +
  coord_flip() +
  facet_wrap(~district) +
  scale_fill_manual(name = "Gender", values = c("pink", "royalblue")) +
  scale_y_continuous(labels = abs,
                     breaks = scales::pretty_breaks(n = 5)) +
  labs(x = "", y = "Population",
       title = "Population Distribution by Age and Gender") +
  theme_few() +
  theme(legend.position="bottom")
```

Last example, it's cooler!



Workshop Outlines



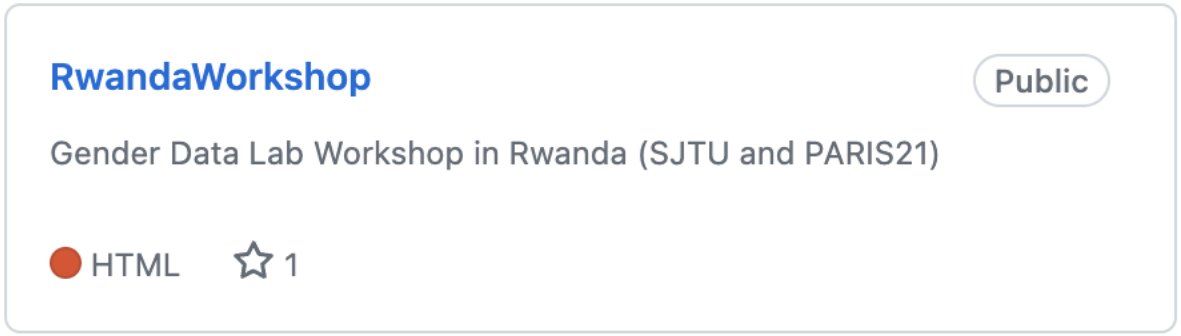
Software and Readings

- For this workshop we will be using `R`.
- `R` is a programming language that is especially powerful for data exploration, visualization, and statistical analysis.
- To interact with `R`, we use RStudio.

- Additional readings for further learning.
  - *R for Data Science*, Hadley Wickham and Garrett Grolemund, https://r4ds.had.co.nz/.
  - *ggplot2: Elegant Graphics for Data Analysis*, by Hadley Wickham, https://ggplot2-book.org/index.html.

Notes, Data and Code

- I will share with you all workshop notes and code on Github.



- https://github.com/fushu18/RwandaWorkshop
  - The notes are in html format, please download them and open in your browser.
  - The data and code will be shared as well.

Day 1: Reproducible Data

- Session 1.a: Course Overview
- Session 1.b: Programming 101 – Hands-on session
- Session 2.a: Reproducible Data Analysis
- Session 2.b: Reproducible Data Analysis – Hands-on session

Day 2: Visualization

- Session 3.a: Data Visualization
- Session 3.b: Data Visualization – Hands-on Session
- Session 4.a: Communicating with Factsheet and Infographic
- Session 4.b: Communicating with Factsheet and Infographic – Hands-on Session

Day 3: Communication

- Session 5 Hands-on Exercises – Creating data factsheets and infographics

- Session 6 Results presentation and discussion

Workshop Etiquette

- Feel free to interrupt, ANYTIME! No question is stupid.
- Feel free to communicate your need, we’re happy to adapt.
- Work on an R script locally, while having the notes (html files) opened up.


Let's
Roll

Thank you!

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

国际与公共事务学院 SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS

Thank you!

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

国际与公共事务学院 SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS