**Workshop Syllabus**

A Gender Data Lab Capacity Building Programme in Rwanda

**INSTRUCTOR**

Dr. **Shu FU** (fushu@sjtu.edu.cn)is an assistant professor in the School of International and Public Affairs at Shanghai Jiao Tong University (SJTU), China. He received his PhD degree in Political Science at the University of Chicago. He specializes in American politics and empirical methodology. His research is published at top journals in political science, including *Journal of Politics*, *World Politics*, *Political Research Quarterly*, and *Presidential Studies Quarterly*. He teaches courses on Introduction to American Politics, Causal Inference, and Social Science Inquiry, and was awarded the 1st Prize in the Young Faculty Teaching Competition at SJTU.

**FACILITATORS**

Jingjing (Krystal) LI

Junliu (Willow) JIANG

Sida (Star) ZHANG

Han (Alex) BAI

Zetong (Zet) ZHANG

Tingting (Ella) XU

To ensure the skills transferred in the workshop is consolidated, the participants will be paired with a facilitator during the workshop, to follow up on the creation of better data visualisations. Each facilitator is expected to support the participant in solving analytical or coding problems, ensuring each participant can produce the targeted report or presentation.

**OBJECTIVE**

The objective of this three-day workshop is to improve the ability of statisticians to convey gender-related content through user-focused visualizations. Participants will learn how to approach a gender-related social scientific claim, measure it by data, visualize the data, and evaluate gender-related social phenomena. The goal of this workshop is to increase your understanding of the most common statistical concepts and to work on statistical software programming and data visualization.

By the end of the workshop, participants will 1) comprehend basic statistical concepts, 2) be able to do some programming and analyze social science phenomena with empirical data, and 3) be able to visualize gender-related data using R packages.

**SOFTWARE AND READING MATERIALS**

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 download R, go to CRAN (the Comprehensive R Archive Network) at this site, <https://cran.r-project.org/>.

To interact with R, we use RStudio. Please install the latest desktop version of RStudio. You can download the latest version of RStudio at this site, <https://rstudio.com/products/rstudio/download/>.

You might find helpful the following resources, which are additional readings for further learning.

* *R for Data Science* by Hadley Wickham and Garrett Grolemund, which is in open access at <https://r4ds.had.co.nz/>.
* *ggplot2: Elegant Graphics for Data Analysis*, by Hadley Wickham, which is available at <https://ggplot2-book.org/index.html>.

**WORKSHOP OUTLINE**

DAY 1 – Monday 12 August

**Session 1.a Course Overview (9:30–10:45)**

The workshop kicks off with a course overview, which introduces the outline of the workshop and includes an ice-breaking activity. Then, the instructor takes participants to get to know R and showcase how statistical software can significantly save time for repetitive activities.

**Session 1.b Programming 101 – Hands-on session (11:00–12:30)**

This session introduces (or reviews for experienced statisticians) the basic syntax in R and take participants to get familiar with the software in a hands-on manner.

**Session 2.a Reproducible Data Analysis (13:30–15:00)**

This session demonstrates how to work with data, clean the data in a tidy format, in preparation to conducting reproducible data analysis.

**Session 2.b Reproducible Data Analysis – Hands-on session (15:15–16:45)**

Participants will work in groups to test and practice data loading, variable recoding, and producing same data files for multiple regions/provinces/districts using the Rwanda census data.

DAY 2 – Tuesday 13 August

**Session 3.a Data Visualization (9:15–10:45)**

This session introduces data visualization tools in R (ggplot2) that can help statisticians and policy officers to produce informative and flexible visual presentation of data and statistics.

**Session 3.b Data Visualization – Hands-on Session (11:00–12:30)**

Participants will work in groups to test and practice data visualization skills introduced in the previous session. Particularly, facilitators will work with participants in groups to generate gender-related figures using the Rwanda census data, which has been cleaned in the previous day.

**Session 4.a Communicating with Factsheet and Infographic (13:30–15:00)**

Extending on the topic of visualization, this session will cover the basic concepts and principles of creating statistical factsheets and infographics for better data communication. Some useful geodata packages will be introduced for demonstrating data on the map of Rwanda.

**Session 4.b Communicating with Factsheet and Infographic – Hands-on Session (15:15–16:45)**

Participants will be split into groups of two or three. Each group needs to come up with a plan to a statistical product in the coming days, and start to generating figures or infographic using the tools learned in the previous sessions.

DAY 3 – Wednesday 14 August

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

**infographics (9:15–10:45)**

Training facilitators will be paired with a group to create the data factsheets and infographics together from raw data.

**Session 5 (continued) Hands-on Exercises – Creating data factsheets and**

**infographics (11:00–12:30)**

Same as above.

**Session 6 Results presentation and discussion (13:30–15:00)**

After the statistical product is produced, every group will present their data products followed by a discussion with training facilitators and other groups.

**Session 6 (continued) Results presentation and discussion (15:15–16:45)**

Same as above.

Conclusions of the workshop and next steps.