

# Week 03

## Intro to Reproducible Research

*API209: Summer Math Camp*

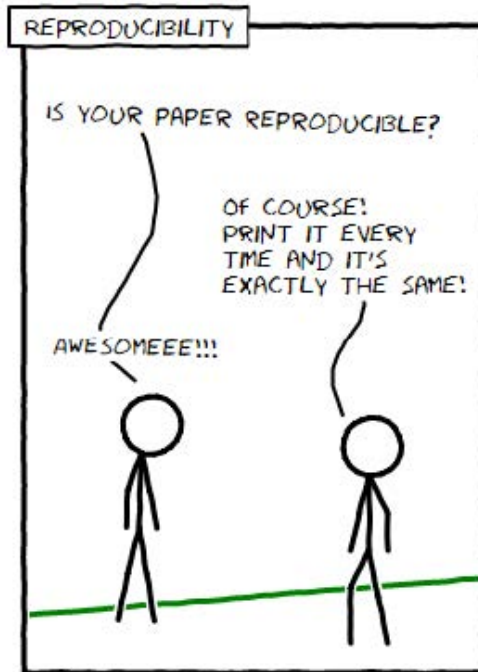
Rony Rodrigo Maximiliano Rodriguez-Ramirez

[rrodriguezramirez@g.harvard.edu](mailto:rrodriguezramirez@g.harvard.edu)

*Harvard University*

August 26, 2024

# Reproducibility



# Why Reproducibility?

## *Are We in a Crisis?*

- The **replication crisis** in social sciences has highlighted significant issues in the credibility of research findings.
- Many high-profile studies have failed to replicate, raising concerns about the **reliability** of published results.
- The crisis has prompted a call for greater **transparency** and **rigor** in research practices.

# The Replication Crisis

## *What Went Wrong?*

- **Selective Reporting:** Only significant findings get published, leading to publication bias.
- **P-Hacking:** Manipulating data and analyses until nonsignificant results become significant.
- **Lack of Transparency:** Opaque methodologies that others cannot replicate or verify.

# The Importance of Reproducibility

## *Building Trust in Research*

- Reproducibility ensures that research findings are not just a result of **chance** or **specific conditions**.
- It allows others to **verify results** and build upon them, fostering cumulative knowledge.
- **Transparent reporting** of data and methods strengthens the credibility and utility of research.

# How Can We Improve Reproducibility?

## *Adopting Best Practices*

- **Pre-registration:** Outlining the study design and analysis plan before data collection.
- **Open Data and Code:** Sharing data and analysis scripts for others to verify and use.
- **Reproducible Workflows:** Using tools like Quarto to create dynamic documents that combine analysis and narrative.

# Reproducibility: The Basics



# Replicability: Expanding the Horizon

# Reproducibility vs. Replicability

- **Reproducibility:**
  - duplication with the same data and procedures;
  - ensuring accuracy and precision.
- **Replicability:**
  - tests the findings using new data *but the same methods*;
  - emphasizing robustness and generalization.

Both concepts are crucial for ensuring the credibility and reliability of research, but they serve different purposes within the scientific process.

**enter Quarto**

# Quarto: A Tool for Reproducible Research

## *What is Quarto?*

Quarto is an open-source scientific and technical publishing system that enables researchers to create dynamic documents, reports, presentations, and websites.

# Why Quarto?

## *The Need for Reproducible Research*

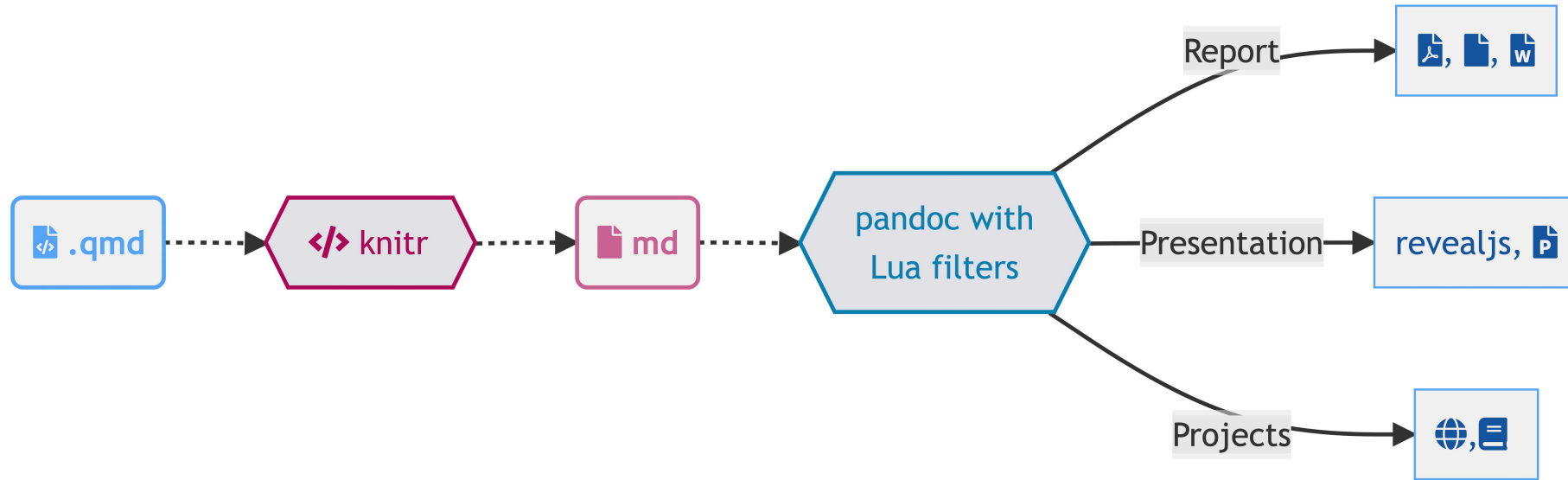
- Quarto ensures that your analysis and outputs (tables, figures, etc.) can be reproduced by others, enhancing the credibility of your work.
- **Integrated with R, Python, Julia, etc.:** Quarto supports multiple languages, making it versatile for various research needs.

# Why Quarto?

HOW STANDARDS PROLIFERATE:  
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)



# Quarto for literate programming



# Key Features of Quarto

1. **Dynamic Documents:** Create documents that are automatically updated with the latest data and analysis.
2. **Multiple Outputs:** Generate reports, presentations, blogs, and books from a single source.
3. **Version Control:** Integrates seamlessly with Git for version control, tracking changes, and collaboration.
4. **Cross-Platform:** Works with RStudio, VSCode, or directly from the command line.

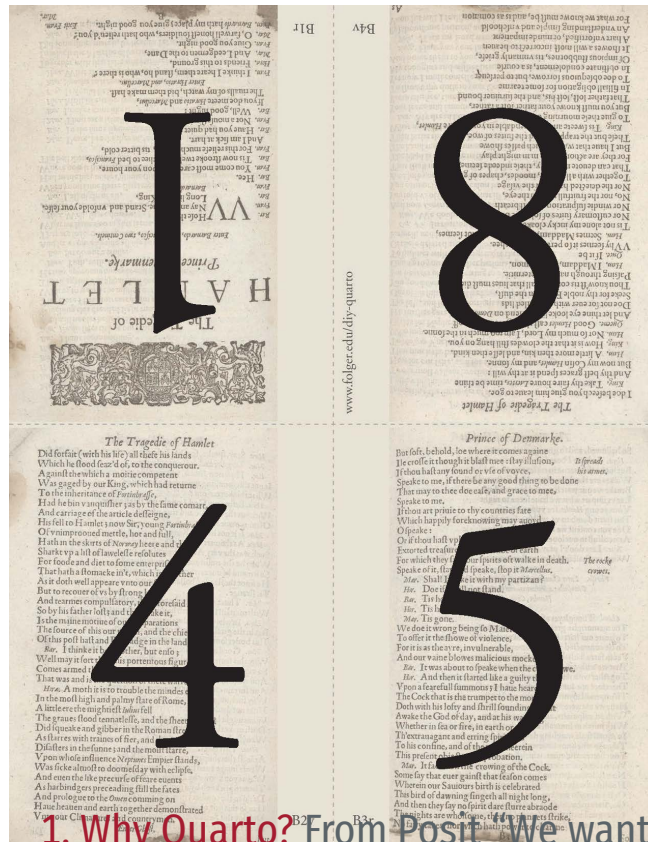


# Why Use Quarto for Your Problem Sets?

## *Consistency and Organization*

- Quarto helps you **organize your code, analysis, and narrative** in a single document.
- It ensures that your problem sets are **well-documented** and **easily understandable**.

# Why the name “Quarto”?<sup>1</sup>



1. Why Quarto? From Posit: We wanted to use a name that had meaning in the history of publishing and landed on Quarto, which is the format of a book or pamphlet produced from full sheets printed with eight pages of text, four to a side, then folded twice to produce four leaves. The earliest known European printed book is a Quarto, the *Sibyllenbuch*, believed to have been printed by *Johannes Gutenberg* in 1452–53.”