Reproducible Research with R

Statistical Analysis with R using RStudio, GitHub, knitr and Shiny

Maja Založnik

Oxford – 19^{th} March 2015

Outline

- **Introduction**
- **Reproducibility**
- **RStudio**
- **GitHub**
- **knitr**
- **RPubs**
- **Shiny**

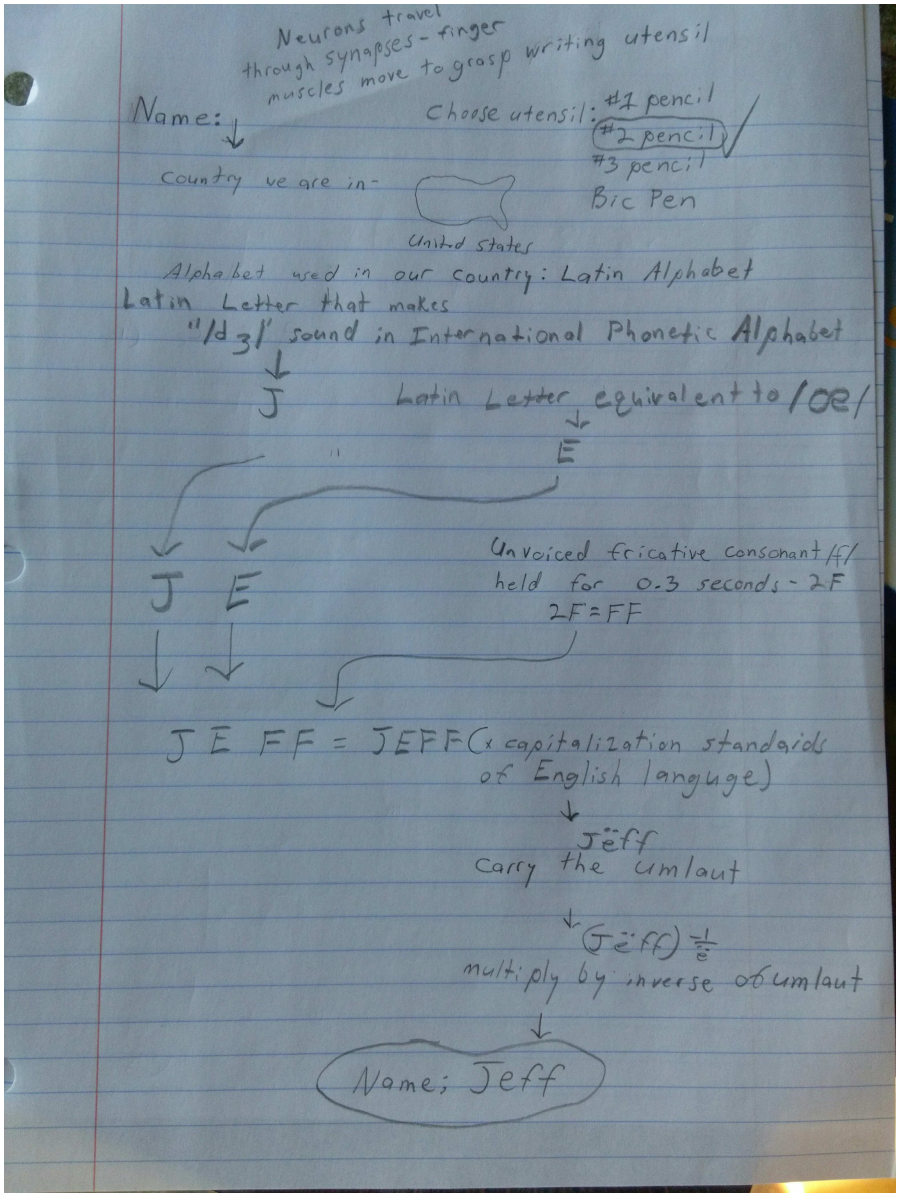
Reproducibility of Research

- **Reproducibility vs Replicability of research?**
- *"The confirmation of results and conclusions from one study obtained independently in another"* (Jasny et al. 2011)
- *"[T]he independent verification of prior findings"* (Santer et al. 2011)
- **Levels of Replication**
    1. Re-ask the question
    2. Re-do the experiment
    3. Re-analyse the data
    4. **Reproduce the analysis**

Reproducibility: Show your work!

imgur user **TVsJeff**:

"A math teacher took points off for not showing all of my work. The next homework assignment i turned in looked like this. It was 45 pages long."



RStudio

- Probably the most popular IDE for R
- Launched February 2011
- January 2012 - Project system and Version control integration (git/SVN)
- May 2012 - knitr & R Markdown publishing tools added
- June 2012 - publish to RPubs integration
- December 2013 - Shiny integration
- October 2014 - direct publishing to shinyapps.io

Version Control: Git

git/GitHub for Reproducible research

- Full documentation
- Collaboration
- Dissemination
- Backup
- RStudio integration
- GitHub - the Facebook of code
- But click here for five free private repos!

Literate programming and knitr

"Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on explaining to human beings what we want a computer to do."

Donald Knuth (1984)

Literate programming and knitr

- Human readable
  - Pure code: WHAT & HOW but not WHY
  - Pure text: WHAT & WHY but not HOW
- Consistent coding style e.g.:
  - Google style guide
  - Hadley Wickham's style guide
- Commenting
- knitting

List of Helpful Links and Free Resources

- Christopher Gandrud's *Reproducible Research with R and RStudio*
- Coursera Data Scientice Specialisation
- GitHub & academic discount link
- R markdown and knitr resources
- RPubs
- Shiny tutorial
- This presentation on github