



Reproducible Research Overview

Jeffrey Leek
Johns Hopkins Bloomberg School of Public Health

Reproducible Research Content

- Structure of a Data Analysis
- Organizing a Data Analysis
- Markdown
- LaTeX
- R Markdown
- Evidence-based data analysis
- RPubS

Steps in a data analysis

- Define the question
- Define the ideal data set
- Determine what data you can access
- Obtain the data
- Clean the data
- Exploratory data analysis
- Statistical prediction/modeling
- Interpret results
- Challenge results
- Synthesize/write up results
- Create reproducible code

Data analysis files

- Data
 - Raw data
 - Processed data
- Figures
 - Exploratory figures
 - Final figures
- R code
 - Raw scripts
 - Final scripts
 - R Markdown files (optional)
- Text
 - Readme files
 - Text of analysis