

# Basic Features of the Case Studies

**Daniel Turek**

# Case Study Template

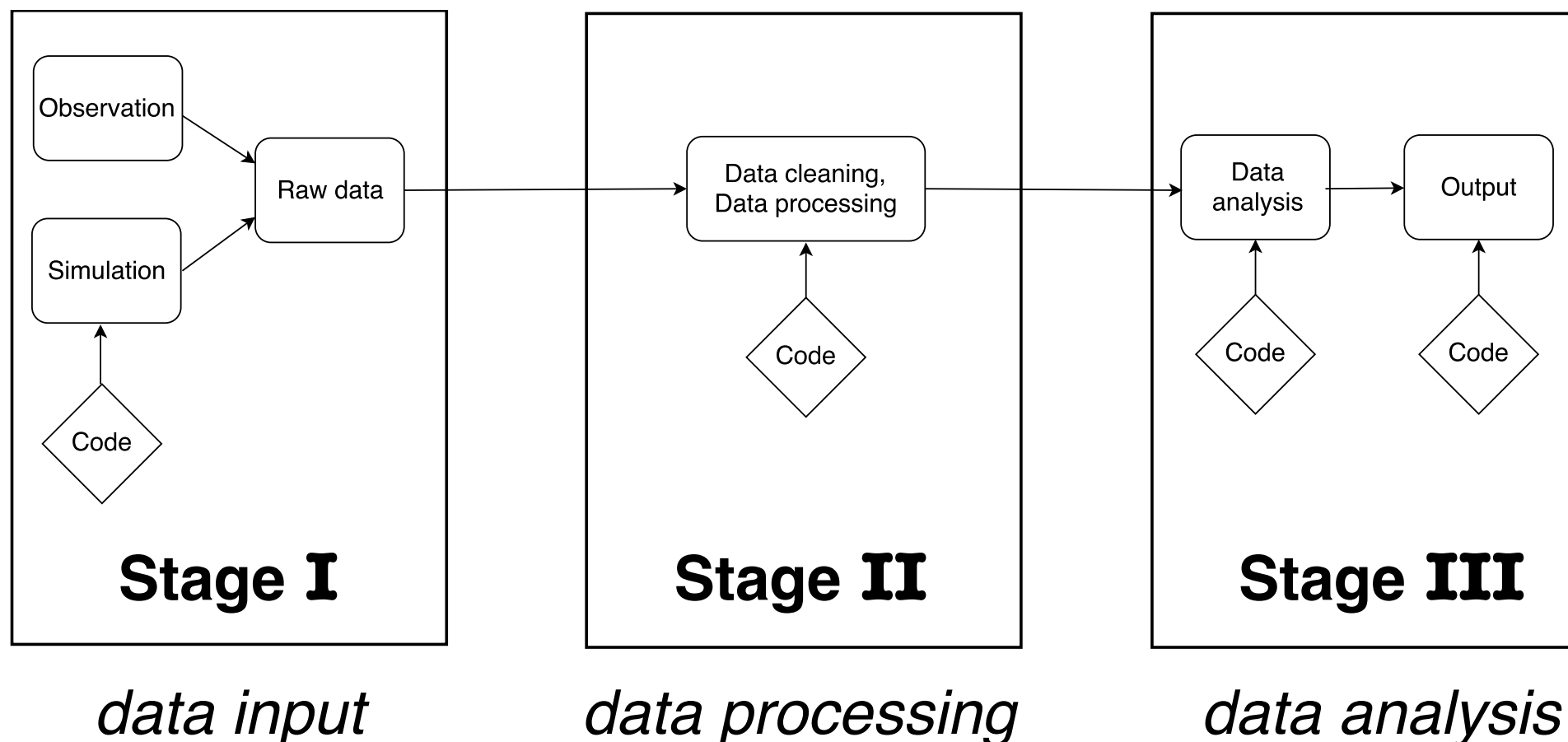
- Introduction
  - Discipline, research context and background
- Core Sections
  - Workflow Diagram
  - Workflow Narrative
- Pain Points
- Key Benefits
- Key Tools
- General questions about Reproducibility
  - e.g.* why is reproducibility important in your domain?

# Stages & Classification



- Collected 31 case studies
  - Partner data science initiatives
  - BIDS Reproducibility Workshop (May 2015)

- Stages:



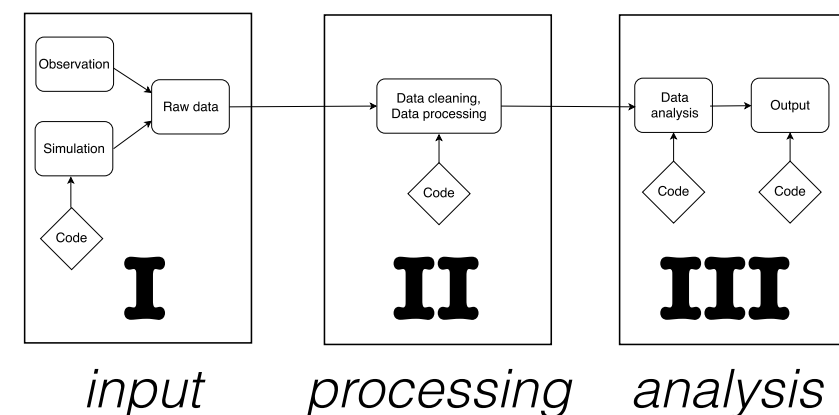
# Stages & Classification



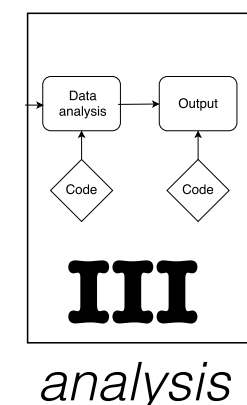
- Collected 31 case studies
  - Partner data science initiatives
  - BIDS Reproducibility Workshop (May 2015)

- Stage Classification:

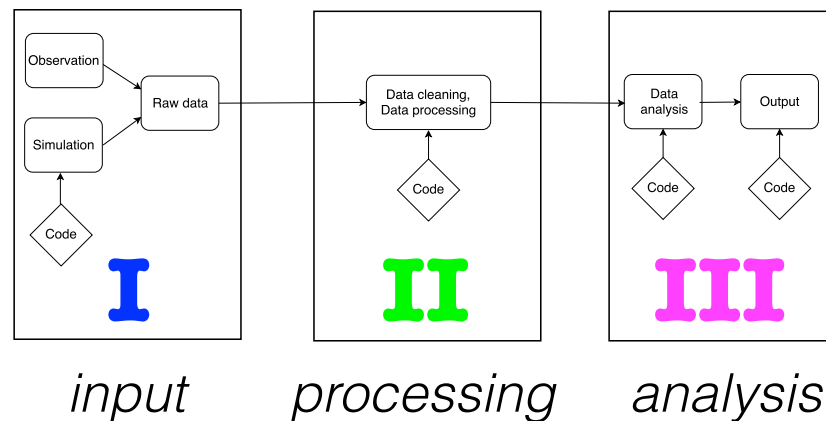
- "High-Level" case studies
  - stages **I, II, III**



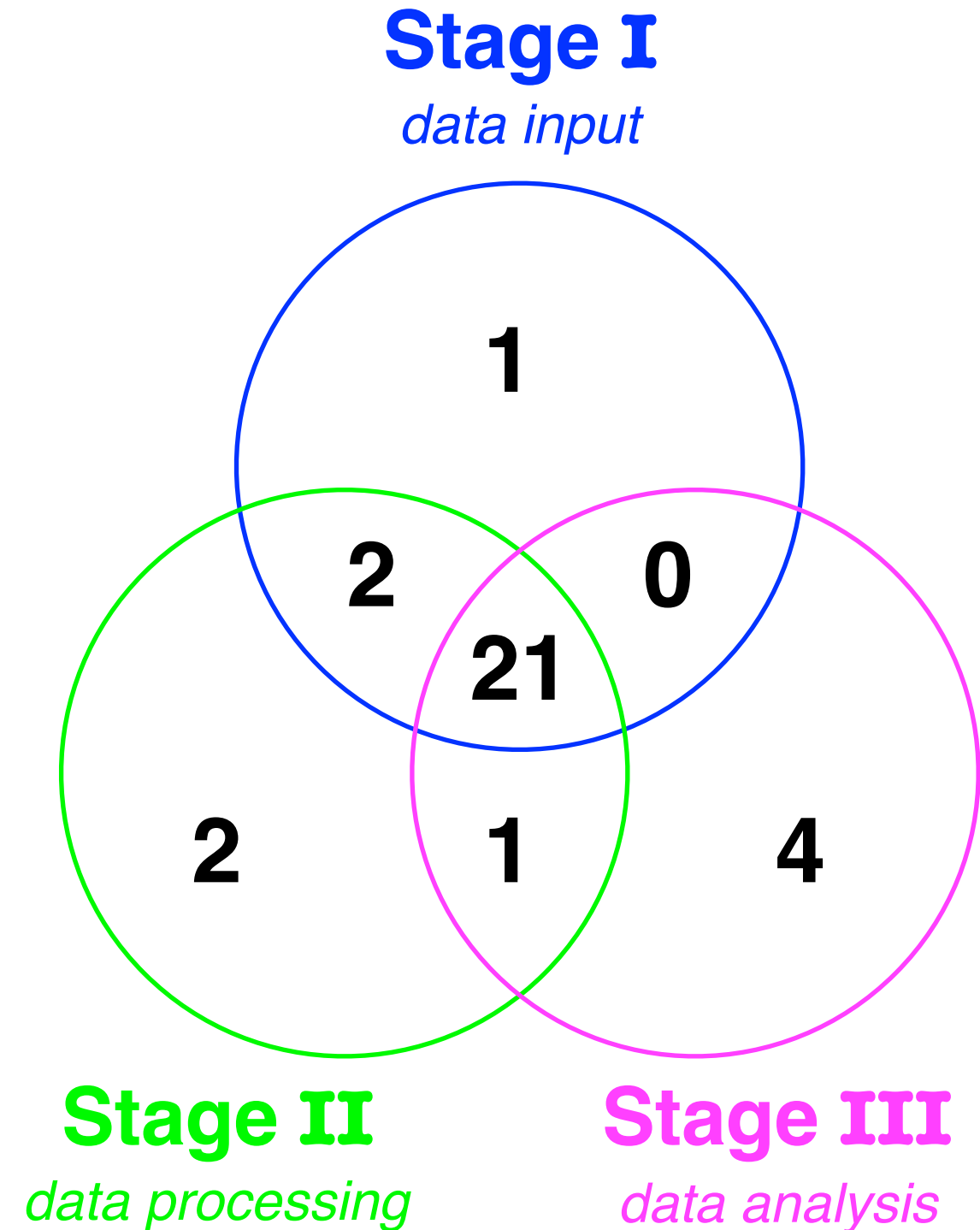
- "Low-Level" case studies
  - *e.g.*, stage **III**



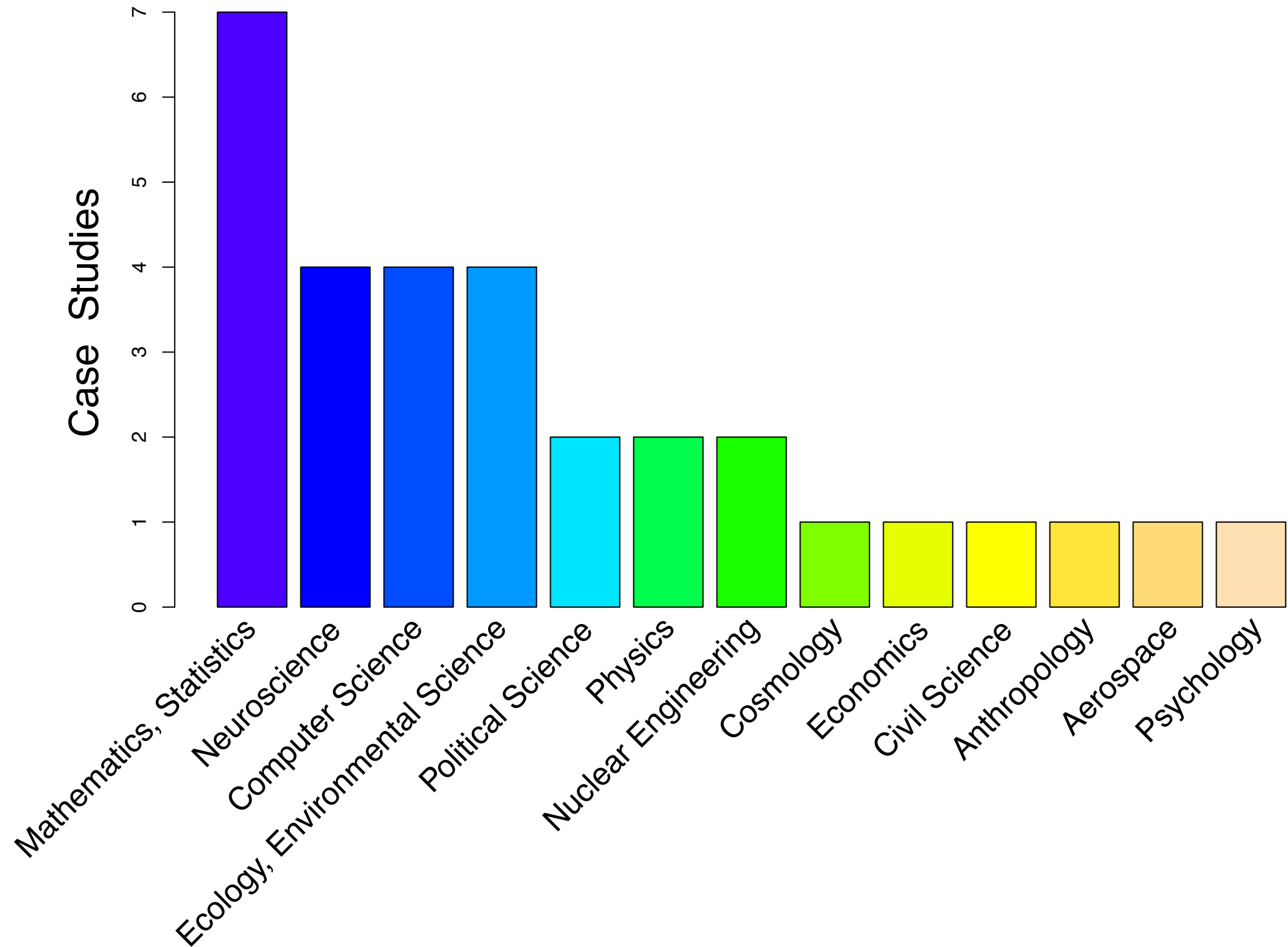
# Stage Breakdown



- High-Level case studies (21)  
(stages **I, II, III**)
- Low-Level case studies (10)  
(e.g., stage **III**)

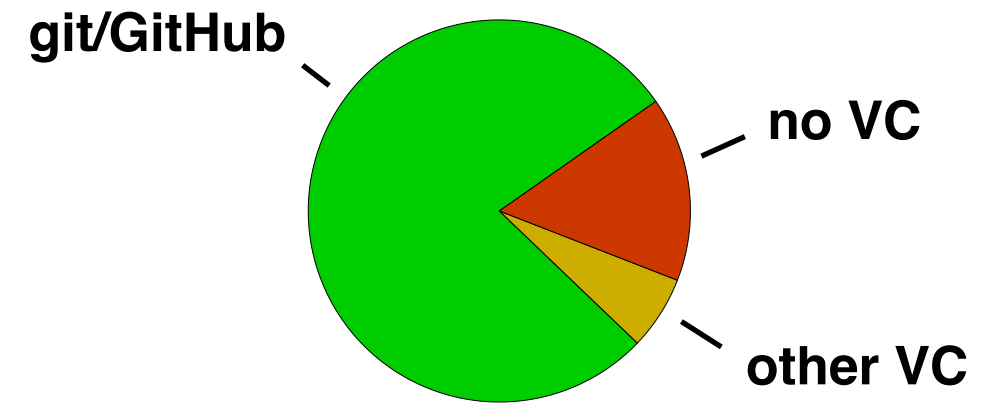


# Discipline



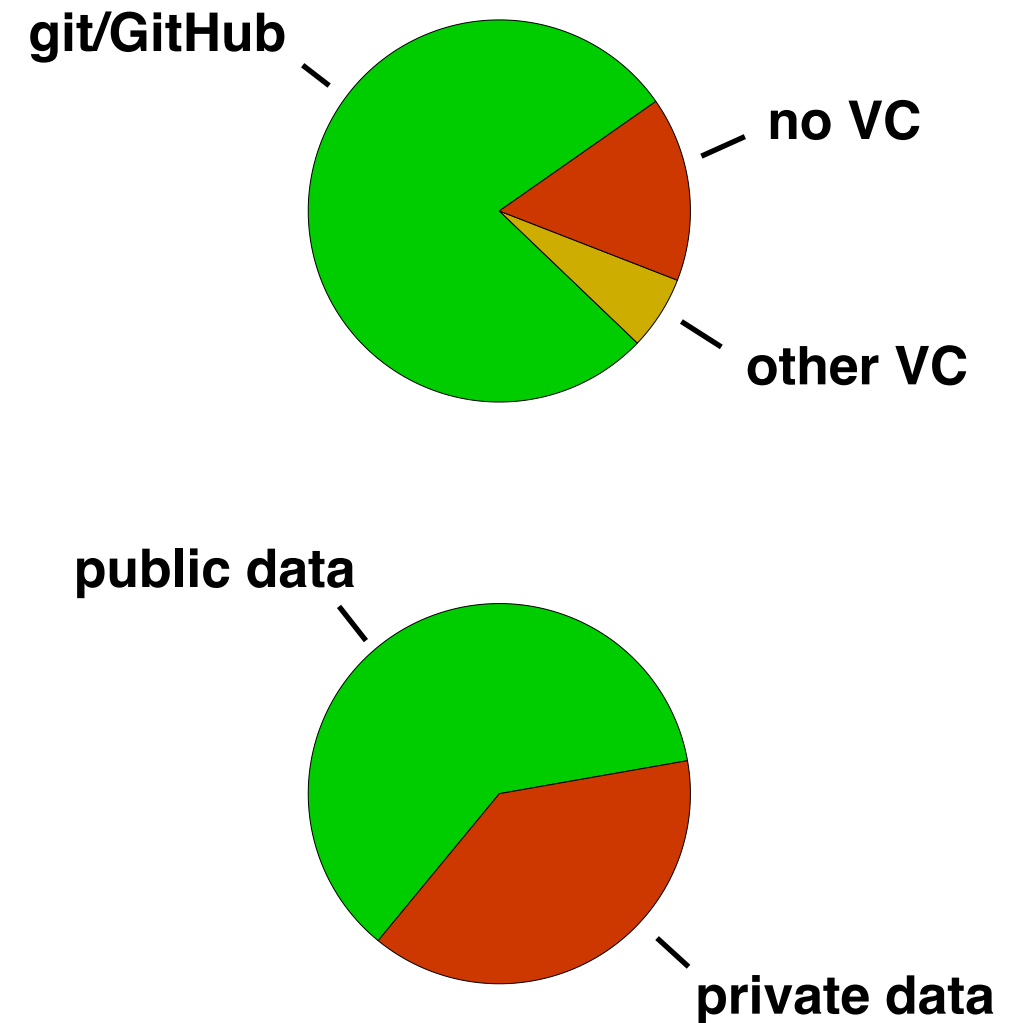
# Case Study Trends

- 80% using git and/or GitHub for version control



# Case Study Trends

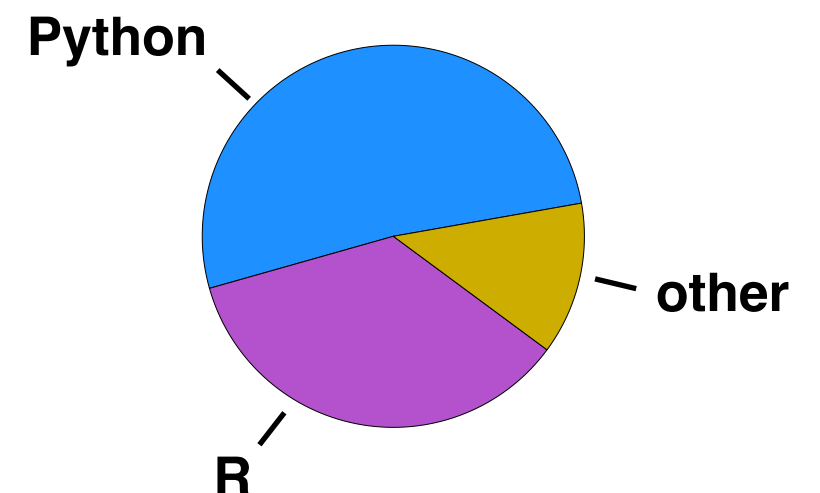
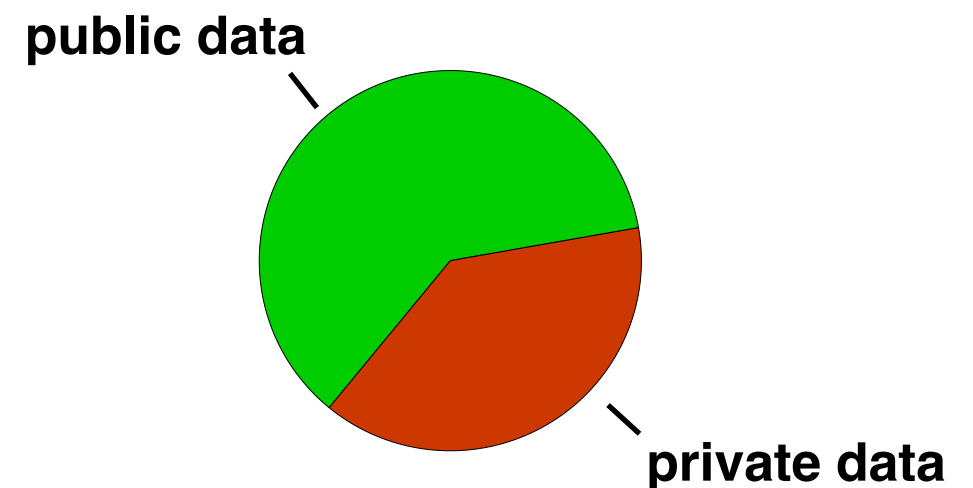
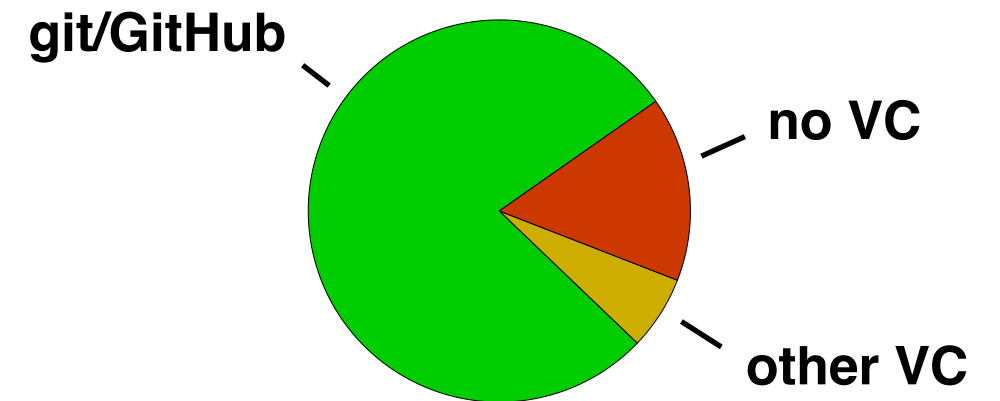
- 80% using git and/or GitHub for version control
- 60% used open data, or made their data publicly available





# Case Study Trends

- 80% using git and/or GitHub for version control
- 60% used open data, or made their data publicly available
- Primary language
  - Python (52%)
  - R (35%)
  - other (13%)



# Case Study Trends

- 80% using git and/or GitHub for version control
- 60% used open data, or made their data publicly available
- Primary language
  - Python (52%)
  - R (35%)
  - other (13%)
- Output
  - Scientific publication
  - Software
  - Workflow

