



**ID529!**

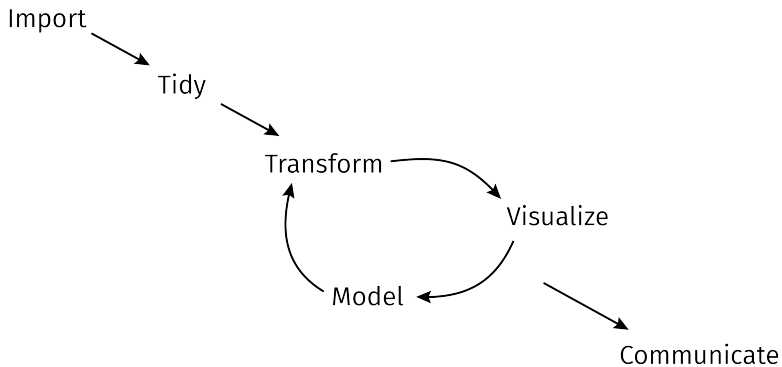
Data Management and  
Analytic Workflows in R  
January 2025



**HARVARD**  
**T.H. CHAN**

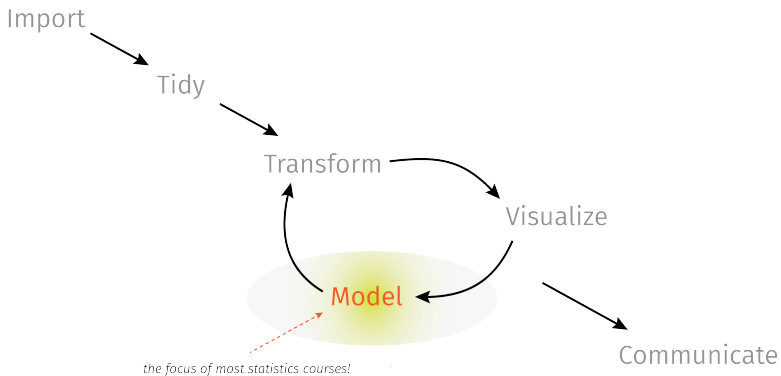
**SCHOOL OF PUBLIC HEALTH**

Why this course?



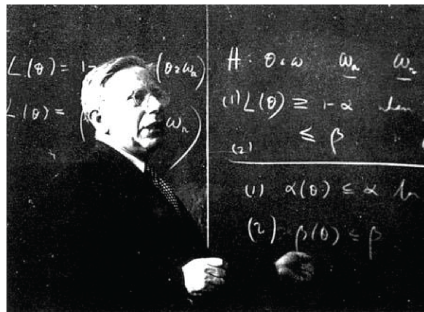
---

Graphic adapted from *Zen And The aRt Of Workflow Maintenance* ([linked](#)) by Jennifer Bryan, RStudio, University of British Columbia



---

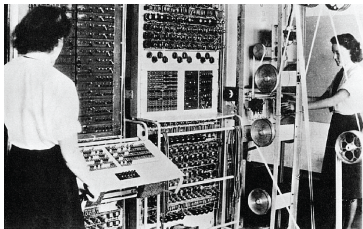
Graphic adapted from [Zen And The aRt Of Workflow Maintenance](#) (linked) by Jennifer Bryan, RStudio, University of British Columbia



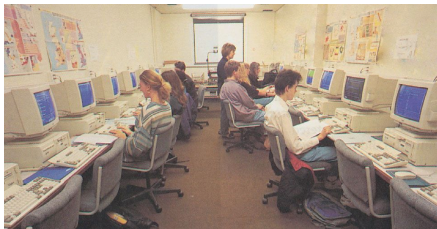
Abraham Wald



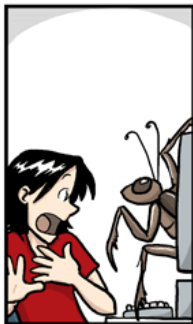
Computing division of the US Veterans Bureau, 1924



Colossus Mark 2 codebreaking computer being programmed, Bletchley Park, 1943



Computer lab at the University of Surrey, 1990s



www.phdcomics.com

JORGE CHAM © 2005

Import

Tidy

ALL OF THESE ARE SKILLS THAT CAN BE LEARNED!  
...AND BENEFIT FROM PRACTICE

Transform

Visualize

Model

Communicate





# Import

readr  
readxl  
haven  
xml2

# Tidy → Transform

tibble  
tidyr

dplyr  
forcats  
hms

lubridate  
stringr

# Visualise

ggplot2

rsample

tidyposterior  
yardstick

# Model

broom  
modelr

recipes

purrr  
magrittr

# Program



[tidyverse.org](https://tidyverse.org)



[r4ds.had.co.nz](https://r4ds.had.co.nz)



**Trevor A. Branch**

@TrevorABranch

Follow



My rule of thumb: every analysis you do on a dataset will have to be redone 10–15 times before publication. Plan accordingly. #Rstats

Import

Tidy

ALL OF THESE ARE SKILLS THAT CAN BE LEARNED!  
...AND BENEFIT FROM PRACTICE



Transform

Visualize

Model

Communicate

DOCUMENT  
TESTING  
VERSION CONTROL  
COLLABORATION

# Generative AI in Coding Education

- **Benefits**
  - Rapid prototyping: quickly generate code snippets for common tasks
  - Inspiration for new approaches and efficient workflow patterns
  - Builds confidence by offering guidance for troubleshooting errors
- **Potential Pitfalls**
  - Overreliance on AI output can hinder skill development
  - Possible inaccuracies ('hallucinations') that require human verification
  - Lack of transparency in complex data or modeling strategies
- **Responsible Usage**
  - Treat AI-generated code as a *draft*, always verify correctness
  - Cross-check results with documentation and your own tests
  - Use AI for learning concepts rather than copy-paste solutions
- **Paradigm Shift in Coding Education**
  - Emphasis shifting from pure syntax memorization to critical evaluation
  - Students learn to collaborate with AI as a coding partner
  - Builds meta-skills: choosing, refining, and validating AI suggestions

---

Disclosure: I asked ChatGPT o1 to make this list!



We've assembled a great instructional team to develop and deliver this course:

Christian Testa

Dean Marengi

Jarvis Chen

The course has also benefitted greatly from the input of:

Amanda Hernandez (TF in January 2023)

ID529 students in 2023 and 2024

We look forward this year to  
working with and learning from

**ALL OF YOU!**