



# ID529!

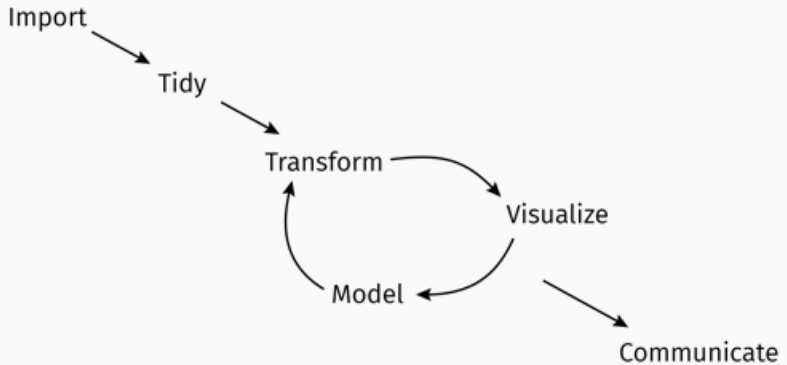
Data Management and  
Analytic Workflows in R  
January 2024



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

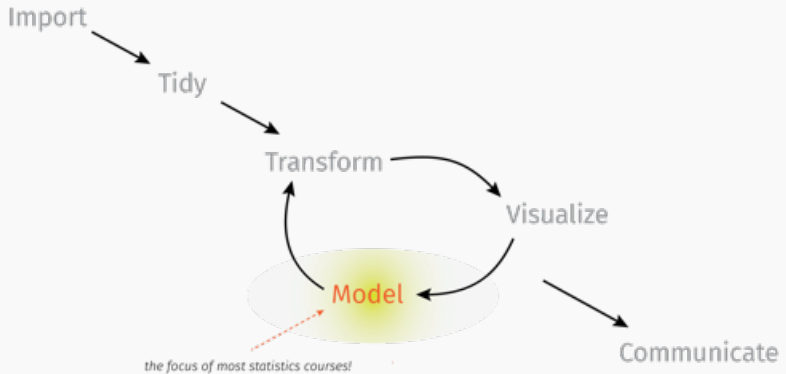
SCHOOL OF PUBLIC HEALTH

Why this course?



---

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



---

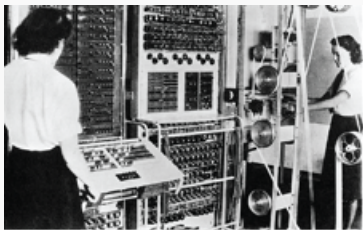
Graphic adapted from [Zen And The aRt Of Workflow Maintenance](#) (linked) by Jennifer Bryan, RStudio, University of British Columbia



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

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

tibble  
tidyr

## Transform

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

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)

Last year's ID529 students

We look forward this year to  
working with and learning from

ALL OF YOU!