Creating interactive shiny dashboards to showcase sociolinguistic research

Justin Davidson, Joseph Roy and Gyula Zsombok

What are dashboards?

- “A **data dashboard** is an information management tool that visually tracks, analyzes and displays key performance indicators (KPI), metrics and key data points to monitor the health of a business, department or specific process.”
  - https://www.klipfolio.com/resources/articles/what-is-data-dashboard

Dashboards for research?

- Dashboards are displays of data(visual and tabular)
  - Facilitate discussion of data applied to a specific set of research questions
  - Showcase a research program in an interactive and visual way

Why R?

- R/Rstudio environment is free and open-source
- There is a large community online and support for newer visualization methods and tools
- R skills can be applied outside of a dashboard building context (i.e. predictive modeling, automatic report generation, complex data wrangling)

Designing Dashboards*

- Update Frequency
  - Real Time
  - Daily/Hourly/Weekly
  - Monthly/Quarterly/semi-annually
  - Annually
- User Expertise

Novice  **Expert**

Adapted from: Few, Stephen. 2013. *Information Dashboard Design*. Analytics Press.

Designing Dashboards*

- Update Frequency
  - ~~Real Time~~
  - ~~Daily/Hourly/Weekly~~
  - Monthly/Quarterly/semi-annually
  - Annually
- User Expertise

This is doable in R, but I will not cover today. In many higher ed data contexts, you want to clean the data and then post it.

Novice

Expert

Adapted from: Few, Stephen. 2013. *Information Dashboard Design*. Analytics Press.

Designing Dashboards

- Audience (size)
  - One person
  - Multiple people (with the same requirements)
  - Multiple stakeholders (with different lenses on the data)
- Technology Platform
  - Desktop App
  - Shared Drive
  - Stand-alone Server

Designing Dashboards **in R**

- Update Frequency
  - ~~Real Time~~
  - ~~Daily/Hourly/Weekly~~
  - Monthly/Quarterly/semi-annually
  - Annually

Novice  **Expert**

- User Expertise

Adapted from: Few, Stephen. 2013. *Information Dashboard Design*. Analytics Press.

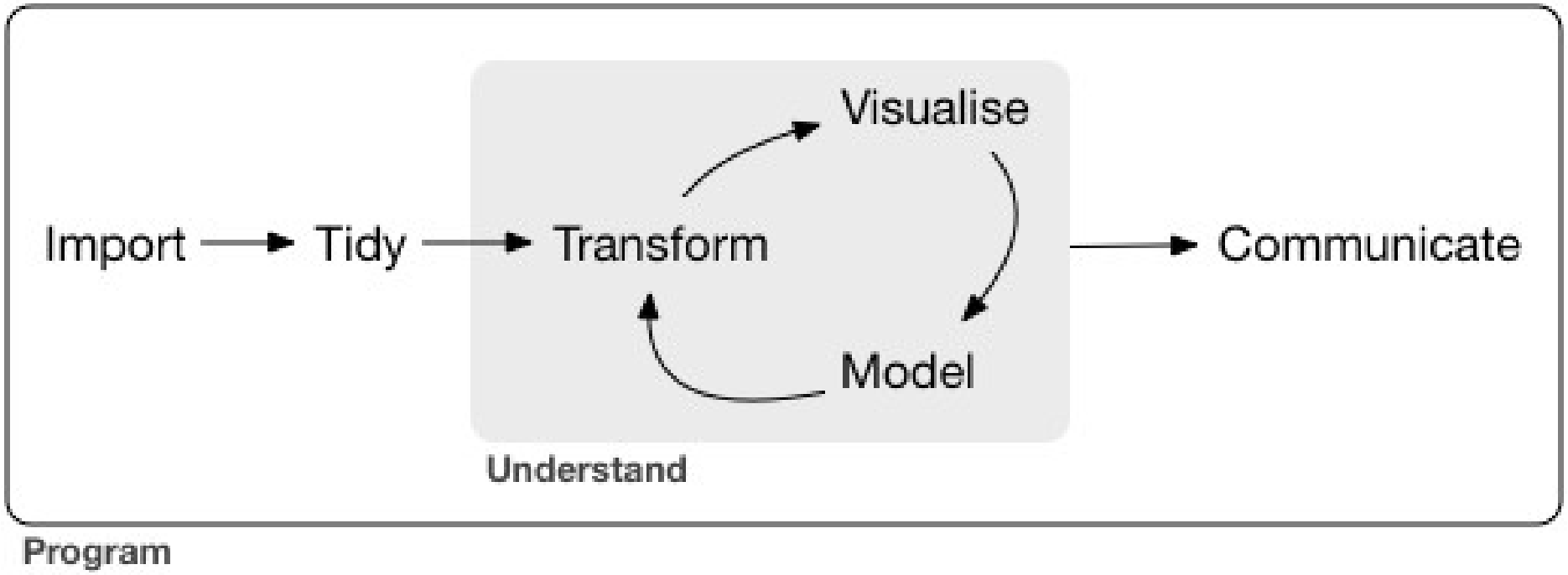
Designing Dashboards **in R**

- Audience (size)
  - One person
  - Multiple people (with same requirements)
  - Multiple stakeholders (with different lenses on the data)
- Technology Platform
  - Desktop App (Rstudio Project)
  - Shared Drive App (Rstudio Project)
  - Stand-alone Server
    - Internal/External Access [E.g. LAMP stack + shiny (free)/1-2k for server – ask your IT staff]
  - Shinyapp.io

Designing Dashboards **in R**

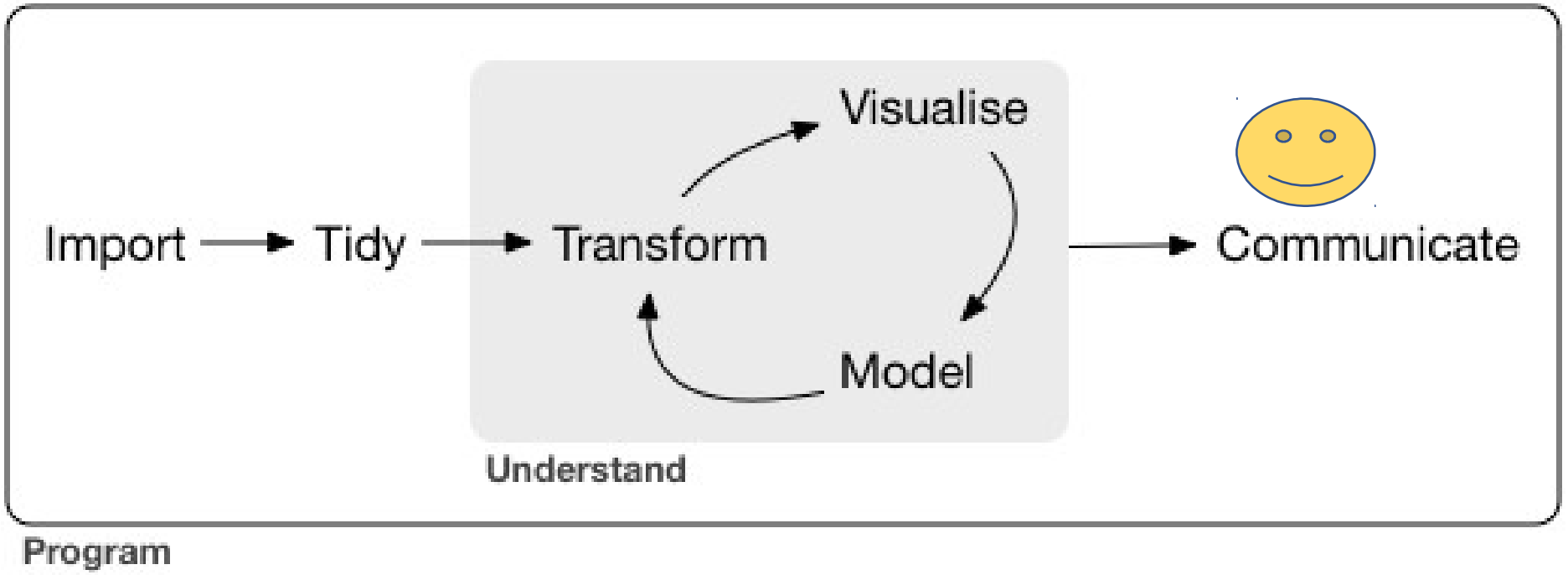
- **Data Wrangling**
  - **Tip**: Do not wrangle the data in the app
  - Producing a clean set of data for use in the app will help with it speed and prevent any issues running
  - Also increases compute time (and reduces utility of free accounts on shinyapps.io)
- **Data Modeling/Visualization**
  - **Tip**: Don't do data exploration/modeling/visualization during the dashboard design or after
- **Use an iterative/cumulative workflow**
  - What does this mean? Make your workflow reproducible, extendable and easily deployable
  - You should script as much as possible using techniques that don't require everything be hand-coded (e.g. **avoid things like setwd()**)

Model of Current Data Science



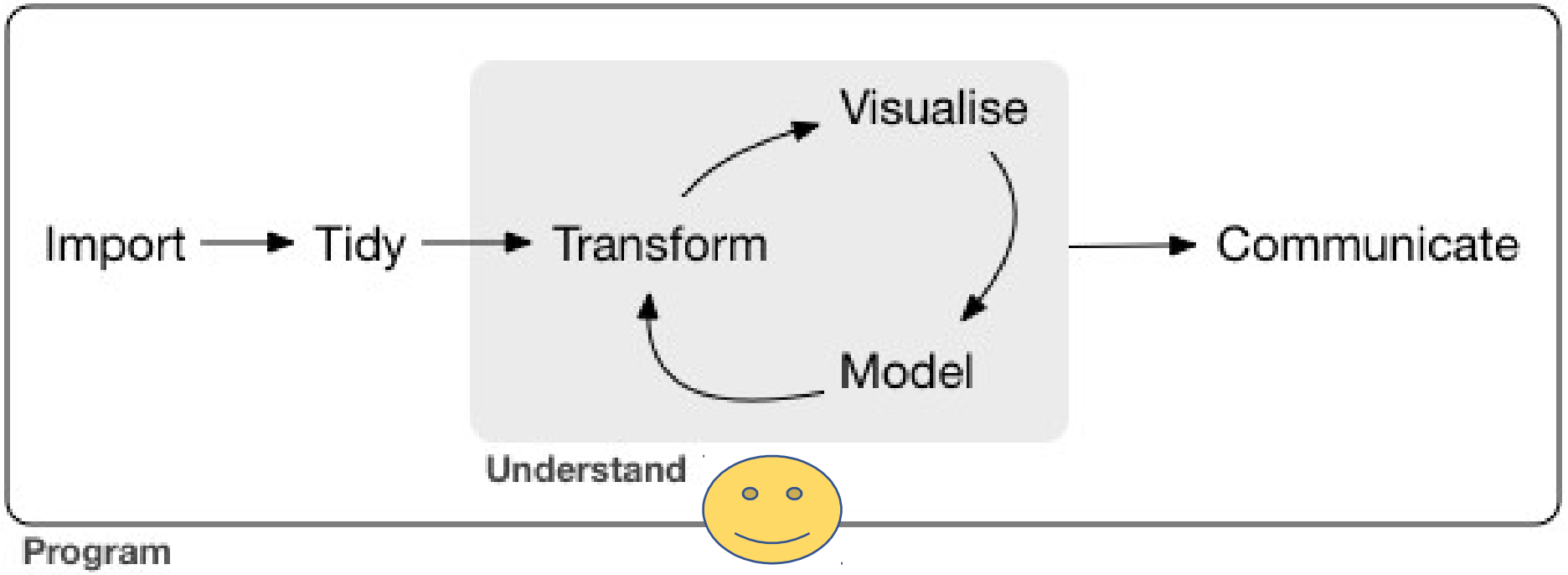
From Grolemund & Wickham, R for Data Science: https://r4ds.had.co.nz/introduction.html

Model of Current Data Science: **Dashboards**



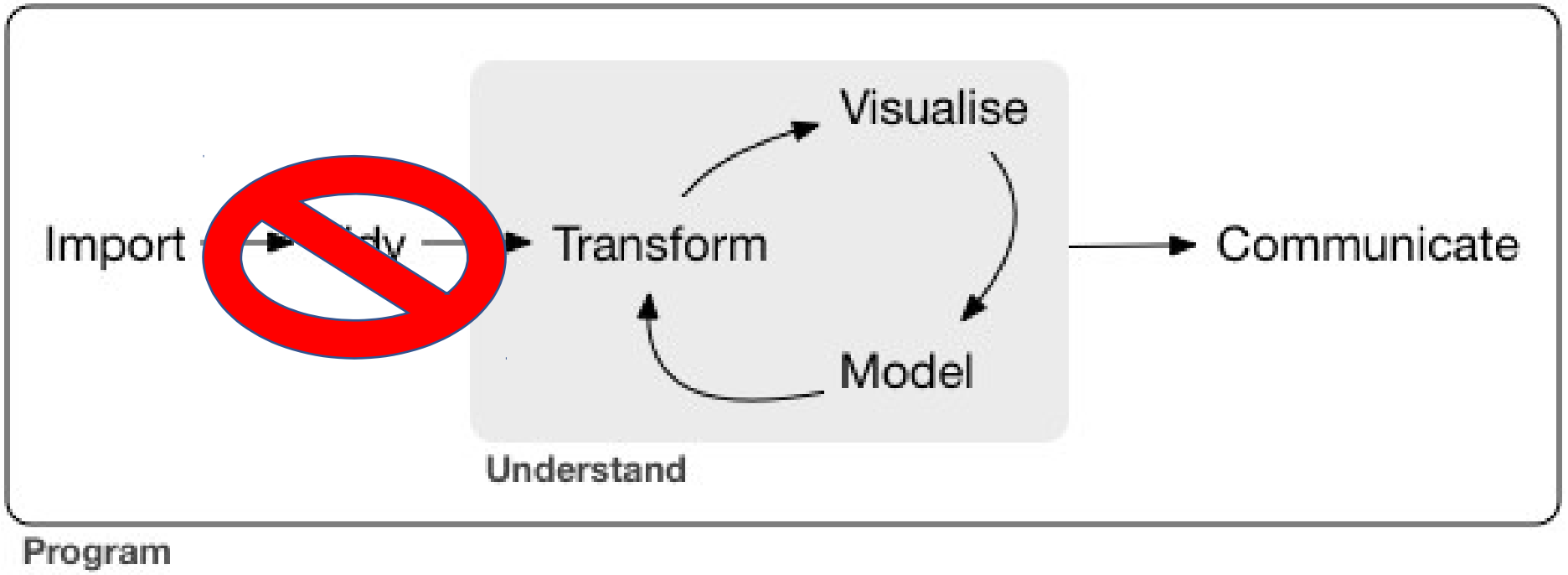
From Grolemund & Wickham, R for Data Science: https://r4ds.had.co.nz/introduction.html

Model of Current Data Science: **Dashboards**



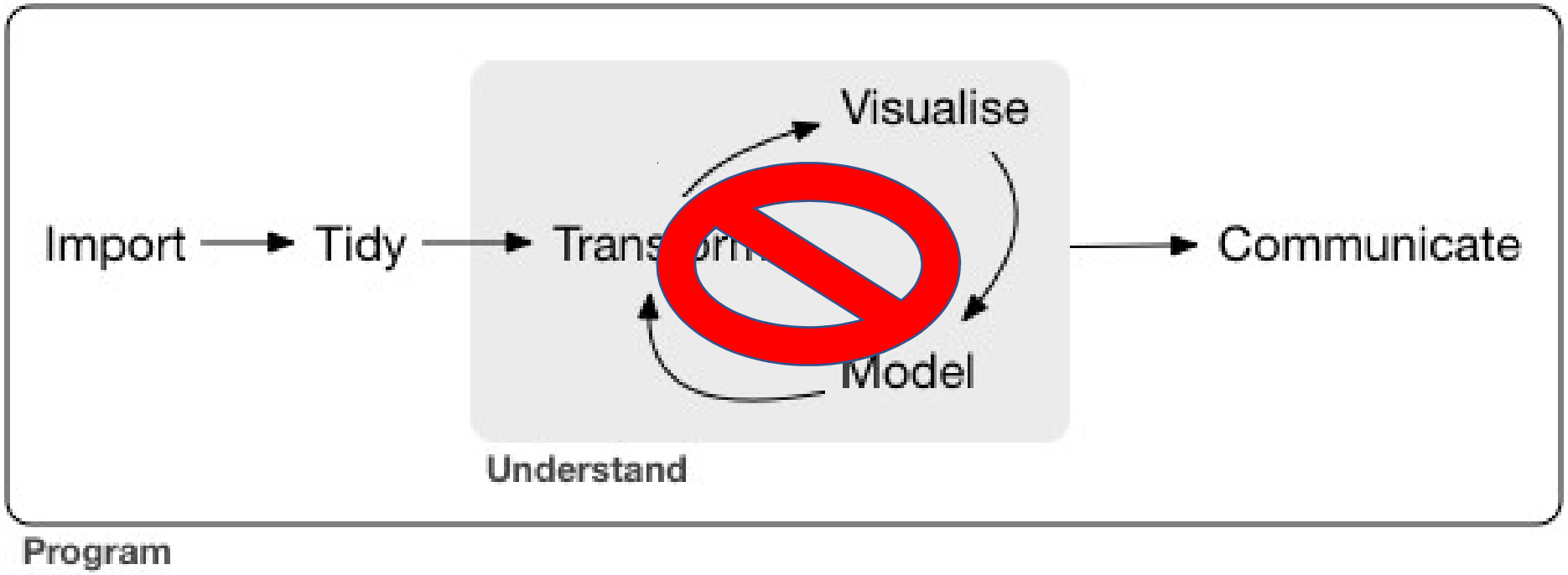
From Grolemund & Wickham, R for Data Science: https://r4ds.had.co.nz/introduction.html

Model of Current Data Science: **Dashboards**



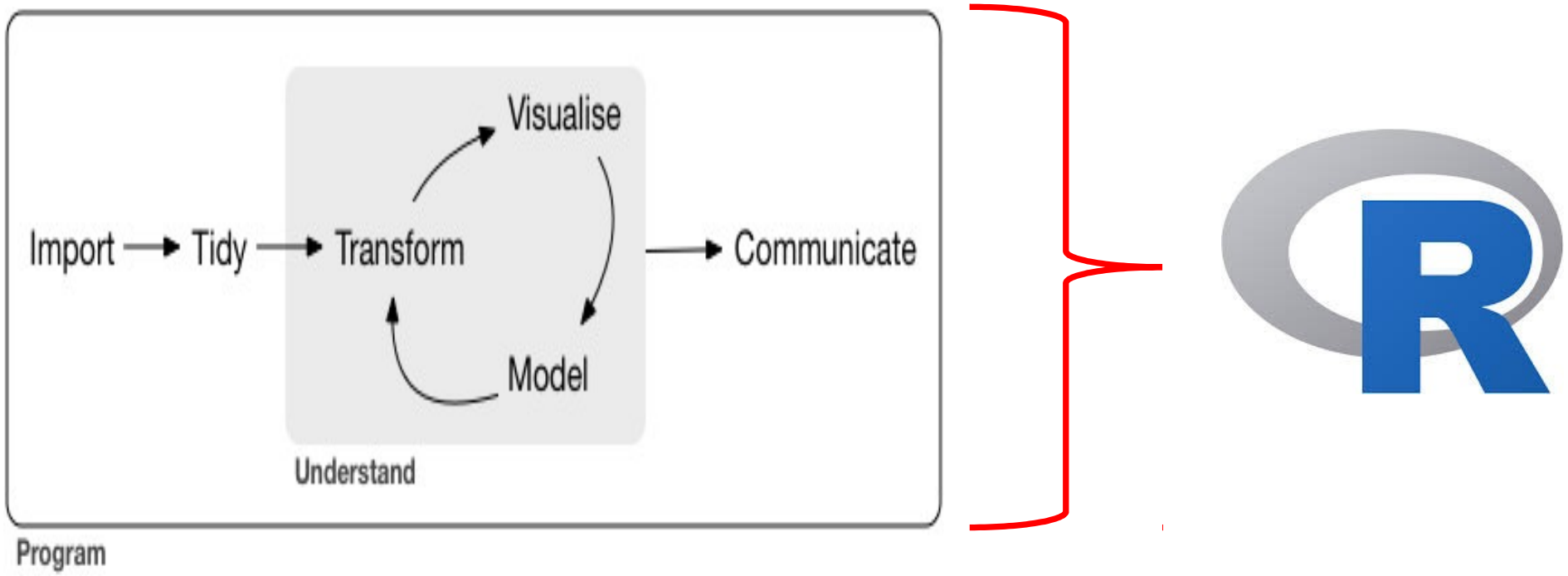
From Grolemund & Wickham, R for Data Science: https://r4ds.had.co.nz/introduction.html

Model of Current Data Science: **Dashboards**



From Grolemund & Wickham, R for Data Science: https://r4ds.had.co.nz/introduction.html

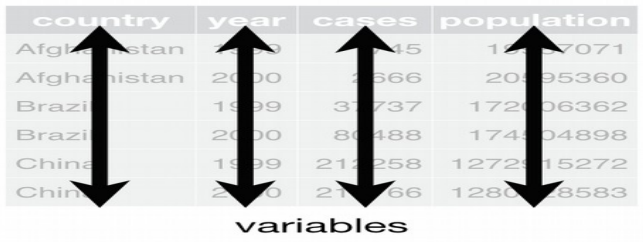
Model of Current Data Science: **Dashboards**

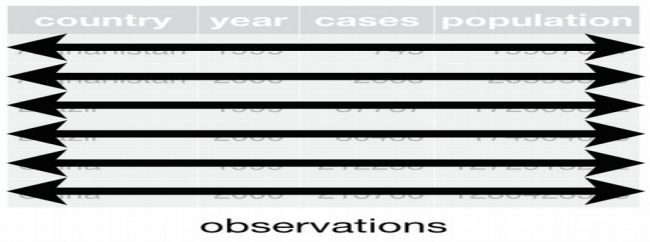


From Grolemund & Wickham, R for Data Science: https://r4ds.had.co.nz/introduction.html

Data Wrangling

- Get the data in a tidy format before putting into a dashboard
- *Following three rules makes a dataset tidy: variables are in columns, observations are in rows, and values are in cells. (from: https://r4ds.had.co.nz/tidy-data.html)*





Why tidy (or near tidy) data?

- Easier to manipulate for R programming language and to use visualization tools
- Easier to follow the logic of a script
- Collaborating with others (and your future self) becomes more straightforward

Visualization Resources in R

- R-Graph Gallery by Prof. Jenny Bryan (UBC):
  - http://shinyapps.stat.ubc.ca/r-graph-catalog/
  - Based on *Creating More Effective Graphs* by Naomi Robbins.
- Ggplot2 Tutorials:
  - https://ggplot2.tidyverse.org/
  - Cheatsheet: https://github.com/rstudio/cheatsheets/raw/master/data-visualization-2.1.pdf

What can be asked of our data?

- Visual pruning
  - What does the data look like when individual levels of a variable are separated out from one another?
  - What does the data look like when it is restricted to select IV levels?
  - What IV effects appear dependent on other IVs (i.e., interactions)?
  - What data points would benefit from being geo-tagged on a map?

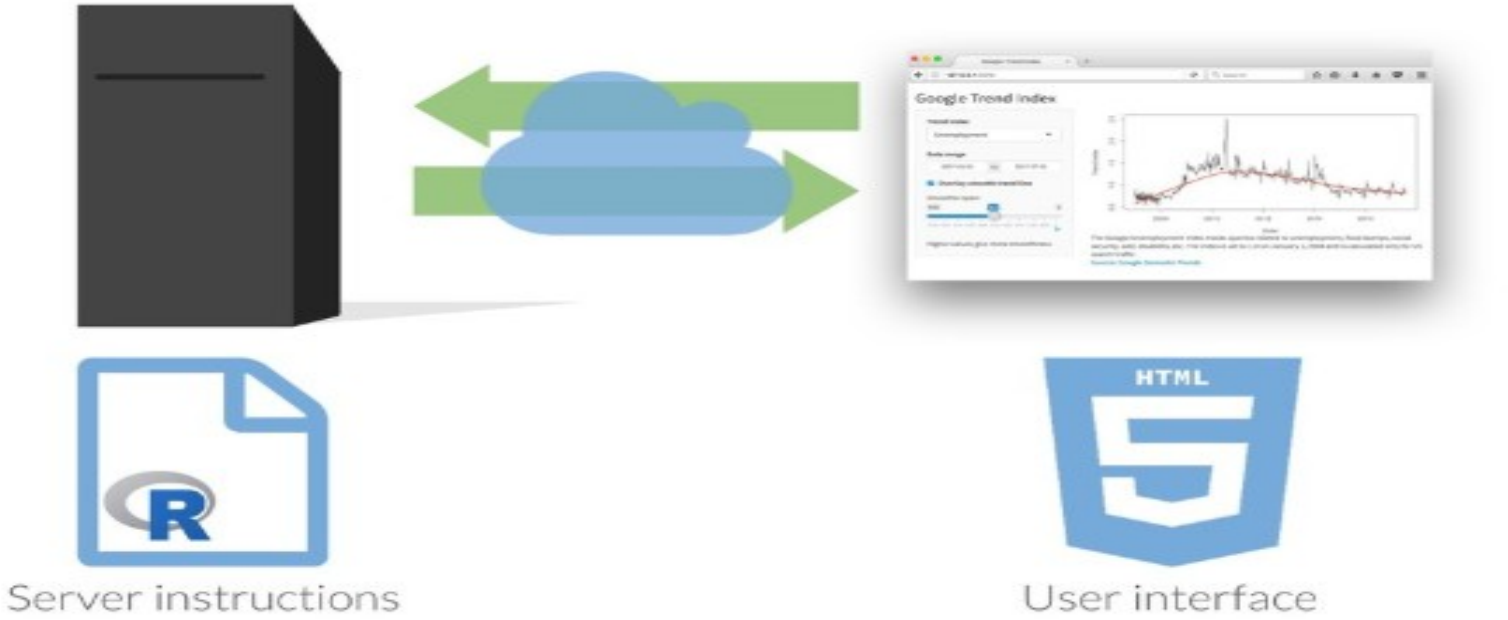
- Visual amplification
  - What IV relationships couldn’t be fully explored/visualized given the constrictions of a particular publication venue?
  - What data points would allow others to best engage with your research and even explore additional research questions?

What can be asked of our data?

- Visual exploration
  - What effects does the use of a series of bins have on the dataset?
  - What kinds of IV comparisons are best shown through select graph types?

- Who will most benefit from exploring/visualizing the data?
  - Local audiences: Run Shiny on your local machine
  - Open Source: Publish using ShinyApp.io

Shiny in R



https://shiny.rstudio.com/images/shiny-cheatsheet.pdf

Packages that extend shiny

- library(shinydashboard): https://rstudio.github.io/shinydashboard/
- library(flexdashboard): https://rmarkdown.rstudio.com/flexdashboard/

Open Rstudio!