Loreen Henry

EPPS 6356 Data Visualization

Hadley Wickham: Data Visualization and Data Science (EMBL)

September 20, 2021

Hadley Wickham, Chief Scientist at RStudio and creator of many packages for the R programming language, in speaking at the EMBL discusses the principles of designing data graphic figures to communicate research findings. He focuses on how it is possible to utilize data visualization and color effectively to help convey information. He explains how data tools such as R can help data scientists better communicate information to audiences and allow for there to be more accurate quantitative judgements by taking advantage of fundamental principles of human perception.

Tools such as R will implement strategies for visualizing complex datasets. Some of the technologies and techniques Wickham introduced are tools for computational and cognitive displays. He explained creating R packages to support data communication; Tidy data; reshaping data; Tidyverse; Tidyr; Purr; and Dplyr. In the video Wickham discusses what came before 2004 when he worked with the lattice, a multivariate data visualization with R. He goes on to discuss what ideas and thoughts helped him to evolve gplot2, a chart-making system for the statistical programming language R and how to learn one package that allows for learning and building on to other packages.

His main points were that he wanted to discuss what visualization means and the parts of visualization he finds most useful. He wanted to present the ideas that data visualization has evolved and continues to change. Graphical methods for data analysis and data presentation

should provide guidelines for communicating useful information and interpretation. Data scientists should present data so that people can extract informative meaning from the visualization.

References

Hadley Wickham. 2019. European Molecular Biology Laboratory (EMBL) EMBL Keynote
    Lecture. *Data Visualization and Data Science.*
    https://www.youtube.com/watch?v=9YTNYT1maa4 (September 20, 2021).