

# Data Literacy for All, with R

Ryan Womack

Data Librarian, Rutgers University, <https://ryanwomack.com>

IAASSIST Annual Conference, Montréal, Canada, May 29, 2018



This work is licensed under a [Creative Commons Attribution  
-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

# Outline

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- Introduction
- Statistical Literacy
- Data Literacy
- Competencies
- Role of NSOs and Academia
- In Practice

# Literacy: What do we mean?

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- **Literate:** “able to read and write”
- Literate: “having knowledge or competence”
- An old concept that has been extended into many new areas

# Multiple Literacies

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- “Literacy” borrows the force of our original reading experience
  - the critical skills to understand, interpret, and create in that domain
- *Media Literacy, Digital Literacy, Social/Emotional Literacy* ([Digital Age Literacies](#))
- *Scientific Literacy*
- *Numeracy, Quantitative Literacy* - in the domain of mathematics

# Which Literacy?

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

We will explore:

- Statistical Literacy
- Information Literacy
- Data Information Literacy
- Data Literacy → our final goal

# Statistics

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

Two meanings:

- *numerical* - a collection of quantitative data
- *conceptual* - a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data

# Statistical Literacy

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

Statistical literacy deals with both numerical and conceptual meanings.

- An old concept: see Walker, “Statistical Literacy in the Social Sciences”, *The American Statistician*, 1951
- The original concept focused on limited examples, basic math
- **Numeracy** and **Quantitative Literacy** are similar - focus on basic math skills
- **Statistical Literacy** requires Quantitative Literacy
- But emphasis is on using the statistical concepts in practice (with statistical sources)

# Statistical Literacy, example

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

Question from **ARTIST**:

Statisticians often use the term “**robust**” when referring to a statistical measure or test. Which of the following is an example of a robust measure of center?

- 1 **Mean**, because it is the most commonly used measure of center.
- 2 **Median**, because it is resistant to outliers.
- 3 **Mode**, because it can be used with observational or quantitative data.



# Statistical Literacy, example

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

The prior example is addressed to a statistical audience, but we can illustrate with examples that would work for any level of audience.

- Dataset: 1,2,2,2,3
- What is Mean, Median, Mode?
- Replace 3 with 18
- What is “robustness”? Develop intuitive understanding.

# Statistical Literacy, defined

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- Gal, “Adults’ Statistical Literacy: Meanings, Components, Responsibilities.” *International Statistical Review*, 2002

“statistically literate behavior depends on the availability and joint activation of multiple knowledge bases”

- **Literacy** (reading, but also including handling graphs, charts, and tables, and other forms of textual evidence)
- **Statistical**
- **Mathematical**
- **Context**
- **Critical Skills**

these are combined with *Beliefs and Attitudes* and *Critical Stance*  
=> this adds up to **Statistical Literacy**

# Statistical Literacy, in practice

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

## Everyday statistical literacy

- Knowing things like mean, median, mode
- Understanding that data points can be distributed in different ways
- Being able to interpret a graph or chart
- Knowing to look for bias, critically analyze

# Statistical Knowledge Base

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

## Statistical Knowledge Base

- 1 Knowing why data are needed and how data can be produced
- 2 Familiarity with basic terms and ideas related to descriptive statistics
- 3 Familiarity with basic terms and ideas related to graphical and tabular displays
- 4 Understanding basic notions of probability
- 5 Knowing how statistical conclusions or inferences are reached

quoted from Gal, 2002

# Information Literacy

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- **Information Literacy** is a major concept in libraries and elsewhere in education
- Interpreting textual material
- Understanding the information environment
- Many library educational standards have been defined around this concept
- For librarians, data literacy can be viewed as an offshoot or subset of information literacy
- see Scheld, “Information Literacy, Statistical Literacy and Data Literacy”, *IASSIST Quarterly*, Summer/Fall 2004

# Research Data Management

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

## Libraries and others have become increasingly involved in Research Data Management

- supporting the data lifecycle for researchers
- creating, analyzing, sharing, and reusing data
- less emphasis on statistical skills and understanding
- MacMillan, "Data Sharing and Discovery: What Librarians need to know" *The Journal of Academic Librarianship*, 2014

# Data *Information* Literacy

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

A hybrid of data management and information literacy, this terminology is new and not as widely used.

- How to Inventory, Store, and Backup Your Data
- How to Create Data That You (and Others) Can Understand
- How to Navigate Rights and Ownership of Your Research Data
- How to Share Your Data and Ethically Reuse Data Created by Others
- How to Digitally Preserve Your Data for the Future
  - see Johnston and Jeffries, “Steal This Idea: A library instructors’ guide to educating students in data management skills”, *College and Research Libraries News*, September 2014

# Data Explosion

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

Larger worldwide trend of interest in data

- Big Data
- Data Science
- Data Analytics
- driving a different kind of focus on data vs. statistics



# Data Science Skills

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- more emphasis on large samples, multivariate description, and data visualization
- understanding and exploring data vs. applying techniques to sample datasets
- students need to learn to think statistically and to develop an aesthetic for data handling and modelling based on solving practical problems
- see Ridgway, “Implications of the Data Revolution for Statistics Education”, *International Statistical Review*, 2016

# Data Literacy, first steps

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- Data literacy, unlike research data management or data information literacy, places focus on the understanding and analysis of data
- As in all “literacy”, competency and knowledge (of data) are important
- Statistical knowledge is required
- A greater mixture of data exploration, data cleaning, “data wrangling”
- Making sense of disparate data through visualization and other methods (e.g., machine learning)

# Data Literacy, a definition

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- An emerging area

- “there is no standard definition for data literacy nor is there a definitive list of competencies” [Dechman and Syms, “Working Together to Maximize the Utilization of Open Data Across Social Science and Professional Disciplines”, *Behavioral & Social Sciences Librarian*, 2014]

- But here is a reasonable start

- “Data literacy can be defined, then, as the component of information literacy that enables individuals to **access, interpret, critically assess, manage, handle and ethically use data**” [Prado and Marzal, “Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents”, *Libri*, 2013]

# Data Literacy and Librarians

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- Viewed in this way, librarians have traditionally trained users in how to access, interpret, handle, and ethically use information
- So extending those roles to data via data literacy seems more natural for librarians than the methodologies emphasized in statistical literacy
- Open Data and online analysis broaden the applicability of data literacy to a wide audience, increasing the potential for librarians to play a role.
- Data literacy can be applied to many disciplinary contexts, not just statistics. Social justice, urban planning, and other fields.
- Just as quantitative literacy is required for statistical literacy, statistical literacy is required for data literacy. It is necessary, but not sufficient.

# Ingredients of Competencies

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

## Beyond statistical literacy...

- Ethics
- Real-world problem-solving context
- Ask questions from data
- Develop hypotheses and identify potential sources of data
- Collect or acquire data
- Analyse and create explanations from data
- Evaluate the validity of explanations based on data and formulate new questions
- more...?

# Data Literacy Competencies, a working definition

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- **Understanding Data** (Role of Data in Society)
- **Finding and/or Obtaining Data** (Sources, Search, Assessment)
- **Reading, Interpreting and Evaluating Data** (Statistical Literacy...)
- **Managing Data** (Metadata, Databases, practices)
- **Using Data** (Handling, Synthesis, Ethics)
- from Prado and Marzal, 2013

# How to Develop Skills

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- Training should be adapted to the level of the user
- Make it Personal
- Keep it Fun
- See MacCuirc, “You Don’t Teach, Students Learn: A Report on a Project on Statistical Literacy in Ireland”, *Austrian Journal of Statistics*, 2015

# Types of Data Literate Citizen

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- **Communicators** – make sense of and tell stories about data for others
- **Readers** – need skills to interpret data as part of their every day life
- **Makers** – need skills to integrate data real-world problem-solving
- **Scientists** – need strong technical data skills with domain expertise
- see Wolff, et. al. “Creating an Understanding of Data Literacy for a Data-driven Society”, *Journal of Community Informatics*, 2017



# Conclusion

Data Literacy  
for All, with R

Ryan Womack

Introduction

Statistical  
Literacy

Data Literacy

Competencies

In Practice

- Data Literacy has been recognized as an essential component of education
- Initial development of official statistics and the discipline of statistics reflects the needs of the industrial age
- Now the information age is upon us
- Why Data Literacy?
- Data literacy lets us seek our own answers in the mass of big data, empowering future development