

**THE**  
**DATA LITERACY**  
**COOKBOOK**

edited by Kelly Getz and Meryl Brodsky



# ***The Data Literacy Cookbook***

*edited by Kelly Getz and Meryl Brodsky*

Association of College and Research Libraries

*A division of the American Library Association*

Chicago • 2022

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences-Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992. ∞

The ACRL Cookbook series was conceived of and designed by Ryan Sittler and Doug Cook.

Other books in this series:

*The Scholarly Communications Cookbook*, edited by Brianna Buljung and Emily Bongiovanni

*The Teaching with Primary Sources Cookbook*, edited by Julie M. Porterfield

*The Library Outreach Cookbook*, edited by Ryan L. Sittler and Terra J. Rogerson

*The Critical Thinking about Sources Cookbook*, edited by Sarah E. Morris

*The Sustainable Library's Cookbook*, edited by Raymond Pun and Gary L. Shaffer

*The Library Assessment Cookbook*, edited by Aaron W. Dobbs

*The First-Year Experience Cookbook*, edited by Raymond Pun and Meggan Houlihan

*The Discovery Tool Cookbook*, edited by Nancy Fawley and Nikki Krysak

*The Embedded Librarian's Cookbook*, edited by Kaijsa Calkins and Cassandra Kvenild

Library of Congress Control Number: (TO COME)

©2022 by the Association of College and Research Libraries, a division of the American Library Association. All rights reserved except those which may be granted by Sections 107 and 108 of the Copyright Revision Act of 1976.  
Printed in the United States of America.

26 25 24 23 22 5 4 3 2 1

# TABLE OF CONTENTS

vii Introduction

## SECTION 1. INTERPRETING POLLS AND SURVEYS

- 3 **[[Ch1]]Survey Literacy: A Skills-Based Approach to Teaching Survey Research**  
*Jesse Klein*
- 6 **[[Ch2]]Setting the Scene with Surveys: Using Polling Software to Demonstrate Primary and Secondary Data**  
*Wendy G. Pothier*
- 9 **[[Ch3]]The Mini-study: A Three-Part Assignment for Original Data Creation, Summation, and Visualization**  
*William Cuthbertson, Lyda Fontes McCartin, and Sara O'Donnell*

## SECTION 2. FINDING AND EVALUATING DATA

- 17 **[[Ch4]]Three-Step Data Searching**  
*Annelise Sklar*
- 21 **[[Ch5]]Transforming Research Questions into Variables: A Recipe for Finding Secondary Data**  
*Alicia Kubas and Jenny McBurney*
- 25 **[[Ch6]]Sweeten the Search: Discover Data for Reuse with a Tool That Links Publications to the Underlying Data**  
*Elizabeth Moss*
- 30 **[[Ch7]]The Most Vital Statistics: Finding and Analyzing Historical Mortality Rates**  
*Alisa Beth Rod and Jennie Correia*

34 **[[Ch8]]Understanding the Enumerated World: Making Sense of Data as an Information Source**

*Alexandra Cooper, Elizabeth Hill, and Kristi Thompson*

38 **[[Ch9]]Looking at Data**  
*Kay K. Bjornen*

43 **[[Ch10]]Interrogating the Data: What Data Sets Can and Cannot Tell Us**  
*Kristin Fontichiaro*

46 **[[Ch11]]Data Zines: A Hands-On Approach to Community Curiosities**  
*Tess Wilson*

49 **[[Ch12]]On the Hunt: Understanding and Analyzing GSS Data Extraction for Incorporation within Sociological Research Projects**  
*Amy Dye-Reeves*

52 **[[Ch13]]Using Statistics to Define the Problem: Data and Service Learning**  
*Amy Harris Houk and Jenny Dale*

56 **[[Ch14]]Data and Statistics in the News and Media**  
*Kaetlyn Phillips*

## SECTION 3. DATA MANIPULATION AND TRANSFORMATION

61 **[[Ch15]]A Kinesthetic Approach to Data: Moving to Understand Nominal, Ordinal, Interval, and Ratio Relationship in Data**  
*Wendy Stephens*

64 **[[Ch16]]Text Mining Charcuterie Board**  
*Yun Dai and Fan Luo*

67 **[[Ch17]]Anyone Can Cook (R)! Open Data with R, a Five-Week Mini-mester**  
*Jay Forrest and Ameet Doshi*

70 **[[Ch18]]Software Carpentry Al Dente: Rendering Tech Training for Online Artisans**  
*Peace Ossom-Williamson, Shiloh Williams, and Hammad Rauf Khan*

73 **[[Ch19]]A Recipe for Improving Online Instruction for the Carpentries**  
*Kay K. Bjornen and Clarke Iakovakis*

## SECTION 4. DATA VISUALIZATION

79 **[[Ch20]]Correlation Does Not Equal Causality: Introducing Data Literacy through Infographics and Statistics in the Media**  
*Nick Ruhs*

83 **[[Ch21]]Pies, Bars, Charts, and Graphs, Oh My! A Data Visualization Appetizer**  
*Haley L. Lott*

86 **[[Ch22]]Data Visualizations: The Good, the Bad, and the Ugly**  
*Kaetlyn Phillips*

89 **[[Ch23]]Seasonal Visual Literacy: Using Current Events to Teach Data and Spatial Literacy Skills with Adaptable LibGuides**  
*Jacqueline Fleming and Theresa Quill*

93 **[[Ch24]]To Visualize Is to Experience Data**  
*Chelsea H. Barrett and Gerard Shea*

97 **[[Ch25]]Upping the Baseline for Data Literacy Instruction**  
*Jessica Vanderhoff*

## Table of Contents

- 101 **[[Ch26]]A Literacy-Based Approach to Learning Visualization with R's ggplot2 Package**

*Angela M. Zoss*

- 104 **[[Ch27]]Build Your Own Data Viz Pizza: A Modular Approach to Data Visualization Instruction**

*Rachel Starry*

- 108 **[[Ch28]]Veggie Pizza: Choosing a Data Visualization Tool**

*Rachel Starry*

- 111 **[[Ch29]]Four-Cheese Pizza: Color and Accessible Design**

*Rachel Starry*

- 114 **[[Ch30]]Data Visualization using Web Apps in a Rainbow Layer Cake**

*Yun Dai and Fan Luo*

- 117 **[[Ch31]]Graphical Abstracts: Creating Appetizing Infographics for Your Research Article**

*Aleshia Huber*

### SECTION 5. DATA MANAGEMENT AND SHARING

- 123 **[[Ch32]]Making File Names for Digital Exhibits**

*Kate Thornhill and Gabriele Hayden*

- 126 **[[Ch33]]Data Management Failures: Teaching the Importance of DMPs through Cautionary Examples**

*Richard M. Mikulski*

- 131 **[[Ch34]]Low-Fat Research Data Management**

*Elizabeth Blackwood*

- 134 **[[Ch35]]Managing Qualitative Social Science Data: An Open, Self-Guided Course**

*Sebastian Karcher and Diana Kapiszewski*

- 136 **[[Ch36]]Seven Weeks, Seven DMPs: Iterative Learning around Data Management Plan Creation**

*Emma Slayton and Hannah C. Gunderman*

- 140 **[[Ch37]]Equitable from the Beginning: Incorporating Critical Data Perspectives into Your Research Design**

*Jodi Coalter, David Durden, and Leigh Amadi Dunewood*

### SECTION 6. GEOSPATIAL DATA

- 147 **[[Ch38]]Challenge Accepted: Introducing Geospatial Data Literacy through an Online Learning Path**

*Joshua Sadvari and Katie Phillips*

- 151 **[[Ch39]]GIS for Success Series: Learning the Basics of QGIS Workshop**

*Kelly Grove*

- 154 **[[Ch40]]GIS for Success Series: Let's Make a Map in QGIS Workshop**

*Kelly Grove*

- 157 **[[Ch41]]Statistical and Geospatial Literacy for Integrative Genetics**

*Jay Forrest and Chrissy Spencer*

- 161 **[[Ch42]]Web Map Layer Cake: Teaching Web Mapping Skills with Leaflet for R**

*Sarah Zhang and Julie Jones*

### SECTION 7. DATA IN THE DISCIPLINES

- 167 **[[Ch43]]Data in Context: How Data Fit into the Scholarly Conversation**

*Theresa Burress*

- 171 **[[Ch44]]Let the Dough Rise! Integrating Library Instruction in a Digital Humanities Course**

*René Duplain and Chantal Ripp*

- 175 **[[Ch45]]Ethics and Biodiversity Data**

*Rebecca Hill Renirie*

- 179 **[[Ch46]]Data Decisions and the Research Process in the Sciences and Social Sciences**

*Nicole Helregel*

- 182 **[[Ch47]]Financial Data for Economics Students**

*Jennifer Yao Weinraub*

- 183 **[[Ch48]]Stuffed Shiny App with Business Intelligence**

*Yun Dai and Fan Luo*

- 189 **[[Ch49]]Fast Casual Marketing Strategies**

*Juliann Couture, Halley Todd, and Natalia Tingle Dolan*

- 192 **[[Ch50]]When and Where: A Framework for Finding and Evaluating Social Science Data for Reuse**

*Ari Gofman*

- 197 **[[Ch51]]Data Literacy Layered Lasagna for Preservice Teachers**

*Brad Dennis and Allison Hart-Young*

### SECTION 8. DATA LITERACY OUTREACH AND ENGAGEMENT

- 203 **[[Ch52]]Data Visualization Day: Promoting Data Literacy with Campus Partners**

*Wenli Gao*

- 206 **[[Ch53]]Getting Messy Ourselves: An Experiential Learning Curriculum for Subject Librarians to Engage with Data Literacy**

*Adrienne Canino*

- 211 **[[Ch54]]Research Data Management Stone Soup: Gauging Team Competencies**

*Michelle Armstrong, Megan Davis, Ellie Dworak, Yitzhak "Yitzy" Paul, and Elisabeth Shook*

<p><b>214</b> <b>[[Ch55]]Data Literacy Family Style: Full-Day Professional Development</b>  <i>Molly Ledermann, Emilia Marcyk, Terence O'Neill, and Dianna E. Sachs</i></p> <p><b>217</b> <b>[[Ch56]]Everyone Is Welcome at the Table: Outreach for Data Management and Data Literacy in Research Assignment Design</b>  <i>Shannon Sheridan and Hilary Baribeau</i></p> <p><b>220</b> <b>[[Ch57]]Seasoning and Simmering: Cultivating Data Literacy Skills through an Open Data Hackathon</b>  <i>Peace Ossom-Williamson</i></p> <p><b>223</b> <b>[[Ch58]]From Soup to Nuts: Finding Your Way around the Data Services Buffet</b>  <i>Jane Fry and Chantal Ripp</i></p> <p><b>226</b> <b>[[Ch59]]Teaching Data Literacy and Computational Thinking in Educational Technology</b>  <i>Lesley S. J. Farmer</i></p> <p><b>SECTION 9. DATA LITERACY PROGRAMS AND CURRICULA</b></p> <p><b>231</b> <b>[[Ch60]]Cooking Up a Data Literacy Course</b>  <i>Claire Nickerson</i></p> <p><b>238</b> <b>[[Ch61]]Baking a Data Layer Cake: Scaffolding Data Skills through Video Vignettes</b>  <i>Shannon Sheridan</i></p> <p><b>241</b> <b>[[Ch62]]Building Data Literacy through Scaffolded Workshops: Experiences and Challenges</b>  <i>Jiebei Luo and Yaqing (Allison) Xu</i></p> <p><b>244</b> <b>[[Ch63]]Data Literacy Appetizers: LibGuide Data Instruction Modules for Undergraduates</b>  <i>Beth Hillemann and Aaron Albertson</i></p>	<p><b>247</b> <b>[[Ch64]]Data as Curation: Framing Data Creation as a Critical Practice through Collections-Based Research Inquiry</b>  <i>Gesina A. Phillips, Tyrica Terry Kapral, Matthew J. Lavin, and Aaron Brenner</i></p> <p><b>253</b> <b>[[Ch65]]Quantitative Data Skills for Undergraduates: A Seminar Series for Social Science Students</b>  <i>Whitney Kramer and Amelia Kallaher</i></p>
--	---

# Seasoning and Simmering

## Cultivating Data Literacy Skills through an Open Data Hackathon

Peace Ossom-Williamson, MLS, MS, AHIP, Associate Director of NNLM National Center for Data Services and Associate Curator, NYU Langone Health, [peace.williamson@nyulangone.org](mailto:peace.williamson@nyulangone.org)

### NUTRITION INFORMATION

This recipe provides the steps for developing data competitions, such as hackathons, where students practice coding and challenge their thinking around open data and their uses. Hackathons and data competitions foster an environment for applying skills learned in the classroom or elsewhere and build upon them by working on a “project marathon.”<sup>1</sup> A hackathon is a project that must be started and completed in a set amount of time and with a given theme. Hackathons may have a lot of structure, including focusing on a specific issue or problem or a particular output, such as an app. However, they often have very little structure, and students are left to develop their own ideas. Open Data Day (<https://opendata-day.org>) is a great opportunity for this event because it is an international celebration of the use and value of open data and the institutions and organizations that support open data. The Open Data Hackathon provides an opportunity for students to challenge themselves to apply or develop new skills, become acquainted with open data sources, and work collaboratively with students outside of their disciplines.

### TARGET AUDIENCE AND NUMBER SERVED

15 to 100 participants. (See Chef’s Notes

regarding larger events.)

### LEARNING OUTCOMES

Participants:

- discover how data are formally produced, organized, and disseminated
- explore information sources; identify the purpose and audience of potential resources
- define or modify the information needed to achieve a manageable focus
- define a realistic overall plan and time line to acquire and build from the needed information
- incorporate the data and supporting information into their creation
- describe what they created in order to garner support for winning the competition

### COOKING TIME

Preparation time depends upon the size of the hackathon. A hackathon that has around 20 participants may take a month or two to plan, whereas a hackathon with 60–100 participants would take 6 months. Planning for larger hackathons, hackathons that span multiple days, or hackathons that include participants who travel to participate should be planned a year in advance. Regarding the

actual event, hackathons can span 2 hours or longer. Common time spans for in-person hackathons are 12 or 24 hours.

### DIETARY GUIDELINES

Depending on participants’ individual experiences, the hackathon applies at some level concepts from all of the frames of ACRL’s *Framework for Information Literacy for Higher Education*; however, it is centered on Information Creation as a Process, Information Has Value, Research as Inquiry, and Searching as Strategic Exploration.

### INGREDIENTS

- A location that allows for various working styles, including round tables for group work and small solo tables and couches
- Dry erase boards and markers
- Handouts with schedules and links for submissions
- Name tags
- (optional) T-shirts that are different colors for participants and volunteers
- An abundance of places to plug in (power strips, outlets, etc.) with the expectation of 2–3 outlets per participant
- Wi-Fi access and a network that can accommodate the number of connected

- devices and data being transferred
- Enough time for the open data sources to be used (Some APIs take 24 hours to be integrated.)
- Food and drinks, including the meals and snacks to be consumed during the event
- Other amenities as needed for size (See Chef’s Notes.)
- Certificates and/or prizes and awards for winners
- (optional) Workshops teaching lessons, such as how to use Git and GitHub and an introduction to machine learning

**PREPARATION**

Trello or another project management tool is recommended for planning and tracking preparation before the event. The Open Data Hackathon has a theme that designs must incorporate, such as the provision of health information or the incorporation of movement. Technology should be set up for instructions, including (1) a general website with details about the competition, including its rules and the judges’ names and information, and (2) a registration form gathering demographic data about participants, including interest in particular workshops and planning information such as dietary guidelines and restrictions and the need for physical accommodations. Recruit volunteers to assist with day-of activities, including registration, handing out name tags and T-shirts, and setting up food, tables, and other ingredients. A place for submissions should be developed and can range from a shared Google Drive folder to using a hackathon platform, such as Devpost (<https://>

[devpost.com](https://devpost.com)), which allows for submissions, voting, judging, and contacting participants and winners. Librarians can also develop a Lib-Guide of resources about available software, open data sources, and files for workshops.

**INSTRUCTIONS**

Most of the preparation involves marketing the event, developing places for registration and submission, and day-of setup and clean-up. Students from a variety of experience levels and majors are encouraged to attend. This event is especially focused on encouraging students from non-STEM fields to participate. There are many ways to participate in the hackathon, not all of which are coding. Some participants develop the design of their creation, others work on front-end development, some gather data and work with APIs, and others provide content knowledge.

Examples of submissions include teams who have built a COVID-19 tracker, a system that provides students and medical professionals evidence-based information about medicines and illicit drugs and their effects, and a 3D model of a new method of personal identification that works as an all-encompassing ID and payment card. During the workshop, there are typically few questions or needs for assistance. Most of the volunteer efforts occur at mealtimes. However, one hour prior to the conclusion, announcements should be made about concluding documentation, uploading files, and completing submissions. Upon completion, submissions are assessed by three outside judges. One week later, winners receive prizes and certificates.

Table 1 shows a set of review criteria from the UTA Open Data Day Hackathon.

**TABLE 1**  
**Review criteria from the UTA Open Data Day Hackathon.**

<b>Usefulness</b>	How useful is your project? In what ways does it improve access to or evaluation of information related to the hackathon theme?
<b>Fidelity</b>	Did you use open data sources in your project? Was the use of the data accurate and true to the original data and compliant with existing standards? Did you thoroughly document your project?
<b>Design</b>	Did you put thought into the user experience? How well designed is the interface? Are the purpose and functionality clear for the user?
<b>Technology</b>	How technically impressive was the hack? Did you use a particularly clever technique or many different components? Did the technology involved make people go “Wow”?
<b>Learning</b>	Did you stretch your skills, learn new tricks, or apply new tools?



### REVIEWS/ASSESSMENT STRATEGY

Data are collected at registration, including demographic information, which college or department participants are from, (optional) team information if the team is already formed, and other details. This information can be used for creating name tags and for check-in using a digital system. After the hackathon, these data, along with the participants who came and submissions received, can be used to assess the success and growth of the annual event.

### CHEF'S NOTES

For large hackathons with over 100 participants who come into town for the event, additional planning will need to be done and addition provisions offered, such as travel buses, hotel and other accommodations, information about the area, security, and authenticating identification of participants. Partnering with the organization Major League Hacking (<https://mlh.io>) for a large hackathon is recommended because MLH provides promotion on its platform, assists with project management, and assists with resources for a successful large hackathon event.

### CLEANUP

A system, typically a rubric, must be in place for judging and selecting winners, and this system should be communicated to participants ahead of the event. Whether there is judging and a presentation at the closing of the competition or whether judges review

submissions and hold an award ceremony at a later date, the winners should be recognized with a certificate and/or prizes.

### ADDITIONAL RESOURCES

Kaggle. "Competitions." <https://www.kaggle.com/competitions>.

Open Knowledge Foundation. "Open Data Day." <https://opendataday.org>.

UTA Open Data Day Hackathon. "Judging Criteria." Devpost. 2020. <https://opendatahackathon.devpost.com/>.

### NOTE

1. Nancy Shin, Peace Ossom Williamson, and Bethany McGowan. "vMLA 2021—On the Edge of Innovation: How Libraries Can Develop Health Datathons and Hackathons for Data Literacy" (immersion session, Medical Library Association 2021 Virtual Conference, May 25, 2021), <https://doi.org/10.17605/OSF.IO/SXFHW>.