DATALAB Workshop Facilitator Guide

Making Sense of Data: A Beginner's Guide

Duration: 90 minutes
Target Audience: Complete beginners to R and data analysis
Prerequisites: None

Pre-Workshop Checklist

One Week Before

- ☐ Send participants the pre-check script to verify their setup
- ☐ Send installation instructions for R and RStudio
- ☐ Share sample datasets so participants can preview them

Day Before

- ☐ Test your own setup with the latest package version
- ☐ Prepare breakout rooms if conducting online
- ☐ Load all datasets and run through examples once

Day Of

- ☐ Open RStudio with a clean workspace
- ☐ Have participant_template.R open and ready
- ☐ Have exercise_answers.R open in another window (for quick reference)

Workshop Schedule

Introduction (10 minutes)

Objectives:

- Set expectations: This is about insights, not coding
- Show the power of data visualization
- Make participants comfortable with making mistakes

Script:

"Welcome! Today we're going to learn how to make data tell stories. You don't need any programming experience. Our goal is simple: take data and turn it into insights you can understand and share. Everything we do today uses simple functions that work like this: you tell the computer WHAT you want to see, and it handles the HOW."

Demo (3 minutes): Show one impressive example:

```r
data(heart_data)
data <- heart_data
piechart(chest_pain_type, sex)
```

Explain what the chart shows in plain language. Ask: "What story does this tell?"

Section 1: Setup & First Visualization (15 minutes)

Learning Goal: Load data and create a basic chart

Teaching Points:

1. Loading the package with `library(datalab)`
2. Loading a dataset and assigning it to `data`
3. Looking at data with `head()`
4. Creating a simple pie chart

Live Demo:

r

```r
library(datalab)
data(heart_data)
data <- heart_data
head(data)
piechart(chest_pain_type)
```

Key Message: "Notice you just typed one line and got a professional chart. That's the power of datalab."

Practice Time (5 minutes):

- Have participants create their first pie chart
- Walk around (or monitor screen shares) to help
- Common issues: Typos in variable names

Debrief (2 minutes): Ask: "What does this chart tell us?" Get 2-3 responses.

Section 2: Comparing Groups (15 minutes)

Learning Goal: Use filters to compare subgroups

Teaching Points:

1. The filter parameter creates multiple charts automatically
2. Colors help distinguish charts
3. Sample sizes (n=) help interpret results

Live Demo:

r

```
piechart(chest_pain_type, sex)
```

Facilitation Tip: Point out loud: "See how we get two charts automatically? One for males, one for females. Notice the different color schemes—that helps us see patterns."

Practice Time (8 minutes):

- Participants create boxplot with grouping
- Walk through common question: "Why use boxplot vs piechart?"
  - Answer: "Boxplots for numbers, piecharts for categories"

Debrief (2 minutes): Show one participant's solution. Ask: "What differences do you see between groups?"

Section 3: Understanding Distributions (15 minutes)

Learning Goal: Read and interpret boxplots

Teaching Points:

1. What each part of a boxplot means (but keep it simple)
   - Middle line = median (middle value)
   - Box = where most data lives
   - Dots = unusual values
2. When to use single vs grouped boxplots

Live Demo:

r

```
boxplot(age)
boxplot(age, group = sex)
```

Interactive Moment: Ask participants: "Is age distributed the same for males and females?" Wait for responses.

Practice Time (8 minutes): Participants explore resting_bp

Common Question: "What does the box show?" **Answer:** "The box shows where the middle 50% of your data lives. If the box is tall, values are spread out. If it's short, values are similar."

Section 4: Finding Patterns (10 minutes)

Learning Goal: Use line charts to see relationships

Switch Dataset:

r

```
data(agriculture_data)
data <- agriculture_data
```

Teaching Point: "We just switched to different data. See how easy that was?"

Live Demo:

r

```
linechart(nitrogen_content, crop_yield)
linechart(nitrogen_content, crop_yield, smooth = TRUE)
```

Key Message: "Line charts help us see if two things move together. Does more nitrogen mean more yield? The smooth line makes the pattern clearer."

Practice Time (5 minutes): Participants add smooth line to their charts

---

Section 5: Summary Statistics (10 minutes)

Learning Goal: Get numerical summaries

Teaching Points:

1. Statistics give exact numbers to back up visual patterns
2. Mean vs Median (briefly)
3. SD tells us about spread

Live Demo:

r

```r
data(heart_data)
data <- heart_data
descriptives(age, resting_bp, cholesterol)
```

Keep It Simple: Don't dive deep into statistics. Say: "These numbers help us describe our data precisely. Mean is the average. SD tells us how spread out values are."

Practice Time (5 minutes): Participants get statistics for one variable

---

Section 6: Counting Categories (10 minutes)

Learning Goal: Create frequency tables

Live Demo:

r

```r
frequencies(chest_pain_type)
frequencies(chest_pain_type, sort = "desc")
```

Teaching Point: "Sometimes we just need counts. How many of each type do we have?"

Practice Time (5 minutes): Quick exercise with frequencies

Exercise 1: Guided Story-Telling (10 minutes)

Setup: "Now you'll answer a real question using what you've learned: Is there a difference in cholesterol levels between males and females?"

Process:

1. Give participants 5 minutes to work
2. Show solution step-by-step
3. Emphasize the interpretation, not just the code

Facilitator Notes:

- This is where learning solidifies
- Emphasize that good analysis combines charts AND statistics
- Model how to write interpretations

Exercise 2 & 3: Open Exploration (10 minutes)

Goal: Build confidence through free exploration

Instructions: "Choose any two variables and explore their relationship. Use any functions you like."

Facilitation Strategy:

- Circulate and ask probing questions: "What surprised you?" "What story does this tell?"
- Highlight interesting discoveries: "Let's see what [name] found!"

Wrap-Up (5 minutes)

Key Takeaways:

1. You can explore data without being a programmer
2. Good analysis combines visuals and statistics
3. The story matters more than the code

Resources:

- Cheat sheet (hand out)
- Vignette for reference
- Encourage them to try with their own data

Final Words:

"You now have tools to turn data into insights. The functions are simple, but the stories you can tell are powerful. Start small, explore often, and remember: every dataset has a story waiting to be discovered."

Common Questions & Answers

Q: "I got an error: variable not found" A: Check three things: (1) spelling, (2) did you load the data?, (3) is it in this dataset?

Q: "How do I know which chart to use?" A: Numbers → boxplot or linechart. Categories → piechart or frequencies. Comparing groups → add filter or group parameter.

Q: "Can I use my own data?" A: Yes! Load with `read.csv()` and assign to `data`. Show them the template section for this.

Q: "Why do some pie charts have different colors?" A: When comparing groups, different colors help you see each chart clearly.

Q: "What if I have too many categories for a pie chart?" A: Use `frequencies()` first to see counts, or use a different variable.

Troubleshooting Tips

Participants Can't Install Package

- Have a backup: Share .R files with functions
- Pair them with someone who has it working

Code Doesn't Run

- Check for typos in variable names (most common)
- Verify `data` is loaded
- Check if they're using the right dataset

Participants Feel Lost

- Reassure: "It's normal to feel confused at first"
- Pair them with a neighbor
- Go back to a working example

Moving Too Fast/Slow

- Watch body language and chat messages
- Ask: "Should I slow down?" or "Ready to move on?"
- Skip Exercise 3 if running short on time

Post-Workshop

Follow-Up Email Template

Subject: Workshop Resources + Next Steps

Hi everyone,

Great session today! Here are your resources:

1. Participant template script (attached)
2. Cheat sheet (attached)
3. Package vignette: [link]

Try this challenge: Take one of your own datasets and create 3 different visualizations. Share what you discover!

Questions? Reply to this email.

Best,
[Your name]

Feedback Collection

- Send a 3-question survey:
    1. What was most helpful?
    2. What was confusing?
    3. Will you use datalab in your work?

Adaptation Notes

For Shorter Workshop (60 min)

- Skip Section 4 (line charts)
- Combine Sections 5 & 6
- Do only Exercise 1

For Longer Workshop (120 min)

- Add a case study with real-world data
- Include a "data interpretation" group activity
- Add more open exploration time

For Domain-Specific Audiences

- Replace sample datasets with domain-relevant data
- Adjust examples to match their context
- Add domain-specific interpretation guidance