📚 Learning Path

This tutorial is designed for progressive learning:

1. **Beginners**: Start with the Introduction and basic ggplot2 sections
2. **Intermediate**: Focus on Advanced ggplot2 and Specialized Visualizations
3. **Advanced**: Explore the interactive Plotly section and ggmosaic examples

Running the Tutorial Section by Section

To get the most out of this tutorial:

1. Start with the first section and run each code block sequentially
2. Study the output of each visualization and read the comments
3. Try modifying parameters to see how they affect the output
4. Use the online environments to experiment without worrying about installation

🧩 Example: Creating Your First Visualization

Here's a quick example of what you'll learn in the first section:

```r
# Load the built-in mtcars dataset
data(mtcars)
head(mtcars)

# Convert cylinders to factor
mtcars$cyl <- as.factor(mtcars$cyl)

# Simple scatter plot with base R
plot(mtcars$wt, mtcars$mpg,
     main = "Car Weight vs. Fuel Efficiency",
     xlab = "Weight (1000 lbs)",
     ylab = "Miles Per Gallon",
     pch = 19,
     col = "blue")

# The same scatter plot with ggplot2
library(ggplot2)
ggplot(mtcars, aes(x = wt, y = mpg)) +
  geom_point(color = "blue") +
  labs(title = "Car Weight vs. Fuel Efficiency",
       x = "Weight (1000 lbs)",
       y = "Miles Per Gallon") +
  theme_minimal()
```

🤝 Contributing

Contributions to improve the tutorial are welcome! Please feel free to submit a pull request or open an issue if you find any errors or have suggestions for improvements.
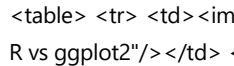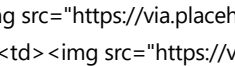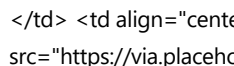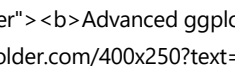
📄 License

This project is licensed under the MIT License - see the LICENSE file for details.

🙏 Acknowledgments

- The R community for creating and maintaining excellent visualization packages
- Authors of the included datasets for making them available for educational purposes

- RStudio (now Posit) for their excellent IDE and R Markdown support
- The Rocker Project for R Docker images used in Codespaces
- Binder for enabling executable environments for R notebooks## 🖼️ Preview Gallery

The R Markdown notebooks include rendered visualizations that you can view directly on GitHub:

<table> <tr> <td></td> <td></td> </tr> <tr> <td align="center">Base R vs ggplot2 Comparison </td> <td align="center">Advanced ggplot2 Techniques</td> </tr> <tr> <td></td> <td> </td> </tr> <tr> <td align="center">Interactive Plotly Visualizations</td> <td align="center">Mosaic Plots for Categorical Data</td> </tr> </table>

Note: Replace these placeholder images with actual screenshots from your tutorial outputs# R Data Visualization Tutorial

Show Image

Show Image

Show Image

Show Image

A comprehensive tutorial on data visualization techniques in R, from basic plots to advanced interactive visualizations.

📊 Overview

This repository contains a hands-on R tutorial covering modern data visualization techniques. The tutorial progresses from simple base R plotting to advanced interactive visualizations using ggplot2, plotly, and specialized packages.

Author: Jennifer Favaloro (jennifer.h.favaloro@gmail.com)
Last Updated: May 19, 2025

🚀 Run This Tutorial Online

You can run this tutorial without installing anything on your local machine:

- **Open in GitHub Codespaces**: Launch a complete development environment in your browser with all required packages pre-installed.
- **Open in RStudio via Binder**: Run the tutorial in an RStudio environment directly in your browser.

- **Open in Posit Cloud**: Run in RStudio through Posit Cloud (formerly RStudio Cloud).

🌟 Features

- **Interactive Execution**: Run the code directly in your browser using Codespaces, Binder, or Posit Cloud

- **Step-by-Step Instructions**: Clear explanations with runnable code examples

- **Progression from Basic to Advanced**: Learn visualization techniques in increasing order of complexity

- **Multiple Visualization Libraries**: Covers base R, ggplot2, plotly, and specialized packages

- **Real-World Datasets**: Uses built-in datasets for practical examples

- **Code Comments**: Extensive documentation within the code

📋 Contents

Each section of the tutorial is available both in the main R script and as separate R Markdown notebooks:

1. **Introduction to Data Visualization in R**
   - R Script Section
   - R Markdown Notebook
   - Covers base R plotting, intro to ggplot2, and comparisons between systems

2. **Getting Started with ggplot2**
   - R Script Section
   - R Markdown Notebook
   - Explains the Grammar of Graphics concept and basic ggplot2 syntax

3. **Advanced ggplot2 Techniques**
   - R Script Section
   - R Markdown Notebook
   - Covers faceting, custom themes, colors, annotations, and transformations

4. **Specialized Visualizations**
   - R Script Section
   - R Markdown Notebook
   - Statistical visualizations, correlation plots, and time series techniques

5. **Interactive Visualizations with Plotly**
   - R Script Section
   - R Markdown Notebook
   - Create interactive plots with tooltips, zooming, and panning

6. **Mosaic Plots with ggmosaic**
   - R Script Section
   - R Markdown Notebook
   - Techniques for visualizing categorical data relationships

🚀 Getting Started

Run Online (No Installation Required)

Choose one of these options to run the tutorial directly in your browser:

1. **GitHub Codespaces**:
   - Click the "Open in GitHub Codespaces" button at the top of this README
   - Wait for the environment to load (includes R, RStudio Server, and all required packages)

- The R scripts will open automatically in RStudio Server
- Run the code chunks by selecting them and pressing Ctrl+Enter

2. **Binder**:
   - Click the "Launch Binder" button at the top of this README
   - Wait for the RStudio environment to load in your browser
   - Open the `Data Viz with RStudio - 05-19-2025.R` file from the file browser
   - Run the code chunks by selecting them and pressing Ctrl+Enter

3. **Posit Cloud** (formerly RStudio Cloud):
   - Click the "Open in Posit Cloud" link at the top of this README
   - Create a free account if you don't have one
   - The project will be copied to your account
   - Run the code just like in regular RStudio

Local Installation

If you prefer to run the tutorial locally:

1. **Clone this repository:**

   bash

   ```bash
   git clone https://github.com/yourusername/r-data-viz-tutorial.git
   cd r-data-viz-tutorial
   ```

2. **Open the R script in RStudio:** Open the file `Data Viz with RStudio - 05-19-2025.R` in RStudio.

3. **Install required packages:** Run the package installation code at the beginning of the script:

   r

   ```r
   install.packages('pacman')
   pacman::p_load('tidyverse', 'ggplot2', 'dplyr', 'RColorBrewer',
                  'reshape2', 'corrplot', 'patchwork', 'devtools',
                  'plotly', 'vcd')

   devtools::install_github("haleyjeppson/ggmosaic", upgrade = "never")
   library(ggmosaic)
   ```

🖥️ Repository Structure

This repository is set up for optimal learning and execution:

```
r-data-viz-tutorial/
├── README.md                          # This file
├── Data Viz with RStudio - 05-19-2025.R # Main R script
├── notebooks/                         # R Markdown versions
│   ├── 01-introduction.Rmd            # Introduction section as Rmd
│   ├── 02-ggplot2-basics.Rmd          # ggplot2 basics as Rmd
│   ├── 03-advanced-ggplot2.Rmd        # Advanced ggplot2 as Rmd
│   ├── 04-specialized-viz.Rmd         # Specialized visualizations as Rmd
│   ├── 05-plotly.Rmd                  # Plotly examples as Rmd
│   └── 06-ggmosaic.Rmd                # ggmosaic examples as Rmd
├── .binder/                           # Binder configuration
│   └── runtime.txt                    # R version specification
├── .devcontainer/                     # GitHub Codespaces config
│   └── devcontainer.json              # Development environment setup
├── install_packages.R                 # Package installation script
└── LICENSE                            # MIT License file
```

The main script contains all code, while the `notebooks/` directory contains the same content split into R Markdown files that you can run section by section with rendered outputs.

📝 Converting to R Markdown for GitHub Viewing

This repository includes R Markdown (`.Rmd`) versions of each section in the `notebooks/` directory. These files:

1. Render nicely on GitHub with code highlighting

2. Show outputs of code chunks with visualizations when viewed on GitHub

3. Can be downloaded and run interactively in RStudio

How to navigate the R Markdown files:

- Start with `01-introduction.Rmd` and work your way through sequentially

- Each notebook includes the same code as the main R script but with added text explanations

- The outputs of code cells (plots, tables) are visible directly on GitHub

Example of a code chunk in R Markdown:

```markdown
```{r scatter-plot-example, fig.width=8, fig.height=5}
# Create a scatter plot with ggplot2
ggplot(mtcars, aes(x = wt, y = mpg, color = cyl)) +
  geom_point(size = 3) +
  labs(title = "Car Weight vs. Fuel Efficiency by Cylinders",
       x = "Weight (1000 lbs)",
       y = "Miles Per Gallon",
       color = "Cylinders") +
  theme_minimal()
```
```

🛠️ Technical Setup Files

For GitHub Codespaces (`.devcontainer/devcontainer.json`):

```json
{
  "name": "R Data Visualization Tutorial",
  "image": "ghcr.io/rocker-org/devcontainer/tidyverse:4.2",
  "features": {
    "ghcr.io/rocker-org/devcontainer-features/r-packages:1": {
      "packages": "RColorBrewer,reshape2,corrplot,patchwork,devtools,plotly,vcd,ggmosaic"
    }
  },
  "postCreateCommand": "Rscript install_packages.R",
  "customizations": {
    "vscode": {
      "extensions": [
        "reditorsupport.r",
        "REditorSupport.r-lsp"
      ]
    }
  }
}
```

For Binder (`.binder/runtime.txt`):

Package Installation Script (`install_packages.R`):

```r
if (!require("pacman")) install.packages("pacman")
pacman::p_load('tidyverse', 'ggplot2', 'dplyr', 'RColorBrewer',
               'reshape2', 'corrplot', 'patchwork', 'devtools',
               'plotly', 'vcd')

if (!require("ggmosaic")) {
  devtools::install_github("haleyjeppson/ggmosaic", upgrade = "never")
}
```

📥 Access Options

This tutorial can be accessed in multiple ways:

1. Run Online (Recommended for Quick Start)

- **GitHub Codespaces**: Instant development environment with all dependencies
- **Binder**: Run in RStudio directly in your browser
- **Posit Cloud**: Professional RStudio experience in the cloud

2. Download Files

- **Complete R Script**: Download R Script
- **R Markdown Notebooks**: Each section is available as a separate .Rmd file in the notebooks directory
- **Clone Repository**: `git clone https://github.com/yourusername/r-data-viz-tutorial.git`

3. View on GitHub

- Browse the R Markdown files directly on GitHub to see code and outputs
- Each section can be viewed independently with rendered visualizations