Applied Data Analysis for Chinese Lending Data Using R & LLMs

Teal Emery

2025-01-14

Table of contents

Preface

This textbook accompanies the intensive month-long course **Applied Data Analysis For Chinese Overseas Lending Data Using R & LLMs** developed for AidData research analysts. The course combines four weekly 90-minute online sessions with a full day of in-person instruction, designed to equip analysts with powerful new tools for data analysis, visualization, and automation.
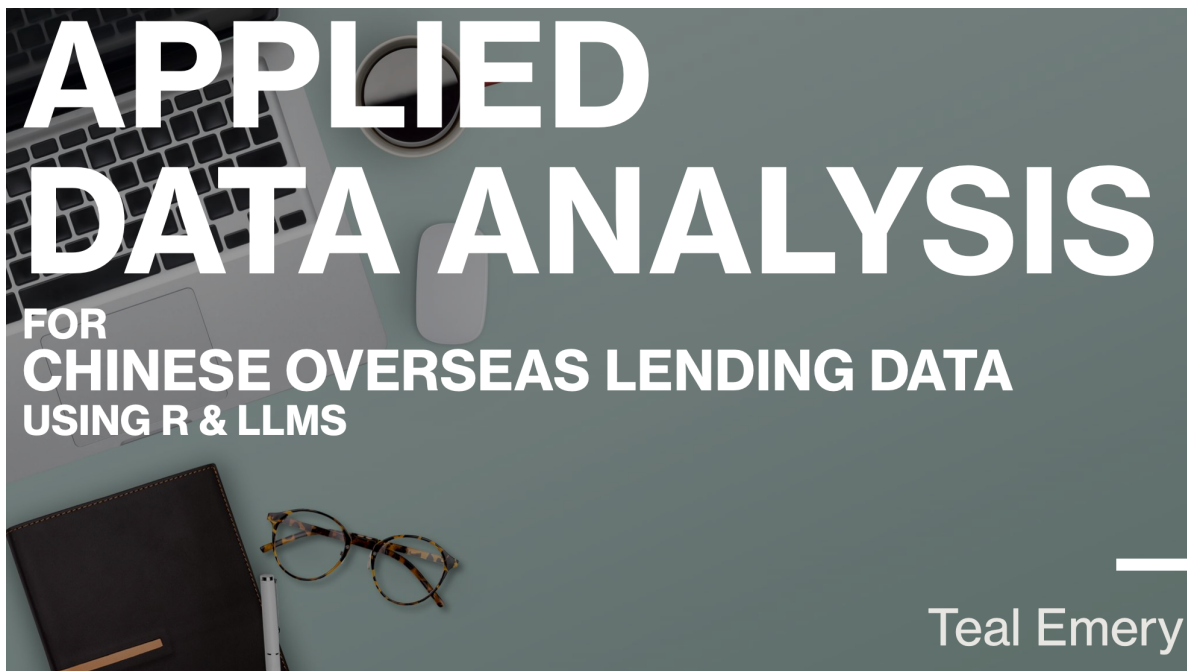
Our Promise to You

By the end of this course, you will be able to:

- Transform complex datasets into compelling visual stories
- Automate those repetitive tasks that consume hours of your week
- Create analyses that update automatically when new data arrives
- Generate professional reports that combine narrative, code, and visuals
- Leverage AI tools to enhance your analytical capabilities

- Solve common data challenges more efficiently

Our Teaching Philosophy

Think of R not as a programming language to master, but as a powerful toolkit that helps you tell stories with data. Just as you don't need to be a mechanic to drive a car effectively, you don't need to be a programmer to use R powerfully. This course focuses on giving you practical tools that will immediately enhance your analytical capabilities.

We embrace two key advantages that make this possible:

1. **Modern Tools**: The tidyverse ecosystem transforms R from a statistical programming language into an intuitive data analysis toolkit
2. **AI Assistance**: Large Language Models (LLMs) act as your personal guide, helping you find and implement the right tools for each task

What Makes This Course Different

Traditional R courses often get bogged down in programming concepts before getting to practical applications. We flip this approach:

- Start with practical tools you can use immediately
- Focus on real AidData challenges and solutions
- Use AI tools to overcome technical hurdles
- Build from practical application to deeper understanding

Week-by-Week Journey

Week 1: First Steps with R

- Set up R & R Studio on your computer
- Create your first data visualization
- Learn to use AI tools for coding assistance
- Begin working with Quarto for reproducible reports

Week 2: Data Visualization Mastery

- Create publication-ready plots with ggplot2
- Master the grammar of graphics
- Build interactive visualizations
- Design effective data presentations

Week 3: Data Transformation

- Clean and reshape real development finance data
- Master key data manipulation verbs
- Replicate analyses from AidData reports
- Automate repetitive data tasks

Week 4: Advanced Topics

- Handle complex data cleaning challenges
- Create reproducible workflows
- Generate automated reports
- Build functions for common tasks

In-Person Session

- Work on your own projects with expert guidance
- Tackle advanced visualization challenges
- Learn to extract structured data from text using AI
- Build confidence through hands-on practice

Being Realistic About Learning

You won't become an R expert in four weeks—and that's okay. What you will achieve:

- Master enough R to make your daily work easier and more efficient
- Get past the steepest part of the learning curve
- Build confidence in your ability to learn more
- Develop a foundation for continued learning

Your AI Learning Assistant

Learning R in 2025 is fundamentally different from even a few years ago. Modern AI tools serve as 24/7 tutors that can:

- Explain complex code in plain English
- Help debug your problems
- Suggest improvements to your code
- Answer your questions any time

Think of these AI tools as having a knowledgeable colleague always ready to help—they won't do the work for you, but they'll help you learn faster and overcome obstacles more efficiently.

How to Use This Book

This book serves multiple purposes:

1. A reference during the course
2. A guide for self-paced learning
3. A resource for future consultation

Each chapter includes: - Clear learning objectives - Practical examples using real data - Exercises to reinforce learning - Tips for using AI tools effectively - Resources for deeper learning

Let's begin this journey together. By the end of the course, you'll have new tools and skills to analyze Chinese development finance data more effectively and efficiently than ever before.

1 Introduction

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

```
1 + 1
```

```
[1] 2
```

2 Summary

In summary, this book has no content whatsoever.

```r
1 + 1
```

```
[1] 2
```

References

Knuth, Donald E. 1984. "Literate Programming." *Comput. J.* 27 (2): 97–111. https://doi. org/10.1093/comjnl/27.2.97.