Introduction to Statistical Computing

Susan Vanderplas

Spring 2022

Contents

Preface

This book is designed to demonstrate introductory statistical programming concepts and techniques. It is intended as a substitute for hours and hours of video lectures - watching someone code and talk about code is not usually the best way to learn how to code. It's far better to learn how to code by ... coding.

I hope that you will work through this book week by week over the semester. I have included comics, snark, gifs, YouTube videos, extra resources, and more: my goal is to make this a collection of the best information I can find on statistical programming.

In most cases, this book includes **way more information** than you need. Everyone comes into this class with a different level of computing experience, so I've attempted to make this book comprehensive. Unfortunately, that means some people will be bored and some will be overwhelmed. Use this book in the way that works best for you - skip over the stuff you know already, ignore the stuff that seems too complex until you understand the basics. Come back to the scary stuff later and see if it makes more sense to you.

How to Use This Book

I've made an effort to use some specific formatting and enable certain features that make this book a useful tool for this class.

Special Sections

Watch Out

Watch out sections contain things you may want to look out for - common errors, etc.

Examples

Example sections contain code and other information. Don't skip them!

My Opinion

These sections contain things you should definitely not consider as fact and should just take with a grain of salt.

Go Read

Sometimes, there are better resources out there than something I could write myself. When you see this section, go read the enclosed link as if it were part of the book.

Try It Out

Try it out sections contain activities you should do to reinforce the things you've just read.

Learn More

Learn More sections contain other references that may be useful on a specific topic. Suggestions are welcome (email me to suggest a new reference that I should add), as there's no way for one person to catalog all of the helpful programming resources on the internet!

Note

Note sections contain clarification points (anywhere I would normally say "note that ....)

Expandable Sections

These are expandable sections, with additional information when you click on the line

This additional information may be information that is helpful but not essential, or it may be that an example just takes a LOT of space and I want to make sure you can skim the book without having to scroll through a ton of output.

Many times, examples will be in expandable sections

This keeps the code and output from obscuring the actual information in the textbook that I want you to retain. You can always look up the syntax, but you do need to absorb the details I've written out.

About This Book

This is a Quarto book. To learn more about Quarto books visit https://quarto.org/docs/books.

I have written this entire book using reproducible techniques, with R and python code and results included within the book's text.

Stat 151 will be offered for the first time in Spring 2022, as I'm writing this in Fall 2021. Initially, my goal is to write the book in R and include python as an additional option/example. Eventually, I hope to teach Stat 151 in R and Python at the same time.

References