Chapter 2: Learning The “Hello World” Of Security Data Analysis

“from one thing, know ten thousand things”

― Miyamoto Musashi, The Book of Five Rings: Miyamoto Musashi

If you’ve ever tried to learn a new programming language there’s a good chance you started of with a “Hello World” example that quickly introduces basic language structure and code execution. The immediate sense of accomplishment as the syntax is verified by the compiler/interpreter and the familiar two-word output is displayed becomes a catalyst for the notion that, soon, you shall have the ability to bend this new language to your will.

This chapter takes the “Hello World” concept and expands it to a walk-through of a self-contained, introductory security data analysis use case that you will be able to follow along and execute and take concepts from as you start to perform your own analyses. There are side-by-side examples for both Python and R to give you an idea of the similarities, strengths and differences between both languages in a real life example context. If you’re not familiar with one or both of those languages you should read Chapter 1.5 first and at least skim some of the external resources it references. Remember, all the source code, sample data and visualizations are on the book’s web site, so no need for transcription, just focus on the flow of the analyses.

Getting Data

Reading In Data

Exploring Data

Augmenting Data