Chapter 1: Start off with Jewish teaching parable: wrap the truth in a story and doors open.

This has two points:

* + 1. We first need to find truth
    2. Make it digestible to the receiver (what we mean when we talk about data storytelling

We can learn from our environment (find the truth) by gathering data

Dataviz is a relatively simple method of analysis

we can find the truth with and without data, but we have a much better chance of learning with it and we have a better chance of discovering when we are wrong.

Our own stories about the “aha” moments when dataviz.

Chatper 2: Learning from our environment

Bill James: why we are ignorant and a career spent finding and asking questions that have objective answers. We don’t just want to find a question to ask, we want to find a question to ask that we have a shot of answering. (example: adding rogue devices onto our networks, “how many have we seen?” “what was the impact of those?”)

DIKW: Environment to better decisions. In a simple environment, our best opinion is probably fine, but that tipping point in complex environments requires more. Discuss DIKW and the theories of going from our environment to “wisdom”

**Intuition**: simple environments

**Statistics**: a tool built for this purpose, it applies to every field of science and we cannot live without it.

**Data viz**: we can do a really good job of faking statistics with dataviz

Chapter 3: Making Sense of Data

The ABC’s of data viz

Beyond the ABC’s of data viz (time, space, ordered, etc)

Chapter 1: Who gives a crap? (what problem are we solving, why should anyone care at all?) Do we need to learn about (and from) data? What is the potential upside? Can we continue on as is? Why would anyone ever change from what they are doing? (any stories of wasted efforts? maybe TGT USB stick)

How do we change over time? How can we change this game of cat and mouse (actually, more like boxer and punching bag)

When did the power of visualization hit you? What changed that made you realize it was worth every effort?

For me: Appendix A was the tipping point, addressing the intelligent adversary with the 3-d bar chart, before that, it was the cumulative probability distribution from FAIR. Maybe discuss the cover art. (cover could be that one dataviz that really hits home for us the power of viz)

Necessary Skills for being a data scientist: if you’re already in I.T or security you probably are already a lifelong learner. it changes so rapidly we could not have it any other way. Data analysis and dataviz is a language, you can learn the semantics but it isn’t until you attempt to speak it and use it as a common lanauge that you’ll really learn it.

Why do it? Within I.T. Security, there is a never ending list of things to fix and audit. There is no shortage of controls we can implement and the industry keeps coming up with better and more controls to implement. (cite the nuclear commission about the runaway costs of security). From a business stand point, we want to know which controls we can skip. We’ve got at least 100 different controls, and we can do 5 (or 20 or 50), so which do we implement and which do we skimp on? In short we want to collect data to improve our data analysis.

chapter 2: you give a crap: learning from our environment

DIKW - what field of research is that from?

Methods of getting from Env, parse it through “Learning” and come out better than before.

Stats versus dataviz (stats and dataviz)

chapter 3: Making Sense of data

Walking with Wade Baker and a huge sign for big data something or other was in front of us and he shook he head at the topic, and said “do you realize how much we have yet to learn from the little data?” Not only was that the most insightful statement I heard the entire week, he also encapsulated this entire chapter in a single breath.

The ABC’s (BAC) of data analysis

Exploratory versus explanatory analysis

Bill James: spent 40 years searching for and asking questions that have objective answers. In other words, don’t ask, “what risks do I have?” ask “what risks have I seen?” and “do I think the future will look like the past?” It won’t, but we can see trends: attackers don’t want to get caught, they want to get as much as they can with a small investment (their lazy), they will generally stick with something that works over and over until they need to change.

How many X have we seen? Within orgs like ours, how many have X versus Y? (there is a question here to encapsulate risk)

Are there examples in history where a single aggregate point was not enough? Have we seen “data sharing” work in any other industry? Maybe medicine with the invention of the CDC? They aggregate the whole of reports and spot trends long before anyone else would be capable of seeing it.

Questions as a key point

Chapter 4: Getting the right data

lying with data

data collection, cleaning, normalizing