Chapter 1: Unleashing The Securing Power Of Data

Standing On The Shoulders Of Giants

Use cases showing brief examples how other fields moved towards being more data drivn, but with different emphases: battling 'expertise', transforming science and applying data to a gut-feel industry)

- Snow

- Fisher

- Demming

Developing the Skills Of A Security Data Scientist

- Curosity

- Domain (security) knowledge

- Programming

- Data munging & management

- Statistics

- Visualization

Recognizing The Importance Exploratory Data Analysis

- Choosing A Good Research Question

\* "What is the…"

\* "Are there similarities…"

\* "Are there differences…"

- Exploratory Security Data Analysis : The Zero Access Botnet

\* Source data: "So what?"

\* Posing a good Question

\* The quest for the answer

WEB CONTENT: ZeroAccess data, visualizations

Chapter 2: Learning The "Hello World" Of Security Data Analysis

Reading in data

Getting a feel for the data

Basic Stats

Basic Graphs

Web scraping

Augmenting your analysis with different, relevant data

Asking a basic question of the data

Which country has the most malicious hosts of type “x”

WEB CONTENT: AlienVault data (w/link to source), python code, R code, visuals

Chapter 3: Analyzing “Badness”

Chapter Use Case: Crunching AlienVault IP Reputation Data

What is an “IP Address”?

32-bit integer (“how does your computer see an IP address?”) + machine info

Part of a subnet / logical layout / MAC addresses

Perhaps has a hostname (DNS)

Larger context: part of a global network organized by ASNs (BGP)

Lager context: Has a physical location

Mapping Outside The Continents

USE CASE: Visualizing AlienVault ASN data

WEB CONTENT: AlienVault data (w/link to source), python code, R code, visuals

Chapter 4: Mapping “Badness”

Chapter Use Case: ZeroAccess Botnet Analysis

What can you learn from just a set of lat/lon pairs?

Getting basic lat/long metadata (lat/lon -> country/city)

Visualizing lat/long data

Choropleth

Dot plot

Getting more advanced metadata (internet user population & income)

The quest for correlation

WEB CONTENT: ZeroAccess code & visuals

Chapter : Visualizing Your Security Data : The Fundamentals

Understanding The Foundations Of Good Visual Communication

- Focused overview of the basics & challenges of human visual perception with references to other books for deeper investigation

- USE CASE: Improving visual defaults in [Excel|Python|R] to enhance communication

Moving From Tables And Spreadsheets To A More Visual Medium

- Why we use tables (and how to use them better)

- Visualizing tabular data

USE CASE: Visualizing Firewall Data

Use this use case to run through basics and then comparisons of core (bar/dot/line/pie) charting techniques

WEB CONTENT: [visual defaults use case] Excel files, R code, python code

WEB CONTENT: Firewall data set, R code, visuals

Chapter : Getting A Handle On Your Security Data With Descriptive Statistics And Descrptive Visualization

Describing Attributes of IP Addresses Over Time

USE CASE: full packet capture analysis (tcpdump) of opportunistic network traffic

- How to use descriptive statistics in a security context

\* Univariate analysis (distributions, central tendency, dispersion, variance, standard deviation)

Re-Orienting Your Analyses With Visualizations

- Time-series

- Scatterplots

- Animation

WEB CONTENT: packet data, R code, scripts, visuals, animations

Chapter : Visualizing Your Security Data : Advanced Techniques

Understanding The Challenges Of Visualizing Lots Of Data

USE CASE: Visualizing Complex Networks To Understand Them Better

Radial Graphs

Force-directed Graphs

Chord Diagrams

Hive Plots

USE CASE: Visualizing Data Breaches (Multivariate Data Analysis)

WEB CONTENT: network data, R code, visuals

WEB CONTENT: breach data, R code, visuals

Chapter : Finding The Order In Chaos

The Challenges Of Security Information Sharing

- Cost-per-datum challenge

Inferential Estimations

- Confidence intervals, fisher-test and chi-square for independence

Comparisons To Other Methods

- Strengths and limitations

Chapter : Breaking Up With Your Relational Database

Realizing The Container Has Constraints

Managing Non-Relational Data (Saying 'Yes' to NoSQL)

Exploring Alternative Data Stores

USE CASE: "Have we seen this IP address?"

- practical example of how a traditional monolotic approach can hinder use of critical threat intelligence and how re-thinking how you intake, crunch and store data can open up new possibilities

WEB CONTENT: sample code for the use case

Chapter : Having The Machine Learn For You

De-mystifying Machine Learning

- Will discuss the surprisingingly straightforward underpinnings of ML and setup the rest of the chapter

Understanding The Security Potential of ML

Unsupervised Learning: Clustering Host Activity

- Applying MDS, KNN & K-Means techniques to security data

- USE CASE: Predicting potential rogue behaviour with security data

Supervised Learning: Classifying Host Activity

- Using logistic regression / random forests to detect network intrusions

Chapter : Building Dynamic Security Dashboards

Designing Dashboards For Effective Security Response

- Dashboards are a call to action

- Making differences stand out

- The never-ending quest for "so what?"

Applying Appropriate Visualizations To Your Security Data Streams

- Knowing when and how to use line graphs, bar charts, maps, etc

Understanding The Importance of Baselines And Thresholds

Communicating With Dashboards

- There is no One Security Dashboard to rule them all; designing security dashboards

USE CASE: The Incident Response Manager's Dashboard

- At-a-glance overview

- current incidents in play

- broad view of incidents over time

- insight about the incident handler team members

- etc …

USE CASE: The Threat Management Dashboard

- Tactical overview of internal threat landscape

- Communication attempts (success/fail) to "badness" (i.e. Matched IoCs)

\* to IPs/ASNs

\* from what (servers/workstations)

\* possibly geo-located

- Successful malware infections trends w/emphasis on known 0-days

- etc …

Chapter : Keeping It Simple

Putting Security Data Analysis Into Perspective

Comparing A "Drilling For Oil" Approach To a "Pan For Gold" Approach

Understanding The Reality Of Our Environments

Re-iterating That Data Analysis Assists Our Thinking, Not Replaces It

What Lies Ahead In Security Analytics?