Title of Your Project: 2013 M3 - Big Marketing

Mini Challenge Number: 15

Team Name: 6

Your Names and UT IDS:

Alex Hoang (agh2398)
Daniel Miao (dm52663)
Jack Hu (hh27683)
Kelly To (kt24854)

Analysis: *Any changes that you made to your analysis*

Our team did not make any changes to the analysis from the previous submission.

Design: *Visualization packages that were used and interactive visuals that were created (given in bulleted form)?*

- Matplotlib.pyplot
    - Average Session Durations Each Day
    - Hostname Usage Over Week 1 & 2
- Seaborn
    - Counts of Flags by Day and Hour
- Plotly.express
    - Frequency of Top 10 Source IPs each hour in NF for week 1
    - Frequency of Top 10 Destination IPs each hour in NF for week 1
    - Number of Unique Destination Ports each hour in NF for week 1
    - Maximum Outbound Payload Size Each Hour for Week 1 From Admin IPs
    - Total Number of Flows per Day
    - Total Number of Flows per Minute
    - Top 10 SrcIp per Hour
    - Top 10 SrcIp per Hour Choropleth
    - Top 10 destIp each hour in IPS for week 2
    - Top 10 messageCode each hour in IPS for week 2

    - Top 10 messageCode each minute between 2013-04-11 11:00:00 and 2013-04-11 13:00:00
    - Number of Unique destPort each hour
    - Port 80 being used each hour
    - Port 3389 being used each hour
    - Port 25 being used each hour
    - Number of times destPort 22 is used each hour
    - Number of Port 22 in Use
- Networkx
    - Connections Between Ports in IPS at 12 PM on 04-12-2013
- Plotly.graph_objects
    - Connections Between Ports in IPS at 12 PM on 04-12-2013

Strengths: *What do you think are the strengths of your project (give in bulleted form)?*

- Abundance of datasets to create visuals from helped identify trends and patterns
- Offers an opportunity to develop technical skills in network analysis tools and data interpretation
- Helps learn how networks operate, providing a deeper understanding of network protocols, traffic patterns, and anomalies
- The skills acquired in this project are highly relevant to various career paths in IT, cybersecurity, network administration, or data analysis

Challenges: *What were the challenges that you faced (give in bulleted form)?*

- Network terminology may require external research to understand dataset and how to create visuals
- Massive amounts of data that was overwhelming to process and analyze
- Lack of suitable hardware caused the whole dataset to take 10-15 minutes to completely load

- Understanding the data visuals required a lot of interpretation and understanding of networks in order to identify anomalies
- Determining the significance of an event is challenging because not all anomalies are threats, which requires an understanding of the network's normal operations

Advice: *What advice would you give the next generation of students doing this project?*
The biggest piece of advice the group would like to give the next generation of students is to be proactive in learning the dataset. The resources provided will help guide the students to understand the dataset and gain a better understanding of the challenge at hand. Also, look into the solution and its resources. This helps fit together the big picture to understand the conclusion the students are striving towards. The last piece of advice would be to make sure the students start on the project early and work at a constant pace. It is a lot of work and if left to the last minute, will not be fun to endure.