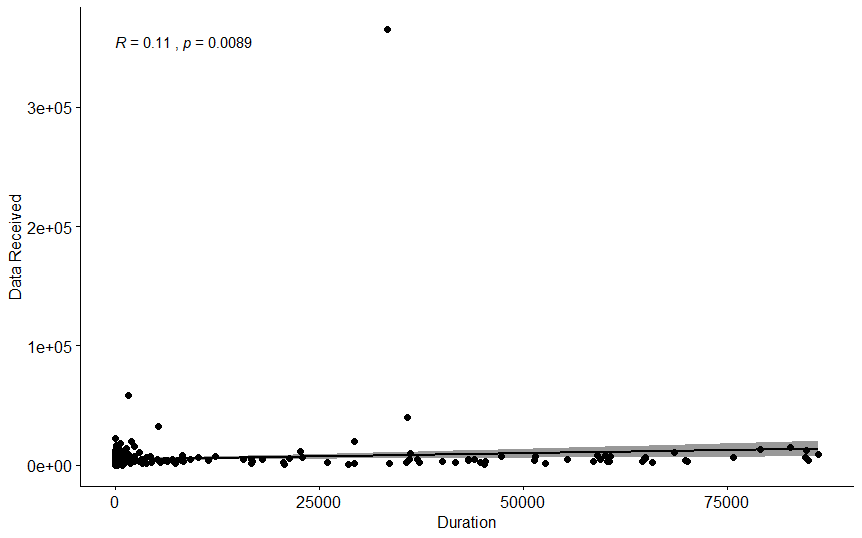
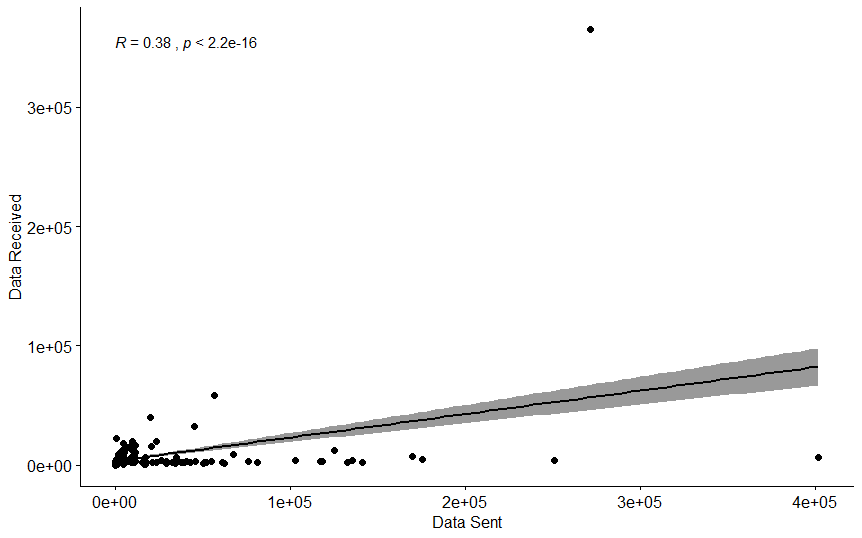
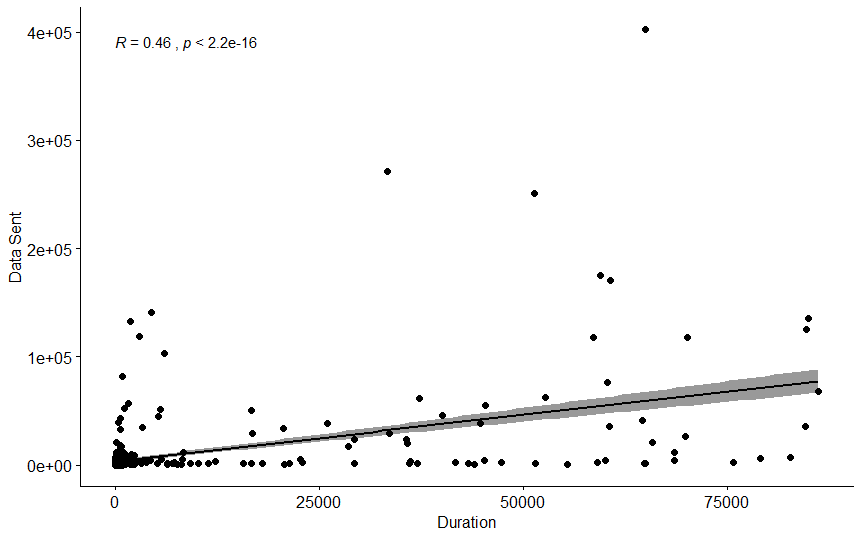
1. How many unique server IPs do you see? How many unique client IPs?
   1. Done By checking all SYN packets, put them in a file and then checking unique server and client from that.
2. How many unique TCP flows do you see? Hint: A TCP flow is
   1. Got all lines having TCP ACK, and then identified flow uniquely by its source IP, destination IP, source port, and destination port.

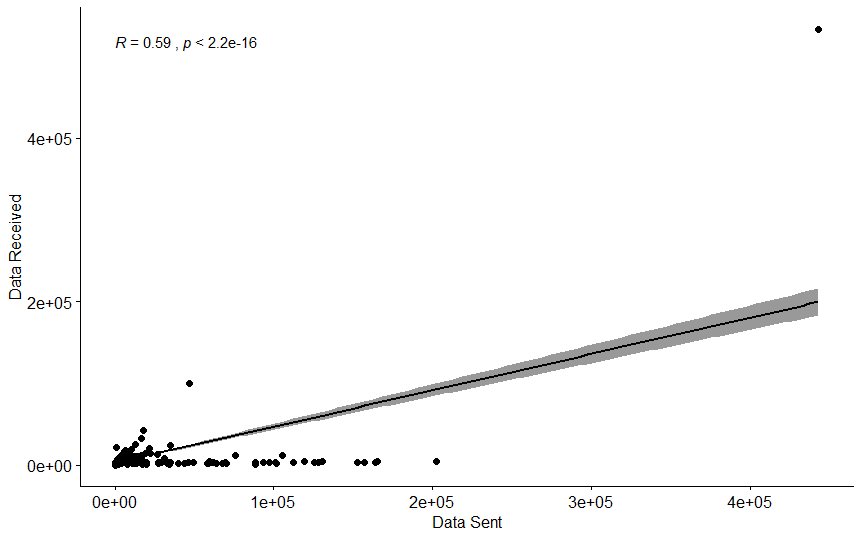
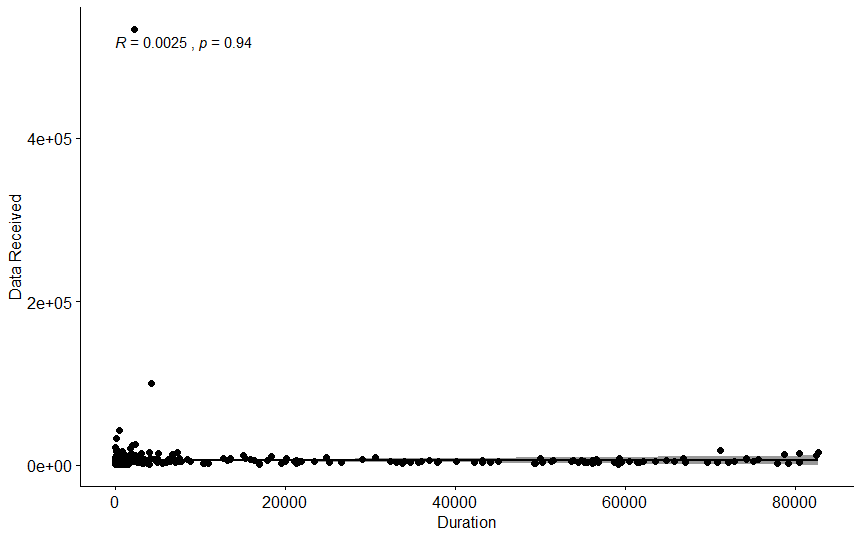
|  |  |  |  |
| --- | --- | --- | --- |
|  | lbnl.anon-ftp.03-01-11.csv | lbnl.anon-ftp.03-01-14.csv | lbnl.anon-ftp.03-01-18.csv |
| No. of clients | 522 | 939 | 510 |
| No. of servers | 45 | 50 | 89 |

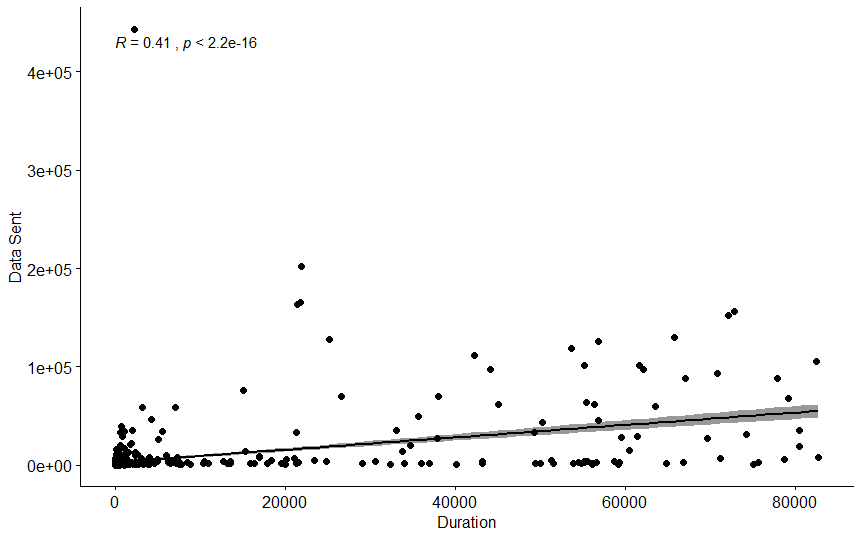
1. Draw a plot of the number of connections (TCP flows) opened to any FTP server within 60min windows over the 24-hour duration, ie. within hours 00:00-01:00, 01:00-02:00, 02:00-03:00, etc. How can you use these traffic profiles to detect if the system is under a DoS attack?
2. For all the connections, find the duration over which a connection was kept open, and plot the CDF (cumulative distribution function) of these connection durations.
   1. Notice how most connections are of a short duration. Why do you think so?
3. On the same lines as above, for all the connections find the number of bytes sent and received over the connection, and check if there is a correlation between the connection duration (as computed in question 4 above, ie. the duration over which the connection was kept open) and the number of bytes sent to the servers. Similarly check if there is a correlation between the number of bytes sent to the servers and number of bytes received from the servers.
   1. Do you find that there is a correlation? Does the scatterplot help you identify any outliers? If you eliminate the outliers, does your correlation improve?

Day1: Plots





Day2: Plots



Day3: Plots