Using R in Security Scenarios

Wilson L. Chua

Why R?

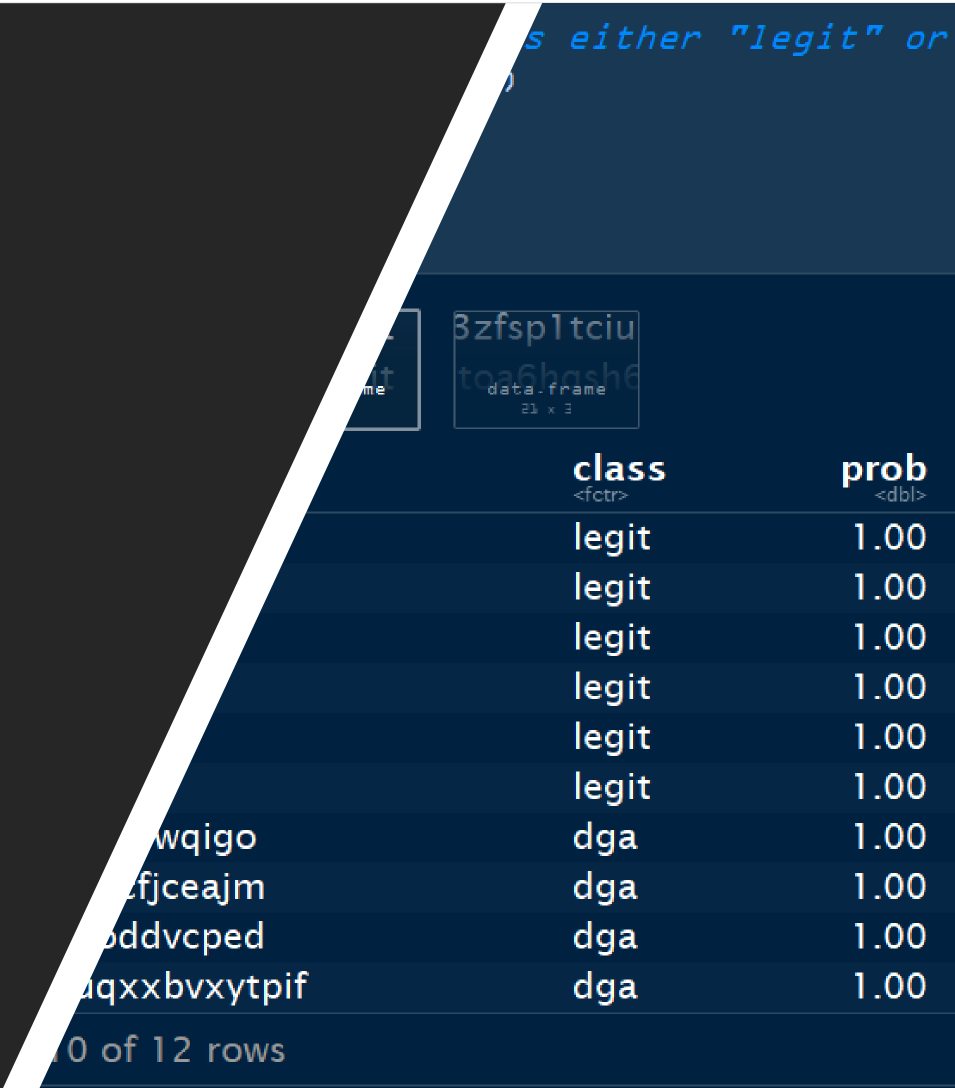
- Powerful yet Free
- Large library for data manipulation, extraction and visualization
- R Notebook and R Markdown output formats include html pages

Security Scenarios

- DGA detection using dgapredict()
- Common Attack detection
- APT activity detection
- Data Exfiltration Attempts
- IoT attacks
- Brute Force Attacks
- Shodan Integration

Malware, C&C & DGA

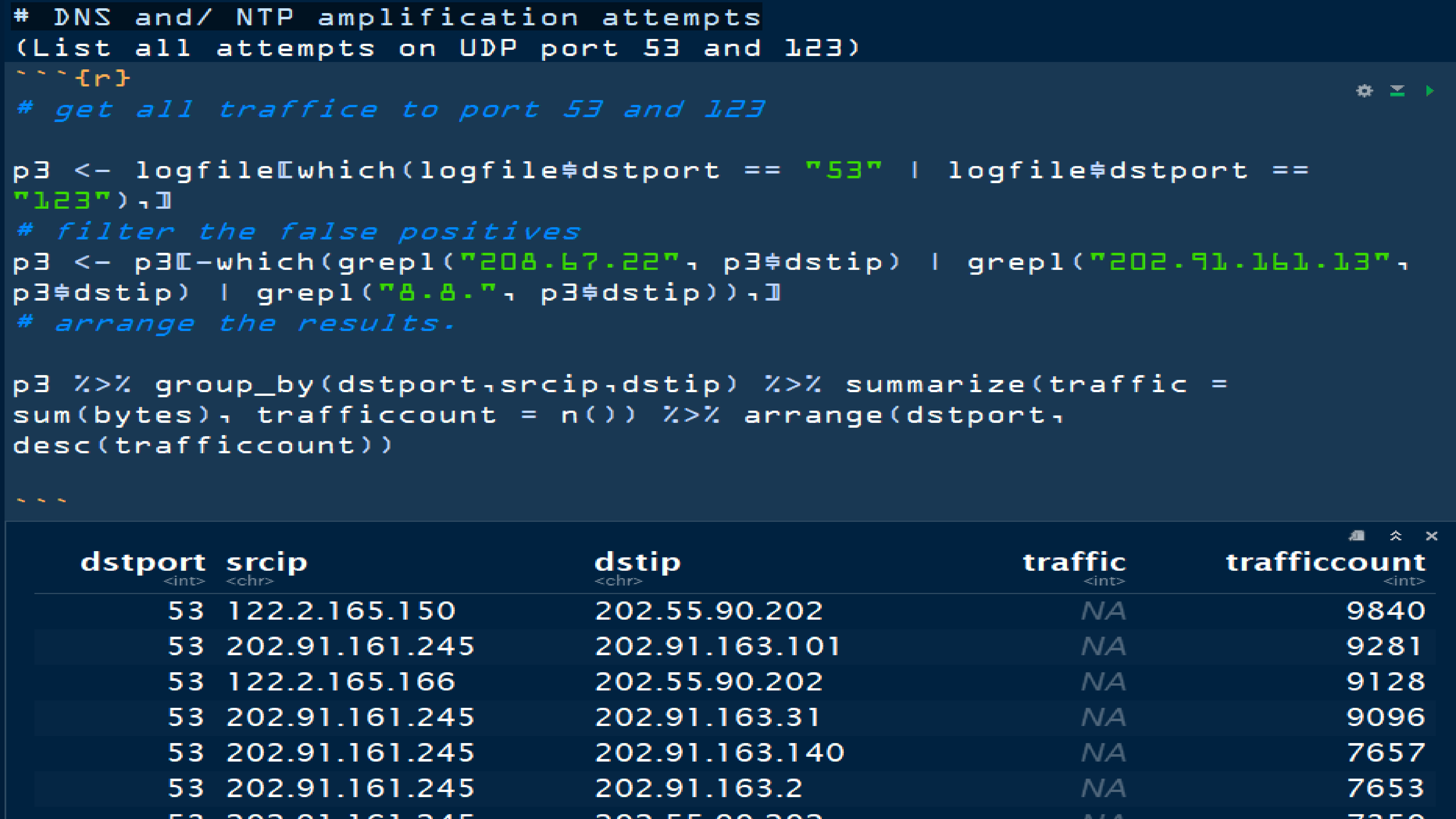
- Dead Giveaway of Malware infected PCs
- Detection based on DNS res req with
  - Low dictionary word matches
  - Low vowel to consonant ratios
  - High DNS Nxerrors
- SrcIP with Highest NXERRORS are likely to be DGA AND proof of infection
- Lookup Srcip NS request with NO errors
- Block the DGA NS with no errors



```
s either "legit" or
)

        3zfsp1tciu
        toa6hqsh6
me      data.frame
        21 x 3

                 class      prob
                 <fctr>     <dbl>
                 legit      1.00
                 legit      1.00
                 legit      1.00
                 legit      1.00
                 legit      1.00
                 legit      1.00
wqigo            dga        1.00
fjceajm          dga        1.00
ddvcped          dga        1.00
uqxxbvxytpif     dga        1.00
0 of 12 rows
```



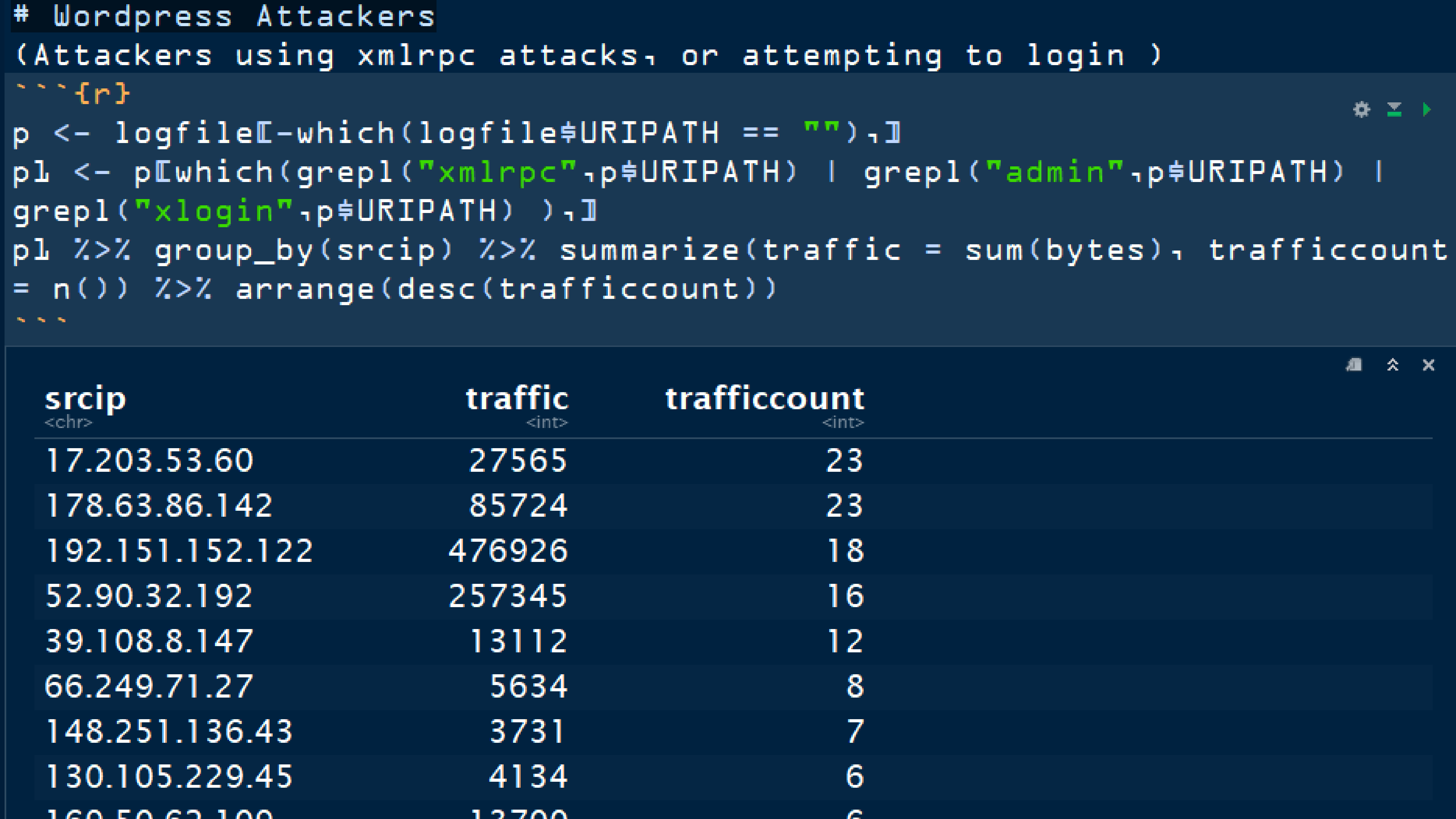
~~~
# DNS and/ NTP amplification attempts
(List all attempts on UDP port 53 and 123)
```{r}
# get all traffice to port 53 and 123

p3 <- logfile[which(logfile$dstport == "53" | logfile$dstport ==
"123"),]
# filter the false positives
p3 <- p3[-which(grepl("208.67.22", p3$dstip) | grepl("202.91.161.13",
p3$dstip) | grepl("8.8.", p3$dstip)),]
# arrange the results.

p3 %>% group_by(dstport,srcip,dstip) %>% summarize(traffic =
sum(bytes), trafficcount = n()) %>% arrange(dstport,
desc(trafficcount))

```
~~~

dstport <int>	srcip <chr>	dstip <chr>	traffic <int>	trafficcount <int>
53	122.2.165.150	202.55.90.202	NA	9840
53	202.91.161.245	202.91.163.101	NA	9281
53	122.2.165.166	202.55.90.202	NA	9128
53	202.91.161.245	202.91.163.31	NA	9096
53	202.91.161.245	202.91.163.140	NA	7657
53	202.91.161.245	202.91.163.2	NA	7653

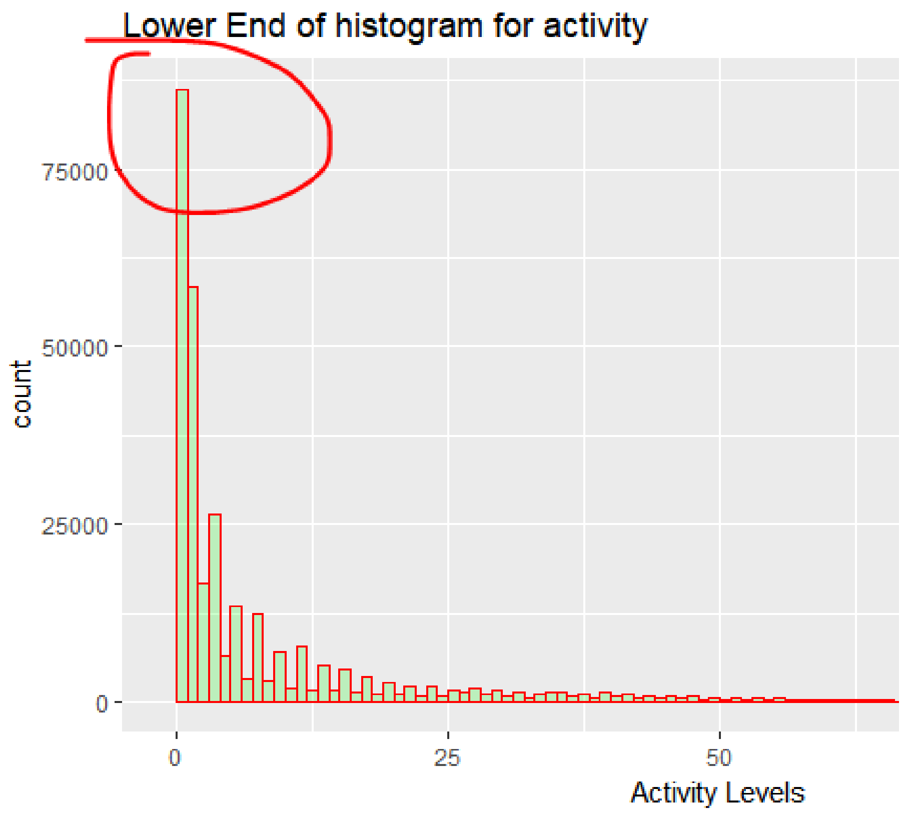


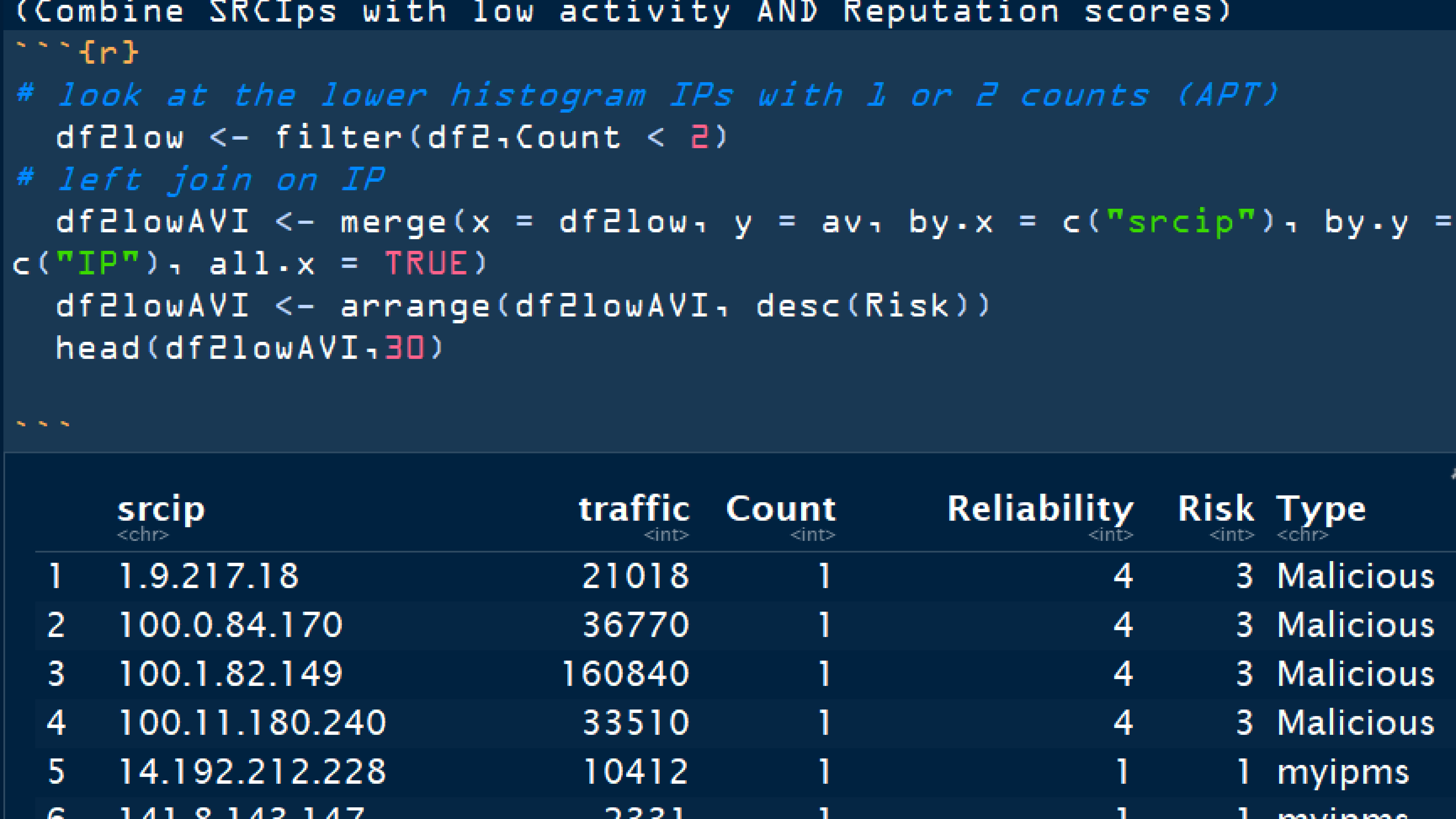
~~~
# Wordpress Attackers
(Attackers using xmlrpc attacks, or attempting to login )
```{r}
p <- logfile[-which(logfile$URIPATH == ""),]
p1 <- p[which(grepl("xmlrpc",p$URIPATH) | grepl("admin",p$URIPATH) |
grepl("xlogin",p$URIPATH) ),]
p1 %>% group_by(srcip) %>% summarize(traffic = sum(bytes), trafficcount
= n()) %>% arrange(desc(trafficcount))
```
~~~

```
srcip                 traffic   trafficcount
<chr>                   <int>          <int>
17.203.53.60            27565             23
178.63.86.142           85724             23
192.151.152.122        476926             18
52.90.32.192           257345             16
39.108.8.147            13112             12
66.249.71.27             5634              8
148.251.136.43           3731              7
130.105.229.45           4134              6
169.50.62.100           13700              6
```

APT Attack detection

- Low density traffic (count and bytes)
- Bad reputation
- Over days or weeks
- Histogram of lower end of traffic count





(Combine SRCIps with low activity AND Reputation scores)

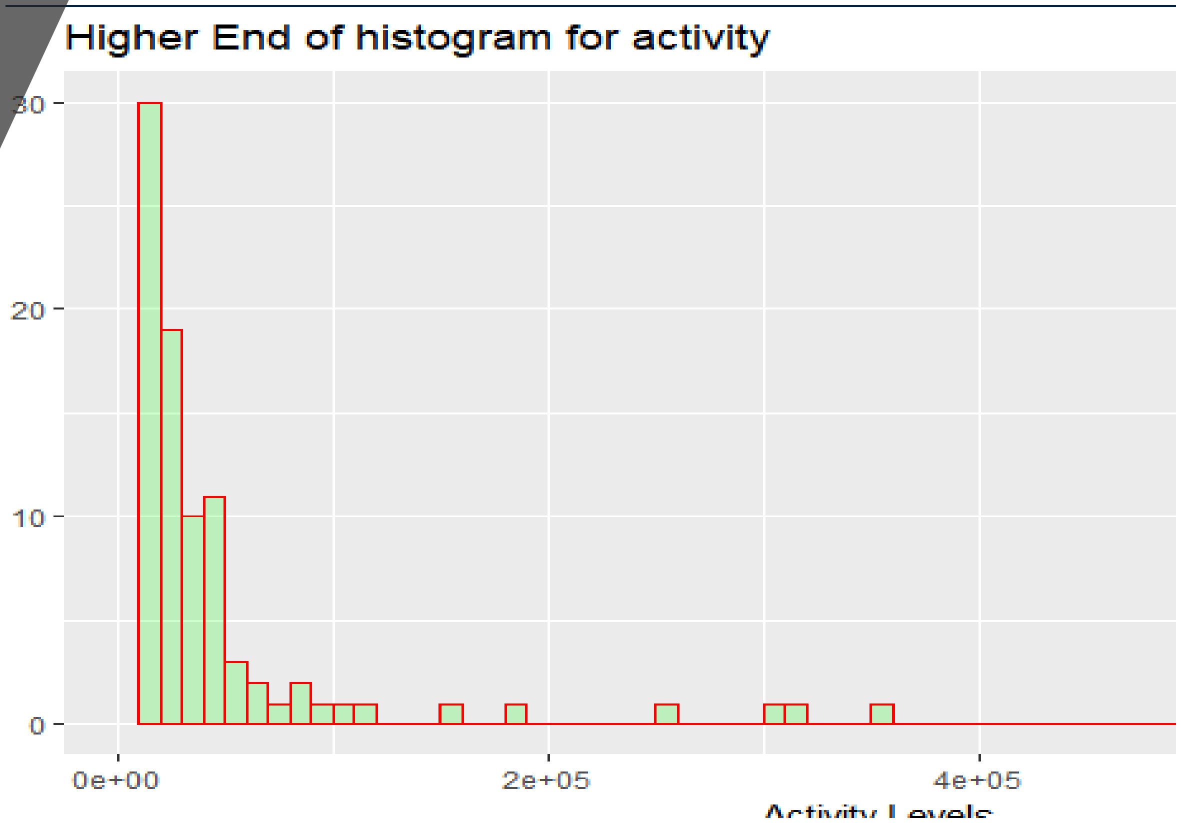
```{r}
# look at the lower histogram IPs with 1 or 2 counts (APT)
  df2low <- filter(df2,Count < 2)
# left join on IP
  df2lowAVI <- merge(x = df2low, y = av, by.x = c("srcip"), by.y =
c("IP"), all.x = TRUE)
  df2lowAVI <- arrange(df2lowAVI, desc(Risk))
  head(df2lowAVI,30)

```

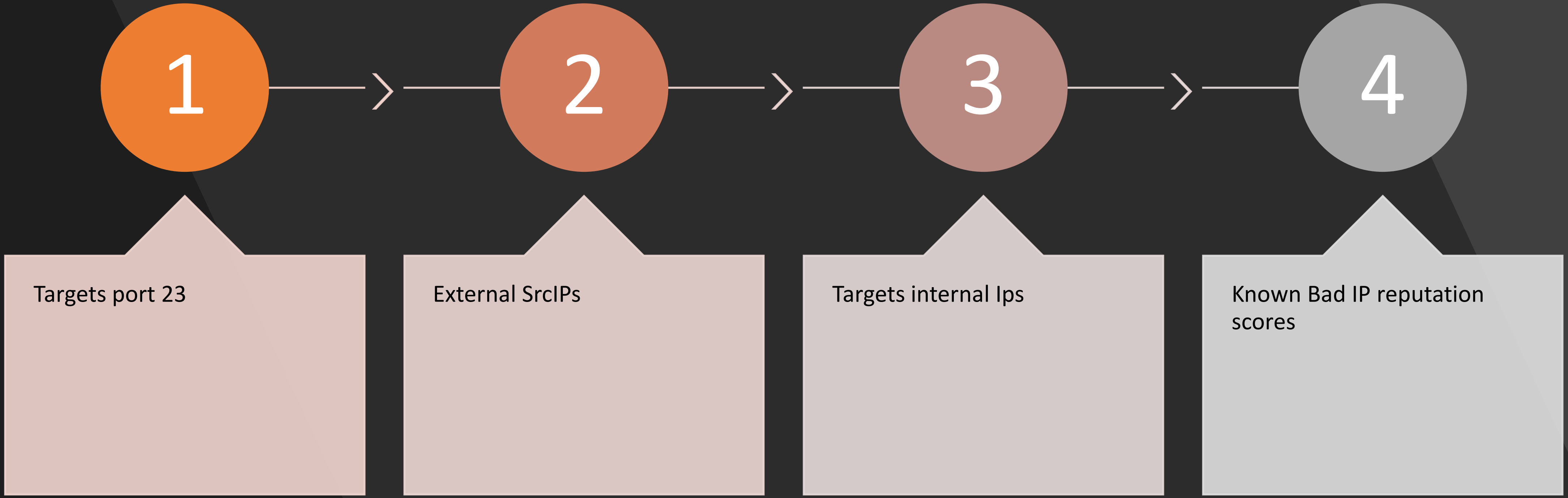
```
   srcip             traffic  Count      Reliability  Risk  Type
   <chr>               <int>  <int>            <int> <int>  <chr>
1  1.9.217.18          21018      1                4     3  Malicious
2  100.0.84.170        36770      1                4     3  Malicious
3  100.1.82.149       160840      1                4     3  Malicious
4  100.11.180.240      33510      1                4     3  Malicious
5  14.192.212.228      10412      1                1     1  myipms
6  141.8.143.147        2331      1                1     1  myipms
```

Data Exfiltration Attempts

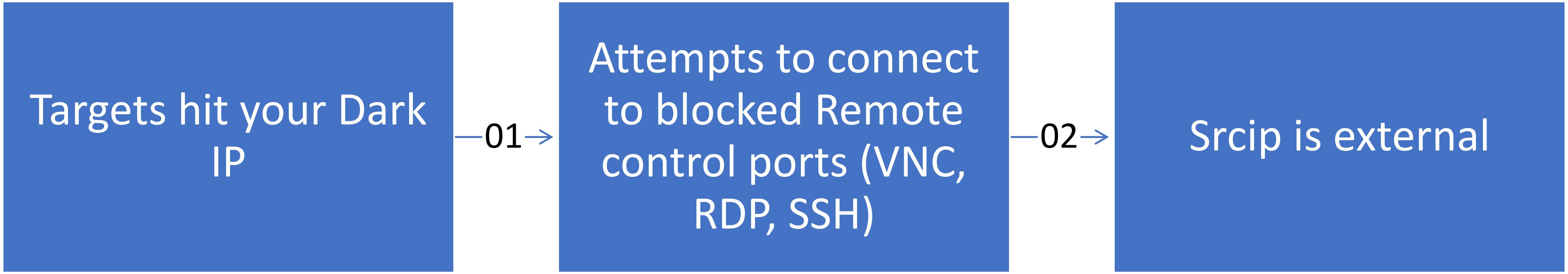
- High end of Traffic (bytes) histogram
- Source is INTERNAL IP
- Destination is EXTERNAL IP



IoT attacks



1. Targets port 23
2. External SrcIPs
3. Targets internal Ips
4. Known Bad IP reputation scores



Targets hit your Dark IP

01

Attempts to connect to blocked Remote control ports (VNC, RDP, SSH)

02

Srcip is external

Brute Force Attacks/script kiddies

Demo

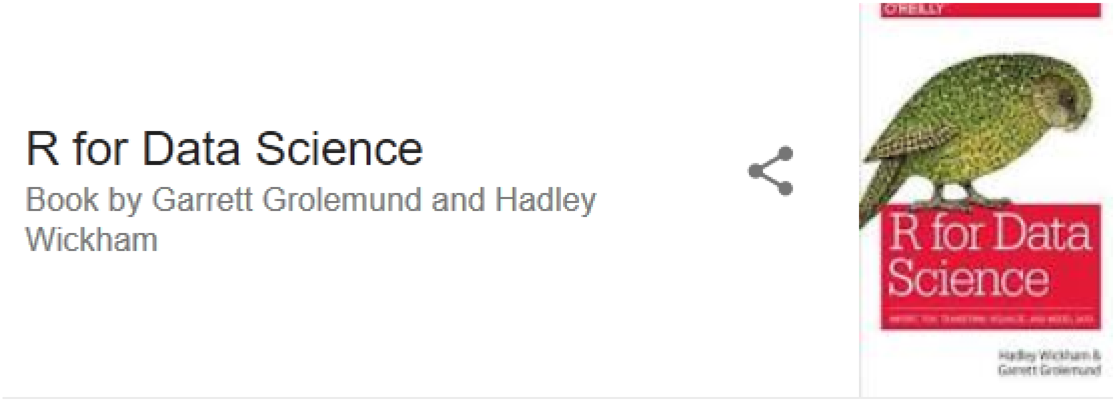
R and Shodan Integration

Credits

- Joseph Tabadero Jr for Maxmind integration
- Eric Reyata for help with DGA samples

Data-Driven Security: Analysis, Visualization and Dashboards

Book by Bob Rudis and Jay Jacobs



R for Data Science

Book by Garrett Grolemund and Hadley Wickham

Contact Info

- Wilson L. Chua
- wilson@futuregen.sg