

# Packet Count by Size across Protocols

Avery Holloman

2024-08-13

```
# Libraries
library(ggplot2)
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
```

```
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
# Loading the data from the specified file
data <- read_excel("C:/Users/jacob/OneDrive/Desktop/R Studio Projects 2024/Datasets/anomaly_detection_network_traffic.xlsx")

# Taking a summary of the packet counts by Packet_Size (log10 data transformation) and grouping by Protocol
data_summary <- data %>%
  mutate(Packet_Size_Log10 = log10(Packet_Size)) %>%
  group_by(Packet_Size_Log10, Protocol) %>%
  summarise(Packet_Count = n(), .groups = 'drop') # Ensures the group is dropped after summarizing

# Utilizing ggplot2 for better visibility
ggplot(data_summary, aes(x = Packet_Size_Log10, y = Packet_Count, color = Protocol)) +
  geom_line(size = 1) +
  scale_y_log10() +
  labs(x = "Packet Size (log10)", y = "Packet Count (log10)", color = "Protocol") +
  theme_minimal() +
  theme(
    text = element_text(size = 12),
    legend.position = "top",
    legend.title = element_blank(),
    plot.title = element_text(hjust = 0.5)
  ) +
  ggtitle("Packet Count by Size across Protocols")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
## i Please use `linewidth` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was  
## generated.
```

## Packet Count by Size across Protocols

