Correlation is a statistical measure that describes the extent to which two variables change together. It quantifies the strength and direction of the linear relationship between two variables. The correlation coefficient, typically denoted by r, ranges from -1 to 1.

if r =1 it indicates a perfect positive correlation, meaning that as one variable increases, the other variable also increases proportionally.

if r = -1 it indicates a perfect negative correlation, meaning that as one variable increases, the other variable decreases proportionally.

if r = 0 it indicates no correlation between the variables.

Correlation analysis is crucial in various fields, including cybersecurity, where it can help identify relationships between different factors or variables in network traffic data. For example, correlating the number of login attempts with the occurrence of security breaches can provide insights into potential attack patterns.

Code:

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load network traffic data
data = pd.read_csv("network_traffic_data.csv")

# Drop non-numeric columns or encode them properly
data_numeric = data.select_dtypes(include=['number'])

# Calculate correlation matrix
corr_matrix = data_numeric.corr()

# Plot heatmap of correlation matrix
plt.figure(figsize=(10, 8))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Correlation Heatmap of Network Traffic Data')
plt.show()
```

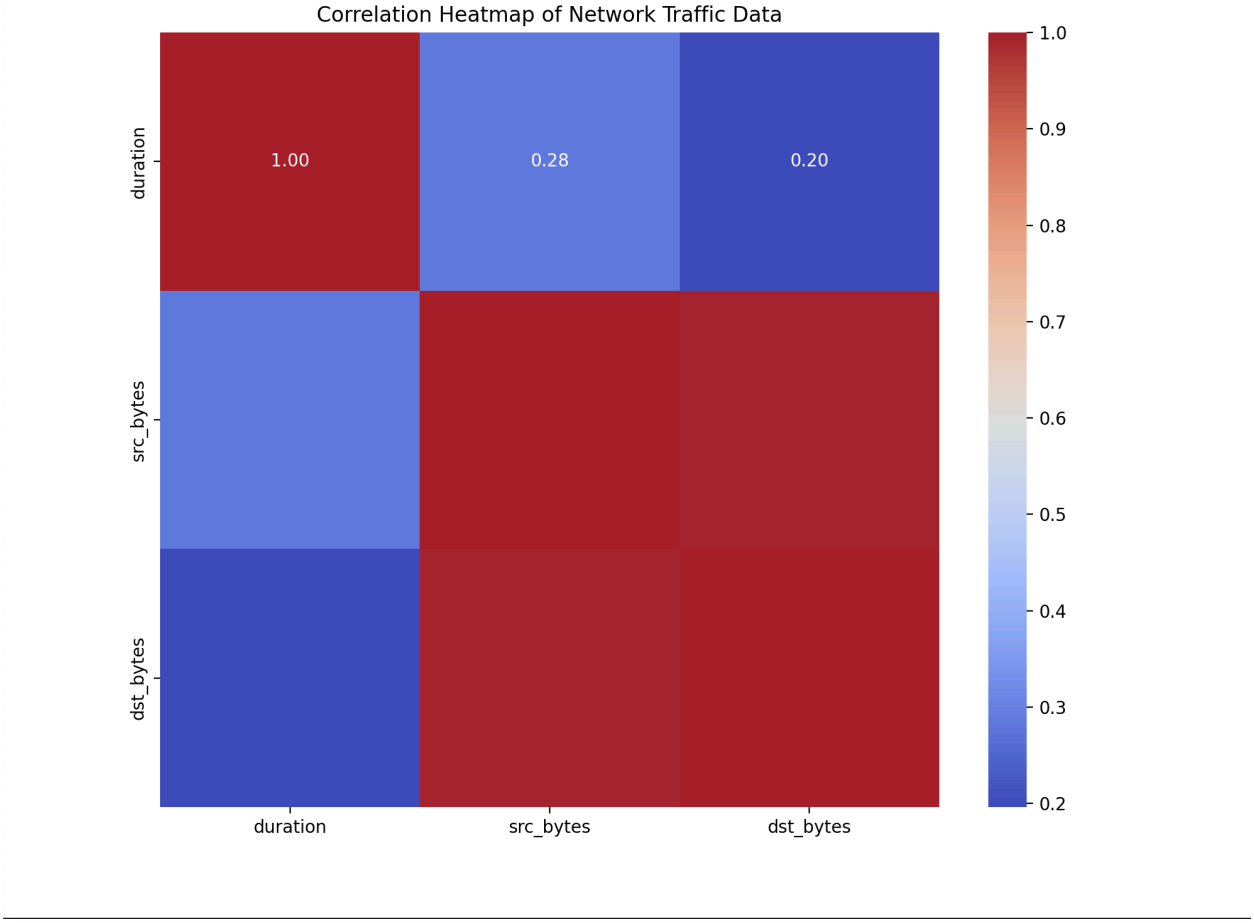
network_traffic_data.csv
Data:

```
duration,protocol_type,service,flag,src_bytes,dst_bytes,label
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
```

90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal

90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack
215,tcp,http,SF,184,0,normal
162,tcp,ftp,SF,2607,1977,normal
90,icmp,dns,SF,146,146,attack
78,udp,domain,SF,146,146,attack

Result:


Correlation Heatmap of Network Traffic Data
duration
src_bytes
dst_bytes
1.00
0.28
0.20
duration
src_bytes
dst_bytes
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2