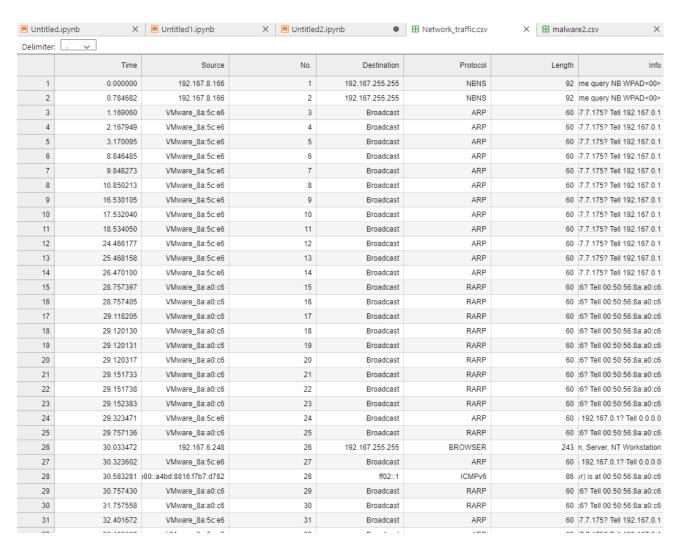
## Task 3

Correlation is a statistical measure that describes the relationship between two variables. It indicates the extent to which changes in one variable are associated with changes in another variable. Correlation quantifies the degree to which two variables move together in a systematic way.

There are several types of correlation measures, but the most commonly used is Pearson's correlation coefficient, which measures the linear relationship between two continuous variables. Other correlation measures include Spearman's rank correlation coefficient and Kendall's tau coefficient, which are used for non-parametric data or when the relationship is not linear.

Correlation analysis is widely used in various applications, including finance, marketing, health. Correlation provides valuable insights into patterns and associations in data, enabling informed decision-making in various fields.



# Importing required libraries

import pandas as pd

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

# Load the dataset
data = pd.read_csv("network_traffic.csv")

# Calculate the correlation matrix
corr_matrix = data.corr()

# Visualize the correlation matrix using a heatmap
plt.figure(figsize=(12, 8))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt=".2f", linewidths=.5)
plt.title('Correlation Matrix of Network Traffic Features')
plt.show()
```