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.

Untitled.ipynb | Untitled1.ipynb | Untitled2.ipynb | Network_traffic.csv | malware2.csv

Delimiter: ,

	Time	Source	No.	Destination	Protocol	Length	Info
1	0.000000	192.167.8.166	1	192.167.255.255	NBNS	92	me query NB WPAD<00>
2	0.784682	192.167.8.166	2	192.167.255.255	NBNS	92	me query NB WPAD<00>
3	1.169060	VMware_8a:5c:e6	3	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
4	2.167949	VMware_8a:5c:e6	4	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
5	3.170095	VMware_8a:5c:e6	5	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
6	8.846485	VMware_8a:5c:e6	6	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
7	9.848273	VMware_8a:5c:e6	7	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
8	10.850213	VMware_8a:5c:e6	8	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
9	16.530105	VMware_8a:5c:e6	9	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
10	17.532040	VMware_8a:5c:e6	10	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
11	18.534050	VMware_8a:5c:e6	11	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
12	24.466177	VMware_8a:5c:e6	12	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
13	25.468158	VMware_8a:5c:e6	13	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
14	26.470100	VMware_8a:5c:e6	14	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1
15	28.757397	VMware_8a:a0:c6	15	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
16	28.757405	VMware_8a:a0:c6	16	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
17	29.118205	VMware_8a:a0:c6	17	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
18	29.120130	VMware_8a:a0:c6	18	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
19	29.120131	VMware_8a:a0:c6	19	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
20	29.120317	VMware_8a:a0:c6	20	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
21	29.151733	VMware_8a:a0:c6	21	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
22	29.151738	VMware_8a:a0:c6	22	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
23	29.152383	VMware_8a:a0:c6	23	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
24	29.323471	VMware_8a:5c:e6	24	Broadcast	ARP	60	192.167.0.1? Tell 0.0.0.0
25	29.757136	VMware_8a:a0:c6	25	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
26	30.033472	192.167.6.248	26	192.167.255.255	BROWSER	243	n, Server, NT Workstation
27	30.323602	VMware_8a:5c:e6	27	Broadcast	ARP	60	192.167.0.1? Tell 0.0.0.0
28	30.583281	80::a4bd:8816:f7b7:d782	28	ff02::1	ICMPv6	86	r) is at 00:50:56:8a:a0:c6
29	30.757430	VMware_8a:a0:c6	29	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
30	31.757558	VMware_8a:a0:c6	30	Broadcast	RARP	60	6? Tell 00:50:56:8a:a0:c6
31	32.401672	VMware_8a:5c:e6	31	Broadcast	ARP	60	7.7.175? Tell 192.167.0.1

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()
```