**CHAPTER-8**

**Source code**

## Comprehensive Study of Cyber Attacks Using the UNSW-NB Dataset for Building Resilient Systems

# importing libraries:

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import warnings

warnings.filterwarnings('ignore')

# importing dataset:

data=pd.read\_csv('cybersecurity\_attacks.csv')

data.head()

## DATA PREPROCESSING:

data.info()

data.shape

data.describe()

data.nunique()

data['Log Source'].unique()

## DATA CLEANING:

data.isnull().sum()

data.duplicated().sum()

data.fillna(np.mean,inplace=True)

data.isnull().sum()

data['Log Source'].value\_counts()

## EDA(EXPLORATORY DATA ANALYSIS)

plt.figure(figsize=(15,6))

sns.countplot(x="Log Source",data=data,palette="Set1")

plt.xticks(rotation = 45)

plt.show()

## Result:

The above plot represents the count plot for the Log Source column of the dataset.Also we can see that the Firewall attacks are happening slightly more compare to the Server attacks.

plt.figure(figsize=(15,6))

sns.countplot(x="Protocol",data=data,palette="Set1")

plt.xticks(rotation = 45)

plt.show()

## Result:

The above plot represents the count plot of protocol column of thr dataset.Also it represents that the protocol column consists of three different classes.Each class consists of similar number of data items.

plt.figure(figsize=(15,6))

sns.countplot(x="Packet Type",data=data,palette="Set1")

plt.xticks(rotation = 45)

plt.show()

## Result:

The above plot represents the count plot of the packet type column of the datset.Also the above plot represents that there are two different classes in the packet type column.And also the control packet type consists of more data items.

plt.figure(figsize=(15,6))

sns.countplot(x="Attack Type",data=data,palette="Set1")

plt.xticks(rotation = 45)

plt.show()

plt.figure(figsize=(15,6))

sns.countplot(x="Traffic Type",data=data,palette="Set1")

plt.xticks(rotation = 45)

plt.show()

# Gender gap in internet usage by area

plt.figure(figsize=(10, 5))

sns.barplot(data=data, x='Packet Length', y='Packet Type', ci=None, palette='viridis')

plt.title('Bar plot between the packet type and length')

plt.xlabel('packet length')

plt.ylabel('packet type')

plt.show()

## Result:

The above plot represents the Bar plot between the type of packet and packet length.And also it represents that the control packet consists of more data points.

# Gender gap in internet usage by area

plt.figure(figsize=(10, 5))

sns.barplot(data=data, x='Log Source', y='Packet Length', ci=None, palette='viridis')

plt.title('Bar plot between the packet type and Log souce')

plt.xlabel('Log source')

plt.ylabel('packet Length')

plt.show()

## Result:

The above plot represents the Bar plot between the Log source column and Packet length columns of the dataset.Also it represents that the both classes of Log Source have almost consists of data points.

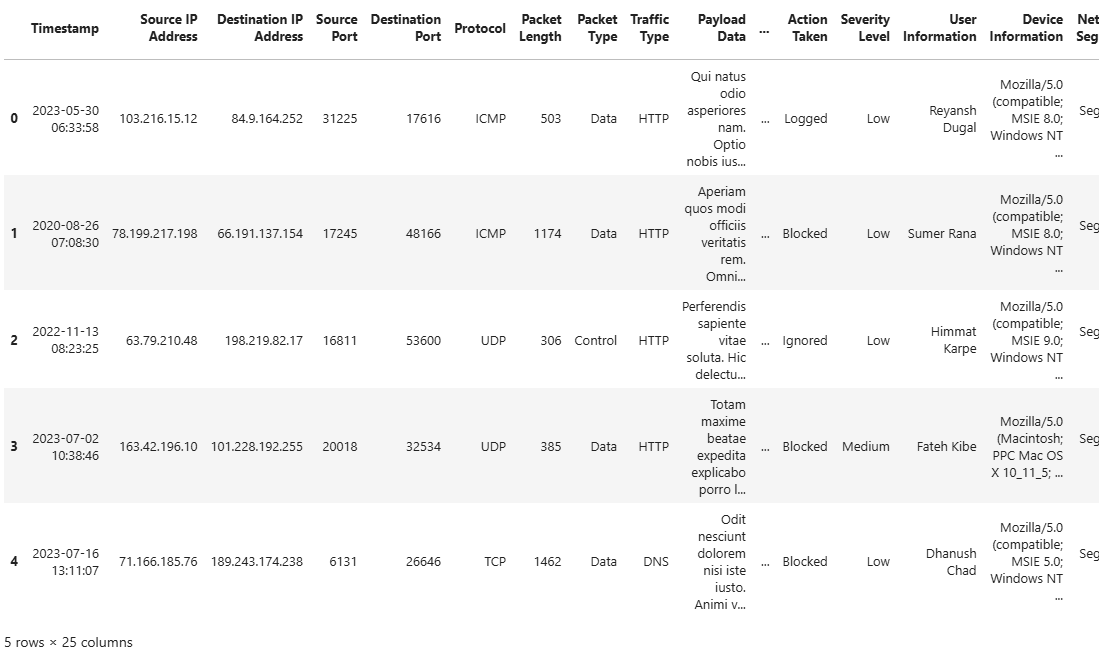
## CONCLUSION:

From the above plots we can conclude that the both the type of attacks are happeneing.Moreover the Firewall cyber attacks are happening more comparerd to the server attacks.

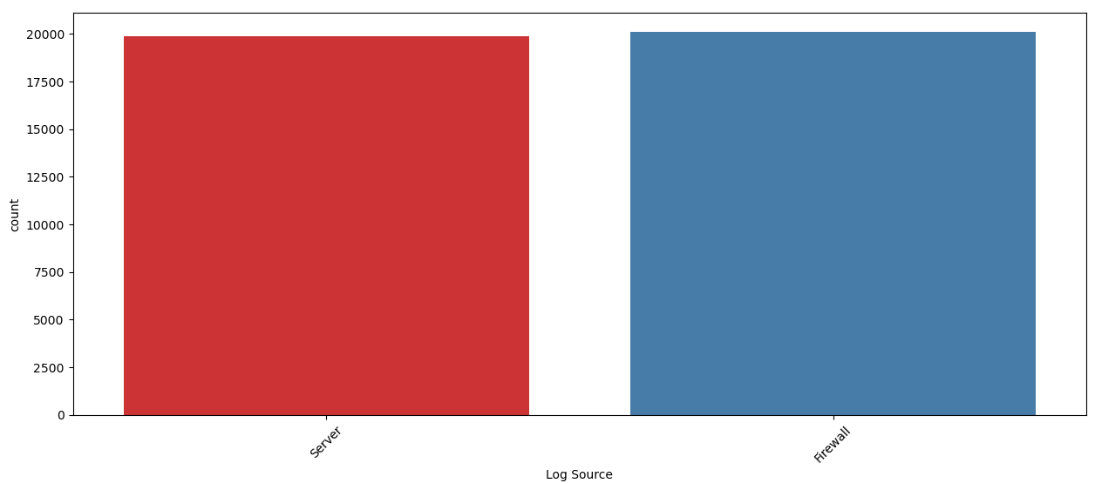
SO the respective authority will precautions to avoid both the attacks. Also we can see that the attacks types are actively involved.

**CHAPTER-9**

**RESULTS**

****

**Figure 1**: The first 5 rows of the dataset.



**Figure 2:** Count plot of the Log Source column of the dataset.

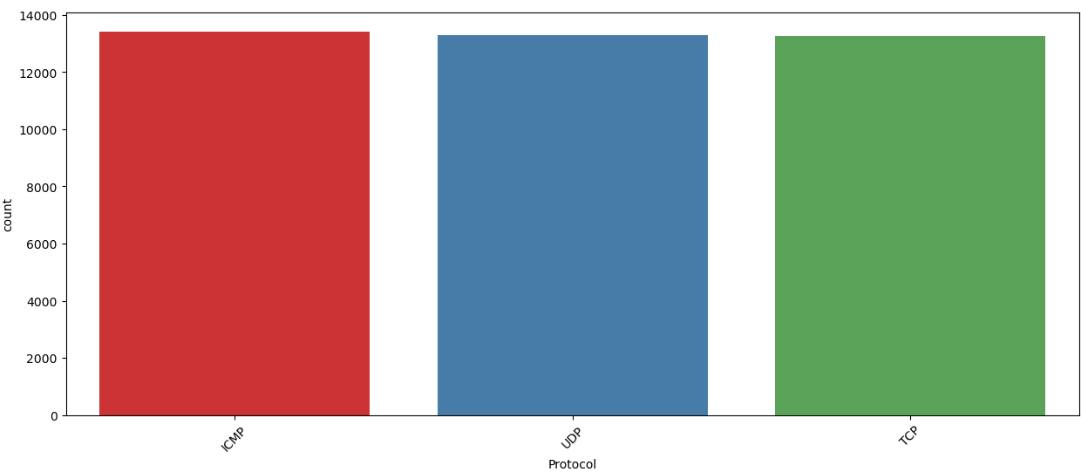


Figure 3: Count plot of protocol column of the dataset.

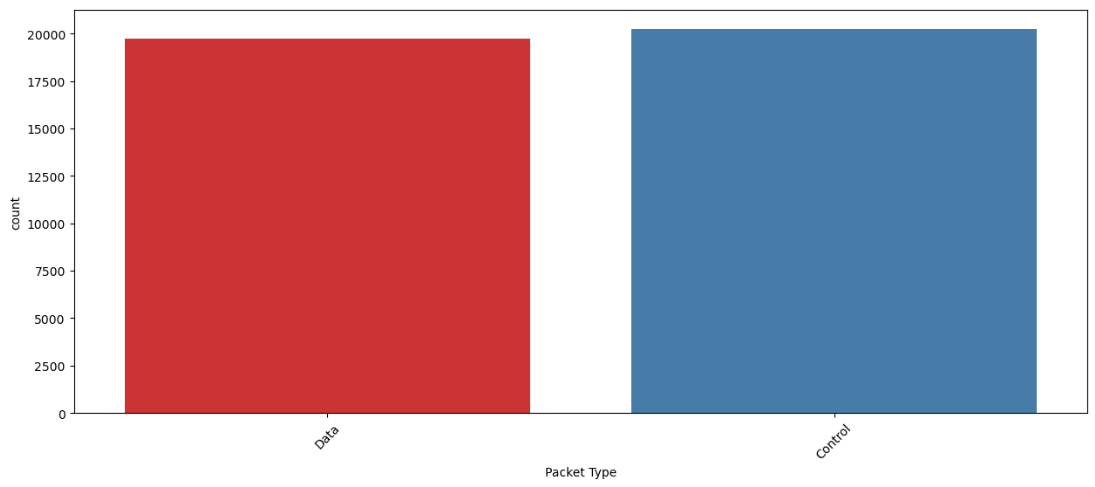
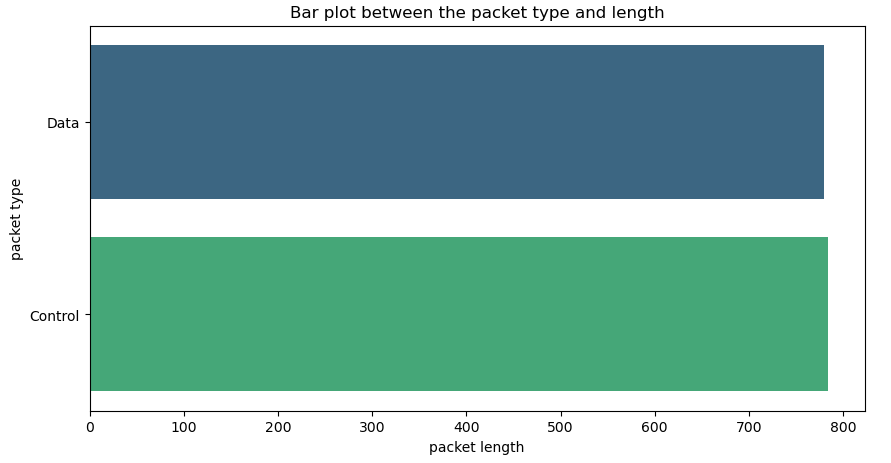


Figure 4: Count plot of packet type column of the dataset.



**Figure 5**: Bar plot between the packet length and type of packet

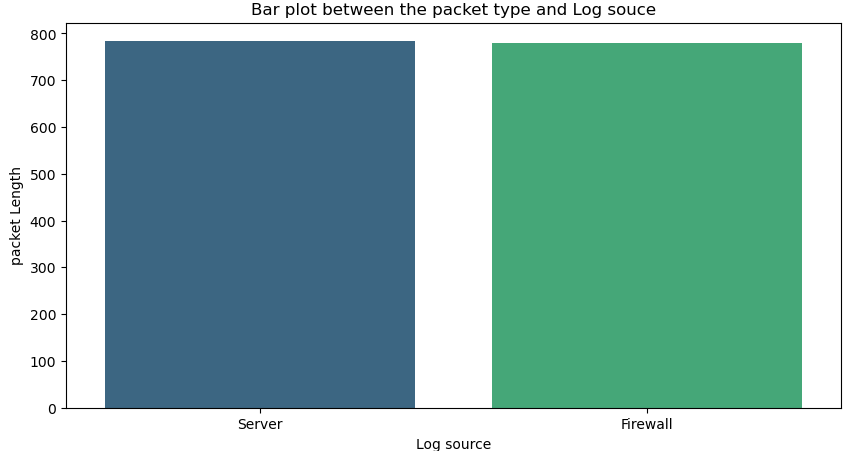


Figure 6: Bar plot between Log source and packet length