



## Can cyber security stop the hack before **it happens?**

Most of our transactions are done online these days, making life easier for us. However, it also poses safety risks, as our confidential information is exposed to fraud and hackers. In fact, hackers detected several ways to steal important information, which makes our lives more difficult, they can access secure information by only entering the phone number connected to the bank account.

Due to rising risks associated with electronic payment systems and the increase in piracy operations around the world, the need to tighten and develop regulatory systems is increasing.

### Questions:

Q 1: Which month has the most penetration?

Q2: How many times has the anonymous recording?

Q3: What number of hacks have there?

Q4: How long has there been no penetration of server?

### Dataset :

I'll work with the [Kaggle](#) dataset of Malicious Server Hacks, I propose that by using knowledge of cybersecurity, we can predict whether a server will be hacked before it happens.

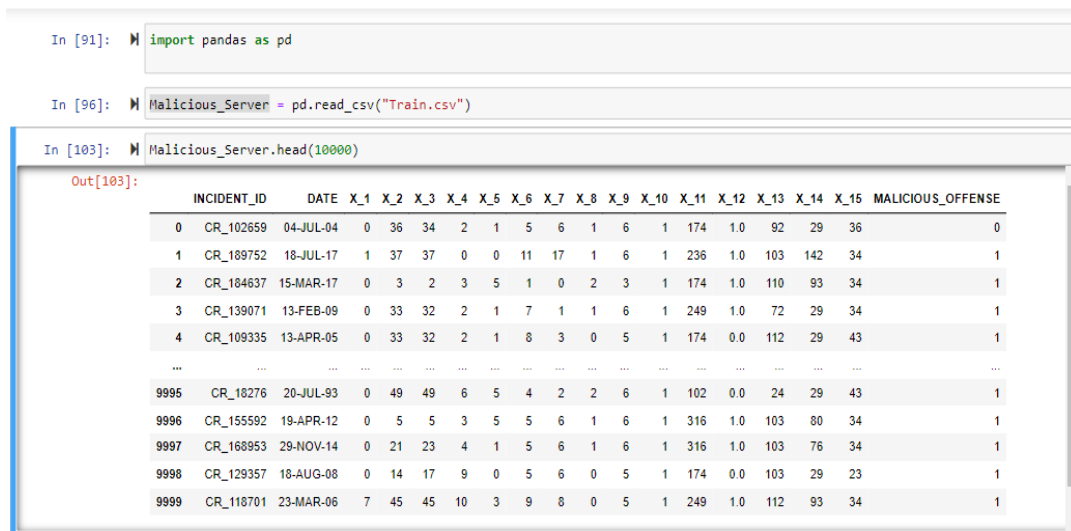


Figure 1

Here is what real data looks like!

This is a subset of the full data loaded as pandas which contains 18 features and 23856 observations

The Features:

- 1 - INCIDENT\_ID represents Unique identifier for an incident log
- 2 - DATE represents the Data of incident occurrence
- 3 - X1 - X15 represents Anonymized logging parameters
- 4 - MALICIOUS\_OFFENSE represents [Target] Indicates if the incident was a hack [1: Yes; 0:No]

## Tools

This study will be utilizing tools such as pandas, matplotlib, numpy for data analysis, and a Jupyter notebook for its execution.

## TO DO

I'll follow this workflow on the dataset step by step:

- 1 - Data cleaning
- 2 - Exploratory data analysis
- 3 - Feature engineering
- 4 - Machine learning
- 5 - Model evaluation and data visualization