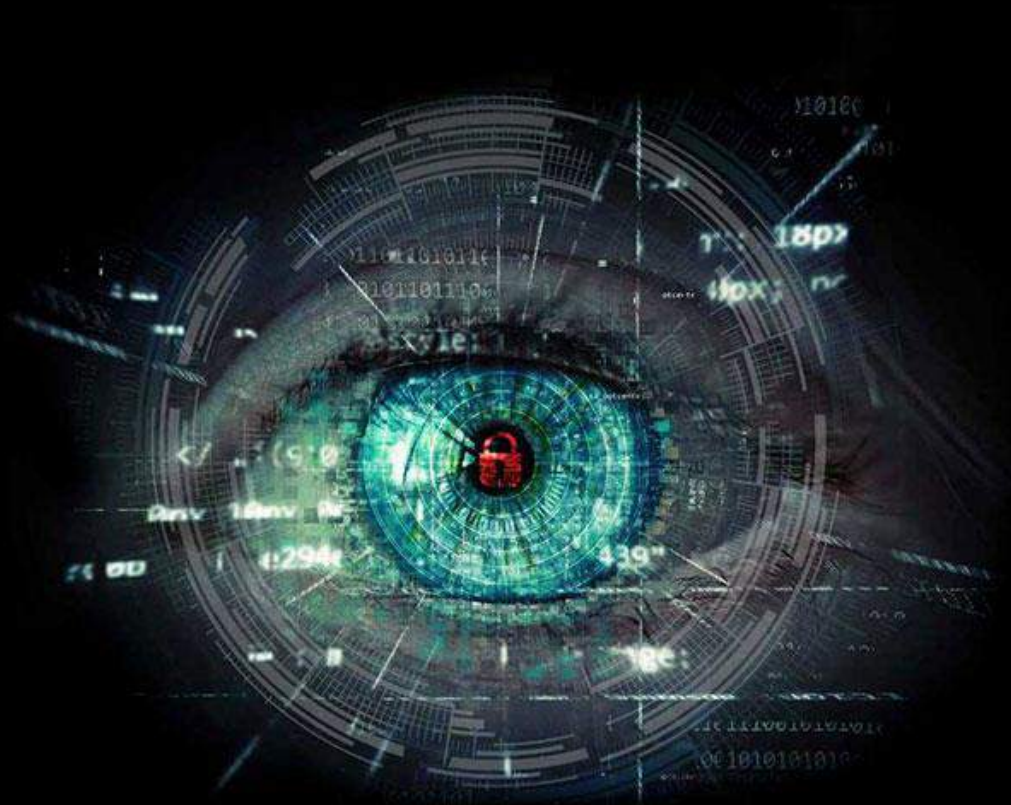
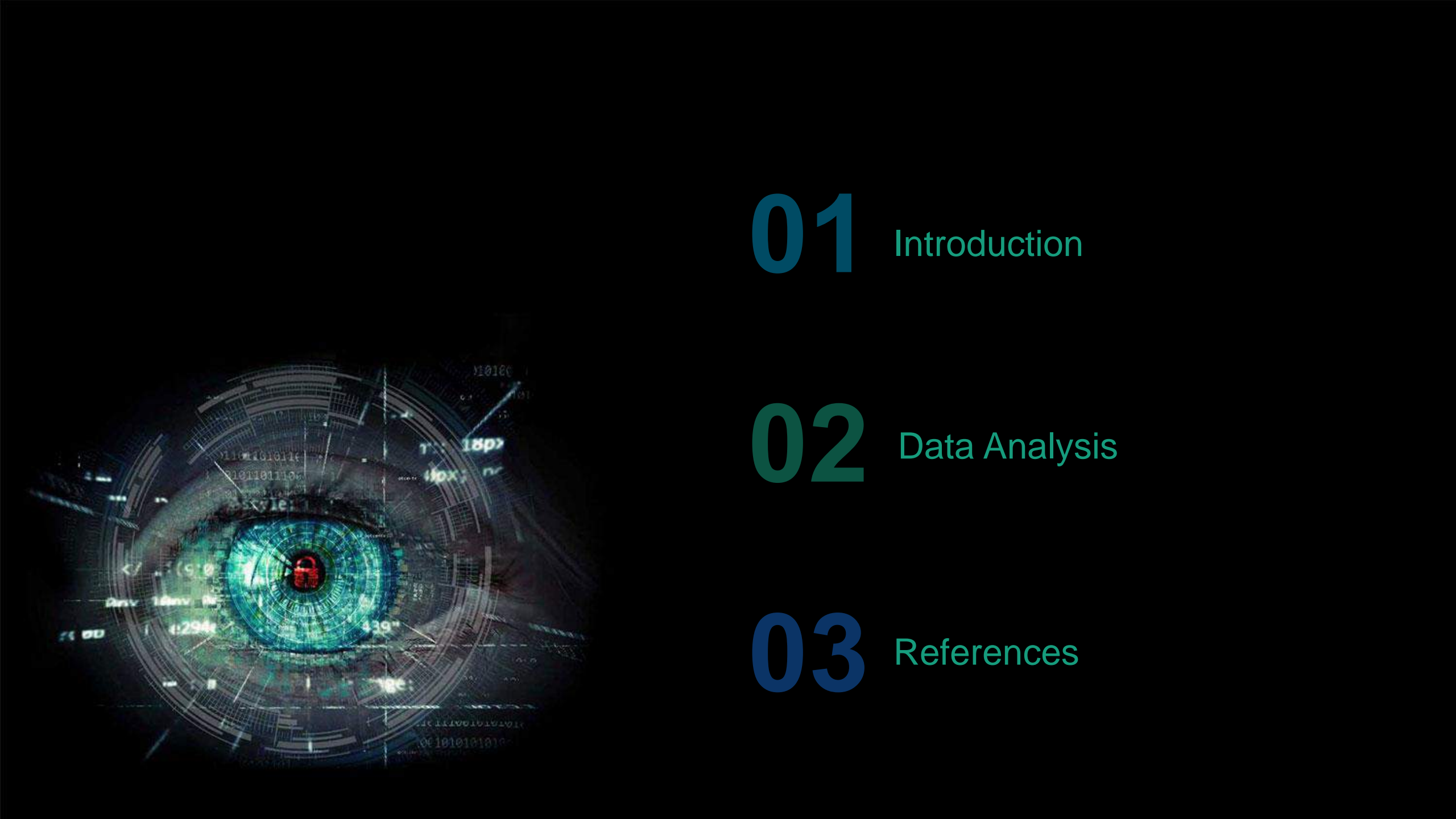


Data Science : Malicious Server Hack



Prepared By : Munira Alzhrani



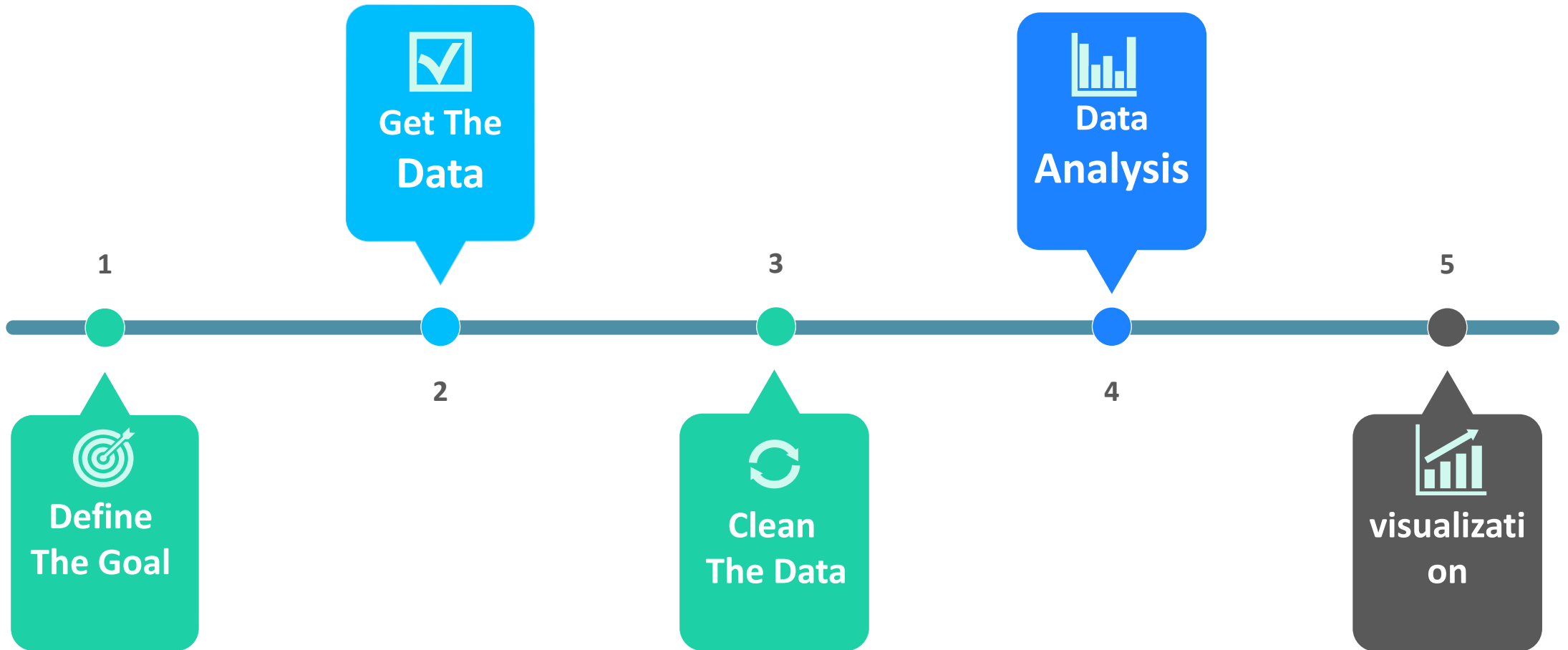
01 Introduction

02 Data Analysis

03 References

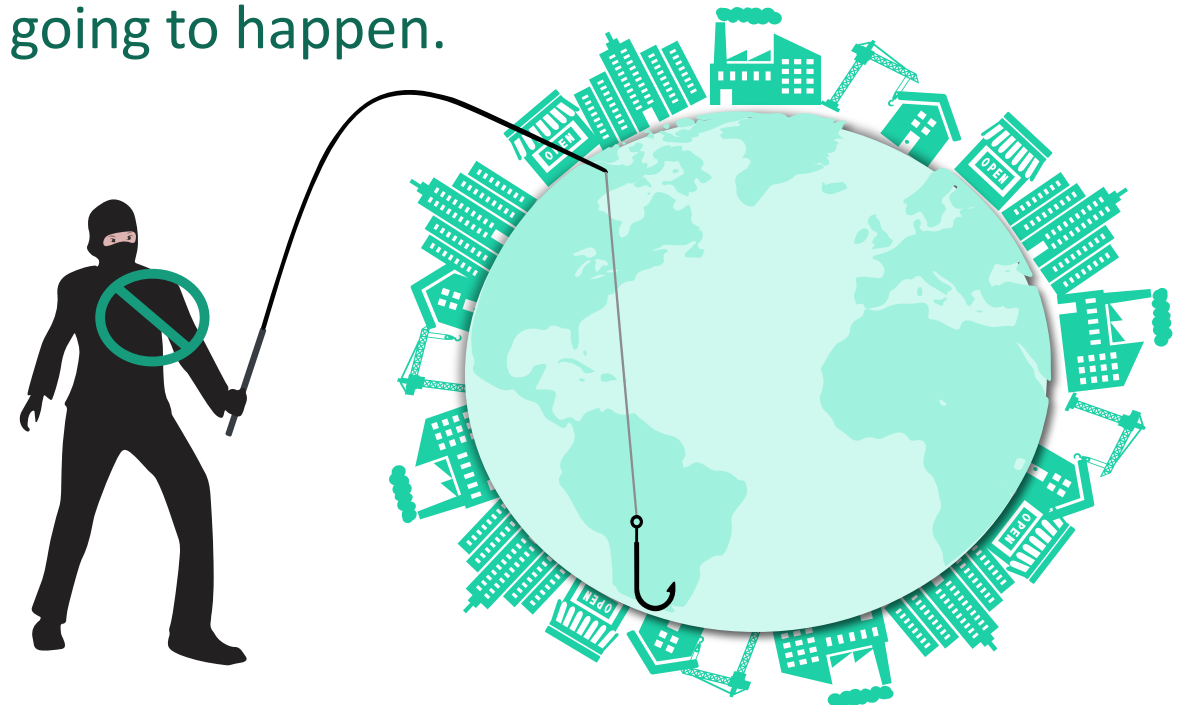
1. Introduction

Data Science Process



1.1 Problem Description :

All the countries across the globe have adapted to means of digital payments and with the increased volume of digital payments, hacking has become a pretty common event wherein the hacker can try to hack your details just with your phone number linked to your bank account. However, there is data with some anonymized variables based on which one can predict that the hack is going to happen.



**Are you Ready
To Discover Our
Data ??**

DS



**Fly With Me To Discover
Our Datasets**

DataSet [1]

```
Malicious_Server= pd.read_csv("Train.csv") # Read data into a dataframe
```

```
Malicious_Server.head()# Display top 5 of dataframe
```

```
]:
```

	INCIDENT_ID	DATE	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_10	X_11	X_12	X_13	X_14	X_15	MALICIOUS_OFFENSE
0	CR_102659	04-JUL-04	0	36	34	2	1	5	6	1	6	1	174	1.0	92	29	36	0
1	CR_189752	18-JUL-17	1	37	37	0	0	11	17	1	6	1	236	1.0	103	142	34	1
2	CR_184637	15-MAR-17	0	3	2	3	5	1	0	2	3	1	174	1.0	110	93	34	1
3	CR_139071	13-FEB-09	0	33	32	2	1	7	1	1	6	1	249	1.0	72	29	34	1
4	CR_109335	13-APR-05	0	33	32	2	1	8	3	0	5	1	174	0.0	112	29	43	1

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

Questions

DS

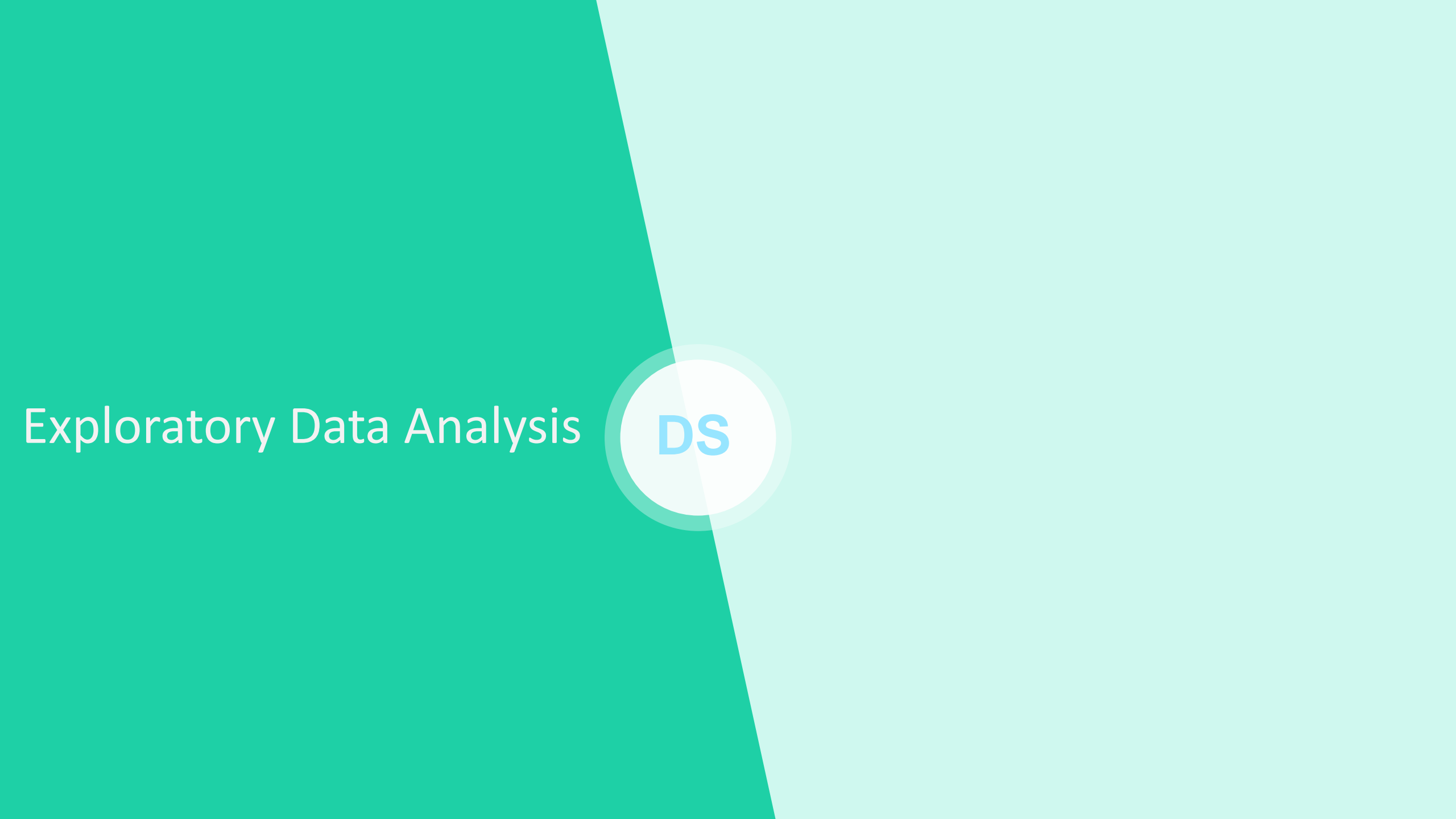
- 1- How many times has the anonymous recording?
- 2- How many hacks are there ?
- 3- Which month has the most penetration?
- 4- How long has there been no penetration of server?

Data Cleaning

A circular logo with a light blue border and a white center, containing the letters 'DS' in blue.

DS

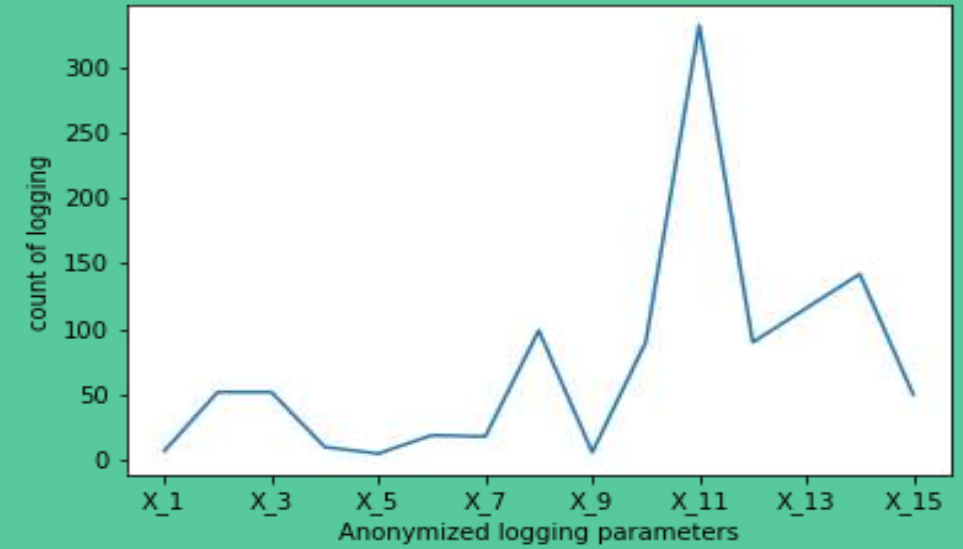
- 1-Remove the column Date, as we don't need it.
- 2-Verify if there are null values in the data.
- 3-Add a new Features.



Exploratory Data Analysis

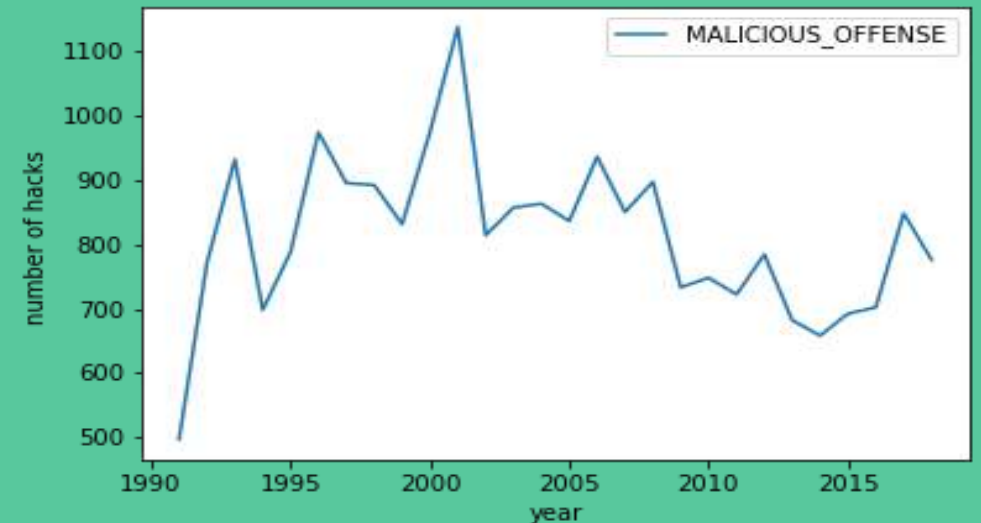


 Ans: How many times has the anonymous recording?



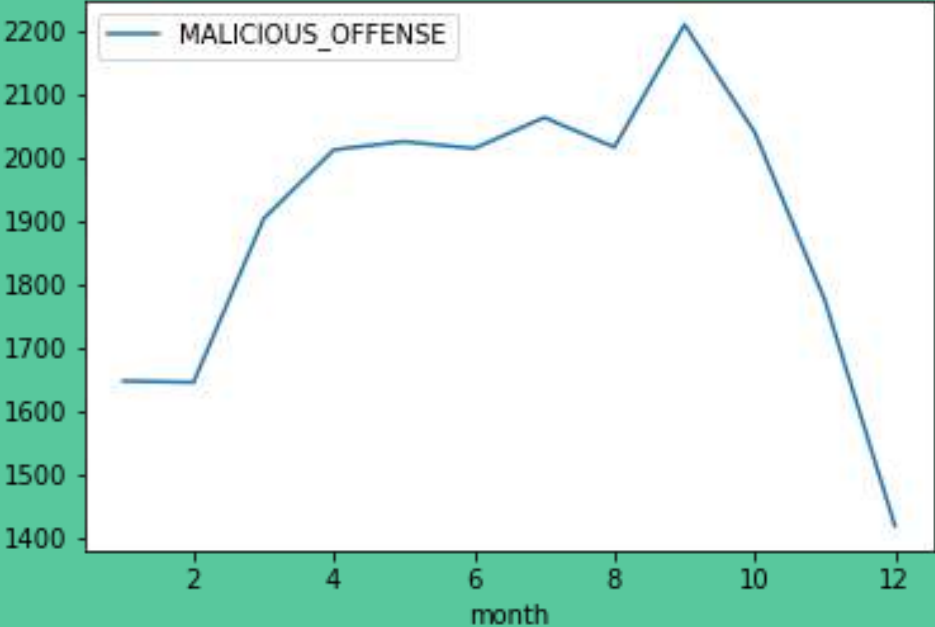
 Ans: How many hacks are there ?

By using Function Value_counts ()
I concluded happened 22788 hacks cases.

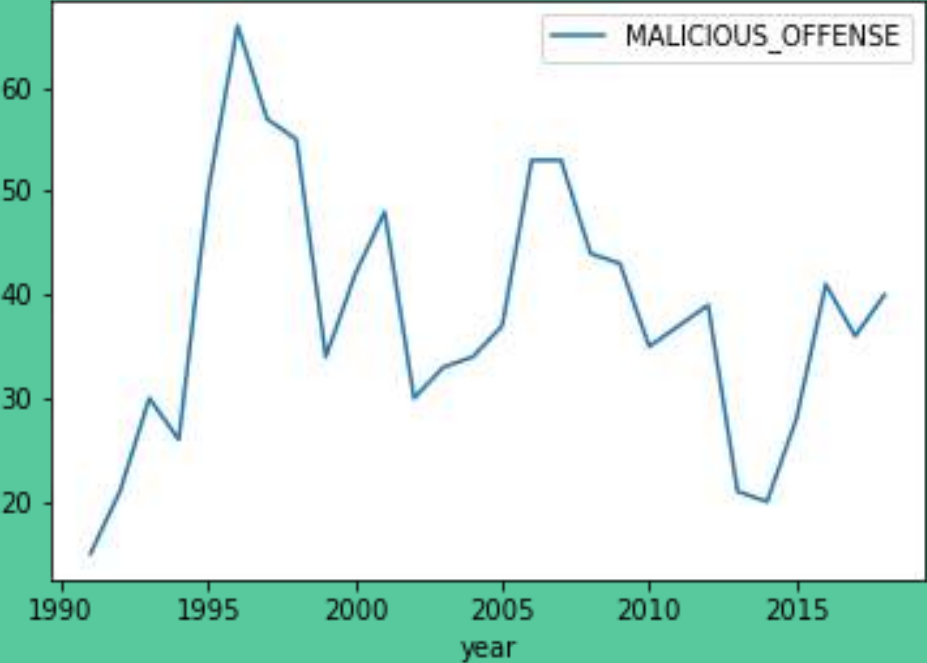




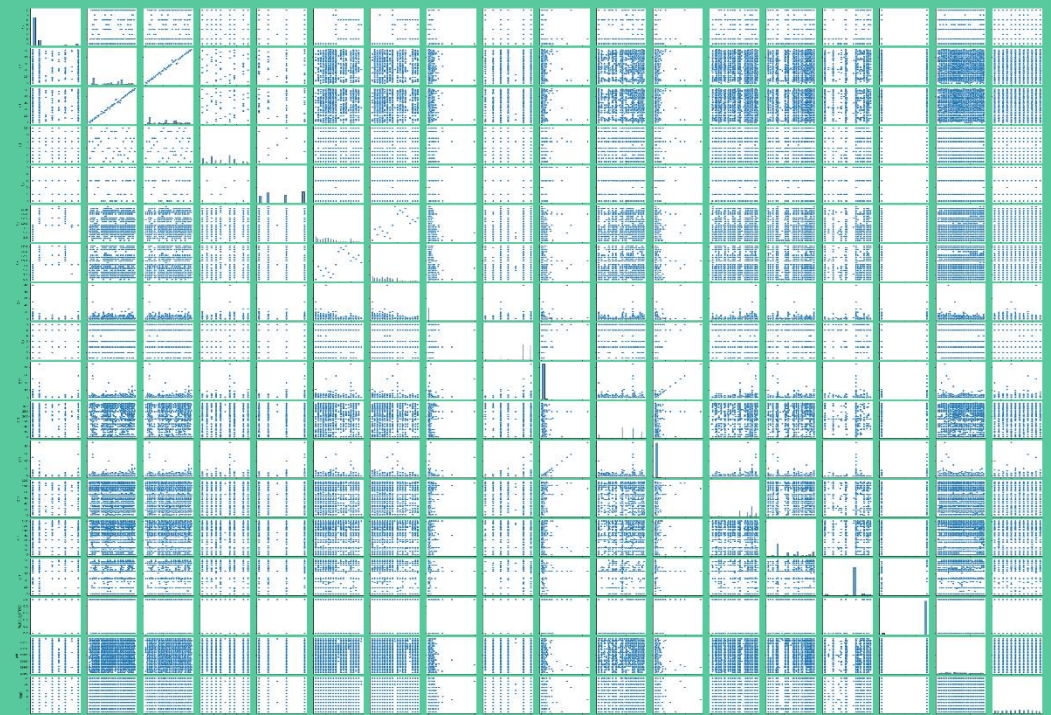
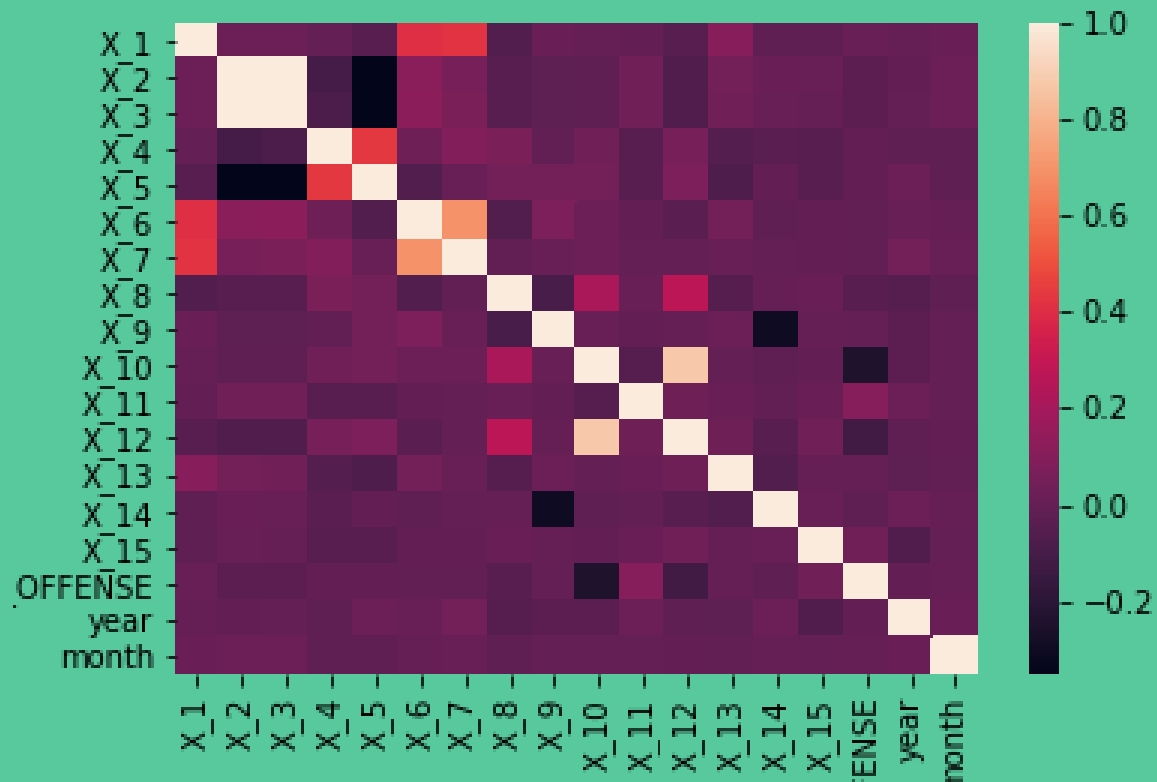
Ans: Which month has the most penetration?



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



Relationships In The Data



Tools

Data Processing **Pandas**

Visualization **Matplotlib , Seaborn**

Jupyter notebook for its execution.

References

[1] <https://www.kaggle.com/lplenka/malicious-server-hack>

kaggle

