

Handling Null Values

Date	21 October 2022
Team ID	PNT2022TMID24961
Project Name	Web Phishing Detection
Maximum Marks	2 Marks

```

In [8]: Analysing the data using pandas and checking if the dataset contains any null values.

In [9]:
df.info()
df.isnull().any() #no null values

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11955 entries, 0 to 11954
Data columns (total 32 columns):
 #   Column                Non-Null Count  Dtype
---  --
 0   index                 11955 non-null  int64
 1   having_phishing_IP_Address  11955 non-null  int64
 2   URLURL_Length         11955 non-null  int64
 3   Shortening_Service     11955 non-null  int64
 4   having_at_symbol       11955 non-null  int64
 5   double_slash_redirecting 11955 non-null  int64
 6   Prefix_Suffix          11955 non-null  int64
 7   having_sub_domain      11955 non-null  int64
 8   SSLfinal_State         11955 non-null  int64
 9   Domain_registration_length 11955 non-null  int64
10   Favicon               11955 non-null  int64
11   port                  11955 non-null  int64
12   HTTPS_token           11955 non-null  int64
13   Request_URL           11955 non-null  int64
14   URL_of_anchor          11955 non-null  int64
15   Links_in_tags          11955 non-null  int64
16   SFA                   11955 non-null  int64
17   Submitting_to_email    11955 non-null  int64
18   Abnormal_URL           11955 non-null  int64
19   Redirect              11955 non-null  int64
20   on_mousedown          11955 non-null  int64
21   on_mouseover           11955 non-null  int64
22   RightClick            11955 non-null  int64
23   popUpOnShow           11955 non-null  int64
24   IFrame                11955 non-null  int64
25   age_of_domain          11955 non-null  int64
26   DNSRecord             11955 non-null  int64
27   web_traffic            11955 non-null  int64
28   Page_Rank              11955 non-null  int64
29   Google_Index           11955 non-null  int64
30   Links_pointing_to_page 11955 non-null  int64
31   Statistical_report      11955 non-null  int64
32   Result                 11955 non-null  int64
dtypes: int64(32)
memory usage: 2.7 MB

Out[9]:
index                False
having_phishing_IP_Address  False
URLURL_Length        False
Shortening_Service    False
having_at_symbol      False
double_slash_redirecting False
Prefix_Suffix         False
having_sub_domain     False
SSLfinal_State        False
Domain_registration_length False
Favicon               False
port                  False
HTTPS_token           False
Request_URL           False
URL_of_anchor          False
Links_in_tags         False
SFA                   False
Submitting_to_email    False
Abnormal_URL           False
Redirect              False
on_mousedown          False
on_mouseover           False
RightClick            False
popUpOnShow           False
IFrame                False
age_of_domain          False
DNSRecord             False
web_traffic            False
Page_Rank              False
Google_Index           False
Links_pointing_to_page False
Statistical_report     False
Result                False
dtypes: bool

```