

Handling Null Values

Date	21 October 2022
Team ID	PNT2022TMID08772
Project Name	Web Phishing Detection
Maximum Marks	2 Marks

```

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

In [9]:
df.info()
df.isnull().any() any null values

select 'pandas.core.frame.DataFrame'
RangeIndex: 11000 entries, 0 to 10999
Data columns (total 33 columns):
 #   column                non-null count  type
---  --
 0   index                 11000 non-null  int64
 1   having_phishing_in_Address 11000 non-null  int64
 2   url_len              11000 non-null  int64
 3   shortening_service    11000 non-null  int64
 4   having_at_symbol      11000 non-null  int64
 5   double_click_redirecting 11000 non-null  int64
 6   prefix_suffix         11000 non-null  int64
 7   having_sub_domain     11000 non-null  int64
 8   suffix_state          11000 non-null  int64
 9   domain_registration_length 11000 non-null  int64
10   favicon              11000 non-null  int64
11   port                 11000 non-null  int64
12   https_flag           11000 non-null  int64
13   Request_URL          11000 non-null  int64
14   url_of_anchor        11000 non-null  int64
15   links_in_tags        11000 non-null  int64
16   src                  11000 non-null  int64
17   submitting_to_email  11000 non-null  int64
18   abnormal_url         11000 non-null  int64
19   Redirect             11000 non-null  int64
20   on_mouseover         11000 non-null  int64
21   on_mousedown         11000 non-null  int64
22   copyandpaste         11000 non-null  int64
23   frame                11000 non-null  int64
24   age_of_domain        11000 non-null  int64
25   DNSRecord            11000 non-null  int64
26   web_traffic          11000 non-null  int64
27   Page_Rank            11000 non-null  int64
28   Google_Index         11000 non-null  int64
29   links_pointing_to_page 11000 non-null  int64
30   structural_report     11000 non-null  int64
31   Result               11000 non-null  int64
Types: int64(31)
memory usage: 2.7 MB

Out[9]:
index                False
having_phishing_in_Address  False
url_len              False
shortening_service    False
having_at_symbol      False
double_click_redirecting False
prefix_suffix         False
having_sub_domain     False
suffix_state          False
domain_registration_length False
favicon              False
port                 False
https_flag           False
Request_URL          False
url_of_anchor        False
links_in_tags        False
src                  False
submitting_to_email  False
abnormal_url         False
Redirect             False
on_mouseover         False
on_mousedown         False
copyandpaste         False
frame                False
age_of_domain        False
DNSRecord            False
web_traffic          False
Page_Rank            False
Google_Index         False
links_pointing_to_page False
structural_report     False
Result               False
dtype: bool

```