Find the missing data

In order to find the missing data in large dataset we use the functions of pandas library, functions is isnull()

- DataFrame_name.isnull()
- it will give the outout as boolean for every element of dataset.
- To get the status of a single comlumn apply the with `.any(axis=1)` parameter like `df[df.isnull().any(axis-1)]`

> NOTE: DataFrame_name.isnull() function do not consider an empty string or string with space as missing value. So we have to check and replace the empty string from our dataset if any exist.

Process to Handle Missing Data

There are mainly two process to handle the missing data in our observation

1. drop the missing data by row/column
2. Imput the missing values using mean/median/mode

- To drop the data row with missing/nan values :

```
df.dropna()
```

- We can imput the data in servel ways using the `df.fillna()` on the dataset.

```
df['col_name'].fillna(1,inplave=True) # It will replace nan with 1 in the col_name
column

df['date'].fillna(method='ffill', inplace=True) # It will fill the previous value
on the next nan observation
```

- We can also use the mean,mode and median to fill the missing value

What if we ignore the Missing Value/data

- Missing Data can lead to bias and inaccurate result and conclusions.