# Handling Missing Values in Dataset

```
In [1]:   import pandas as pd
          import numpy as np
```

**To Nan**

```
In [86]:  def to_nan(df):
              df1 = df.copy()
              df1['Values'] = df1['Values'].replace('?',np.nan)
              return df1
```

**Removing Rows Entirely**

```
In [87]:  def remove_rows(df):
              df1 = df.copy()
              df1['Values'] = df1['Values'].replace('?',np.nan)
              df1 = df1.dropna()
              return df1
```

**Replacing by mean**

```
In [100…  def replace_by_mean(df):
              df1 = df.copy()
              df1 = to_nan(df)
              df1 = df1.fillna(np.mean(np.array(remove_rows(df)['Values'],dtype=float)))
              return df1
```

```
In [37]:  n = int(input('Enter the number of elements: '))

          print('Enter the element names in comma separated manner: ')
          names= input()
          X_names = names.split(',')

          print('Enter the values in comma separated manner: ')
          values = input()
          X_values = values.split(',')

          X = {'Name': X_names,
               'Values': X_values}
          df = pd.DataFrame(X)
          df
```

```
Enter the number of elements: 10
Enter the element names in comma separated manner:
A,B,C,D,E,F,G,H,I,J
Enter the values in comma separated manner:
45,37,59,?,47,39,?,43,53,?
```

Out[37]:

| | Name | Values |
|---|---|---|
| 0 | A | 45 |
| 1 | B | 37 |
| 2 | C | 59 |
| 3 | D | ? |
| 4 | E | 47 |
| 5 | F | 39 |
| 6 | G | ? |
| 7 | H | 43 |
| 8 | I | 53 |
| 9 | J | ? |

```
In [109…  choice = int(1)

          choices = {1 : to_nan(df),
                     2 : remove_rows(df),
                     3 : replace_by_mean(df)}
          while choice != 0:
              print("----------------------------MENU-----------------------------")
              print("\t\t 1. Replace by NaN ")
              print("\t\t 2. Remove the rows ")
              print("\t\t 3. Replace by Mean ")
              print("\t\t 0. EXIT")
              print("-------------------------------------------------------------")
              choice = int(input("Enter your choice: "))
              if choice == 0:
                  break
              print(choices.get(choice))
```

```
----------------------------MENU-----------------------------
                 1. Replace by NaN
                 2. Remove the rows
                 3. Replace by Mean
                 0. EXIT
-------------------------------------------------------------
Enter your choice: 1
  Name Values
0    A     45
1    B     37
2    C     59
3    D    NaN
4    E     47
5    F     39
6    G    NaN
7    H     43
8    I     53
9    J    NaN
----------------------------MENU-----------------------------
                 1. Replace by NaN
                 2. Remove the rows
                 3. Replace by Mean
                 0. EXIT
-------------------------------------------------------------
Enter your choice: 2
  Name Values
0    A     45
1    B     37
2    C     59
4    E     47
5    F     39
7    H     43
8    I     53
----------------------------MENU-----------------------------
                 1. Replace by NaN
                 2. Remove the rows
                 3. Replace by Mean
                 0. EXIT
-------------------------------------------------------------
Enter your choice: 3
  Name   Values
0    A       45
1    B       37
2    C       59
3    D  46.1429
4    E       47
5    F       39
6    G  46.1429
7    H       43
8    I       53
9    J  46.1429
----------------------------MENU-----------------------------
                 1. Replace by NaN
                 2. Remove the rows
                 3. Replace by Mean
                 0. EXIT
-------------------------------------------------------------
Enter your choice: 0
```

```
In [ ]:
```