

'Output':

Out[246]:

| | First | Second | Third | First | Second | Third |
|---|-------|--------|-------|-------|--------|-------|
| 0 | 2.0 | 3.0 | 9.0 | 3 | 6 | 3 |
| 1 | 8.0 | 7.0 | 7.0 | 2 | 2 | 1 |
| 2 | 8.0 | 6.0 | 4.0 | 9 | 3 | 8 |
| 3 | NaN | NaN | NaN | 2 | 9 | 2 |

Handling Missing Data

Dealing with missing data is an integral part of the data cleaning/data analysis process. Moreover, some machine learning algorithms will not work in the presence of missing data. Let’s see some simple Pandas methods for identifying and removing missing data, as well as imputing values into missing data.

Identifying Missing Data

In this section, we’ll use the **isnull()** method to check if missing cells exist in a DataFrame.

```
# let's create a data frame with missing data
my_DF = pd.DataFrame({'age': [15,17,np.nan,29,25], \
                        'state_of_origin':['Lagos', 'Cross River', 'Kano', \
                        'Abia', np.nan]})
```

my_DF

'Output':

| | age | state_of_origin |
|---|------|-----------------|
| 0 | 15.0 | Lagos |
| 1 | 17.0 | Cross River |
| 2 | NaN | Kano |
| 3 | 29.0 | Abia |
| 4 | 25.0 | NaN |

Let’s check for missing data in this data frame. The **isnull()** method will return **True** where there is a missing data, whereas the **notnull()** function returns **False**.