```
dataframe = [[1,2,3],[4,5,6],[7,8,9]]

dataframe
```

```
[[1, 2, 3], [4, 5, 6], [7, 8, 9]]
```

```
dataframe1 ={
    'c1':[1,2,3],
    'c2':[2,3,4]
}

dataframe1
```

```
{'c1': [1, 2, 3], 'c2': [2, 3, 4]}
```

```
import pandas as pd
import numpy as np


dataframe2 ={
    'c1':[1,2,3],
    'c2':[2,3,4]
}

df=pd.DataFrame(dataframe2)

df
```

```
   c1  c2
0   1   2
1   2   3
2   3   4
```

```
df.describe()
```

```
        c1   c2
count  3.0  3.0
mean   2.0  3.0
std    1.0  1.0
min    1.0  2.0
25%    1.5  2.5
50%    2.0  3.0
75%    2.5  3.5
max    3.0  4.0
```

```
df.mean()
```

```
c1    2.0
c2    3.0
dtype: float64
```

```
df.corr()
```

```
     c1   c2
c1  1.0  1.0
c2  1.0  1.0
```

```
df.median()
```

```
c1    2.0
c2    3.0
dtype: float64
```

```
df.count()
```

```
c1    3
c2    3
dtype: int64
```

```
df.max()
```

```
c1    3
c2    4
dtype: int64
```

```
df.min()
```

```
c1    1
c2    2
dtype: int64
```

```
df.std()
```

```
c1    1.0
c2    1.0
dtype: float64
```

```
df.T
```

```
    0  1  2
c1  1  2  3
c2  2  3  4
```

```
df.sort_index(axis=1,ascending=False)
```

```
   c2  c1
0   2   1
1   3   2
2   4   3
```

```
df.sort_index(axis=0,ascending=True)
```

```
   c1  c2
0   1   2
1   2   3
2   3   4
```

```
import pandas as pd
import numpy as np

data = {
    'C1':[1,2,3,4,5],
    'C2':[2,3,4,5,6],
    'C3':[5,6,7,8,9]
}

df=pd.DataFrame(data)

df
```

```
   C1  C2  C3
0   1   2   5
1   2   3   6
2   3   4   7
3   4   5   8
4   5   6   9
```

```
df.head()
```

```
   C1  C2  C3
0   1   2   5
1   2   3   6
2   3   4   7
3   4   5   8
4   5   6   9
```

```
df.tail()
```

```
   C1  C2  C3
0   1   2   5
1   2   3   6
2   3   4   7
3   4   5   8
4   5   6   9
```

```
df.describe()
```

```
             C1        C2        C3
count  5.000000  5.000000  5.000000
mean   3.000000  4.000000  7.000000
std    1.581139  1.581139  1.581139
min    1.000000  2.000000  5.000000
25%    2.000000  3.000000  6.000000
50%    3.000000  4.000000  7.000000
75%    4.000000  5.000000  8.000000
max    5.000000  6.000000  9.000000
```

```
df.count()
```

```
C1    5
C2    5
C3    5
dtype: int64
```

```
df.T
```

```
    0  1  2  3  4
C1  1  2  3  4  5
C2  2  3  4  5  6
C3  5  6  7  8  9
```

```
df.sort_index(axis=1,ascending=True)
```

```
   C1  C2  C3
0   1   2   5
1   2   3   6
2   3   4   7
3   4   5   8
4   5   6   9
```

```
df.sort_index(axis=0,ascending=False)
```

```
   C1  C2  C3
4   5   6   9
3   4   5   8
2   3   4   7
1   2   3   6
0   1   2   5
```

```
df.sort_index(axis=1,ascending=False)
```

```
   C3  C2  C1
0   5   2   1
1   6   3   2
2   7   4   3
3   8   5   4
4   9   6   5
```

```
df.loc[0,0]=34
```

```
df
```

```
   C1  C2  C3     0
0   1   2   5  34.0
1   2   3   6   NaN
2   3   4   7   NaN
3   4   5   8   NaN
4   5   6   9   NaN
```

```
df.loc[0,'C1']=56
```

```
df
```

```
   C1  C2  C3     0
0  56   2   5  34.0
1   2   3   6   NaN
2   3   4   7   NaN
3   4   5   8   NaN
4   5   6   9   NaN
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   C1      5 non-null      int64
 1   C2      5 non-null      int64
 2   C3      5 non-null      int64
dtypes: int64(3)
memory usage: 252.0 bytes
```

```
# drop row
df.drop([0])
```

```
   C1  C2  C3
1   2   3   6
2   3   4   7
3   4   5   8
4   5   6   9
```

```
#drop multiple rows
df.drop([1,4,3])
```

```
   C1  C2  C3
0  56   2   5
2   3   4   7
```

```
df.drop(['C1'], axis=1)
```

```
   C2  C3
0   2   5
1   3   6
2   4   7
3   5   8
4   6   9
```

```
df.drop([0],axis=0)
```

```
   C1  C2  C3
1   2   3   6
2   3   4   7
```

```
3   4   5   8
4   5   6   9
```

```
df.drop([1,3,4], inplace=True)
```

```
df
```

```
   C1  C2  C3
0  56   2   5
2   3   4   7
```

```
df.shape
```

```
(2, 3)
```

```
# drop missing values or we can say remove from dataset
df.dropna(subset=['C1'])
```

```
   C1  C2  C3
0  56   2   5
2   3   4   7
```

```
# check is there any missing value or not
df.isna().sum()
```

```
C1    0
C2    0
C3    0
dtype: int64
```

```
# drop rows containing strings
# only_str = df.select_dtypes(exclude='numbers')
# df.drop(only_str,axis=1)
```