```
import pandas as pd
import numpy as np
from sklearn import preprocessing


dic = {
    "Name":["abc","lmn","stu","xyz","efg","klm","rst","uvw",np.nan,"def"],
    "Maths":[96,88,np.nan,73,44,88,72,84,93,92],
    "Reading":[77,200,88,np.nan,np.nan,91,78,np.nan,84,93],
    "Writing":[86,np.nan,170,87,69,91,np.nan,88,92,83],
    "Placement":[88,76,91,84,96,84,80,78,np.nan,93],

"Gender":["Male","Female",np.nan,"Female","Female","Male","Female","Male",
"Male", "Male"],
    "Joining Date":[2018,2019,2022,2018,2017,2023,2016,2020,2020,2019]
}


df = pd.DataFrame(dic)
df


df.isnull()


df.notnull()


df['Maths'] = df['Maths'].fillna(df['Maths'].mean())
df


df['Writing'] = df['Writing'].fillna(df['Writing'].median())
df


df['Placement'] = df['Placement'].fillna(df['Placement'].mode())
df


df['Writing'] = df['Writing'].fillna(df['Writing'].max())
df


df = pd.DataFrame(dic)
df
df.dropna(how='all')


df.dropna(how='any')




df
```

```
df.dropna(axis=0, how='any')
```

```
df.dropna(axis=0, how='all')
df
```

```
df.dropna(axis=1, how='any')
```

```
df.replace(to_replace=np.nan, value=60)
```

```
dic = {
    "Name":["abc","lmn","stu","xyz","efg","klm","rst","uvw",np.nan,"def"],
    "Maths":[96,88,np.nan,73,44,88,72,84,93,92],
    "Reading":[77,200,88,np.nan,np.nan,91,78,np.nan,84,93],
    "Writing":[86,np.nan,170,87,69,91,np.nan,88,92,83],
    "Placement":[88,76,91,84,96,84,80,78,np.nan,93],
    "Gender":["Male","Female",np.nan,"Female","Female","Male","Female",
"Male", "Male", "Male"],
    "Joining Date":[2018,2019,2022,2018,2017,2023,2016,2020,2020,2019],

"Region":["Pune","Mumbai","Delhi",np.nan,"Surat","Solapur","Bengaluru","Kol
kata",np.nan,"Pune"]
}
```

```
df = pd.DataFrame(dic)
df
```

```
df['Region'].replace(to_replace=np.nan,value='Goa')
```

```
col = ['Maths', 'Writing','Placement']
df.boxplot(col)
```

```
df
```

```
df.dropna(inplace=True)
df
```

```
col1 = ['Maths', 'Writing','Placement']
df.boxplot(col1)
```

```
df
```

```
rscore = df['Writing']
q1 = np.percentile(rscore, 25)
q3 = np.percentile(rscore, 75)
print(q1, q3)


iqr = q3 - q1
print(iqr)


lower_bound = q1 - 1.5*iqr
upper_bound = q3 + 1.5*iqr
print(lower_bound, upper_bound)


r_outlier = []
for i in rscore:
    if i < lower_bound or i > upper_bound:
        r_outlier.append(i)
        print(r_outlier)


median = np.median(rscore)
median
df['Writing'] = np.where(df['Writing'] > upper_bound, median,
df['Writing'])
df['Writing']


x = df.drop(axis=1, columns=['Region', 'Gender', 'Joining Date'])
x

min_max_scaler = preprocessing.MinMaxScaler()
df[['Maths', 'Reading', 'Writing', 'Placement']] =
min_max_scaler.fit_transform(df[['Maths', 'Reading', 'Writing',
'Placement']])
df
```