```python
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

import math

import warnings

warnings.filterwarnings("ignore")

%matplotlib inline


rollno = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]

name = ["a", "b", "c", "d", "e", "f", "g", "h", "i", "j", np.nan, np.nan, "k", "l", "m"]

marks = [40, 23, 50, 78, 48, 89, 90, 67, 84, 96, 76, np.nan, 97, np.nan, 65]

grade = ["F", "F", "P", "P", "P", "P", "P", "P", "P", "P", "P", "F", "P", np.nan, np.nan]


df = pd.DataFrame({"rollno" : rollno, "name" : name, "marks" : marks, "grade" : grade})


df


print(df)


df.info()


df.describe()


df.dtypes


df.columns


df.isna().sum()


df.to_csv("academic_performance.csv")
```

```python
df.isna().sum()

df["marks"] = df["marks"].fillna(df["marks"].mean())

df

def fun1(value):
    return int(math.floor(value))

df["marks"] = df["marks"].apply(fun1)

df

df = df[df['name'].notna()]

df

for index, row in df.iterrows():
    # print(row['marks'], row['grade'])
    if (row['marks'] > 40):
        df.loc[index, 'grade'] = 'P'
    else:
        df.loc[index, 'grade'] = 'F'

df

first_outlier = [16, 'n', 200, 'P']
second_outlier = [17, 'o', -100, 'F']

df.loc[15] = first_outlier
df.loc[16] = second_outlier
```

```
df
```

```
sns.countplot(data=df, x=df['marks']);
```

```
plt.show()
```

```
sns.boxplot(data=df, x='marks');
```

```
plt.show()
```

```
from matplotlib.cbook import boxplot_stats
```

```
outliers = boxplot_stats(df['marks']).pop(0)['fliers']
```

```
outliers
```

```
df
```

```
df = df.drop([15,16], axis=0)
```

```
df
```

```
from sklearn.preprocessing import MinMaxScaler
```

```
scaler = MinMaxScaler()
```

```
df[['marks']] = scaler.fit_transform(df[['marks']])
```

```
df
```