```
import pandas as pd
import numpy as np
```

```
df = pd.read_csv('/content/AcademicPerformance.csv')
```

```
df.head()
```

	Gender	Math Score	Reading Score	Writing Score	Placement Score	Placement Offer Count	Region
0	Female	85.0	75	65.0	27	2	Pune
1	Male	90.0	80	60.0	30	3	Mumbai
2	Female	75.0	70	67.0	25	1	Amravati
3	Male	92.0	82	70.0	40	4	Pune
4	Female	70.0	62	55.0	20	1	Mumbai

```
from sklearn import preprocessing
le = preprocessing.LabelEncoder()
```

```
from sklearn.preprocessing import LabelEncoder
```

```
df['Gender'] = LabelEncoder().fit_transform(df['Gender'])
```

```
df.head()
```

	Gender	Math Score	Reading Score	Writing Score	Placement Score	Placement Offer Count	Region
0	0	85.0	75	65.0	27	2	Pune
1	1	90.0	80	60.0	30	3	Mumbai
2	0	75.0	70	67.0	25	1	Amravati
3	1	92.0	82	70.0	40	4	Pune
4	0	70.0	62	55.0	20	1	Mumbai

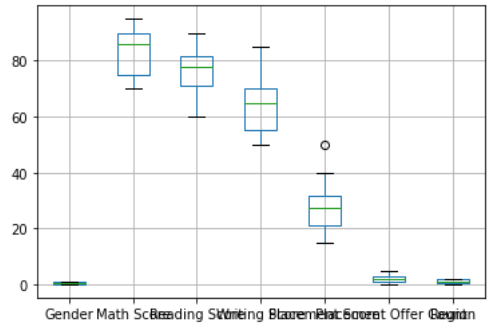
```
df['Region'] = LabelEncoder().fit_transform(df[' Region'])
```

```
df.head()
```

	Gender	Math Score	Reading Score	Writing Score	Placement Score	Placement Offer Count	Region	Region
0	0	85.0	75	65.0	27	2	Pune	2
1	1	90.0	80	60.0	30	3	Mumbai	1
2	0	75.0	70	67.0	25	1	Amravati	0
3	1	92.0	82	70.0	40	4	Pune	2
4	0	70.0	62	55.0	20	1	Mumbai	1

```
df.boxplot()
```

```
<Axes: >
```



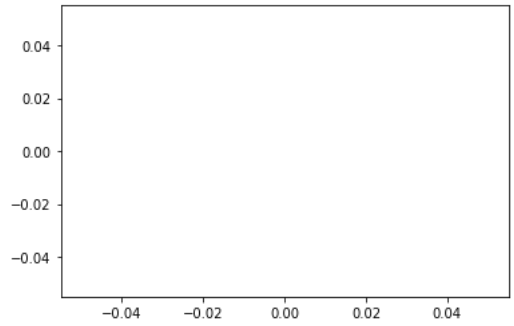
```
import seaborn as sns


import matplotlib.pyplot as plt



plt.plot()
```

```
[]
```




```
from scipy import stats


z = np.abs(stats.zscore(df['Placement Score']))
print(z)
```

```
0    0.15087
1    0.15087
2    0.35203
3    1.15667
4    0.85493
5    0.05029
6    1.05609
7    0.35203
8    2.16247
9    1.35783
Name: Placement Score, dtype: float64
```

```
threshold = 2


print(np.where(z > 1))
```

```
(array([3, 6, 8, 9]),)
```

```
df['Math Score'].mean()
```

```
83.66666666666667
```

```
df['Math Score'].std()
```

```
9.578622030334008
```

```
zs = (95 - 83.66666666666667)/ 9.578622030334008
print(zs)
```

```
1.1831903688695955
```

```
mean = df['Math Score'].mean()
std = df['Math Score'].std()
threshold = 70
outlier = []
for i in df['Math Score']:
    zs = (i - mean)/std
    print(zs)
    if zs  > threshold :
      outlier.append(i)
print("Ouliers", outlier)
```

```
0.13919886692583433
0.6611946178977148
-0.9047926350179268
0.8699929182864671
-1.4267883859898074
0.24359801712021042
nan
```

```
0.6611946178977148
1.1831903688695955
-1.4267883859898074
Ouliers []
```