

Five Number Summary and Box Plot Report

This report covers the five-number summary and box plots for a given dataset of marks.

Five Number Summary

```
import numpy as np
lst_marks = [45, 32, 56, 75, 89, 54, 32, 89, 90, 87, 67, 54, 45, 98, 99, 67, 74]
minimum, Q1, median, Q3, maximum = np.quantile(lst_marks, [0, 0.25, 0.50, 0.75, 1.0])
IQR = Q3 - Q1
lower_fence = Q1 - 1.5 * IQR
higher_fence = Q3 + 1.5 * IQR
```

The five-number summary consists of the minimum, first quartile (Q1), median, third quartile (Q3), and maximum values of the dataset. These statistics provide a quick overview of the distribution of the data.

The five-number summary for the dataset is as follows:

Minimum: 32.0

Q1: 54.0

Median: 67.0

Q3: 89.0

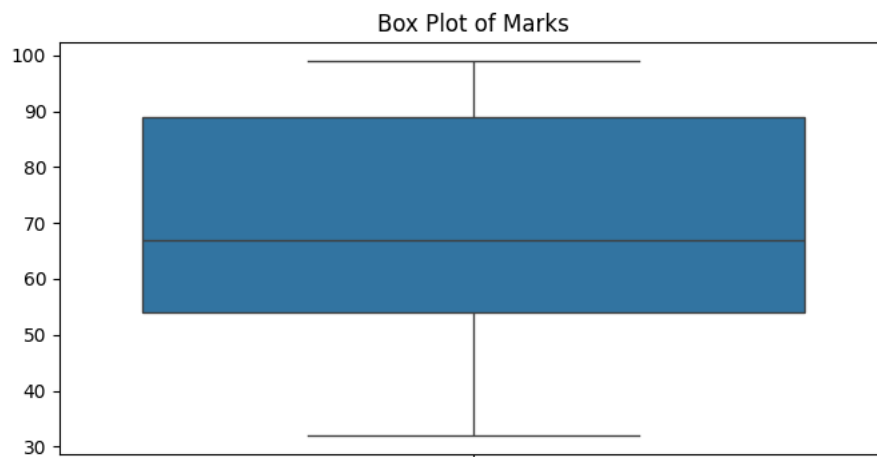
Maximum: 99.0

Box Plot

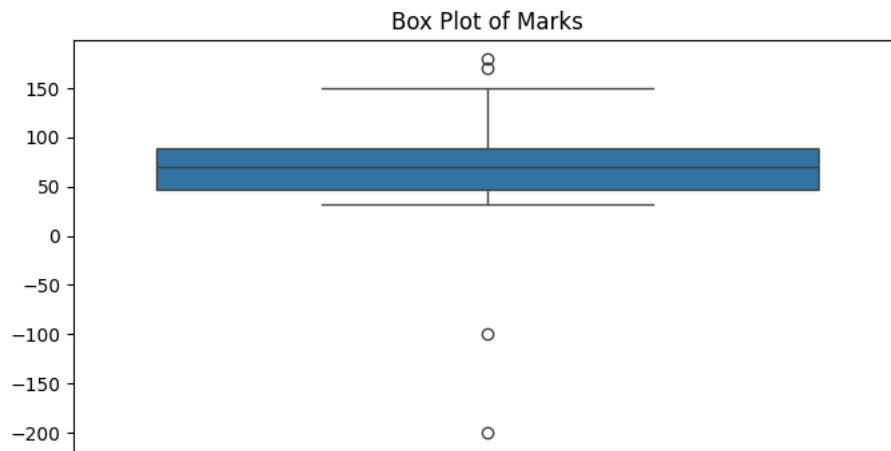
```
import seaborn as sns
lst_marks_with_outliers = [-100, -200, 45, 32, 56, 75, 89, 54, 32, 89, 90,
                           87, 67, 54, 45, 98, 99, 67, 74, 150, 170, 180]
sns.boxplot(lst_marks_with_outliers)
```

A box plot visually represents the five-number summary of the dataset. It displays the minimum, Q1, median, Q3, and maximum values, along with any potential outliers. The box represents the interquartile range (IQR), and the line inside the box indicates the median.

Box Plot of Marks (without outliers):



Box Plot of Marks (with outliers):



Conclusion

This report provided an overview of the five-number summary and box plots for a dataset of marks. These tools are essential for understanding the distribution and identifying outliers in the data.

Glossary

1. Five Number Summary: A descriptive statistic that provides information about a dataset's minimum, first quartile (Q1), median, third quartile (Q3), and maximum.
2. Box Plot: A graphical representation of the five-number summary that displays the distribution of data based on a five-number summary.
3. Outlier: A data point that differs significantly from other observations in the dataset.
4. Interquartile Range (IQR): A measure of statistical dispersion, calculated as the difference between the third quartile (Q3) and the first quartile (Q1).