

Box Plots

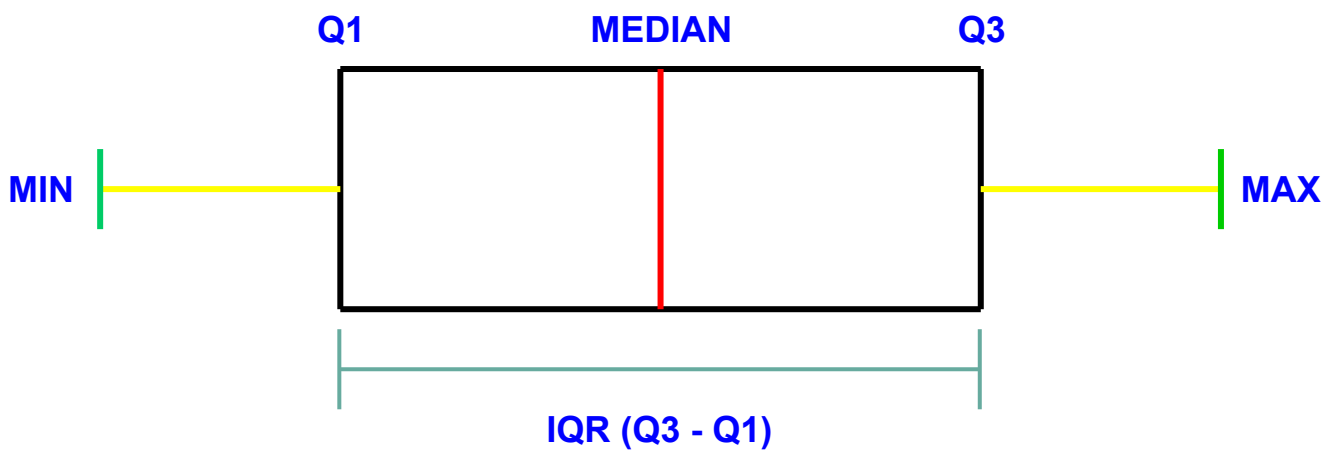
In this lesson, box plots with the seaborn package are discussed.

We'll cover the following ^

- Introduction
- Example

Introduction

This visualization is used to display groups of statistical or numerical data using their [quantiles](#). These plots have lines extended from them known as *whiskers*. The yellow lines in the below figure represent these *whiskers*. These lines may represent any outliers that occur due to variance in data. The outliers for this plot are the values that lie below or above the first and third quartiles, respectively.



Box Plot

Example

The following code example generates some random numbers and plots them on a box plot.

```
import numpy as np

import seaborn as sns

set1 = np.random.randn(220) # Generating random data
```



```
sns1 = sns.boxplot(set1) # plotting box plot
```



The box plot is clearly visible in the output. The whiskers denote the maximum variance that can be in the data; the points that do not lie in this range are considered outliers. It can be seen that there are some points that lie outside the plot range. These are the extreme outliers, and with a box plot, they are clearly visible.

The following example uses a parameter that can cover these outliers in the box plot.

```
import numpy as np

import seaborn as sns

set1 = np.random.randn(220)

sns1 = sns.boxplot(set1, whis=np.inf)
```



The `whis` parameter is used and assigned a value of `np.inf`, meaning *infinity*. This elongates the whiskers of the box plot to cover these extreme variances of data. As can be seen in the output, no extreme outliers are present anymore because all the data points lie between `-inf` and `+inf`.

In the next lesson, regression plots are discussed.