

# Data Science

Quarto [↗](#)

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see <https://quarto.org>.

## Write text to answer the question:

### 1. What do you think being a data scientist is about?

A data scientist is an analytics professional who is responsible for collecting, analyzing, and interpreting data to help drive decision making in an organization. Data Scientist do many tasks like to find the pattern and trends in dataset to uncover insights.

### 2. What differences/similarities do you see between data scientists and statisticians?

**Data scientist** organize, extract and analyze data with the goal of addressing specific problems or answering questions. They focus on creating systems that perform these analyses and produce useful results whereas statisticians is using mathematical and statistical models to analyze data.

**Similarity** Both collect the information and use it to perform the analysis.

### 3. How do you view yourself in relation to these two areas?

As a data scientist I can analyze the data, extract insights and provide the actionable information. For this first of all I will perform EDA, then build model and at last predict the model. And if I know the statistics then I can apply many statistical concepts like hypothesis testing, probability etc for making good decision.

## Running Code

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

## R code

The Iris dataset:- comprises measurements of iris flowers from three different species: **Setosa**, **Versicolor**, and **Virginica**. Each sample consists of four features: sepal length, sepal width, petal length, and petal width. Additionally, each sample is labeled with its corresponding species.

```
# Get a vector of values from iris dataset
y <- density(iris$Petal.Width)
```

You can add options to executable code like this

## Create R code chunk to determine the class, type, and structure of the object y

```
# Define the class of y object
class(y)
```

[1] "density"

```
# show the type of class y
typeof(y)
```

[1] "list"

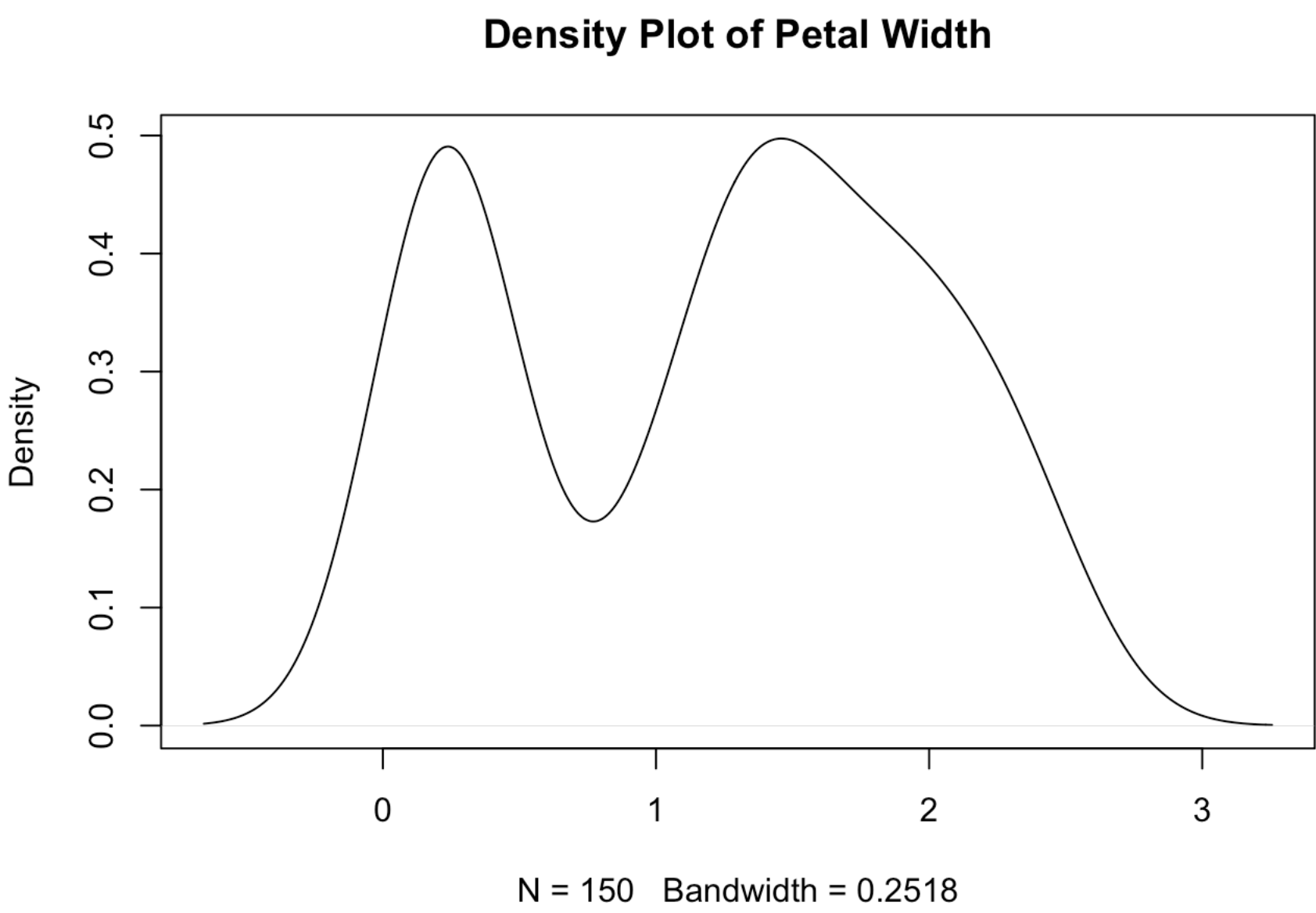
```
# show the structure of class y
structure(y)
```

Call:
 density.default(x = iris\$Petal.Width)

Data: iris\$Petal.Width (150 obs.); Bandwidth 'bw' = 0.2518

x	y
Min. :-0.6555	Min. :0.0005219
1st Qu.: 0.3222	1st Qu.:0.0830968
Median : 1.3000	Median :0.2650517
Mean : 1.3000	Mean :0.2551522
3rd Qu.: 2.2778	3rd Qu.:0.4190136
Max. : 3.2555	Max. :0.4975026

## Create an R code chunk that uses the plot function on y. Hide the R code in the final document by setting echo to false



[1] 4

The `echo: false` option disables the printing of code (only output is displayed).