

Module 8 - Creating a Histogram

Anik Chakraborty, (Intern) R Team, FOSSEE, Department of Mathematics, IIT Bombay

December 29, 2022

Steps to create a histogram in R

This module illustrates the procedure to create a histogram in R.

Step 1. Select a continuous data column and save it in the variable **smv**. In this module, we shall use the data column 'SMV'. We will draw the histogram of *smv* by using the **hist()** function of R.

```
smv = Indian_Agriculture_Data$SMV  
hist(smv)
```

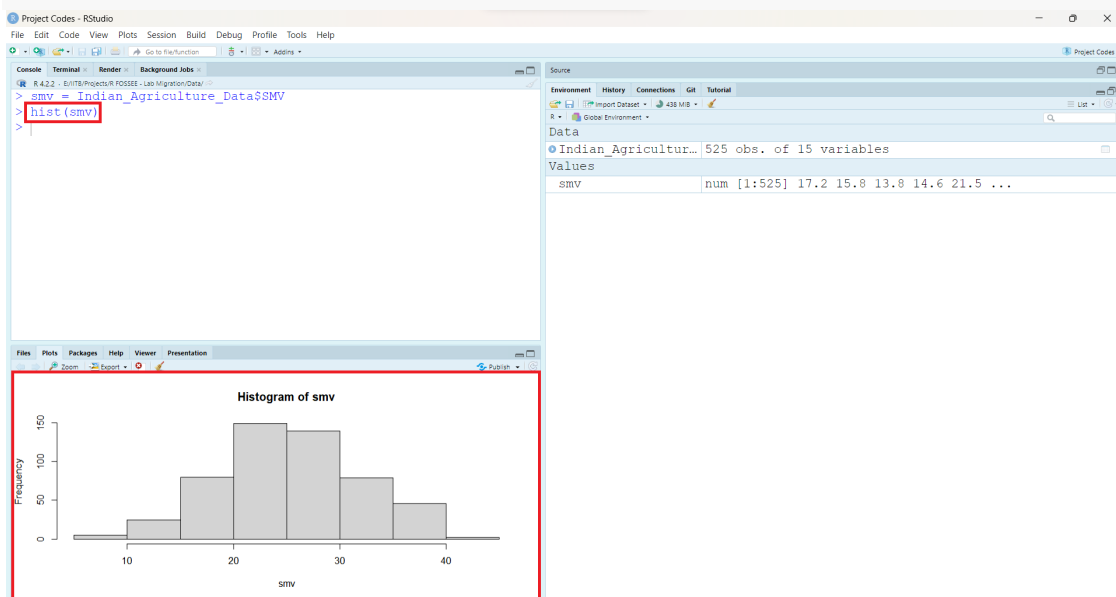


Figure 1: Histogram of **SMV** column data

Step 2. Draw the histogram of a continuous variable showing *density* in the vertical axis by setting the *probability* argument as **TRUE** of **hist()** function of R.

```
hist(smv,  
     probability = TRUE)
```

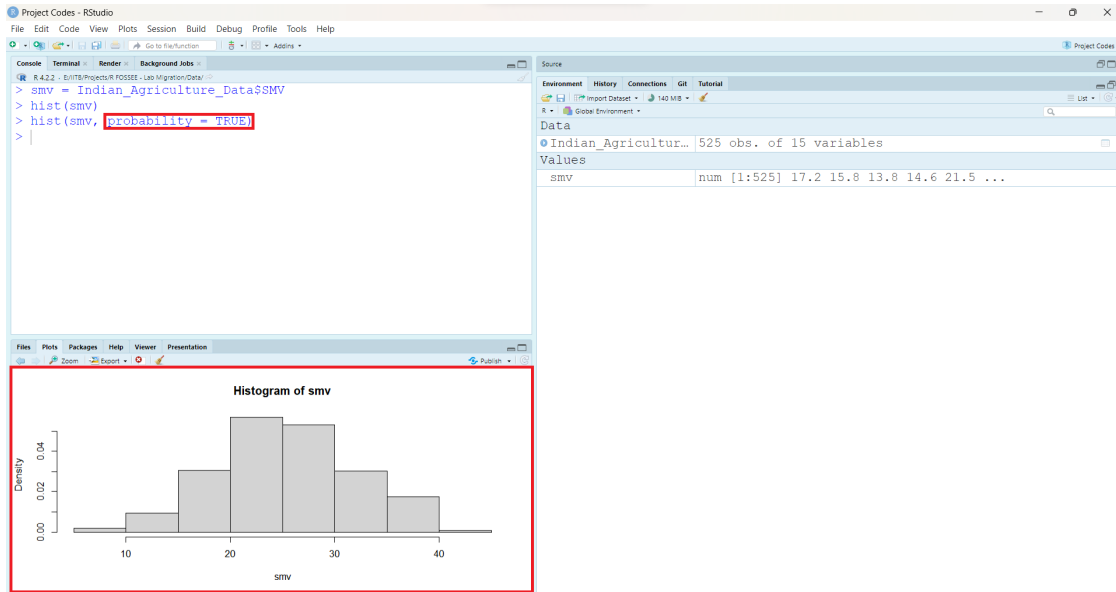


Figure 2: Histogram of **SMV** column data showing density

Step 3. Add a title to the histogram by passing text input to the *main* argument of the **hist()** function.

```

hist(smv,
     probability = TRUE,
     main = "Histogram of Aggregate Soil Moisture Volume Percentage")

```

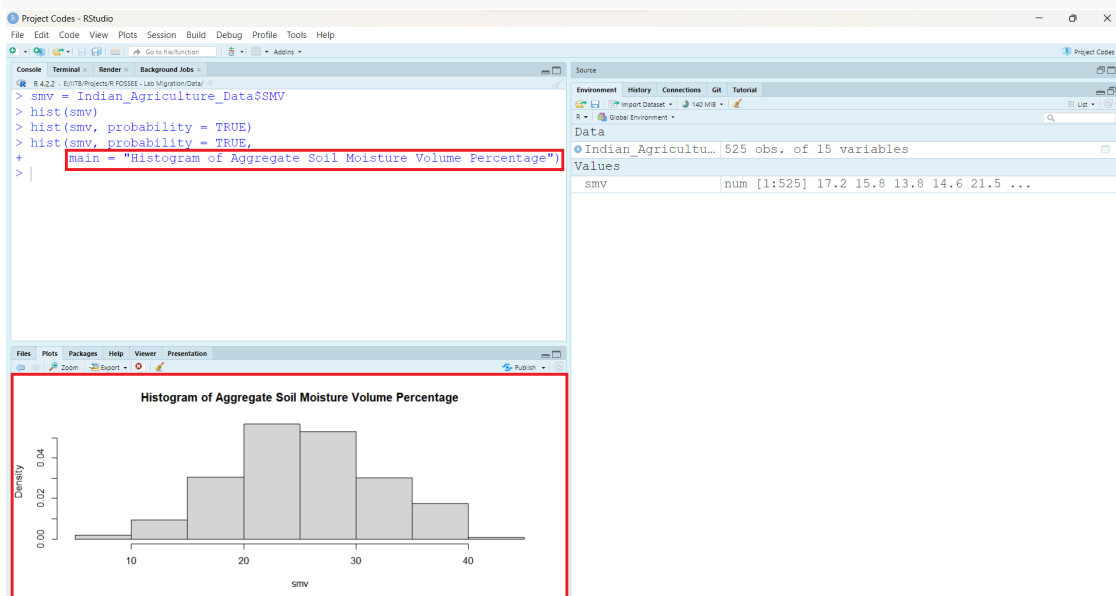


Figure 3: Histogram with added **title**

Step 4. Add *x-label* and *y-label* to the histogram by passing text input to the *xlabs* and *ylabs* argument of the **hist()** function, respectively.

```
hist(smv,
     probability = TRUE,
     main = "Histogram of Aggregate Soil Moisture Volume Percentage",
     xlab = "Aggregate Soil Moisture Volume Percentage",
     ylab = "Frequency Density")
```

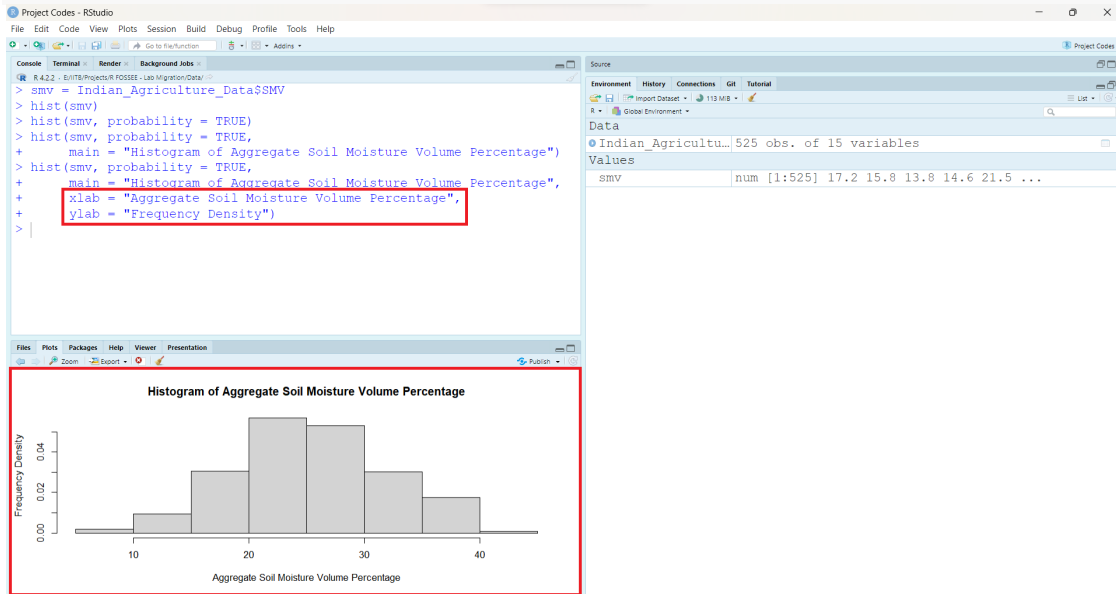


Figure 4: Histogram with added x label and y label

Step 5. Increase the size of x-axis labels and y-axis labels by passing a numerical value to the `cex.lab` argument of the `hist()` function.

Default value of `cex` is 1

```
hist(smv,
     probability = TRUE,
     main = "Histogram of Aggregate Soil Moisture Volume Percentage",
     xlab = "Aggregate Soil Moisture Volume Percentage",
     ylab = "Frequency Density",
     cex.lab = 1.2)
```

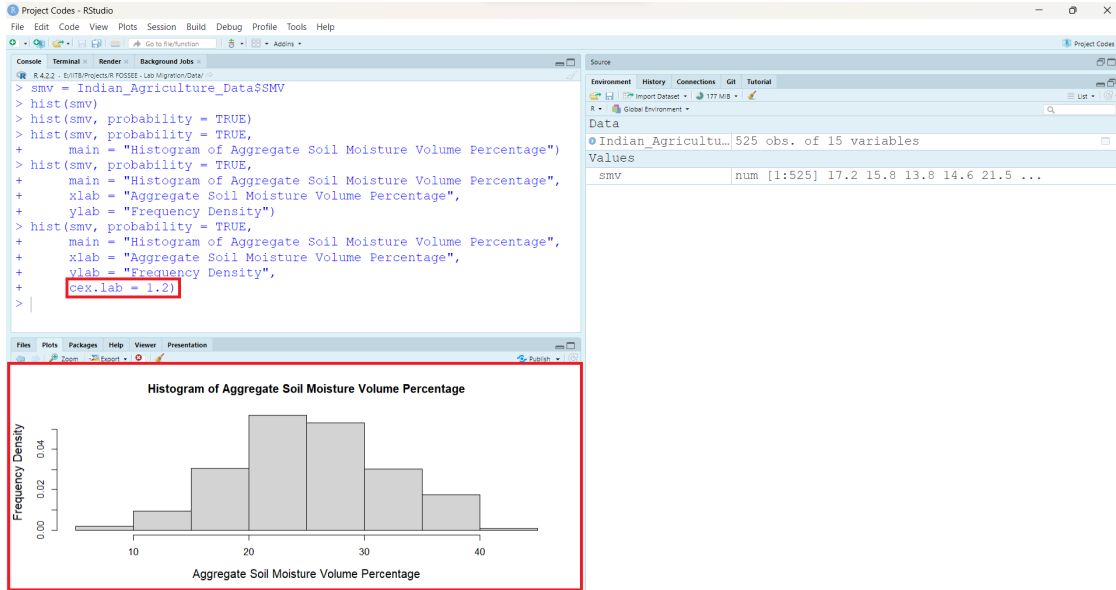


Figure 5: Histogram with increased *axis labels*

Step 6. Add fill and border color to the histogram by passing color name to the *col* argument and the *border* argument of the **hist()** function, respectively.

```

hist(smv,
     probability = TRUE,
     main = "Histogram of Aggregate Soil Moisture Volume Percentage",
     xlab = "Aggregate Soil Moisture Volume Percentage",
     ylab = "Frequency Density",
     cex.lab = 1.2,
     col = "skyblue", border = "brown")

```

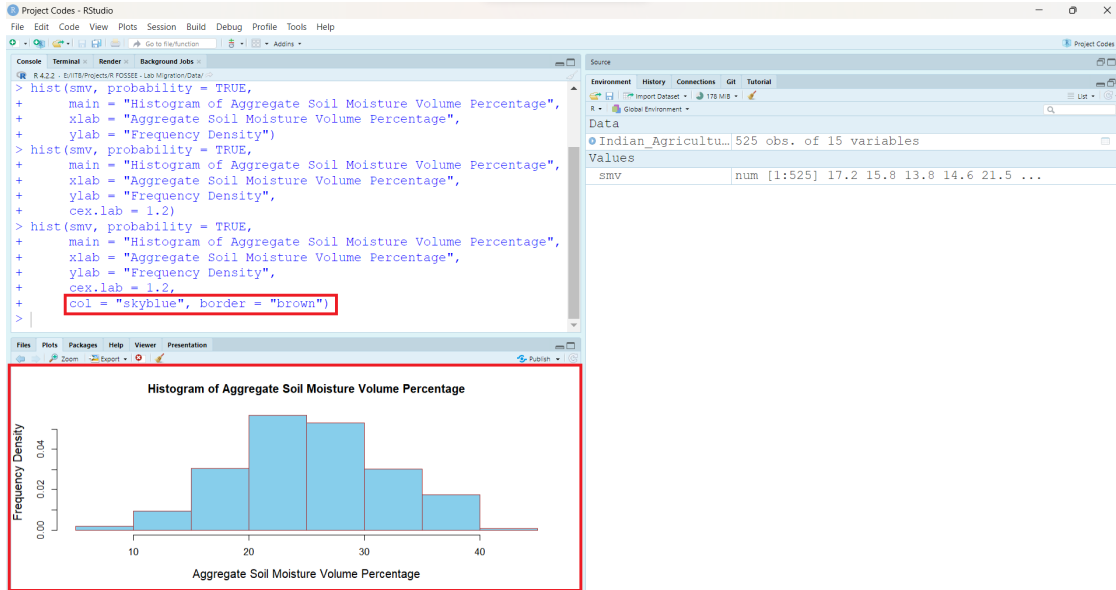


Figure 7: Histogram with customized color and border

Step 7. Set the y-axis limit by passing a numeric vector of length 2 to the *ylim* argument of the **hist()** function.

```

hist(smv,
     probability = TRUE,
     main = "Histogram of Aggregate Soil Moisture Volume Percentage",
     xlab = "Aggregate Soil Moisture Volume Percentage",
     ylab = "Frequency Density",
     cex.lab = 1.2,
     col = "skyblue", border = "brown",
     ylim = c(0, 0.06))

```

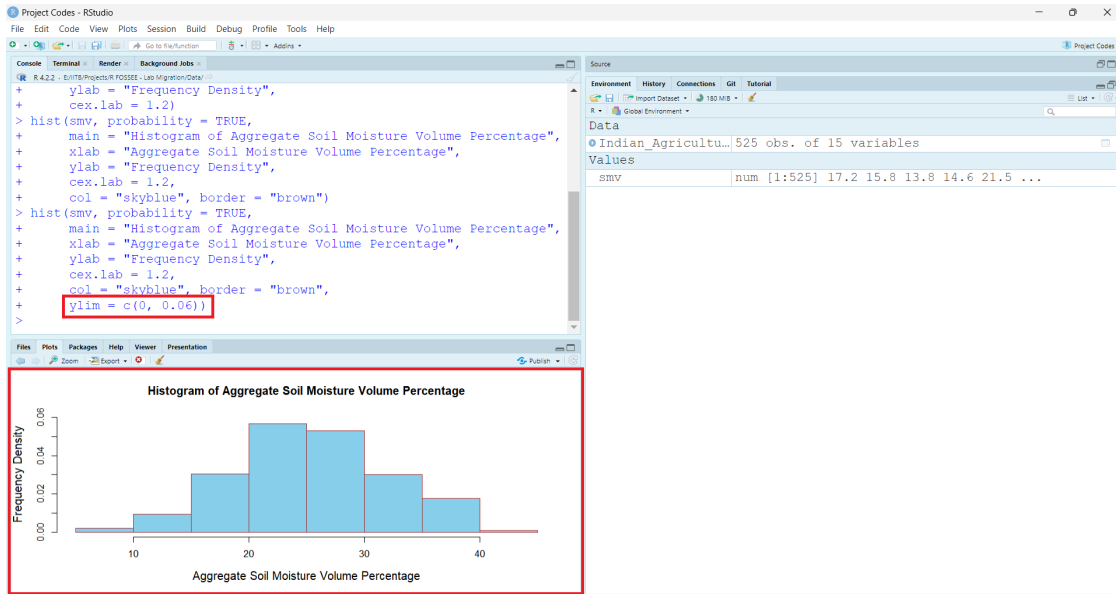


Figure 8: Histogram with *customized y-axis limit*

Spoken Tutorials

For more details, refer to the [Plotting Histograms and Pie Chart Spoken Tutorial](#) video.