

Module 2: Reading data files in R

created by

R Team, FOSSEE, IIT Bombay

and

contributed by

Mr. Digvijay Singh,
Project Research Associate,
R Team, FOSSEE, IIT Bombay

Mrs. Usha Viswanathan,
Sr. Project Manager,
FOSSEE, IIT Bombay

under the guidance of

Prof. Radhendushka Srivastava,
Mathematics Department, IIT Bombay

24 October 2022

Steps to read data files in R

This module explains the methods for importing data from a file present in the working directory to the R software. Before importing data, set the folder containing the data file as the working directory by following the instructions from Module 1.

Data from a **.csv**, **.xlsx** or **.xls** file can be read in R. Let us read data from a **.csv** file in R.

Using RStudio

Step 1. Left-click on the name of the data file in the **Files** tab of the **Plots and Files** window.

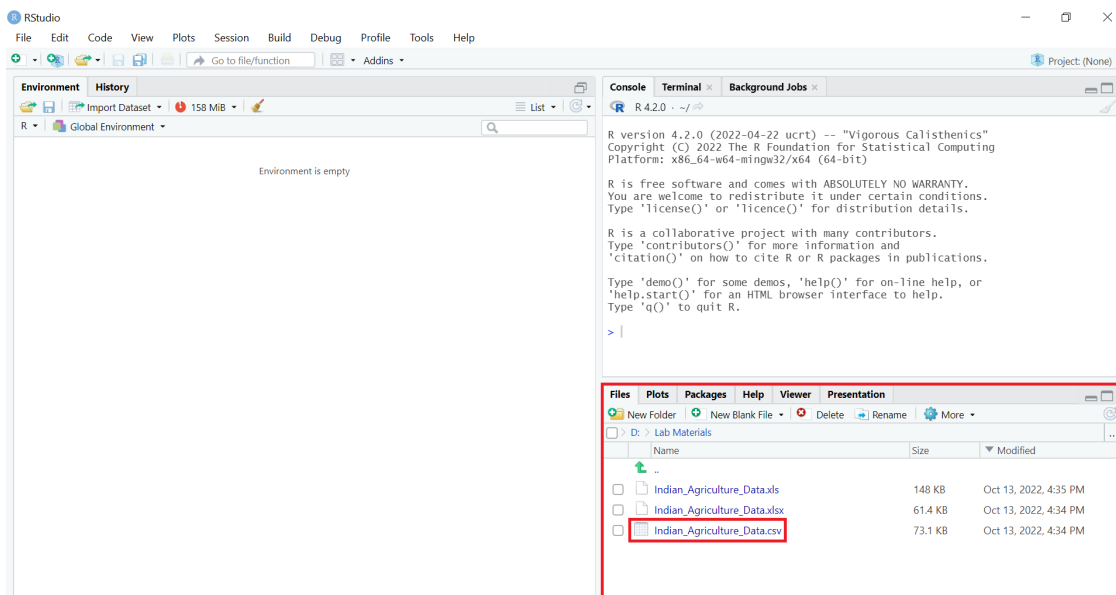


Figure 1: Left-click on the name of the data file to be read.

Step 2. Select the option **Import Dataset...** from the dropdown menu.

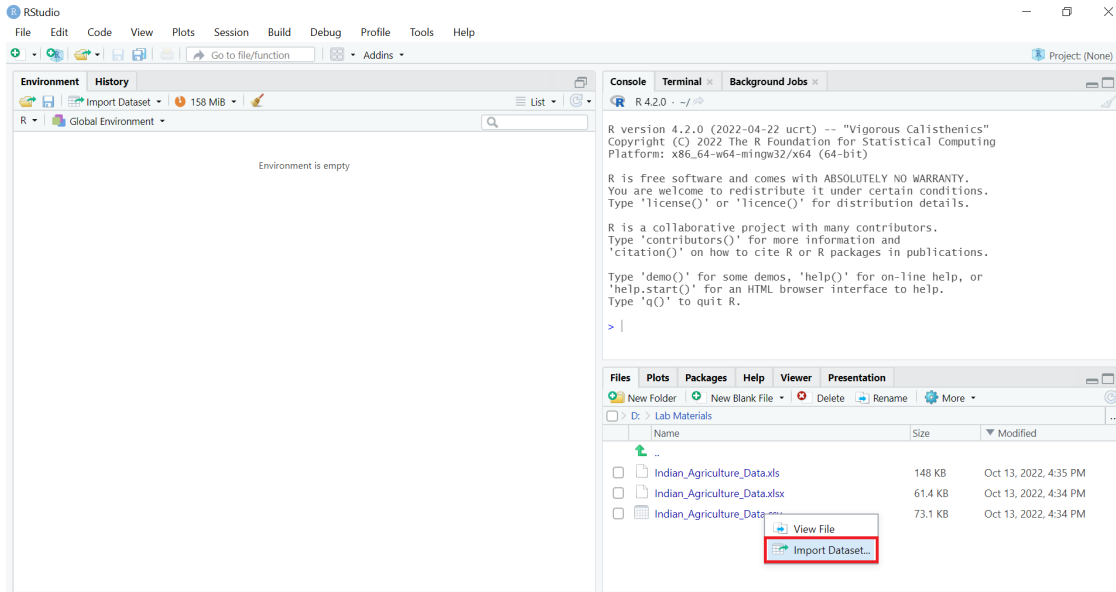


Figure 2: Select the **Import Dataset...** option.

A window will appear on the screen. It will display the data to be imported along with default data types for each column.

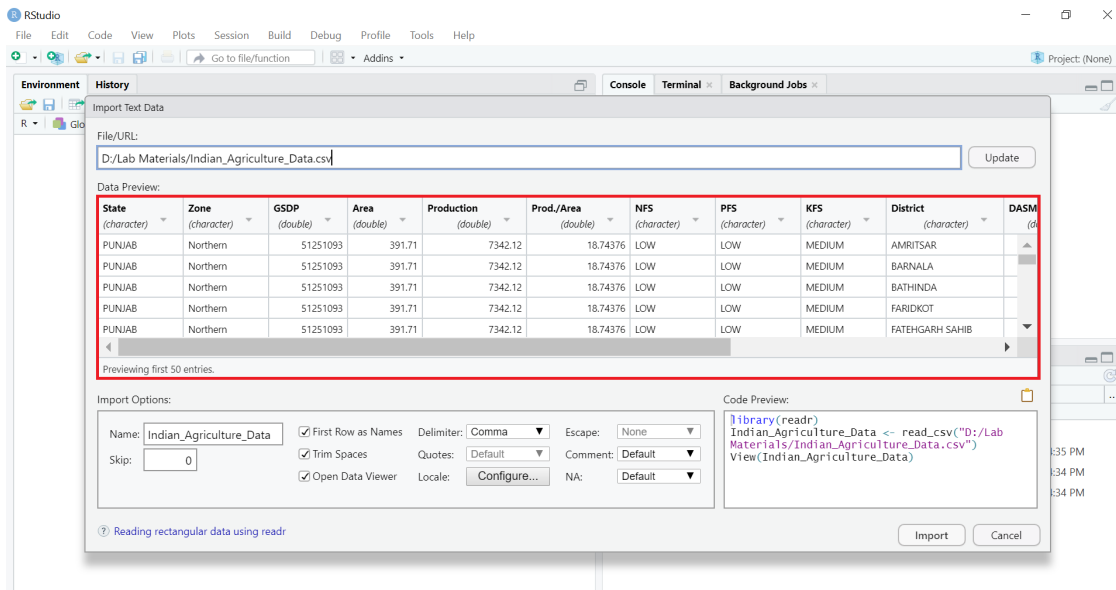


Figure 3: **Import Text Data** window.

Step 3. To change the default data type of a column, left-click anywhere in the box containing the column header. A drop down list will appear. Select the desired data type from the drop down list.

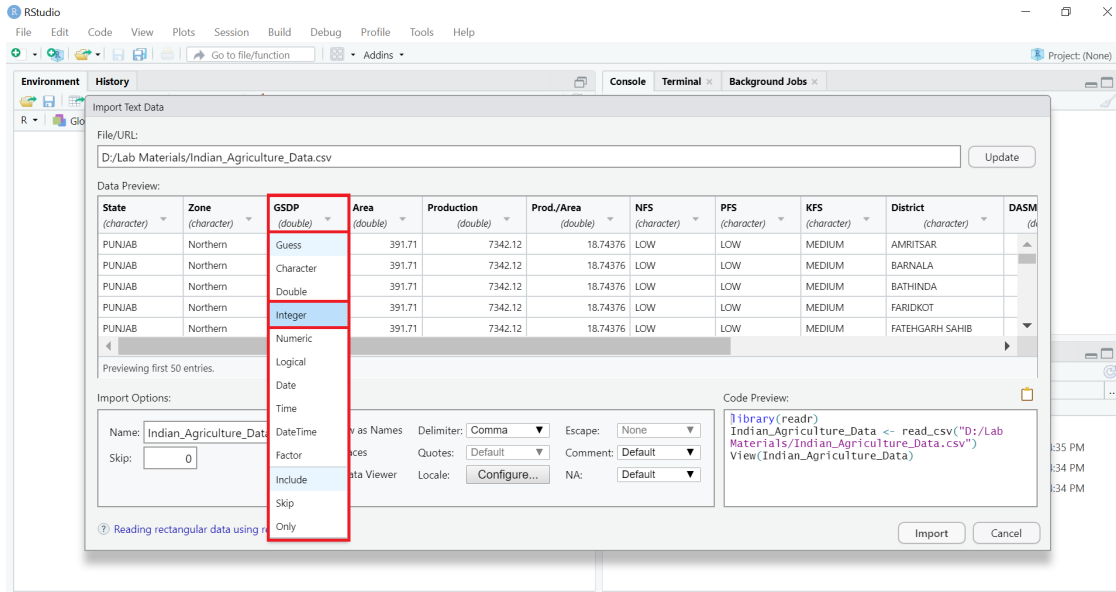


Figure 4: Change default data type of a column.

In the above figure the data type of **GSDP** column is changed from *double* to *integer*.

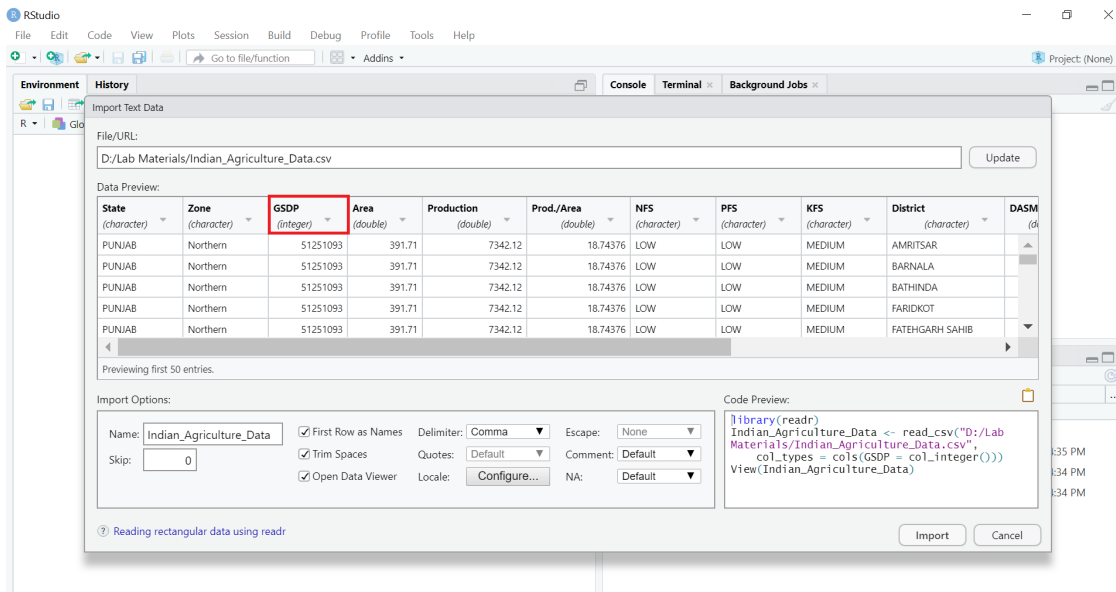


Figure 5: **GSDP** column with integer data type.

Step 4. Left-click on the **Import** button present at the bottom-right corner of the window.

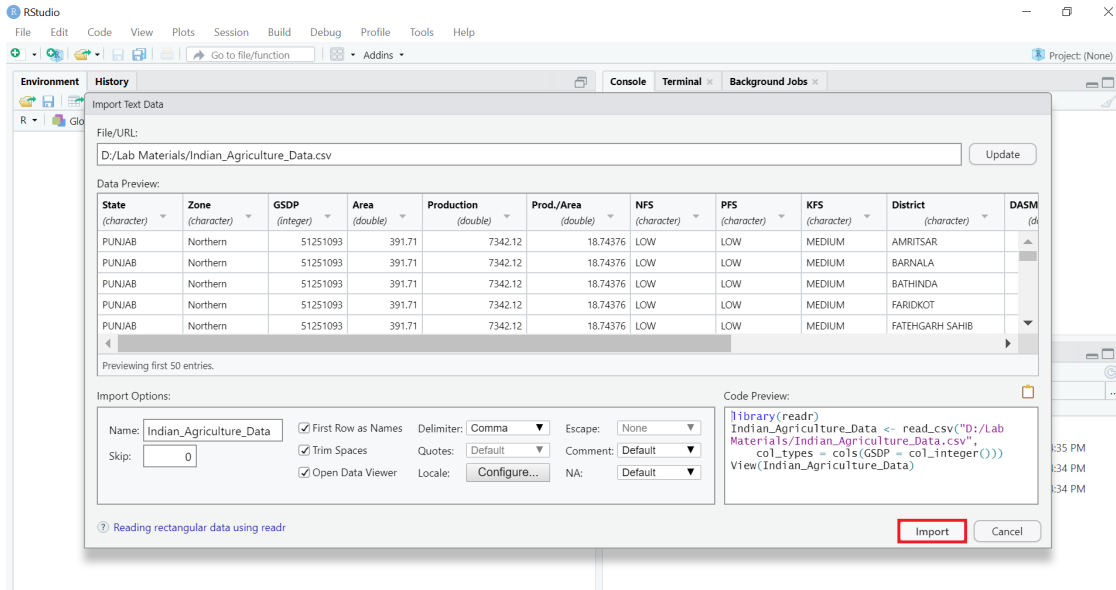


Figure 6: Select **Import** button to import the data.

The data after importing will automatically get displayed on the **Source** window.

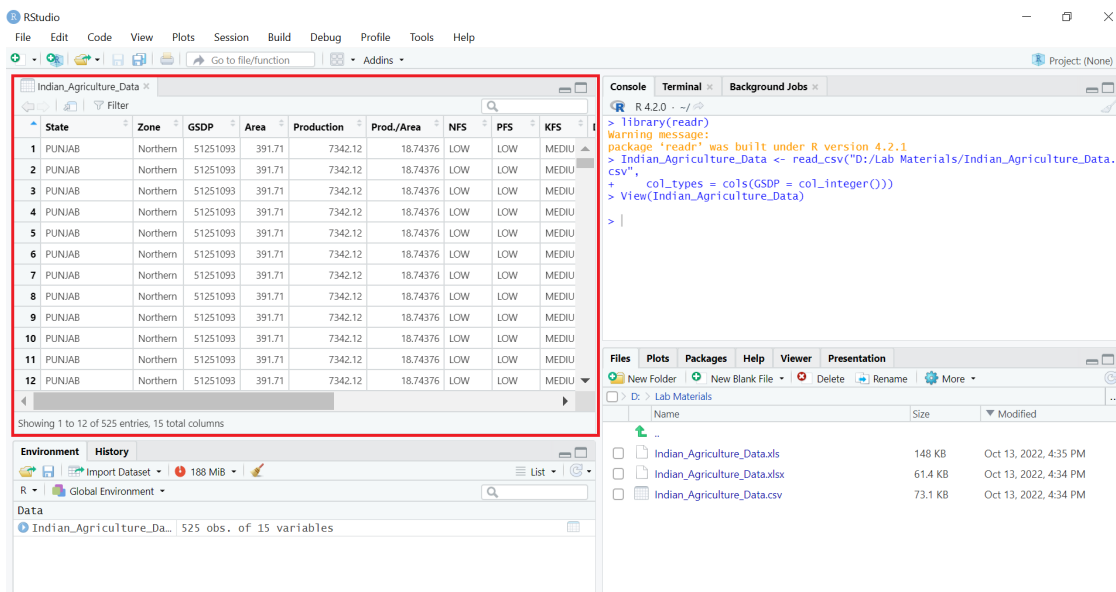


Figure 7: Data imported and viewed in R.

Follow the same steps to read data from an **.xlsx** or **.xls** file.

Alternatively, using R console to read data files

To read and view data from **.csv** file execute the following commands:

```
# Loading "readr" library to use its "read_csv()" function
library(readr)
# Reading "Indian_Agriculture_Data.csv" file's data using the "read_csv()"
function and storing it in the "Indian_Agriculture_Data" object
Indian_Agriculture_Data <- read_csv("D:/Lab
Materials/Indian_Agriculture_Data.csv")
# Viewing the contents of "Indian_Agriculture_Data" object
View(Indian_Agriculture_Data)
```

The above commands shall read and display the data present in the **"Indian_Agriculture_Data.csv"** file.

To read and view data from **.xlsx** or **.xls** file execute the following commands:

```
library(readxl)
Indian_Agriculture_Data <- read_excel("D:/Lab
Materials/Indian_Agriculture_Data.xls")
View(Indian_Agriculture_Data)
```

The above commands shall read and display the data present in the **"Indian_Agriculture_Data.xls"** file. Same commands can be used to read data from **"Indian_Agriculture_Data.xlsx"** file by changing the input data file name.

Spoken Tutorials

For more details, refer to the [Introduction to Data Frames in R Spoken Tutorial](#) video.