## Module 3 - Extracting a specific column from a data frame

created by

Mr. Digvijay Singh, Project Research Associate, R Team, FOSSEE, IIT Bombay Ms. Usha Viswanathan, Sr. Project Manager, FOSSEE, IIT Bombay

24 October 2022

#### Steps to extract a column from a data frame in R

This module shows the procedure to extract a specific column from a data frame in R software. In this module we shall extract a column from the data used in Module 2. To do so, first read data from a file after setting its location as the working directory using the instructions given in Modules 1 and 2.

### **Using RStudio**

**Step 1.** Type the first three letters of the data frame's name from which a column has to be extracted, either in the R console window or in an R script in the Source window. A drop down list will appear.

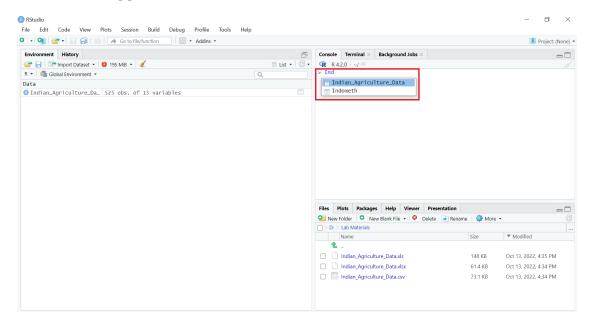


Figure 1: Drop down list from the first three letters of the data frame's name.

**Step 2.** Select the name of the data frame from the drop down list. It will auto-complete the name of the data frame.

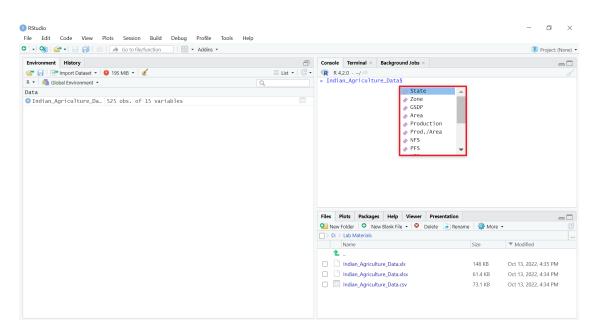
RStudio File Edit Code View Plots Session Build Debug Profile Tools Help		- 0 ×
•     • </th <th></th> <th>Project: (None) •</th>		Project: (None) •
Environment History	Console Terminal × Background Jobs ×	-0
🐨 📊 🖙 Import Dataset 👻 🚯 195 MiB 🔹 🔏 📃 List 🔹 🔞 🖛	R 4.2.0 · ~/	
R 💌 🍓 Global Environment 👻 🔍	> Ind	
Data O Indian_Agriculture_Da_ 525 obs. of 15 variables	Indometh	
	Files         Plots         Packages         Help         Viewer         Presentation           Image: Second S	e 🏘 More 👻 🕝
	Name	Size Modified
	1	
	Indian_Agriculture_Data.xls	148 KB Oct 13, 2022, 4:35 PM
	Indian_Agriculture_Data.xlsx	61.4 KB Oct 13, 2022, 4:34 PM
	Indian_Agriculture_Data.csv	73.1 KB Oct 13, 2022, 4:34 PM

Figure 2: Selecting data frame name from the drop down list.

🔹 🧐 🚰 🔹 🔚 📑 🛛 🦽 Go to file/function 🔤 🛛 🔛 🔹 Addins 🔹					
vironment History	Ð	Console Terminal	Background Jobs	×	-
🕨 📄 🐨 Import Dataset 🔹 🍪 195 MiB 👻 💰		😱 R 4.2.0 · ~/ 🗇			
📕 Global Environment 👻	٩,	> Indian_Agricu	ture_Data		
Ta la					
ndian_Agriculture_Da 525 obs. of 15 variables					
		Files Plots Pack	ages Help Viewer	Presentation	
				Presentation Delete 🖷 Rename 🏼 🆓 Mo	re •
		New Folder     New Folder     D: > Lab Materia	New Blank File 🔹 🚨	Delete 💽 Rename 🏼 🎲 Mo	re •
		New Folder     New Folder     D: > Lab Materia     Name	New Blank File 🔹 🚨		
		New Folder     New Folder     D: > Lab Materia	New Blank File 🔹 🚨	Delete 💽 Rename 🏼 🎲 Mo	re •
		New Folder  New Folder  D: > Lab Materia  Name  L	New Blank File • 3	Delete 💽 Rename   🏟 Mo	re ▼ ▼ Modified
		<ul> <li>New Folder</li> <li>D: &gt; Lab Materia</li> <li>Name</li> <li>Lab Materia</li> <li>Name</li> <li>Indian_Agr</li> </ul>	New Blank File - O	Delete Rename 🔯 Mo	re ▼ ▼ Modified Oct 13, 2022, 4:35 PM
		<ul> <li>New Folder</li> <li>D: &gt; Lab Materia</li> <li>Name</li> <li>Lab Materia</li> <li>Name</li> <li>Indian_Agr</li> </ul>	New Blank File • 3	Delete 💽 Rename   🏟 Mo	re ▼ ▼ Modified
		<ul> <li>New Folder</li> <li>D: &gt; Lab Materia</li> <li>Name</li> <li>Lab Indian_Agr</li> <li>Indian_Agr</li> </ul>	New Blank File - O	Delete Rename 🔯 Mo	re ▼ ▼ Modified Oct 13, 2022, 4:35 PM

*Figure 3: Auto-completed name of the data frame.* 

**Step 3.** Type a dollar sign, **\$**, after the name of the data frame. It will display the list of its column names/headers.



*Figure 4: List of the column names/headers.* 

**Step 4.** Select a column for extraction by left-clicking on its name in the drop down list. Press **Enter** to view its contents.

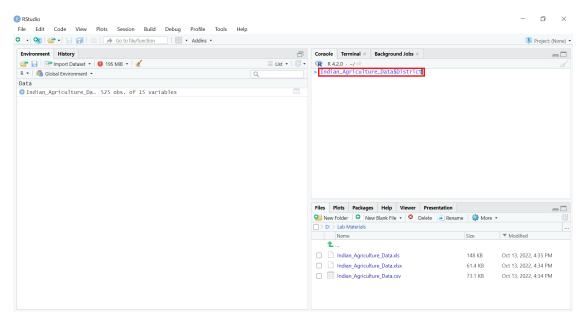


Figure 5: Selected column for extraction.

🔹 🚳 🐨 🐑 🕞 👘 🦾 👘 Go to file/function 🔤 🛛 🔀 🔹 Addins 👻			\$	Project: (None)
invironment History	ð	Console Terminal × Background Jobs ×		
🚰 📊 🖙 Import Dataset 👻 😃 195 MiB 👻 🔏	≡ List • 🛛 🕑 •	😨 R 4.2.0 · -/ 🖗		ß
🧧 🛑 Global Environment 👻	Q,	<pre>&gt; Indian_Agriculture_Data\$District [1] "AMRITSAR" "BARNALA"</pre>		4
Data ● Indian_Agriculture_Da_ 525 obs. of 15 variables		13         "BATHLINDA"         "FARIDOS           15         "FATELOSARI SANIPA"         "CURDASPI           16         "HOSHLARPUR"         "LUDHLAN           17         "HOSHLARPUR"         "LUDHLAN           18         "KAPURASA"         "LUDHLAN           131         "MANSA"         "PATLALA"           133         "MURTSAR"         "SANGRUF           131         "MURTSAR"         "SANGRUF           131         "MURTSAR"         "BIRBHLM"           121         "MADAH"         "MURSHLD           121         "ADALA"         "MURSHLD           123         "MANACHUR"         "BIRGUSAR"           123         "MANACHUR"         "BIRGUSAR"           123         "ADANCHUR"         "BIRGUSAR"           123         "BANACHUR"         "DOPALCA"           123         "ANANCHUR"         "DOPALCA"           123         "JAMUI"         "COPALCA"           123         "ANANCHUR"         "DOPALCA"           123         "ANANCHUR"         "DOPALCA"           123         "ANANCHUR"         "DOPALCA"           123         "ANANCHUR"         "DOPALCA"           124         "ANACHUR"	R" " BAD" AD" T" T" J" J" J" T" AT"	
		Files Plots Packages Help Viewer Presentat		
		New Folder     New Blank File      Delete     Delete	Rename – 🎲 More 👻	
		Name	Size Modified	
		1.		
		Indian Agriculture Data.xls	148 KB Oct 13, 2022,	4:35 PM
		Indian Agriculture Data.xlsx	61.4 KB Oct 13, 2022,	4:34 PM

Figure 6: Selected column's data.

# Alternatively, using R console to extract a specific column from a data frame

To extract a specific column from a data frame execute the following command with the column name placed after the **\$** sign:

```
Indian_Agriculture_Data$District
```

### **Spoken Tutorials**

For more details, refer to the Introduction to Data Frames in R Spoken Tutorial video.