

Welcome to “Introduction to R and RStudio.”

After watching this video, you will be able to

explain what is R,

list R capabilities,

describe RStudio environment, and

list the R libraries for data science.

R is a statistical programming language.

It is a powerful tool for data processing and manipulation, statistical inference, data analysis, and machine learning algorithm.

Based on 2017 analysis, it was found that R is used most by academics, healthcare, and the government.

R supports importing of data from different sources like flat files, databases, web, and statistical software such as SPSS and STATA.

R is a preferred language for some data scientists because R functions are easy to use.

It is also known for producing great visualizations and contains packages to handle data analysis without the need to install additional libraries.

A popular integrated development environment for developing and running the R language source code and programs is RStudio.

It improves and increases productivity with the R language.

R studio includes:

a syntax-highlighting editor that supports direct code execution and

a place where you can keep a record of your work,

a Console for typing R commands,

a workspace and History tab that shows the list of R objects you created during your R session and the history of all previous commands, and finally, Files, Plots, Packages, and Help tabs.

The Files tab shows files in your working directory.

The Plots tab displays the history of plots you have created.

You can also export plots to PDF or image files.

The Packages tab displays external R packages available on your local computer.

And, the Help tab provides help on R resources, R studio support,

packages, and many more.

If R is your tool choice for data science,

here are some popular R libraries available in the Data Science community:

dplyr for manipulating data,

stringr for manipulating strings,

ggplot for visualizing data, and

caret for machine learning.

To get you up and learning quickly, we have provided you with an R Studio

virtual environment as part of the Skills Network Labs.

This virtual lab environment is designed to assist you to easily practice what you learn

in the course and skip the need to create an account or download or install anything.

In this video, you learned

the capabilities of R and its uses in Data Science,

the RStudio interface for running R codes, and

popular R packages for Data Science.