# Introduction to R

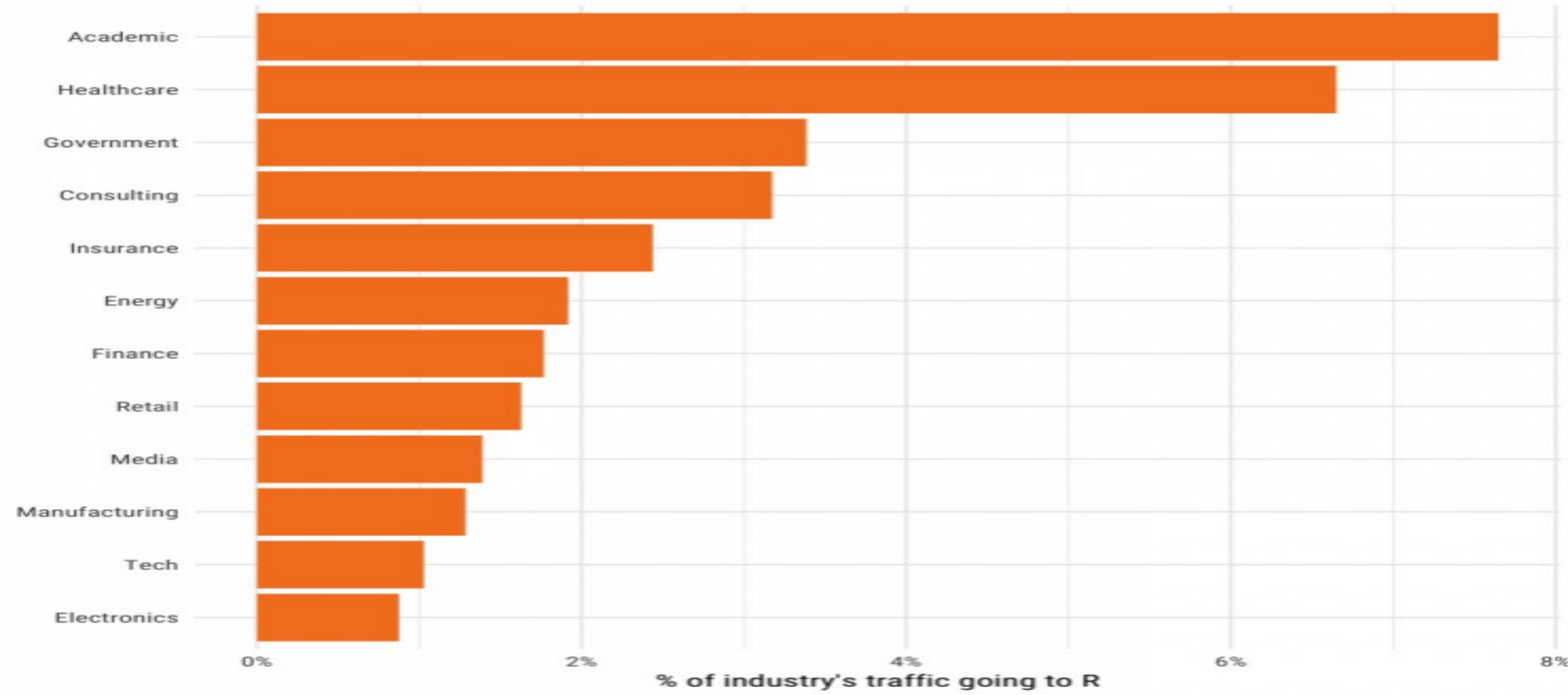
- What is R ?

- R is a programming language and a free software environment

- Primarily used for statistical computing and graphics.

- It provides a wide variety of statistical and graphical techniques, including linear and nonlinear modelling, time-series analysis, classification, clustering, and more.

# Who is use R ?


Academic
Healthcare
Government
Consulting
Insurance
Energy
Finance
Retail
Media
Manufacturing
Tech
Electronics
0%
2%
4%
6%
8%
% of industry's traffic going to R

# Why learning R?

- R is open source, so it's free.
- R is cross-plateform compatible, so it can be installed on Windows, MacO and Linux
- R provides a wide variety of statistical techniques and graphical capabilities.
- R provides the possibility to make a reproducible research by embeddin script and results in a single file.
- R is highly extensible and it has thousands of well-documented extension (named R packages) for a very broad range of applications in statistica analysis, health care, bioinformatics and more

# Rstudio

➔R Studio is an integrated development environment (IDE) that supports the R programming language.

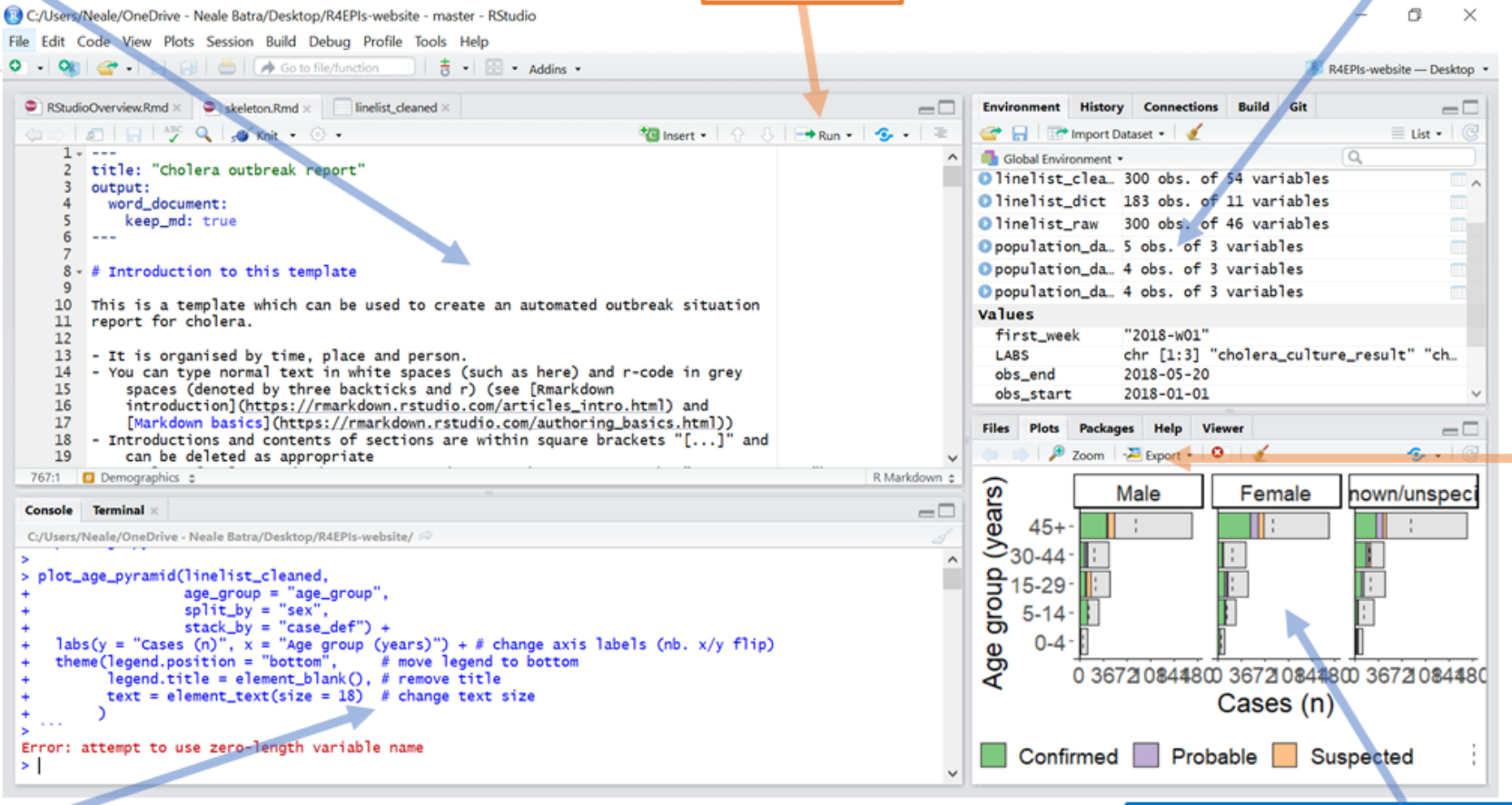
➔It has a user-friendly interface for writing, running, debugging, and visualizing R code

**Source Pane**
Edit and run scripts (e.g. Rmarkdown templates), and view datasets

*Tip*: Run script

**Environment Pane**
Overview of objects (datasets, parameters, lists, etc.) you have imported or created.

*Tip*:
Start new script




*Tip*: Zoom and export plots

**R Console Pane**
R commands run are shown here, and non-graphic output and errors are displayed

**Plots, Packages, and Help Pane**
Commonly used to view graphics, install packages, and view help

# R Packages

- R packages consist of R functions, compiled code, and sample data. In the R environment, they are placed in a directory named "library"
- R automatically installs a set of packages during installation
- More packages are can be added later, as they are required for a specific purpose
- R packages could be easily installed using the default function
- # install.packages(“package_name”)
- Or using the pacman package and it’s p_load(package_name) function

# Most used packages in R and why ?

➔**ggplot2**: A powerful package for creating high-quality and customizable graphs, It provides a high-level interface for creating a wide range of plots.

➔**dplyr**: Another package by Hadley Wickham, dplyr provides a set of functions for data manipulation and transformation. It offers intuitive verbs like **filter**(), **select**(), **mutate**(), **group_by**(), and **summarize**() for efficiently handling data manipulation tasks

➔**tidyverse**: A collection of packages designed to work together seamlessly for data manipulation, visualization, and analysis in R

# R and R-studio installation

R installation instruction

- Windows : https://cran.mirror.ac.za/bin/windows/
- Linux : https://cran.mirror.ac.za/bin/linux/
- MacOS : https://cran.mirror.ac.za/bin/macosx/
- R studio installation
- https://posit.co/download/rstudio-desktop/

# Descriptive Statistics and Graphics

An introduction, using R:

http://www.sthda.com/english/wiki/descriptive-statistics-and-graphics