

Introduction to R and Tidyverse Cheatsheet

The tables below consist of valuable functions and commands that will help you through this module.

Each table represents a different library/tool and the corresponding commands.

Please note that these tables are not intended to tell you all the information you need to know about each command.

The hyperlinks found in each piece of code will take you to the documentation for further information on the usage of each command.

Please be aware that the documentation will generally provide information about the given function's most current version (or a recent version, depending on how often the documentation site is updated).

This will usually (but not always!) match what you have installed on your machine.

If you have a different version of R or other R packages, the documentation may differ from what you have installed.

Table of Contents

- [Base R](#)
- [tidyverse](#)
 - [dplyr](#)
 - [ggplot2](#)
 - [readr](#), [fs](#), [tibble](#), [tidyr](#)

Base R

Read the [Base R](#) documentation.

Library/Package	Piece of code	What it's called	What it does
Base R	<code>library()</code>	Library	Loads and attaches additional packages to the R environment.
Base R	<code><-</code>	Assignment operator	Assigns a name to something in the R environment.
Base R	<code> ></code>	Pipe operator	Funnels an object from the output of one function to input of the next function
Base R	<code>c()</code>	Combine	Combines values into a vector or list.
Base R	<code>%in%</code>	"in" logical operator	Checks if the given value(s) on the left side of the operator are in the vector or other R object defined on the right side of the operator. It returns a logical <code>TRUE</code> or <code>FALSE</code> statement. This resource also provides a helpful explanation about its usage.
Base R	<code>rm(x)</code>	Remove	Removes object(s) <code>x</code> from your environment.
Base R	<code>==, <=, >=, !=</code>	Relational Operators	These are binary operators which allow for the comparison of values in an object.
Base R	<code>str(x)</code>	Object Structure	Gets a summary of the object <code>x</code> structure.
Base R	<code>class(x)</code>	Object Class	Returns the type of the values in object <code>x</code> .
Base R	<code>nrow(x); ncol(x)</code>	Number of Rows; Number of Columns	Get the number of rows and the number of columns in an object <code>x</code> , respectively.
Base R	<code>length(x)</code>	Length	Returns how long the object <code>x</code> is.
Base R	<code>min(x)</code>	Minimum	Returns the minimum value of all values in an object <code>x</code> .
Base R	<code>sum(x)</code>	Sum	Returns the sum of all values (values must be integer, numeric, or logical) in object <code>x</code> .
Base R	<code>mean(x)</code>	Mean	Returns the arithmetic mean of all values (values must be integer or numeric) in object <code>x</code> or logical vector <code>x</code> .
Base R	<code>log(x)</code>	Logarithm	Gives the natural logarithm of object <code>x</code> . <code>log2(x)</code> can be used to give the logarithm of the object in base 2. Or the base can be specified as an argument.
Base R	<code>head(); tail()</code>	Head; Tail	Returns the top 6 (<code>head()</code>) or bottom 6 (<code>tail()</code>) of an object in the environment by default. You can specify how many rows you want by including the <code>n =</code> argument.

Library/Package	Piece of code	What it's called	What it does
Base R	<code>factor(x)</code> or <code>as.factor(x)</code>	Factor	Coerces object <code>x</code> into a factor (which is used to represent categorical data). This function can be used to coerce object <code>x</code> into other data types, i.e., <code>as.character</code> , <code>as.numeric</code> , <code>as.data.frame</code> , <code>as.matrix</code> , etc.
Base R	<code>levels(x)</code>	Levels attributes	Returns or sets the value of the levels in an object <code>x</code> .
Base R	<code>summary(x)</code>	Object summary	Returns a summary of the values in object <code>x</code> .
Base R	<code>data.frame()</code>	Data Frame	Creates a data frame where the named arguments will be the same length.
Base R	<code>sessionInfo()</code>	Session Information	Returns the R version information, the OS, and the attached packages in the current R session.
Base R	<code>file.path()</code>	File path	Constructs the path to a desired file.
Base R	<code>dir()</code>	Directory	Lists the names of the files and/or directories in the named directory.
Base R	<code>getwd()</code>	Get working directory	Finds the current working directory.
Base R	<code>setwd()</code>	Set working directory	Changes the current working directory.
Base R	<code>dir.exists()</code>	Directory exists	Checks the file path to see if the directory exists there.
Base R	<code>dir.create()</code>	Create directory	Creates a directory at the specified path.
Base R	<code>apply()</code>	Apply	Returns a vector or list of values after applying a specified function to values in each row/column of an object.
Base R	<code>round()</code>	Round	Rounds the values of an object to the specified number of decimal places (default is 0).
Base R	<code>names()</code>	Names	Gets or sets the names of an object.
Base R	<code>colnames()</code>	Column names	Gets or sets the column names of a matrix or data frame.
Base R	<code>all.equal()</code>	All equal	Checks if two R objects are nearly equal.
Base R	<code>all()</code>	All	Checks if all of the values are <code>TRUE</code> in a logical vector.
Base R	<code>t()</code>	Transpose	Returns the transpose of a matrix or data frame. If given a data frame, returns a matrix.

tidyverse

Read the [tidyverse package documentation](#), as well as the [philosophy behind the tidyverse](#).

dplyr

Read the [dplyr package documentation](#), and a [vignette on its usage](#).

Library/Package	Piece of code	What it's called	What it does
dplyr / magrittr	<code>%>%</code>	Pipe operator	Funnels an object from the output of one function to input of the next function (used like the base pipe <code> ></code> found in R 4.1 and later, but can be used in earlier versions of R)
dplyr	<code>filter()</code>	Filter	Returns a subset of rows matching the conditions of the specified logical argument
dplyr	<code>arrange()</code>	Arrange	Reorders rows in ascending order. <code>arrange(desc())</code> would reorder rows in descending order.
dplyr	<code>select()</code>	Select	Selects columns that match the specified argument
dplyr	<code>mutate()</code>	Mutate	Adds a new column that is a function of existing columns
dplyr	<code>summarize()</code>	Summarize	Summarizes multiple values in an object into a single value. This function can be used with other functions to retrieve a single output value for the grouped values. <code>summarize</code> and <code>summarise</code> are synonyms in this package. However, note that this function does not work in the same manner as the base R <code>summary</code> function.
dplyr	<code>rename()</code>	Rename	Renames designated columns while keeping all variables of the data.frame
dplyr	<code>group_by()</code>	Group By	Groups data into rows that contain the same specified value(s)
dplyr	<code>inner_join()</code>	Inner Join	Joins data from two data frames, retaining only the rows that are in both datasets.

ggplot2

Read the [ggplot2 package documentation](#), an [overall reference for ggplot2 functions](#), and a [vignette on the usage of the ggplot2 aesthetics](#).

Additional vignettes are available from the "Articles" dropdown menu on this webpage.

Library/Package	Piece of code	What it's called	What it does
ggplot2	<code>ggplot()</code>	GG Plot	Begins a plot that is finished by adding layers.
ggplot2	<code>aes()</code>	Aesthetic Mappings	Designates how variables in the data object are mapped to the visual properties of the ggplot.
ggplot2	<code>geom_boxplot()</code>	Boxplot	Creates a boxplot when added as a layer to a <code>ggplot()</code> object.
ggplot2	<code>geom_density()</code>	Density Plot	Creates a smoothed plot when added as a layer to a <code>ggplot()</code> object based on the computed density estimate.
ggplot2	<code>geom_point()</code>	Scatterplot	Creates a scatterplot when added as a layer to a <code>ggplot()</code> object.
ggplot2	<code>geom_line()</code>	Line plot	Creates a line plot when added as a layer to a <code>ggplot()</code> object by connecting the points in order of the x axis variable.
ggplot2	<code>geom_hline()</code>	Horizontal line	Annotates a plot with a horizontal line when added as a layer to a <code>ggplot()</code> object
ggplot2	<code>geom_vline()</code>	Vertical line	Annotates a plot with a vertical line when added as a layer to a <code>ggplot()</code> object
ggplot2	<code>theme_classic()</code>	Classic Theme	Displays <code>ggplot</code> without gridlines. The ggtheme documentation has descriptions on additional themes that can be used.
ggplot2	<code>labs()</code>	Labels	Modify labels (axis, title, legends) on a <code>ggplot()</code> object.
ggplot2	<code>xlab(); ylab(); ggtitle()</code>	X Axis Labels; Y Axis Labels; GG Title	Alternative individual functions to add individual plot labels: x-axis, y-axis, and title, respectively.
ggplot2	<code>facet_wrap()</code>	Facet Wrap	Plots individual graphs using specified variables to subset the data.
ggplot2	<code>ggsave()</code>	GG Save	Saves the last plot in working directory.
ggplot2	<code>last_plot()</code>	Last plot	Returns the last plot produced.

readr, fs, tibble, tidyr

Read the [readr package documentation](#) and a [vignette on its usage](#).

Read the [fs package documentation](#).

Read the [tibble package documentation](#) and a [vignette on its usage](#).

Read the [tidyr package documentation](#) and a [vignette on its usage](#).

<https://fs.r-lib.org/reference/index.html>

Library/Package	Piece of code	What it's called	What it does
readr	<code>read_tsv()</code>	Read TSV	Reads in a TSV file from a specified file path. This function can be tailored to read in other common types of files, e.g. <code>read_csv()</code> , <code>read_rds()</code> , etc.
fs	<code>dir_create()</code>	Create directory	Create a directory, unless the directory already exists.
tibble	<code>column_to_rownames()</code>	Column to Rownames	Transforms an existing column called by a string into the rownames.
tibble	<code>rownames_to_column()</code>	Rownames to Column	Transforms the rownames of a data frame into a column (which is added to the start of the data frame). The string supplied as an argument will be the name of the new column.
tidyr	<code>pivot_longer()</code>	Pivot Longer	Lengthens a data frame by increasing the number of rows and decreasing the number of columns.