

Intro 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.

Base R

Read the Base R documentation [here](#).

Library/Package	Piece of code	What it's called	What it does
Base R	<code><-</code>	Assignment operator	Assigns a name to something in the R environment
Base R	<code>c()</code>	Concatenate	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.
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)</code>	Number of Rows	Gets the number of rows in an object <code>x</code> . Similarly, <code>ncol(x)</code> gets the number of columns in an object <code>x</code> .
Base R	<code>length(x)</code>	Number of Rows	Returns how long the vector <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 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.
Base R	<code>head()</code>	Head	Returns the top 6 rows of an object in the environment by default. You can specify how many rows you want by including the <code>n =</code> argument.
Base R	<code>tail()</code>	Tail	Returns the bottom 6 rows of an object in the environment by default. You can specify how many rows you want by including the <code>n =</code> argument.
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>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>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 file 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>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> for a set of logical vectors
Base R	<code>t()</code>	Transpose	Returns the transpose of a matrix or data frame

dplyr

Read the `dplyr` package documentation [here](#).

A vignette on the usage of the `dplyr` package can be found [here](#).

Library/Package	Piece of code	What it's called	What it does
<code>dplyr</code>	<code>%>%</code>	Pipe operator	Funnels a data.frame through tidyverse operations
<code>dplyr</code>	<code>filter()</code>	Filter	Returns a subset of rows matching the conditions of the specified logical argument
<code>dplyr</code>	<code>arrange()</code>	Arrange	Reorders rows in ascending order. <code>arrange(desc())</code> would reorder rows in descending order.
<code>dplyr</code>	<code>select()</code>	Select	Selects columns that match the specified argument
<code>dplyr</code>	<code>mutate()</code>	Mutate	Adds a new column that is a function of existing columns
<code>dplyr</code>	<code>summarise()</code>	Summarise	Summarises 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.
<code>dplyr</code>	<code>group_by()</code>	Group By	Groups data into rows that contain the same specified value(s)
<code>dplyr</code>	<code>inner_join()</code>	Inner Join	Joins data from two data frames, retaining only the rows that are in both datasets.

Docker

Read documentation on the Docker container [here](#). *Note: This section is not relevant for our virtual training workshop participants.*

Piece of code	What it's called	What it does
<code>docker pull</code>	Docker Pull	Pulls an image from a docker container
<code>docker run</code>	Docker Run	Runs processes in a docker container

ggplot2

Read the `ggplot2` package documentation [here](#).

A vignette on the usage of the `ggplot2` package can be found [here](#).

Library/Package	Piece of code	What it's called	What it does
<code>ggplot2</code>	<code>ggplot()</code>	GG Plot	Begins a plot that is finished by adding layers.
<code>ggplot2</code>	<code>xlab()</code> : <code>ylab()</code>	X Axis Labels; Y Axis Labels	Modifies the labels on the x axis and on the y axis, respectively
<code>ggplot2</code>	<code>geom_boxplot()</code>	Boxplot	Creates a boxplot when combined with ggplot()
<code>ggplot2</code>	<code>geom_density()</code>	Density Plot	Creates a smoothed plot when combined with ggplot() based on the computed density estimate
<code>ggplot2</code>	<code>theme_classic()</code>	Classic Theme	Displays <code>ggplot</code> without gridlines
<code>ggplot2</code>	<code>facet_wrap()</code>	Facet Wrap	Plots individual graphs using specified variables to subset the data
<code>ggplot2</code>	<code>ggtitle()</code>	GG Title	Sets the title of a <code>ggplot</code>
<code>ggplot2</code>	<code>ggsave()</code>	GG Save	Saves the last plot in working directory
<code>ggplot2</code>	<code>last_plot()</code>	Last plot	Returns the last plot produced

readr, tibble and tidyr

Read the `readr` package documentation [here](#) and the package vignette [here](#).

Read the `tibble` package documentation [here](#) and the package vignette [here](#). Read the `tidyr` package documentation [here](#) and the package vignette [here](#).

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. i.e. <code>readcsv()</code> , <code>readrds()</code> , etc.
tibble	<code>column_to_rownames()</code>	Column to Rownames	Transforms an existing column 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)
tidyr	<code>pivot_longer()</code>	Pivot Longer	Lengthens a data frame by increasing the number of rows and decreasing the number of columns