

Machine Learning Cheatsheet

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

Each table represents a different library/tool and its 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.

AnnotationDbi

Read the `AnnotationDbi` package vignette [here](#).

Library/Package	Piece of Code	What it's called	What it does
<code>AnnotationDbi</code>	<code>keytypes(.)</code>	Keytypes	Returns a character vector of column names/keytypes (eg., type of gene identifiers) available in an <code>AnnotationDbi</code> package.
<code>AnnotationDbi</code>	<code>mapIDs(.)</code>	Mapped IDs	Extracts the mapped ids for a set of keys (e.g., gene identifiers) of a specific keytype

Base R

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

Library/Package	Piece of Code	What it's called	What it does
Base R	<code>lapply()</code>	Apply over a list	Applies a function to specified elements of a list, and returns a list with these results
Base R	<code>duplicated()</code>	Duplicated	Returns a logical vector, where <code>TRUE</code> represents elements of the object that are duplicates
Base R	<code>any()</code>	Any	Checks to see if at least one of the elements are <code>TRUE</code> when given a logical vector
Base R	<code>is.na()</code>	Is it "Not Available"	Returns a logical vector where <code>TRUE</code> represents an instance of a missing value
Base R	<code>cor()</code>	Correlation	Computes correlation between columns using a specified correlation method, and returns a correlation matrix
Base R	<code>table()</code>	Create Table	Creates a contingency table of counts for each combination of factor levels
Base R	<code>cbind()</code>	Cbind	Combines vectors, matrices, or data.frames by columns
Base R	<code>paste0()</code>	Concatenate Strings	Joins together strings with no separator after they have been converted to character vectors
Base R	<code>pairwise.wilcox.test()</code>	Pairwise Wilcoxon Rank Sum Tests	Calculates the pairwise comparisons between group levels

PLIER

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

A `PLIER` package vignette can be found [here](#) and can also serve as documentation for the commands in the table below.

Library/Package	Piece of Code	What it's called	What it does
<code>PLIER</code>	<code>combinePaths()</code>	Combine Pathways	Combines the pathway data obtained from <code>PLIER</code> and returns the result as a matrix
<code>PLIER</code>	<code>commonRows()</code>	Common Rows	Determines the rows (genes) that are common to the specified data matrices and returns them as a character vector
<code>PLIER</code>	<code>rowNorm()</code>	Row Normalize	Normalizes each row (gene) by z-scoring the expression values
<code>PLIER</code>	<code>num.pc()</code>	Number of Principal Components	Returns the number of significant principal components
<code>PLIER</code>	<code>PLIER()</code>	Main PLIER Function	Main function of the Pathway-Level Information Extractor.
<code>PLIER</code>	<code>plotU()</code>	Plot U Matrix	Plots the U matrix obtained from the <code>PLIER</code> function results, allowing insight into the pathways or cell types captured by the latent variables

ComplexHeatmap

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

Library/Package	Piece of Code	What it's called	What it does
<code>ComplexHeatmap</code>	<code>Heatmap()</code>	Complex Heatmap	Constructs a heatmap whose graphics and features can be defined
<code>ComplexHeatmap</code>	<code>HeatmapAnnotation()</code>	Heatmap Annotation Constructor	Creates an annotation object to be used in conjunction with a Heatmap

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>geom_jitter()</code>	Jittered Points	Adds a small amount of random variation at each point's location on a plot
<code>ggplot2</code>	<code>labs()</code>	Labels	Sets the axis, legend, and plot labels if specified
<code>ggplot2</code>	<code>theme()</code>	Theme	Sets the specified non-data elements of a plot (i.e. plot title, legend spacing, text size, etc.)

`data.table` , `reshape2` , `RColorBrewer` , `ConsensusClusterPlus` ,
`ggsignif`

Documentation for each of these packages can be accessed by clicking the package name in the table below.

Library/Package	Piece of Code	What it's called	What it does
data.table	fread()	F read	Reads in data faster than base R
reshape2	melt()	Melt	Converts an object into a data.frame in 'long' format
RColorBrewer	brewer.pal()	Brewer Pal	Makes specified color palettes from ColorBrewer available in the R environment
ConsensusClusterPlus	ConsensusClusterPlus()	Consensus Clustering	Finds the consensus across multiple runs of the clustering algorithm
ggsignif	geom_signif()	Create Significance Layer	Adds significance information to the plot. It can be used to run statistical tests and display the significance information from those tests. We use it differently, in a way that gives us more control, in the notebook.