Cheatography

R Basic Cheat Sheet

by Dipakk (Dipakk) via cheatography.com/67199/cs/16837/

Util Functions

getwd()	gets the working directory
setwd("C:/file/path")	sets the working directory
data = read.csv(file.choose())	opens file explorer to get data
ls()	lists the variables
str(var)	structure of the variable
rm(var)	removes the variable
help.start()	opens help
install.packages("package name")	installs the package
library("package name")	makes the contents available to use
detach("package name")	detaches the package

Strings

toString(x)	helper function to produce a single character string
toupper()/tolower()	converts text to upper/-lower case
substring(chr,n,n)	retrieve or replaces the substring of the charachter
paste (…, sep = " ", collapse = NULL)	Concatenate vectors after converting to character

Arrays and Matrix

1D = array(1:24)	1 dimensional array
2D = array(1:24, dim = c(6,4))	2 dimensional array

Arrays and Matrix (cont)

3D = array(1:24, dim = c(4,3,2))	3 dimensional array
mat = matrix(1:12, nrow=4, ncol=3)	matrix
cbind(mat1,mat2)	column bind
rbind(mat1,mat2)	row bind

Vector

num = c(1,2,3,4,5,6)	numeric vector
chr = c("aaa","bbb")	character vector
log = c(TRUE,TRUE,FALSE)	logical vector
mean(vec)	mean
sd(vec)	standard deviation
var(vec)	variance
range(vec)	range
which.min(vec)/which.max(vec)	position of the min/max value
rep(1:5,times=3)	replicate elements of vector

DataFrame

df = data.frame(subjectID=1:5,gender=c("M","F","M","M","F"),score=c(8,3,6,5,5))	dataframe
view(df)	opens editor
head(df)/tail(df)	displays top/bottom n rows
summary(df)	returns descriptive statistics of data

Descriptive Statistics

rowMeans(data[])	row mean
rowSums(data[])	row sum
colMeans(data[])	column mean
colSums(data[])	column sum

Loops

for (variable in sequence){ Do something }	for loop
while (condition){ Do something }	while loop
if (condition){ Do something } else { Do something different }	ifelse statement

Hypothesis

t.test(data)	1 sample t test
t.test(data1,data2)	2 sample t test
t.test(pre,post,paired=TRUE)	paired sample t test
wilcox.test(data)	Wilcox test
cor.test(data1,data2)	correlation test
chisq.test(data)	Chi square test
shapiro.test(data)	Shapiro test
aov()	ANOVA

Visualization

qplot(data, line=TRUE,...)	produces quantile-quantile plot
ggplot(data = NULL, mapping = aes(), ...)	initializes a ggplot object
geom_bar()	bar graph
coord_flip()	flip x and y coordinates
facet_grid()	lay out panels in a grid
geom_density	density plot
geom_hist	histogram
geom_point	scatter plots

Probability	
rbinom(n, size, prob)	Binomial distribution
rpois(n,size)	Poisson distribution
runif(n, min = 0, max = 1)	Uniform distribution
rnorm(n,mean,sd)	Normal distribution
rexp(n)	Exponential distribution

Statistics	
summary(lm(y ~ x1 + x2 + x3, data=mydata))	multiple regression
summary(glm(y ~ x1 + x2 + x3, family="", data=mydata))	classification
cluster = kmeans(data)	kmeans cluster analysis

By **Dipakk** (Dipakk)
cheatography.com/dipakk/

Published 29th August, 2018.
Last updated 29th August, 2018.