Cheatography

Introduction to Package R_Sheet Cheat Sheet

by Noella via cheatography.com/67160/cs/16856/

Util Functions

getwd()	gets working dir
setwd("C:/file/path")	set working dir
help.start()	open help
install.packages("package")	install package
library("package")	make content available
detach("package")	detach package
x=read.csv(file.choose())	import data
ls()	list the variables
str(var)	structure of variable
rm(var)	remove variable

Arrays and Matrix

1D = array(1:24)	1-D array
2D=array(1:24,dim=c(6,4))	2-D array
3D=array(1:24,dim=c(4,3,2))	3-D array
matrix(1:12,nrow=4,ncol=3)	matrix
rbind/cbind(mat1,mat2)	row/col bind
t(mat)	transpose

Descriptive Statistics

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

Graphical Plots

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

Graphical Plots (cont)

facet_grid()	lay out panels in a grid
geom_density/hist/point	density/histogram/scatter plot

Strings

toString(x)	produces a single character string
toupper()/tolower()	converts to upper/lower case
substring(chr,n,n)	retrieves/replaces the substring
paste (…, sep= " ", collapse=NULL)	Convert to character + Concatenate

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

Probability Distributions

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

Data Frames

df = data.frame(subjectID=1:5,gender=c("M","F","M","M","F"),score=c(8,3,6,5,5))	Created data frames in R

Data Frames (cont)

fw = read.csv(file.choose())	Importing data by choosing a file
grass = read.csv('C:/Users/Downloads/grass.csv')	Importing data by specifying paths
view(df)	opens editor
rbind(a_data_frame, another_data_frame)	Bind rows/ columns of frames
merge(frame1, frame2, by = "x")	Merge 2 data frames
summary(df)	returns descriptive statistics of data

Loops

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

Hypothesis testing

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

By **Noella**
cheatography.com/noella/

Published 30th August, 2018.
Last updated 30th August, 2018.