

Operators

=	Assigns a value to an object
<-	
>	greater than
<	less than
>=	greater than or equal to
<=	less than or equal to
==	exactly equal to
!=	not equal to
!x	not x
x y	x OR y
x & y	x AND y
%>%	Sends something (e.g., dataframe or function output) to a tidyverse function. i.e., it is an elegant way to nest tidyverse functions.

Example:

```
library(tidyverse)
msleep %>%
  filter (couse rvation == 'domes tic ated') %>%
  summar ise(m = mean(b rainwt, na.rm = TRUE),
            s = sd(bra inwt, na.rm = TRUE),
            med = median (br ainwt, na.rm = TRUE)
  )
```

Note that `summar ise()` is nested within `filter()`, which is using info from the `msleep` dataframe.

C

By **Jeffrey M. Pisklak**
(non_human_entity)
cheatography.com/non-human-entity/

Published 28th December, 2023.
Last updated 28th December, 2023.
Page 1 of 6.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish
Yours!
<https://apollopad.com>

Basic R Functions

Access a function's help file

```
?[function name]
```

Load a csv file

```
read.csv( " fi l e n a m e . c s v ", header = TRUE )
```

Install a library

```
install.packages( " l i b r a r y n a m e " )
```

Load an installed library

```
library( l i b r a r y n a m e )
```

Resize images in Google Collab

```
options( re p r . p l o t . w i d t h = x , re p r . p l o t . h e i g h t = y )
```

Return the amount of values in x

```
length(x)
```

Return the number of rows in a dataframe

```
nrow(d a t a f r a m e )
```

Return the absolute value(s) in x

```
abs(x)
```

Return the sum of all the values in x

```
sum(x)
```

Return the square-root of the value(s) in x

```
sqrt(x)
```

Return the mean of the values in x with optional arguments for trimming and removing NAs

```
mean(x, tr = 0, na.rm = FALSE)
```

Return the median of the values in x with optional arguments removing NAs

```
median(x, na.rm = FALSE)
```

Return the sample standard deviation of values in x with optional argument for removing NAs

```
sd(x, na.rm = FALSE)
```



By **Jeffrey M. Pisklak**
(non_human_entity)
cheatography.com/non-human-entity/

Published 28th December, 2023.
Last updated 28th December, 2023.
Page 2 of 6.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish
Yours!
<https://apollopad.com>

Basic R Functions (cont)

Return the sample variance of values in x with optional argument for removing NAs

```
var(x, na.rm = FALSE)
```

Return the quartiles for x with optional argument for removing NAs

```
quantile(x, na.rm = FALSE)
```

Sort the values of x into ascending order

```
sort(x)
```

Compute the median absolute deviation of x with optional argument to remove NAs

```
mad(x, na.rm = FALSE)
```

Find NA values in x (returns TRUE/FALSE)

```
is.na(x)
```

Paste things together into a single string

```
paste(x, y, z, sep = " ")
```

Create a table of counts

Examples:

```
table(x)
```

```
table(x, y)
```

Data Frames

Create a new data frame

```
Column_1 <- c("A ", " B", " C")
Column_2 <- c(21, 22, NA)
new_df <- data.frame(Column_1, Column_2)
```

Add a column

```
new_df$Column_3 <- c(51, 52, 53)
```

Select a specific value (e.g., 52 = row 2, column 3)

```
new_df[2, 3]
```

Select a series of values (e.g., all of row 2)

```
new_df[2, c(1,2,3)]
```

or

```
new_df[2, ]
```

Select an entire column (e.g., column 2)

```
new_df$Column_2
```

or

```
new_df[, 2]
```

Isolate values that are not NAs

```
new_df$Column_2[!is.na(new_df$Column_2)]
```



Filter Function

Used to select specific observations from a dataframe according to a rule you specify.

```
filter (dataframe, subset rule)
```

Example 1:

```
filter (heightData, Father < 60.1 | Father > 75.3)
```

Example 2:

```
heightData %>% filter (Father < 60.1 | Father > 75.3)
```

Subset Function

Used to select specific observations from a dataframe according to a rule you specify.

```
subset (dataframe, subset rule, select = ("columns to keep"))
```

Example:

```
outliers <- subset (heightData, Father < 60.1 | Father > 75.3, select = c("Father"))
```

Library Functions

library(tidyverse) or library(dplyr)

Aggregate data sets into a new dataframe.

For example ...

```
msleep %>%
  group_by(vore, conservation) %>%
  summarise(m = mean(brainwt, na.rm = TRUE),
            s = sd(brainwt, na.rm = TRUE)
  )
```

library(rcompanion)

Calculates lambda for Tukey's ladder of powers

```
transformTukey(x, plotit = FALSE, returnLambda = TRUE)
```

library(WRS2)

Winsorized variance of x

```
winvar(x, tr = .2)
```



By Jeffrey M. Pisklak
(non_human_entity)
cheatography.com/non-human-entity/

Published 28th December, 2023.
Last updated 28th December, 2023.
Page 4 of 6.

Sponsored by [ApolloPad.com](https://apollopad.com)
Everyone has a novel in them. Finish
Yours!
<https://apollopad.com>

Distribution Functions

Return the the corresponding *quantile* for a given probability

Normal Distribution

```
qnorm (probability, mean, sd)
```

T Distribution

```
qt(probability, df, lower.tail)
```

F Distribution

```
qf(probability, df1, df2, lower.tail)
```

Chi-Square Distribution

```
qchisq (probability, df, lower.tail)
```

Return the the corresponding *probability* for a given quantile.

Normal Distribution

```
pnorm (quantile, mean, sd)
```

T Distribution

```
pt(quantile, df, lower.tail)
```

F Distribution

```
pf(quantile, df1, df2, lower.tail)
```

Chi-Square Distribution

```
pchisq (quantile, df, lower.tail)
```

Note:

- z-scores and t-scores (e.g. critical T and test statistics) are types of quantiles.

- The calculations are all performed from left to right by default unless you specify `lower.tail = FALSE`.

Plotting: library(ggplot2)

Histogram

```
ggplot (dataFrame, aes(x = Dep_Var)) +  
  geom_histogram (colour = "black",  
                  fill = "white")
```

Density Plot

```
ggplot (dataFrame, aes(x = Dep_Var)) +  
  geom_density (colour = "black",  
                fill = "pink",  
                adjust = 1)
```



Plotting: library(ggplot2) (cont)

Boxplot - for one sample

```
ggplot (dataFrame, aes(y = Dep_Var,)) +
  geom_boxplot()
```

Boxplot - for two or more samples

```
ggplot (dataFrame, aes(x = Indep_Var, y = Dep_Var)) +
  geom_boxplot()
```

Barplot with errorbars

```
ggplot (plotData, aes(x = Indep_Var, y = Dep_Var,
                      fill = Indep_Var)) +
  geom_bar(stat = "identity", colour = "black") +
  geom_errorbar(aes(ymin = bottom_values,
                   ymax = top_values),
               width = .25)
```

Q-Q Plot For two independent samples

Remove + facet_wrap() for a single sample

```
ggplot (dataFrame, aes(sample = Dep_Var)) +
  stat_qq() +
  stat_qq_line() +
  facet_wrap(~ Indep_Var)
```

Line Plot of Means with Two Predictors

```
ggplot (plotData, aes(x = PredictorA, y = Means, group = PredictorB, colour = PredictorB)) +
  geom_line(position = position_dodge(width = 0.4)) +
  geom_point(position = position_dodge(width = 0.4))
```

Scatterplot with Regression Line

```
ggplot (dataFrame, aes(x = predictor, y = response)) +
  geom_point() +
  geom_abline(intercept = b0, slope = b1)
```

Note:

Indep_Var = Independent Variable

Dep_Var = Dependent Variable

plotData = Dataframe of aggregated values

R Style Guide (from the Tidyverse)

<https://style.tidyverse.org/>

C

By **Jeffrey M. Pisklak**
(non_human_entity)
cheatography.com/non-human-entity/

Published 28th December, 2023.
Last updated 28th December, 2023.
Page 6 of 6.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
<https://apollopad.com>