PSYCH 213 R Cheat Sheet

1   Operators

```
x <-  5       assign 5 to x
x  = 5        also assigns, but <- is preferred
x  >  y       greater than
x  <  y       less than
x  >= y       greater or equal
x  <= y       less or equal
x  == y       equal
x  != y       not equal
!x            NOT x
x  |  y       x OR y
x  &  y       x AND y
x  |> f()     pipe: send x into a function f()
```

2   Vectors

Create a vector

```r
nums <- c(10, 20, 30, 40)
words <- c("a", "b", "c")
logical <- c(TRUE, FALSE, T, F)
```

Indexing a vector

```r
nums[1]      # first value (10)
nums[2:3]    # values 2 through 3 (20, 30)
nums[-2]     # all but the 2nd value
nums[c(1,4)] # first and fourth values (10, 40)
```

3   Useful Functions

```
length(x)            number of values in x
nrow(df)             number of rows in a data frame
sum(x)               sum of numbers
mean(x)              average
median(x)            middle value
sd(x)                standard deviation
sort(x)              sort smallest to largest
is.na(x)             check for missing values
paste0("a","b")      combine text
table(x)             count values
rep(x, times = n)    repeat values in a vector
seq(from, to, by)    generate a sequence of numbers
```

4 Working With Data

Read data

Base R:

```r
data <- read.csv("mydata.csv")
```

Tidyverse:

```r
library(tidyverse)
data <- read_csv("mydata.csv")
```

4.1 Make a data frame

```r
df <- data.frame(
    name=c("A","B","C"),
    score=c(10, 20, 30)
    )
```

Subset data frames with indexing

```r
df$score     # full column
df[2, ]      # row 2
df[ ,2]      # column 2
df[2, 2]     # row 2, column 2
```

Subset data frames with functions

Base R:

```r
subset(df, score > 15 & name == "B")
```

Tidyverse:

```r
filter(df, score > 15 & name == "B")
```

5 Plots (with ggplot2)

```r
library(tidyverse)

# Histogram
ggplot(df, aes(x = score)) +
  geom_histogram()

# Barplot
ggplot(df, aes(x = name, y = score)) +
  geom_bar(stat = "identity")

# Boxplot
ggplot(df, aes(x = name, y = score)) +
  geom_boxplot()
```

```
# Scatterplot
ggplot(df, aes(x = name, y = score)) +
  geom_point()
```

6   Aggregating with Pipes and the Tidyverse

```
library(tidyverse)

msleep |>
  group_by(vore) |>
  summarise(
    n = length(brainwt),
    m = mean(brainwt, na.rm = TRUE),
    s = sd(brainwt, na.rm = TRUE)
  )
```

7   Probability and Quantile Functions

```
pnorm(q, mean = 0, sd = 1)    probability below z-score (area to left)
qnorm(p, mean = 0, sd = 1)    z-score (quantile) for given probability
pt(q, df)                     probability below t-score
qt(p, df)                     t-score value for given probability
pchisq(q, df)                 probability below chi-square value
qchisq(p, df)                 chi-square value for given probability
```

The argument `lower.tail = FALSE` gives upper-tail probabilities or values.

8   T-testing

One-sample

```
# Compares single mean to null hypothesized mean
t.test(x,
    alternative = c("two.sided", "less", "greater"), # pick only one
    mu = 0
)
```

Paired Two-Sample

```
# Compares the difference between two means for paired/dependent conditions.
t.test(x, y,
    alternative = c("two.sided", "less", "greater"), # pick only one
    mu = 0,
    paired = TRUE
)
```

Independent Two-Sample

```r
# Compares the difference between two means for independent conditions.
t.test(x, y,
  alternative = c("two.sided", "less", "greater"), # pick only one
  mu = 0,
  var.equal = TRUE
)
```

9   Linear Regression

Simple Linear Regression

```r
# y = b0 + b1*x
model <- lm(y ~ x, data = df)
summary(model)
confint(model)
```

10   Goodness-of-Fit and Chi-Square Testing

Goodness-of-Fit Test

```r
# Tests whether observed frequencies match expected proportions
observed <- c(20, 50, 30)
expected <- c(0.25, 0.5, 0.25)   # proportions sum to 1
chisq.test(x = observed, p = expected)
```

One-Way (Pearson) Chi-Square Test of Independence

```r
# Tests whether two categorical variables are independent
tbl <- table(df$var_A, df$var_B)
chisq.test(tbl)
```

Remember: R is case-sensitive!
Also, if you spot any typos on this sheet, please let the instructor know.