

Base R — Midterm 1 Cheat Sheet

Data Frames & Variables

A data frame called `df` has columns (variables) and rows (observations). Name of the dataframe could be different.

```
# Access columns
df$x          # variable "x" from dataframe called "df"
df[, 3]       # third column from dataframe called "df"

# Create/assign variables
df$new_var <- df$a + df$b
v <- df$x     # assign column into vector

# In the functions below, you can use v or df$x
# e.g. table(v) or table(df$x)
# e.g. mean(v) or mean(df$x)
# e.g. cor(v,w) or cor(df$x, df$y) if w=df$y
# e.g. plot(v,w) or plot(df$x, df$y)

# Subset rows
df_rows <- df[df$x > 0, ]      # keep rows where x > 0
df_rows <- df[df$x == "a", ]  # keep rows where x == "a"
```

Frequency Tables

```
# Frequency table
table(c)
table(df$c)

# Relative frequencies
prop.table(table(c))

# Two-way frequency table
table(g, c)

# Conditional frequency tables
prop.table(table(g, c), margin = 1) # row-wise P(c|g)
prop.table(table(g, c), margin = 2) # column-wise P(g|c)
```

Basic Descriptives (Numeric x)

```
mean(x, na.rm=TRUE)    # mean
median(x, na.rm=TRUE)  # median
var(x, na.rm=TRUE)     # variance
sd(x, na.rm=TRUE)      # standard deviation
max(x, na.rm=TRUE)     # maximum
min(x, na.rm=TRUE)     # minimum
quantile(x, 0.8, na.rm=TRUE) # 80th quantile
```

Covariance & Correlation

```
cov(x, y, use="complete.obs") # covariance
cor(x, y, use="complete.obs") # correlation
```

Distributions

`p` = CDF, `q` = quantile

```
# Normal (mean=mu, sd=s)
pnorm(q, mean=mu, sd=s)
qnorm(p, mean=mu, sd=s)

# Student t (df=m)
pt(tval, df=m)
qt(p, df=m)

# Chi-square (df=k)
pchisq(x2, df=k)
qchisq(p, df=k)

# F (df1=d1, df2=d2)
pf(fval, df1=d1, df2=d2)
qf(p, df1=d1, df2=d2)
```

Plots (Base R)

```
plot(x, y)          # scatterplot
hist(x, breaks=30)  # histogram
barplot(table(c))   # bar plot
boxplot(x)          # boxplot (single variable)
boxplot(x ~ g)      # grouped boxplot
```

Arithmetic & Rounding

```
a <- 10; b <- 3
a + b      # addition
a - b      # subtraction
a * b      # multiplication
a / b      # division
a ^ b      # power
round(pi, 2) # rounding
```