DATALAB CHEAT SHEET

Quick Reference for Data Visualization & Analysis

GETTING STARTED

```r
# Load the package
library(datalab)

# Load your data
data(heart_data)        # Load sample data
data <- heart_data      # Assign to 'data'

# Look at your data
head(data)              # See first few rows
```

VISUALIZATION FUNCTIONS

Line Chart

When: Show relationship between two numbers

```r
linechart(x_variable, y_variable)
linechart(age, cholesterol)
linechart(age, cholesterol, smooth = TRUE)  # Add trend line
```

Boxplot

When: Show distribution of numbers

```r
```

```
boxplot(variable)                # Single boxplot
boxplot(age)

boxplot(variable, group = category)  # Compare groups
boxplot(age, group = sex)
```

Pie Chart

When: Show proportions of categories

```
r

piechart(category)               # Single pie chart
piechart(chest_pain_type)

piechart(category, filter)       # Multiple pie charts
piechart(chest_pain_type, sex)   # One chart per sex
```

STATISTICS FUNCTIONS

Descriptive Statistics

When: Get numerical summaries

```
r

descriptives(variable)           # Stats for one variable
descriptives(age)

descriptives(var1, var2, var3)   # Multiple variables
descriptives(age, cholesterol, resting_bp)
```

Output: Mean, SD, Median, Min, Max, N

Frequency Table

When: Count categories

```
r
```

```
frequencies(category)              # Count each category
frequencies(chest_pain_type)

frequencies(category, sort = "desc") # Sort by count
```

QUICK TIPS

Using Variables

```
r

# Assign to short names for easy typing
x = nitrogen_content
y = crop_yield
linechart(x, y)

# Reuse same letters for different variables
a = chest_pain_type
b = sex
piechart(a, b)
```

Loading Your Own Data

```
r

# CSV files
data <- read.csv("myfile.csv")

# Excel files (requires readxl)
library(readxl)
data <- read_excel("myfile.xlsx")
```

Getting Help

```
r

?linechart    # Help for linechart function
?piechart     # Help for any function
```

CHOOSING THE RIGHT FUNCTION

Your Goal	Use This
Compare two numbers	`linechart(x, y)`
See distribution of one number	`boxplot(var)`
Compare number across groups	`boxplot(var, group = category)`
Show category proportions	`piechart(category)`
Compare categories across groups	`piechart(category, filter)`
Get exact statistics	`descriptives(var)`
Count categories	`frequencies(category)`

COMMON ERRORS & FIXES

Error Message	Problem	Fix
`Variable 'x' not found`	Typo or wrong variable	Check spelling
`No data frame found`	Data not loaded	Run `data <- your_data`
Plot looks strange	Too many categories	Try `frequencies()` first

INTERPRETATION GUIDE

Boxplot

- **Middle line** = median (middle value)
- **Box** = where middle 50% of data lives
- **Lines (whiskers)** = range of typical values
- **Dots** = unusual values (outliers)

Pie Chart

- **Bigger slice** = more frequent
- **Percentages** = proportion of total
- Compare slices within one pie, or same slices across pies

Line Chart

- **Line goes up** = positive relationship

- **Line goes down** = negative relationship
- **Line is flat** = no relationship
- **Smooth line** = overall trend

Descriptive Statistics

- **Mean** = average value
- **SD** = how spread out (bigger = more variation)
- **Median** = middle value
- **N** = sample size (how many data points)

EXAMPLE WORKFLOW

```r
# 1. Load and explore
library(datalab)
data(heart_data)
data <- heart_data
head(data)

# 2. Visualize
piechart(chest_pain_type, sex)

# 3. Get statistics
descriptives(age, cholesterol)

# 4. Dive deeper
boxplot(cholesterol, group = sex)

# 5. Interpret and share your findings!
```

Need More Help? Type `vignette("datalab-intro")` for detailed examples

Questions? Contact your workshop facilitator