

R Commands Quick Reference

Essential R commands for daily data science work

Table of contents

| | |
|---------------------------------------|----------|
| 1 Data Import/Export | 1 |
| 2 Data Manipulation (dplyr) | 2 |
| 3 Data Visualization (ggplot2) | 2 |
| 4 Statistical Functions | 2 |
| 5 String Operations | 2 |
| 6 Package Management | 3 |
| 7 Workspace Management | 3 |
| 8 Data Types & Structure | 3 |
| 9 Missing Values | 3 |
| 10 Quick Tips | 4 |

1 Data Import/Export

| Task | Command | Example |
|------------|-----------------------------------|--|
| Read CSV | <code>read.csv()</code> | <code>read.csv("data.csv")</code> |
| Read Excel | <code>readxl::read_excel()</code> | <code>read_excel("data.xlsx")</code> |
| Write CSV | <code>write.csv()</code> | <code>write.csv(df, "output.csv")</code> |
| Save RDS | <code>saveRDS()</code> | <code>saveRDS(data, "data.rds")</code> |
| Load RDS | <code>readRDS()</code> | <code>readRDS("data.rds")</code> |

2 Data Manipulation (dplyr)

| Task | Command | Example |
|----------------|--------------------------|--|
| Filter rows | <code>filter()</code> | <code>filter(df, age > 18)</code> |
| Select columns | <code>select()</code> | <code>select(df, name, age)</code> |
| Create columns | <code>mutate()</code> | <code>mutate(df, age_months = age * 12)</code> |
| Group data | <code>group_by()</code> | <code>group_by(df, category)</code> |
| Summarize | <code>summarise()</code> | <code>summarise(df, mean_age = mean(age))</code> |
| Sort | <code>arrange()</code> | <code>arrange(df, desc(age))</code> |

3 Data Visualization (ggplot2)

| Task | Command | Example |
|--------------|-------------------------------|---|
| Scatter plot | <code>geom_point()</code> | <code>ggplot(df, aes(x, y)) + geom_point()</code> |
| Line plot | <code>geom_line()</code> | <code>ggplot(df, aes(x, y)) + geom_line()</code> |
| Bar plot | <code>geom_bar()</code> | <code>ggplot(df, aes(x)) + geom_bar()</code> |
| Histogram | <code>geom_histogram()</code> | <code>ggplot(df, aes(x)) + geom_histogram()</code> |
| Box plot | <code>geom_boxplot()</code> | <code>ggplot(df, aes(x, y)) + geom_boxplot()</code> |

4 Statistical Functions

| Task | Command | Example |
|--------------------|-----------------------|--|
| Mean | <code>mean()</code> | <code>mean(x, na.rm = TRUE)</code> |
| Median | <code>median()</code> | <code>median(x, na.rm = TRUE)</code> |
| Standard deviation | <code>sd()</code> | <code>sd(x, na.rm = TRUE)</code> |
| Correlation | <code>cor()</code> | <code>cor(x, y, use = "complete.obs")</code> |
| Linear model | <code>lm()</code> | <code>lm(y ~ x, data = df)</code> |
| ANOVA | <code>aov()</code> | <code>aov(y ~ group, data = df)</code> |

5 String Operations

| Task | Command | Example |
|-------------|----------------------|--------------------------------------|
| Concatenate | <code>paste()</code> | <code>paste("Hello", "World")</code> |

| Task | Command | Example |
|-----------------|-------------------------|-------------------------------------|
| Split string | <code>strsplit()</code> | <code>strsplit("a,b,c", ",")</code> |
| Find pattern | <code>grep()</code> | <code>grep("pattern", x)</code> |
| Replace pattern | <code>gsub()</code> | <code>gsub("old", "new", x)</code> |
| String length | <code>nchar()</code> | <code>nchar("hello")</code> |

6 Package Management

| Task | Command | Example |
|-----------------|-----------------------------------|--|
| Install package | <code>install.packages()</code> | <code>install.packages("dplyr")</code> |
| Load package | <code>library()</code> | <code>library(dplyr)</code> |
| Update packages | <code>update.packages()</code> | <code>update.packages()</code> |
| List packages | <code>installed.packages()</code> | <code>installed.packages()</code> |

7 Workspace Management

| Task | Command | Example |
|-------------------|------------------------------|------------------------------------|
| List objects | <code>ls()</code> | <code>ls()</code> |
| Remove objects | <code>rm()</code> | <code>rm(x, y)</code> |
| Clear workspace | <code>rm(list = ls())</code> | <code>rm(list = ls())</code> |
| Working directory | <code>getwd()</code> | <code>getwd()</code> |
| Set directory | <code>setwd()</code> | <code>setwd("/path/to/dir")</code> |

8 Data Types & Structure

| Task | Command | Example |
|--------------|------------------------|---------------------------|
| Data type | <code>class()</code> | <code>class(x)</code> |
| Structure | <code>str()</code> | <code>str(df)</code> |
| Dimensions | <code>dim()</code> | <code>dim(df)</code> |
| Column names | <code>names()</code> | <code>names(df)</code> |
| Summary | <code>summary()</code> | <code>summary(df)</code> |
| First rows | <code>head()</code> | <code>head(df, 10)</code> |
| Last rows | <code>tail()</code> | <code>tail(df, 10)</code> |

9 Missing Values

| Task | Command | Example |
|----------------|-------------------------------|---------------------------------|
| Check for NA | <code>is.na()</code> | <code>is.na(x)</code> |
| Remove NA | <code>na.omit()</code> | <code>na.omit(df)</code> |
| Complete cases | <code>complete.cases()</code> | <code>complete.cases(df)</code> |

10 Quick Tips

- Use `?function_name` to get help
- Use `Tab` for auto-completion in RStudio
- Use `Ctrl+Shift+M` for pipe operator `%>%`
- Use `Ctrl+Shift+C` to comment/uncomment code