

50 Essential R Commands for Beginners

1. Getting Started

> `getwd()`

Returns the current working directory.

Example: `getwd()`

> `setwd("path")`

Sets the current working directory.

Example: `setwd("C:/BioProjects")`

> `install.packages("dplyr")`

Installs the 'dplyr' package.

Example: `install.packages("dplyr")`

> `library(dplyr)`

Loads the 'dplyr' package.

Example: `library(dplyr)`

> `?mean`

Opens the help page for the 'mean' function.

Example: `?mean`

Practice: Check your current working directory and set it to a folder named 'BioData'.

2. Data Structures and Types

> `c()`

Combines values into a vector.

Example: `genes <- c("BRCA1", "TP53", "EGFR")`

> `class(genes)`

Returns the class of an object.

Example: `class(genes)`

> `as.numeric()`

Converts values to numeric.

Example: `as.numeric(c("1", "2", "3"))`

> `length(genes)`

Returns number of elements in the vector.

Example: `length(genes)`

> `str(genes)`

Displays the structure of the object.

Example: `str(genes)`

Practice: Create a vector of five gene names and check its class and length.

50 Essential R Commands for Beginners

3. Biological Dataset Example

```
sample_id <- c("S1", "S2", "S3", "S4", "S5")
gene_expr <- c(12.5, 9.8, 15.2, 7.3, 11.0)
treatment <- c("Control", "Treated", "Control", "Treated", "Control")
data <- data.frame(sample_id, gene_expr, treatment)
```

Practice: Use this dataset to practice filtering samples where gene_expr > 10.

4. Working with Data Frames

> data.frame()

Creates a data frame from vectors.

Example: `df <- data.frame(sample_id, gene_expr, treatment)`

> nrow(df)

Returns number of rows.

Example: `nrow(df)`

> ncol(df)

Returns number of columns.

Example: `ncol(df)`

> df\$gene_expr

Accesses a specific column.

Example: `df$gene_expr`

> df[1,]

Accesses the first row.

Example: `df[1,]`

> df[,1]

Accesses the first column.

Example: `df[,1]`

Practice: Extract the third sample's treatment status and gene expression value.

5. Data Manipulation

> subset()

Returns rows that meet a condition.

Example: `subset(df, gene_expr > 10)`

> rbind()

Adds rows to a data frame.

Example: `df2 <- rbind(df, new_row)`

> cbind()

Adds columns to a data frame.

Example: `df3 <- cbind(df, new_col)`

50 Essential R Commands for Beginners

> **merge()**

Merges two data frames by common columns.

Example: `merge(df1, df2, by = 'sample_id')`

> **colnames(df)**

Returns or sets column names.

Example: `colnames(df) <- c('ID', 'Expression', 'Treatment')`

Practice: Add a new column 'age' to the dataset and assign random values.

6. Descriptive Statistics

> **mean()**

Calculates the mean of numeric values.

Example: `mean(df$gene_expr)`

> **median()**

Returns the median value.

Example: `median(df$gene_expr)`

> **sd()**

Returns the standard deviation.

Example: `sd(df$gene_expr)`

> **var()**

Returns the variance.

Example: `var(df$gene_expr)`

> **table()**

Creates a frequency table.

Example: `table(df$treatment)`

Practice: Calculate mean and standard deviation for gene expression in treated samples.

7. Visualization

> **plot()**

Creates a scatterplot or line plot.

Example: `plot(df$gene_expr ~ df$sample_id)`

> **hist()**

Creates a histogram.

Example: `hist(df$gene_expr)`

> **boxplot()**

Creates a boxplot.

Example: `boxplot(gene_expr ~ treatment, data = df)`

> **barplot()**

50 Essential R Commands for Beginners

Creates a bar chart from a table.

Example: `barplot(table(df$treatment))`

Practice: Create a boxplot comparing gene expression between Control and Treated samples.

8. Bonus Tips

> `write.csv()`

Writes data frame to CSV file.

Example: `write.csv(df, 'data.csv')`

> `read.csv()`

Reads data from CSV file.

Example: `read.csv('data.csv')`

> `any()`

Checks if any condition is TRUE.

Example: `any(df$gene_expr > 20)`

> `all()`

Checks if all conditions are TRUE.

Example: `all(df$gene_expr > 5)`

> `which()`

Returns index of TRUE conditions.

Example: `which(df$gene_expr > 10)`

Practice: Export your dataset and load it again in a new session.