

X. READING, WRITING DATA IN R AND WORKING WITH INBUILT DATA SETS

Exercise 1: Reading a CSV File

```
> # Step 1: Create a sample data frame for student scores
> student_scores <- data.frame(
+   Name = c("John", "Alice", "Bob", "Clara", "Eve"),
+   Subject = c("Math", "Science", "Math", "English", "Science"),
+   Score = c(85, 92, 78, 88, 90)
+ )
>
> # Step 2: Write the data to a CSV file
> write.csv(student_scores, "student_scores.csv", row.names = FALSE)
>
> # Step 3: Read the CSV file
> data <- read.csv("student_scores.csv")
>
> # Step 4: Display the first 5 rows
> head(data, 5)
```

	Name	Subject	Score
1	John	Math	85
2	Alice	Science	92
3	Bob	Math	78
4	Clara	English	88
5	Eve	Science	90

Exercise 2: Writing a Data Frame to a CSV File

```
> # Create a data frame
> employee_data <- data.frame(
+   ID = c(1, 2, 3, 4),
+   Name = c("John", "Alice", "Bob", "Clara"),
+   Salary = c(50000, 60000, 55000, 58000)
+ )
> # Write to CSV
> write.csv(employee_data, "employee_data.csv", row.names = FALSE)
> print("File written successfully") [1] "File written
successfully"
```

Exercise 3: Load an Inbuilt Dataset

```
> # Load the iris dataset
> data(iris)
> # Display its structure
> str(iris)

'data.frame': 150 obs. of      5 variables:
 $ Sepal.Length: num  5.1    4.9 4.7 4.6 5 5.4 4.6    5 4.4 4.9 ...
 $ Sepal.Width : num  3.5    3 3.2 3.1 3.6 3.9 3.4    3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1
1 1 1 1 ...
```

Exercise 4: Summary Statistics of an Inbuilt Dataset

```
> # Load the mtcars dataset
```

```
> data(mtcars)
```

```
> # Display summary statistics
```

```
> summary(mtcars)           mpg
cylMin.           :10.40
Min.              :4.000
1st Qu.:15.43      1st Qu.:4.000
Median :19.20      Median :6.000
Mean     :20.09     Mean     :6.188
3rd Qu.:22.80      3rd Qu.:8.000
Max.      :33.90    Max.      :8.000
disp              hp          Min.      :
71.1    Min.      : 52.0
1st Qu.:120.8      1st Qu.: 96.5
Median :196.3      Median :123.0
Mean     :230.7     Mean     :146.7
3rd Qu.:326.0      3rd Qu.:180.0
Max.      :472.0    Max.      :335.0
drat              wt          Min.
:2.760    Min.      :1.513
1st Qu.:3.080      1st Qu.:2.581
Median :3.695      Median :3.325
Mean     :3.597     Mean     :3.217
3rd Qu.:3.920      3rd Qu.:3.610
Max.      :4.930    Max.      :5.424
qsec              vs          Min.
:14.50    Min.      :0.0000
1st Qu.:16.89      1st Qu.:0.0000
Median :17.71      Median :0.0000
Mean     :17.85     Mean     :0.4375
```

3rd Qu.:18.90	3rd Qu.:1.0000
Max. :22.90	Max. :1.0000
am	gear Min.
:0.0000 Min. :3.000	
1st Qu.:0.0000	1st Qu.:3.000
Median :0.0000	Median :4.000
Mean :0.4062	Mean :3.688
3rd Qu.:1.0000	3rd Qu.:4.000
Max. :1.0000	Max. :5.000
carb Min. :1.000	
1st Qu.:2.000	
Median :2.000	
Mean :2.812	
3rd Qu.:4.000	
Max. :8.000	

Exercise 5: Writing a Data Frame to Excel

```
> if(!require("writexl")) install.packages("writexl") Loading required
package: writexl
> library(writexl)
> # Write data to Excel
> write_xlsx(employee_data,"employee_data.xlsx")
> print("Excel file written successfully") [1] "Excel file written
successfully"
```

Exercise 6: Filtering Data from a Dataset

```
> # Filter rows where mpg > 20
```

```
> filtered_data <- subset(mtcars, mpg > 20)
```

```
> print(filtered_data)                                mpg
```

cyl	disp	hp	drat	Mazda RX4	21.0	
6	160.0	110	3.90	Mazda RX4 Wag	21.0	
6	160.0	110	3.90			
Datsun 710		22.8	4	108.0	93 3.85	
Hornet 4 Drive	21.4		6	258.0	110 3.08	
Merc 240D		24.4	4	146.7	62 3.69	
Merc 230		22.8	4	140.8	95 3.92	
Fiat 128		32.4	4	78.7	66 4.08	
Honda Civic		30.4	4	75.7	52 4.93	
Toyota Corolla		33.9	4	71.1	65 4.22	
Toyota Corona		21.5	4	120.1	97 3.70	
Fiat X1-9		27.3	4	79.0	66 4.08	
Porsche 914-2		26.0	4	120.3	91 4.43	
Lotus Europa		30.4	4	95.1	113 3.77 Volvo	
142E		21.4	4	121.0	109 4.11	
wt	qsec	vs	am	gear	carb Mazda RX4	2.620
16.46	0	1	4	4		
Mazda RX4 Wag	2.875	17.02	0	1	4	4
Datsun 710	2.320	18.61	1	1	4	1
Hornet 4 Drive	3.215	19.44	1	0	3	1
Merc 240D	3.190	20.00	1	0	4	2
Merc 230	3.150	22.90	1	0	4	2
Fiat 128	2.200	19.47	1	1	4	1
Honda Civic	1.615	18.52	1	1	4	2
Toyota Corolla	1.835	19.90	1	1	4	1
Toyota Corona	2.465	20.01	1	0	3	1

Fiat X1-9	1.935	18.90	1	1	4	1
Porsche 914-2	2.140	16.70	0	1	5	2
Lotus Europa	1.513	16.90	1	1	5	2
Volvo 142E	2.780	18.60	1	1	4	2

Exercise 7: Importing Data from a URL

```
> # Read data from a URL1
```

Error! Bookmark not defined.

```
2
3
4
5
```

3
3
3
3

```
> url <- "https://people.sc.fsu.edu/~jburkardt/data/csv/hw_200.csv" > data <- read.csv(url)
```

Warning message:

In read.table(file = file, header = header, sep = sep, quote = quote, : incomplete final line found by
readTableHeader on 'https://people.sc.fsu
.edu/~jburkardt/data/csv/hw_200.csv'

```
> # Display first 5 rows
```

```
> head(data,5)
```

Index

```
Height.Inches...Weight.Pounds..1.65.78..112.99.2..71.52..136.49.3..69.4
0..153.03.4..68.22..142.34.5..67.79..144.30.6..68.70..123.30.7..69.80..141
.49.8..70.01..136.46.9..67.90..112.37.10..66.78..120.67.11..66.49..127.45.
12..67.62..114.14.13..68.30..1 ...
1
71.52
2
69.40
3
68.22
4
```

67.79

5

68.70

Weight.Pounds.

1	136.49
2	153.03
3	142.34
4	144.30
5	123.30

Exercise 8: Appending Rows to a Data Frame

```
> # Create new rows
> new_rows <- data.frame(
+   ID = c(5, 6),
+   Name = c("Eve", "Mark"),
+   Salary = c(61000, 53000)
+ )
> # Append rows
> updated_data <- rbind(employee_data, new_rows)
> print(updated_data)
```

	ID	Name	Salary
1	1	John	50000
2	2	Alice	60000
3	3	Bob	55000
4	4	Clara	58000
5	5	Eve	61000
6	6	Mark	53000

Exercise 9: Saving and Loading Data in RDS Format

```
> # Save data
> saveRDS(employee_data, "employee_data.rds")
> # Load data
> loaded_data <- readRDS("employee_data.rds")
> print(loaded_data)
```

	ID	Name	Salary
1	1	John	50000
2	2	Alice	60000
3	3	Bob	55000
4	4	Clara	58000

Exercise 10: Merge Two Data Frames

```
> # Define two data frames
> df1 <- data.frame(ID = c(1, 2, 3), Name = c("John", "Alice", "Bob"))
> df2 <- data.frame(ID = c(1, 2, 3), Department = c("HR", "IT", "Finance"))
)
```

```
> # Merge
> merged_data <- merge(df1, df2, by = "ID")
> print(merged_data)
```

	ID	Name	Department
1	1	John	HR
2	2	Alice	IT
3	3	Bob	Finance

Exercise 11: Exporting Data

```
> # Export iris dataset to CSV
> write.csv(iris, "iris_data.csv", row.names = FALSE)
> # Verify export by reading back the exported CSV file
> iris_exported <- read.csv("iris_data.csv")
> # Display structure and summary of exported data
> str(iris_exported)
```

'data.frame': 150 obs. of 5 variables:

```
$ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
$ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
$ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
$ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
$ Species      : chr  "setosa" "setosa" "setosa" "setosa" ...
```

```
> summary(iris_exported)
```

```

Sepal.Length      Sepal.Width Min.
      :4.300      Min.      :2.000

1st Qu.:5.100      1st Qu.:2.800

Median :5.800      Median :3.000

Mean   :5.843      Mean   :3.057 3rd
Qu.:6.400 3rd Qu.:3.300 Max.   :7.900

Max.   :4.400

Petal.Length      Petal.Width Min.
      :1.000      Min.      :0.100

1st Qu.:1.600      1st Qu.:0.300

Median :4.350      Median :1.300

Mean   :3.758      Mean   :1.199 3rd
Qu.:5.100 3rd Qu.:1.800 Max.   :6.900

Max.   :2.500

Species
Length:150
Class:character Mode
:character

```

```
> # Display first few rows of exported data
```

```
> head(iris_exported)
```

```

      Sepal.Length  Sepal.Width  Petal.Length
1             5.1           3.5           1.4
2             4.9           3.0           1.4
3             4.7           3.2           1.3
4             4.6           3.1           1.5
5             5.0           3.6           1.4
6             5.4           3.9           1.7

Petal.Width  Species

```

1	0.2	setosa
2	0.2	setosa
3	0.2	setosa
4	0.2	setosa
5	0.2	setosa
6	0.4	setosa

Exercise 12: Exploring Inbuilt Datasets

```
> # Load datasets package
> library(datasets)
> # Load mtcars dataset
> data(mtcars)
> # Display structure and summary of the dataset
> str(mtcars)
```

'data.frame': 32 obs. of 11 variables:

```
$ mpg: num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
$ cyl:  num   6 6   4 6   8 6 8 4   4 6 ...
$ disp: num  160 160 108 258 360 ...
$ hp :  num  110 110  93 110 175 105 245 62 95      123 ...
$ drat: num   3.9  3.9  3.85 3.08 3.15 2.76 3.21      3.69 3.92      3.92 ...
$ wt :  num   2.62 2.88 2.32 3.21 3.44 ...
$ qsec: num  16.5 17 18.6 19.4 17 ...
$ vs :  num   0 0   1 1   0 1 0 1   1 1 ...
$ am :  num   1 1   1 0   0 0 0 0   0 0 ...
$ gear: num   4 4   4 3   3 3 3 4   4 4 ...
$ carb: num   4 4   1 1   2 1 4 2   2 4 ...
```

```
> summary(mtcars)      mpg
```

```
cyl      Min.      :10.40
```

```
Min.      :4.000
```

```
1st Qu.:15.43      1st Qu.:4.000
```

Median :19.20	Median :6.000
Mean :20.09	Mean :6.188
3rd Qu.:22.80	3rd Qu.:8.000
Max. :33.90	Max. :8.000
disp	hp Min. :
71.1 Min. : 52.0	
1st Qu.:120.8	1st Qu.: 96.5
Median :196.3	Median :123.0
Mean :230.7	Mean :146.7
3rd Qu.:326.0	3rd Qu.:180.0
Max. :472.0	Max. :335.0
drat	wt Min.
:2.760 Min. :1.513	
1st Qu.:3.080	1st Qu.:2.581
Median :3.695	Median :3.325

```

      Mean      :3.597      Mean      :3.217
      3rd Qu.:3.920      3rd Qu.:3.610
      Max.      :4.930      Max.      :5.424
qsec          vs          Min.
:14.50      Min.      :0.0000
      1st Qu.:16.89      1st Qu.:0.0000
      Median :17.71      Median :0.0000
      Mean      :17.85      Mean      :0.4375
      3rd Qu.:18.90      3rd Qu.:1.0000
      Max.      :22.90      Max.      :1.0000
am          gear          Min.
:0.0000      Min.      :3.000
      1st Qu.:0.0000      1st Qu.:3.000
      Median :0.0000      Median :4.000
      Mean      :0.4062      Mean      :3.688
      3rd Qu.:1.0000      3rd Qu.:4.000
      Max.      :1.0000      Max.      :5.000
carb          Min.      :1.000
      1st Qu.:2.000
      Median :2.000
      Mean      :2.812 3rd
      Qu.:4.000      Max.
      :8.000

```

```
> # Display first few rows of the dataset
```

```
> head(mtcars)
```

```

              mpg cyl disp  hp drat Mazda
RX4          21.0   6  160 110 3.90
Mazda RX4 Wag 21.0   6  160 110 3.90
Datsun 710    22.8   4  108  93  3.85
Hornet 4 Drive 21.4   6  258 110 3.08
Hornet Sportabout 18.7   8  360 175 3.15

```

Valiant	18.1	6	225	105	2.76
wt qsec vs am gear					
Mazda RX4	2.620	16.46	0	1	4
Mazda RX4 Wag	2.875	17.02	0	1	4
Datsun 710	2.320	18.61	1	1	4
Hornet 4 Drive	3.215	19.44	1	0	3
Hornet Sportabout 3.440		17.02	0	0	3
Valiant	3.460	20.22	1	0	3
carb Mazda RX4		4			
Mazda RX4 Wag	4				
Datsun 710	1				
Hornet 4 Drive	1				
Hornet Sportabout	2				
Valiant	1				