

Week 1 Tasks - Data Science Programming

[Nguyen Dat Thanh]

2025-09-17

Week 1 Tasks

This R Markdown document contains solutions to all Week 1 tasks using the `student_scores.csv` dataset.

Task 1 – Hello World

Write a program that prints: “Welcome to Data Science Programming!”

```
# Task 1: Hello World  
# Write your code here
```

```
print(“Welcome to Data Science Programming!”) # Print greeting message
```

Task 2 – Simple Math

Calculate and print the result of: $-15 + 23 - 120 / 6$
 -5^3 (5 to the power of 3)

```
# Task 2: Simple Math  
# Write your code here
```

```
print(15 + 23) # Add 15 and 23 print(120 / 6) # Divide 120 by 6 print(5 ** 3) # 5 to the power of 3
```

Task 3 – Variables and Vectors

Create a vector of 5 numbers: `c(10, 20, 30, 40, 50)`.

Print the vector and the sum of its elements.

```
# Task 3: Variables and Vectors  
# Write your code here
```

```
vec <- c(10, 20, 30, 40, 50) # Create a numeric vector print(vec) # Print the vector print(sum(vec)) # Print  
the sum of elements
```

Task 4 – Plot Squares

Create a vector of numbers from 1 to 10.

Plot the numbers against their squares using the `plot()` function.

```
# Task 4: Plot Squares  
# Write your code here
```

```
nums <- 1:10 # Numbers 1 to 10
```

```
plot(nums, nums^2, # Plot x = nums, y = nums squared main = "Numbers vs. Their Squares", # Title  
xlab = "Numbers", # X-axis label ylab = "Squares", # Y-axis label type = "b", # "b" = both points and  
lines col = "blue", # Color of points/lines pch = 19) # Solid circle points
```

Task 5 – Load Dataset

Load `student_scores.csv` into your program.

Display the first 5 rows.

```
# Task 5: Load Dataset  
# Write your code here
```

```
df <- read.csv("/Users/datthanhnguyen/Documents/GitHub/Vinuni/DS Prog/Lab1/student_scores.csv")  
# Read CSV file head(df, 5) # Display first 5 rows
```

Task 6 – Summary Statistics

Print the average (mean) of the `Score` column.

Print the minimum and maximum scores.

```
# Task 6: Summary Statistics  
# Write your code here
```

```
mean_score <- mean(df$Score) # Calculate mean of Score column print(mean_score)
```

```
min_score <- min(df$Score) # Find minimum score print(min_score)
```

```
max_score <- max(df$Score) # Find maximum score print(max_score)
```

Task 7 – Filtering Data

Find and print the names of students who scored greater than 80.

```
# Task 7: Filtering Data  
# Write your code here
```

```
high_scorers <- dfName[df$Score > 80] # Select names where Score > 80 print(high_scorers)
```

Task 8 – Sorting Data

Sort the dataset by **Score** in descending order.

Display the top 3 students.

```
# Task 8: Sorting Data  
# Write your code here
```

```
sorted_scores <- df[order(-df$Score), ] # Sort dataset by Score (highest first) head(sorted_scores, 3) #  
Show top 3 students _____
```

Task 9 – Visualization

Create a histogram of scores.

Add appropriate title and axis labels.

```
# Task 9: Visualization  
# Write your code here
```

```
hist(df$Score, main = "Distribution of Student Scores", # Title xlab = "Scores", # X-axis label ylab =  
"Frequency", # Y-axis label col = "lightblue", # Fill color border = "black") # Border color _____
```

Submission Notes

- Ensure all code chunks run without errors
- Include appropriate comments in your code
- Make sure plots are properly displayed
- Knit this document to HTML/PDF before submission
- Save and submit this file as **Week1_Tasks.Rmd**

End of Week 1 Tasks