Problem Set 1

Introduction to R | University of Oxford Sociology

Casey Breen

Problem Set 1

This problem set contains exercises from Session 1 that were originally done in small-groups. To reinforce your understanding, please complete these exercises independently. Answer the following questions using `R` in a Quarto document.

Exercise 1: Assignment, Arithmetic, Logical Expressions

1.1 Assign `x` and `y` to take values 3 and 4.

1.2 Assign `z` as the product of `x` and `y`.

1.3 Calculate the square of 3 and assign it to `three_squared`.

1.4 Write a logical expression to check if `three_squared` is greater than 10.

1.5 Write a logical expression to test whether `three_squared` is *not* greater than 10. Use the negate (`!`) operator.

Exercise 2: Sequencing

2.1 Generate vectors containing the numbers 100, 101, 102, 103, 104, and 105 using 3 different methods (e.g., `c()`, `seq()`, `:`). In what scenarios might each method be most convenient?

2.2 Generate a sequence of all even numbers between 0 and 100. Use the `seq()` function.

2.3 Create a descending sequence from 100 to 1 and assign it to a variable. Use the `seq()` function.

Exercise 3: Data Generation and Basic Statistical Analysis

3.1 Generate a sample of 100 observations from a normal distribution with a mean of 10 and a standard deviation of 2. Use the `rnorm()` function.

3.2 What are the 1st, 10th, and 100th elements of this `vector`?

3.3 Calculate the mean of this `vector`. How does this *sample* mean relate to the theoretical *population* mean (hint: population mean = 10) of the distribution?

3.4 Calculate the difference between the *sample* mean and the *population* mean. Discuss the reason for the discrepancy.

3.5 Repeat steps 1 and 3 with a sample size of 10,000. Did the difference between the sample mean and the population mean decrease? Why?