DS636 Lab1

1. Install UsingR and ISwR packages to your own directory

2. Using R for Introductory Statistics Page 18: 1.1-1.12

Problems

1.1 Use R as you would a calculator to find numeric answers to the following:

1. $1+2(3+4)$
2. 4^3+3^{2+1}
3. $\sqrt{(4+3)(2+1)}$
4. $\left(\frac{1+2}{3+4}\right)^2$

1.2 Rewrite these R expressions as math expressions, using parentheses to show the order in which R performs the computations:

1. `2 + 3 - 4`
2. `2 + 3 * 4`
3. `2/3/4`
4. `2^3^4`

1.3 Use R to compute the following

$$\frac{1+2\cdot 3^4}{5/6-7}.$$

1.4 Use R to compute the following

$$\frac{0.25-0.2}{\sqrt{0.2\cdot(1-0.2)/100}}.$$

1.5 Assign the numbers 2 through 5 to different variables, then use the variables to multiply all the values.

1.6 The `rivers` data set is loaded when R is. View the data by typing its name and then the return key. What is the last value listed?

1.7 The `exec.pay` (`UsingR`) data set is available from the command line after loading the package `UsingR`. Load the package, and inspect the data set. Scan the values to find the largest one.

1.8 For the `exec.pay` (`UsingR`) data set, apply the functions `mean`, `min`, and `max`. What are the values found?

1.9 The basic `mean` function has an additional argument `trim`. When given, the specified proportion of the data is trimmed from the sorted data before the mean is taken. Compare the difference between `mean(exec.pay)` and `mean(exec.pay, trim=0.10)`.

1.10 The `Orange` data set is stored as a data frame with three variables. What are the three variables?

1.11 Compute the average age of the trees in the `Orange` data set using `mean`.

1.12 Compute the largest circumference of the trees in the `Orange` data set.

3. Using R for Introductory Statistics Page 45: 2.1-2.4

Problems

2.1 Enter the following data into a variable `p` using `c`

```
2 3 5 7 11 13 17 19
```

Use `length` to check its length.

2.2 Al recorded his car's mileage at gust last eight fill-ups:

```
65311 65624 65908 66219 66499 66821 67145 67447
```

Enter these numbers into the variable `gas`. Use the function `diff` on the data. What does it give? Interpret what both of these commands return: `mean(gas)` and `mean(diff(gas))`.

2.3 Let our small data set be

```
2 5 4 10 8
```

1. Enter this data into a data vector `x`.
2. Find the square of each number.
3. Subtract 6 from each number.
4. Subtract 9 from each number and then square the answers.

Use the vectorization of functions to do so.

2.4 Create the following sequences:

1. `"a", "a", "a", "a", "a"`
2. 1, 3, ..., 99 (the odd numbers in [1,100])
3. 1, 1, 1, 2, 2, 2, 3, 3, 3
4. 1, 1, 1, 2, 2, 3
5. 1, 2, 3, 4, 5, 4, 3, 2, 1

using `:`, `seq`, or `rep` as appropriate.