
Course: Basics of R programming language for statistical analysis

Instructor: Marina FERENT [marinaferent@gmail.com]

Chapter 2: CONTROL STRUCTURES AND FUNCTIONS | Statistical measures

Meeting 3: FOR LOOPS| Challenge: Mean values

Exercises

PRODUCE Tasks:

1.

a) Find an R built-in function that computes the variance. Compute variance for <<number of enrolled students>> in <<Campus crime.csv[i]>> dataset.

[Estimated time: 10 min]

b) Compute the variance by yourself through the usage of a for loop (or another loop of your choice).

[Structure suggestion – inside the for loop:

- construct a vector of squared differences (between each value and the mean value)
- sum up all the differences
- divide the sum to n or n-1 according to your formula choice (population or sample)].

Compare your result here with the result in point a). Is it the same? Check the description of the built-in function. Which variance formula have you used (population or sample)? Which variance formula is used by the built-in function you chose in point a) (population or sample)?

Consider the standard deviation and coefficient of variation. What other lines would you add to your code to compute them?

[Estimated time: 30 min]

c) Find an R built-in function that computes the variance for each column in a dataset. Compute variance for each variable in Campus crime.csv dataset. Reproduce this function through the usage of a for loop (or another loop of your choice). [Suggestion: add an extra for loop in your code in point b). Pay attention to indentation. How would your code look like from indentation perspective?]

[Estimated time: 15 min]

COMMENT Tasks:

1. Comment the <<yahooFinanceExtract.r>> code in rBasics_Meeting3_COMMENT > Exercise 2 folder. To run the code, use <<Yahoo Symbols.csv>>.

[Estimated time: 20 min]

Challenge 1_REPRODUCE: Change the yahooData column names to <<Name>> of stock instead of <<Symbol>>.

[Estimated time: 5 min]

Challenge 2_DEBUG: Add symbol << KTCG.VI>> -<< Kapsch TrafficCom AG>> to your <<YahooSymbols.csv>> and run the code. You receive a warning message. What does it mean? Why do you receive it?

[Estimated time: 10 min]

---

[i] The data set comprises part of *campus* data from *Wooldridge, Jeffrey M. (2013). Introductory econometrics: a modern approach. Mason, Ohio: South-Western Cengage Learning. Wooldridge Source: These data were collected by Daniel Martin, a former MSU undergraduate, for a final project. They come from the FBI Uniform Crime Reports and are for the year 1992.* The original data set is available for download at:
(1)https://www.cengage.com/cgi-wadsworth/course_products_wp.pl?fid=M20b&product_isbn_issn=9781111531041 Or
(2) https://cran.r-project.org/web/packages/wooldridge/wooldridge.pdf