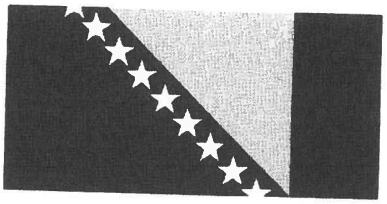
TWINNING CONTRACT

BA 17 IPA ST 01 20

Further Support to the Reform of Statistics System in Bosnia and Herzegovina



MISSION REPORT

Activity 1.4A: R software training - Basic training
Subcomponent 1.4 - R software training

Mission carried out by
Jolien Cremers, Statistics Denmark
Fenja Søndergaard Møller, Statistics Denmark

23 - 26 May 2023

Version: Final

Funded by the European Union

FURTHER SUPPORT TO THE REFORM OF STATISTICS SYSTEM IN BIH

Expert contact information

Jolien Cremers
Statistics Denmark
Copenhagen
joc@dst.dk

Fenja Søndergaard Møller
Statistics Denmark
Copenhagen
fsm@dst.dk

Table of contents

List of Abbreviations

BHAS	Agency for Statistics of Bosnia and Herzegovina
BiH	Bosnia and Herzegovina
CBBH	Central Bank of Bosnia and Herzegovina
EC	European Commission
EU	European Union
FBiH	Federation of Bosnia and Herzegovina
FIS	Institute for Statistics of Federation of Bosnia and Herzegovina
MS	EU Member State
RSIS	Institute for Statistics of Republika Srpska
RTA	Resident Twinning Adviser
ToR	Terms of Reference

1. General comments

This mission report was prepared within the EU Twinning Project "Further Support to the Reform of the Statistics System in Bosnia and Herzegovina". It was the first mission to be devoted to the Subcomponent 1.4, R software training, of the Project.

The purposes of the mission were:

- R computer software - basic training

The consultants would like to express their thanks to all officials and individuals from Bosnia and Herzegovina met for the kind support and valuable information, and which highly facilitated the work of the consultants.

The views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of EU, CBBH, Statistics Denmark and Statistics Finland.

2. Assessment and results

The mission was scheduled as a 3.5-day beginner course in statistical programming in R. The topics covered was as follows:

- R and RStudio
- Basic R
- Packages
- Reading and exporting data
- Data transformation using tidyverse
- Summary statistics
- Introduction to visualisation using ggplot2

The topics "Strings and dates" and "Making reports using markdown" were originally part of the course. Nonetheless, it was decided to skip these parts to adjust the pace of the course to the level of the course participants.

A significant part of the first day was spent on making sure that all participants had the correct software installed on their laptops. An improvement for next time is for all participants to make sure to have the (upfront announced) required software installed before the start of the course, as well as to make sure they have administrator rights on their laptops for installation of additional software packages. Some participants did not have a (functioning) laptop during the (first days of the) course. In a next mission, participants should make sure to take a functioning laptop with them.

The format of the coursed was changed from lectures and individual practicals to a format in which a short lecture was followed by an interactive session. In this session, the students and instructor together set up an R-project, read, transformed, summarised and visualised data, and the students expressed that this was a success. We recommend a next course also adopts this format (if it is a follow-up course, at least start with a recap day in this format).

The course covered a lot of contents and it will require continuous practice by the beneficiaries to get comfortable with programming in R and reap the benefit of using R instead of existing tools in the production of statistics.

The topics were introductory and could be expanded and elaborated on for future missions. Furthermore, future missions should focus on the direct relevance of using R in specific case studies for statistical production. In order to do so, dialogue with course participants is needed up front in terms of what programs they currently use and what requirements they have in terms of functionalities/desired output/input formats etc. During the current mission, we only discovered these things during the course itself. This was not optimal and made it difficult to tailor the course to specific requirements.

3. Conclusions and recommendations

Part of the participants in the course acquired the skills listed in the section above. It is recommended that time and resources will be allocated to practice and use R after the course. This will ensure more familiarity with the R environment.

Since the participants come from BHAS, FIS and RSIS, we would recommend to set up a channel of communication where they can exchange experiences and problems with using R in production, this could be a Microsoft Teams channel, a mailing list or a physical forum.

4. What to do before the next mission for the BC Counterpart

See the recommendations in sections 2 and 3

Annex 1. Terms of Reference for the current mission

Terms of Reference

EU Twinning Project BA 17 IPA ST 01 20

Subcomponent 1.4 - R software training

Tuesday 23th May 10 hrs - Friday 26th May 2023 12:00 hrs.

On-site meeting

Venue: Agency for Statistics of BiH, **Zelenih beretki 26, 1st Floor, Sarajevo.**

Activity 1.4A: R software training - Basic training

1. **Mandatory results**

Staff skills in the use of R software increased through training sessions

2. **Purpose of the activity**

R computer software - basic training

For specific topics expected to be covered, please see the link https://joliencremers.github.io/R-for-BHAS-2023 and the page 6.

3. **Expected output of the activity**

Basic R level knowledge gained to be utilised in the daily work

Topics expected to be covered on the training

Day 1

Morning:

- The role of R (the engine) and RStudio (the control center)
- R-projects, workspace and history
- Using R help function (including built-in help function, online help, and "cheat sheets")

Afternoon:

- Types of elements (numerical, character, logical, factor)
- Types of objects (vectors, matrices, data frames, and lists)
- Special values (NA, NaN, Inf)
- Logical operators

DAY 2

Morning:

- Packages (to extend functionality of base R)
- Reading from and writing to data in R-format (RData and RDS) and delimited files (csv and related); to a lesser extent we have covered working with databases and file formats native to other programs (Excel, SAS, SPSS, Stata etc.)

Afternoon + DAY 3

- *Functional flow and the pipe-operator*
- **Main focus:** Data wrangling using the tidyverse (dplyr, purrr and tidyr)
- Basic summary statistics

DAY 4

Morning:

- Working with the institution's own data
- Discussion on potential additional course

If time allows it: a brief demonstration of more advanced R features:

- R-reports (either compiled from plain R files or knitted from R Markdown)
- Visualisation using graphical functions from base R (plot) and introduction to the grammar of graphics with ggplot2
- Maps on a NUTS2-level.

Annex 2. Persons met

BHAS
Maja Hadzi-Stojanov
Emina Mehanovic
Anita Brkovic
Mirza Agic

RSIS
Andrea Erak Latinovic
Danica Babic
Jelena Komljenovic
Dragana Mandic
Vladimir Koprivica

FIS
Merima Beganovic
Ademira Veličanin
Nasiha Imsirovic
Damir Omanovic
Adis Mizdrak

RTA Team:
Niels Madsen
Larisa Muslimovic

Signatures

For the approval of the contents of the report, representatives from BHAS, RSIS and FIS as well as the MS experts and the RTA sign here:

Date: 26/6 - 2023

BHAS

RSIS

FIS

RTA

MS Expert

MS Expert