TWINNING CONTRACT

BA 15 IPA SR 01 17

Support to the reform of the statistics system in Bosnia and Herzegovina

MISSION REPORT

Activity 2.D.1:

Statistical programming in R – Basis training I

Component 2: Business Statistics

Mission carried out by
Peter Stoltze, Statistics Denmark
Anne Vinkel Hansen, Statistics Denmark

September 17-20, 2019

Version: Final

Expert contact information

Peter Stoltze
Statistics Denmark
Copenhagen, Denmark
Email: psl@dst.dk

Anne Vinkel Hansen
Statistics Denmark
Copenhagen, Denmark
Email: aih@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

Executive Summary

Only if core text of mission exceeds 4 pages – otherwise delete

Main conclusions and highlights from findings.

1. General comments

This mission report was prepared within the EU Twinning Project " Support to the reform of the statistics system in Bosnia and Herzegovina". It was the first mission to be devoted to training in the use of the R software within Component 2 of the project.

The purpose of the mission was:
- Training of non-IT staff to programming in R

The consultants would like to express their thanks to all officials and individuals met for the kind support and valuable information which they received during the stay in Bosnia-Herzegovina and which highly facilitated the work of the consultants.

This views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of EU, BHAS, FIS, RSIS, CBBH, Statistics Denmark, INSEE, Statistics Finland and Croatian Bureau of Statistics.

2. Assessment and results

The Terms of Reference mentions the following expected output of the activity:
- Introduction to the statistical programme R has been made.
- R and the Rstudio interface have been demonstrated, introductions have been made to data structures, data management, programming, data exploration and basic statistics plus visualization.
- Input provided to the ToR of next activity.

The consultants have prepared a set of lectures (including both slides and exercises) which were presented to the participants. At least for the duration of this project the material will be available online.[1] The content of the lectures and the exercises have given the participants an introduction to programming in R and how to use the extended set of possibilities provided by RStudio.

The topics include (but are not restricted to):
- The role of R (the engine) and RStudio (the control center)
- Using R help function (including built-in help function, online help, and "cheat sheets")
- R-projects
- R-reports (either compiled from plain .R files or knitted from R Markdown)
- Types of elements (numerical, character, logical, factor)
- Types of objects (vectors, matrices, data frames, and lists)
- Special values (NA, NaN, Inf)
- Logical operators
- Packages (to extent functionality of base R)
- Work spaces, history
- 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.)
- Control structures like loops (for) and conditionals (if-then)

---

[1] http://peterstoltze.dk/RfOS/

- The apply function and variations (sapply and lapply)
- Writing functions including parameters and returning values
- Data wrangling using the dplyr and tidyr packages
- Visualisation using graphical functions from base R (plot) and introduction to the grammer of graphics with ggplot2 (including maps)
- Basic summary statistics

A course with 26 participants and only two instructors is a challenge, but the consultants were duly impressed by the enthusiasm and eagerness to learn a displayed by the participants. The IT environment was working to a satisfactory level, and all participants have had the possibility to work on a functioning version of R/RStudio.

It is only natural that some of the techniques demonstrated seem strange and unnecessarily complicated when presented for the first time. However, if the participants are given time to review the entire syllabus it is the firm belief of the consultants, that the participants will find all techniques relevant as a foundation for their further venture into programming with R. We acknowledge that very little actual statistics beyond means and variances has been presented, but a solid knowledge of the basic functioning of R is a prerequisite before entering into more advanced topics.

3. Conclusions and recommendations

- The participants should be allowed time to review the entire material once again;
- The participant should try to replicate in R work already done in Excel or otherwise;
- The participant should form small groups for exchange of examples of applications of R.

Two missions building upon the result achieved within this mission has been discussed. One of the consultant will also participate in the next mission, thus in the planning of the next mission there will be a thorough understanding of what has already been taught and what will be relevant new topics.

Actions needed for moving forward as well as for preparing the next mission –add rows as needed.

Action	Deadline	Responsible person

Annex 1. Terms of Reference

<div>

Terms of Reference
EU Twinning Project BA 15 IPA ST 01 17

Component 2: Business Statistics
17-20 September 2019
Institute for Statistics of the Federation of Bosnia and Herzegovina
Sarajevo Zelenih beretki 26

</div>

Activity 2.D.1: Statistical programming in R – Basis training I

1. **Result for the activity**
   - Knowledge gained on programming in R

2. **Purpose of the activity**
   - Training of non-IT staff to programming in R

3. **Expected output of the activity**
   - Introduction to the statistical programme R has been made
   - R and the Rstudio interface have been demonstrated, introductions have been made to data structures, data management, programming, data exploration and basic statistics plus visualization.
   - Input provided to the ToR of next activity

4. **Participants**

Agency for Statistics of BiH (BHAS)
   - Edin Šabanović, Assistant Director in The Sector for Statistical Methodology, Standards, Planning, Quality and Coordination
   - Jasna Isaković, Head of Department for Sample Design and Methodology
   - Aida Zukanović, Senior Advisor in Department for Sample Design and Methodology
   - Ljiljana Okuka, Senior Advisor for Setting Quality Standards
   - Boro Kovačević, Senior Advisor for Forestry Statistics
   - Gorana Knežević, Head of Department for Society Development Statistics
   - Alen Bajramović, Head of Department for Services Statistics
   - Adisa Čubro, Senior Advisor for International Trade Statistics
   - Fahir Kanlić, Head of Department for Industry and Construction Statistics
   - Dajana Mitrović, Senior Advisor for Living Conditions

Institute for Statistics of Federation of BiH (FIS)

- Asmira Selimović, Head of Department for statistical methods and coordination
- Edina Mehidić, Associate for sampling and statistical analysis

Institute for Statistics of Republika Srpska (RSIS)

- Stana Kopranović, Senior Statistician for Environmental Statistics
- Vladimir Koprivica, Senior Statistician for Labour Force Survey
- Aleksandra Đonlaga, Senior Statistician for Service Statistics
- Vladimir Lambeta, Senior Associate for Standard Implementation and Methodology
- Lazo Šegrt, Senior Statistician for Agriculture and Fishery Statistics
- Tatjana Banjac Živković, Senior Statistician for Economic Accounts in Agriculture
- Dijana Maleš, Senior Statistician for Gross Fixed Capital Formation (GFCF)
- Nada Malinović, SBR Coordinator, Head of Registers, Samples and Classifications Division
- Dragana Mandić, Senior Statistician for Migration Statistics
- Danica Babić, Senior Statistician for SBS

Central Bank of Bosnia and Herzegovinia (CBBH)

- Maša Kamenica Jordamović, Senior Specialist in BoP Section
- Aida Kalčo, Specialist in BoP Section
- Ervin Zolić, Head of GFS and Financial Account Section
- Emina Milišić, Senior Specialist in MFS Section
- Bojan Savić, Head of Section for IT Applications

MS Experts

- Peter Stolze, Statistics Denmark
- Anne Vinkel Hansen, Statistics Denmark

Twinning Project Administration

- Katja Møller Hjelvang, RTA
- Đemka Šahinpašić, RTA Assistant
- Haris Imamović, Interpreter

Agenda

Location: Institute for Statistics of the Federation of Bosnia and Herzegovina

Time	Day	Place	Event	Purpose / Details
9:00-11:00	17 September Tuesday	Hotel	MS experts briefing with the RTA	Discuss the programme Preparation of activity
12:00-15:30	17 September Tuesday	FIS	Meeting with BC Experts	**Introduction to R and Rstudio** What is R and Rstudio, R scripts, the command line, the help function and the R community. How is R organized.
9:00-12:00	18 September Wednesday	FIS	Meeting with BC Experts	**Reading in and exporting data** Data types in R. Vectors, matrices, data frames and lists, R projects, R classes. Importing and exporting data from spreadsheets, delimited files, SAS, Oracle databases. Saving R data and exporting data (MS experts)
13:00-15:30	18 September Wednesday	FIS	Meeting with BC Experts	**Loops and functions** If statements, for loops, functions, the apply family, writing your own functions (MS experts)
9:00-12:00	19 September Thursday	FIS	Meeting with BC Experts	**Data manipulation and basic statistics** Data manipulation, pipes, basic analysis, merging, filtering, sorting, aggregating and sampling data Data cleaning (MS experts)
13:00-15:30	19 September Thursday	FIS	Meeting with BC Experts	**Data visualisation** Visualising in base R and ggplot, visualizing data on maps (BC and MS experts)
9:00-12:00	20 September Friday	FIS	Meeting with BC Experts	Agreement on summary mission report Conclusions and next steps