



ევროკავშირი
საქართველოსთვის
The European Union for Georgia



Twinning Project

Contract: GE 16 ENI ST 06 18

Strengthening the Capacity of the Georgian Statistical System

Component 1: Development of External Sector Statistics

Sub-component 1.4: “Statistical Programming in “R” – Basic training”

MISSION REPORT

Activity: 1.4.A (RS) “Statistical programming in R – Basic training I”

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

February 16th – March 10th 2021

Version: Final



ევროკავშირი
საქართველოსთვის
The European Union for Georgia



Expert contact information

Ms Anne Vinkel Hansen
Statistics Denmark,
Copenhagen, Denmark
Tel: +45 26111383
Email: aih@dst.dk

Table of contents

1. General comments	4
2. Assessment and results	4
3. Conclusions and recommendations	4
Annex 1. Terms of Reference	4
Annex 2. Persons met	7

1. General comments

This mission report was prepared within the EU Twinning Project "Strengthening the Capacity of Georgian Statistical System". This was the first mission within the sub-component 1.4: "Statistical Programming in "R" – Basic training provided". Due to COVID-19 the mission was carried out as a Remote Session in the period from February 16th to March 10th 2021. The mission was mainly devoted to introduction to the R language, the Rstudio environment, and basic data manipulation in R.

The purposes of the mission were:

- To introduce the participants to the statistical programme R and the Rstudio interface
- To introduce data structures, data management, programming, data exploration and basic statistics, plus visualisation.

The consultants would like to express their gratitude to the Geostat staff for their dedicated participation in the training.

The views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of the European Union, Geostat, Statistics Denmark, or other statistical institutions involved in the implementation of the project.

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 extend 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.)

¹ <https://annevinkel.github.io/RGeorgia/>

- Data wrangling using the dplyr and tidyr packages
- Visualisation using graphical functions from base R (plot) and introduction to the grammar of graphics with ggplot2 (including maps)
- Basic summary statistics
- Loops, conditional and writing functions in R

An online course with a large number of participants is a challenge, particularly when doing exercises, but the consultants were impressed by the enthusiasm and eagerness to learn displayed by the participants

3. Conclusions and follow up

- The participants should be allowed time to review course material and extra material linked on the course web page as relevant to their work
- The participants should try to replicate in R work already done in Excel or otherwise;
- The participants should form small groups for exchange of examples of applications of R.

Two missions building upon the result achieved within this mission have been discussed. The consultants will participate in the planning of the next mission in order to ensure that planning is based on a thorough understanding of what has already been taught and what will be relevant new topics.

Annex 1. Terms of Reference

EU Twinning Project GE 16 ENI ST 06 18

February 16th – March 10th 2021

Component 1: Development of External Sector Statistics

Sub-component 1.4: Statistical Programming in “R” – Basic training

Mandatory results and benchmarks for sub-component 1.4:

- Statistical Programming in “R” – Basic training provided

Indicators of Achievement (baseline and targets):

- Theoretical knowledge and practical use of the statistical software package “R”.
 - **Baseline:** 2020 – Use of “R” in the statistical production process is very limited in Geostat.
 - **Target:** June 2021 – at least 2 IT-staff and 10 statistical staff can perform basic statistical procedures in “R”

Activity 1.4.A (RS): Statistical programming in R – Basic training I

1. Purpose of the activity

- An introduction to the R environment.
- Getting data into and out of R
- Basic to advanced programming skills: Loops and functions,
- Data manipulation and basic statistics
- Data visualization

2. Expected output of the activity

- The participants have been introduced to the above mentioned topics.
- The participants have been practicing the above mentioned topics.
- Mission report written
- ToR and action plan for the next activity prepared



ევროკავშირი
საქართველოსთვის
The European Union for Georgia



Annex 2. Persons met

Geostat

22 employees from 5 different departments

RTA Twinning Team

Mr. Steen Bielefeldt Pedersen, Resident Twinning Advisor

Ms. Eka Lobzanidze, Resident Twinning Adviser Assistant

Ms. Nino Grdzlishvili, Resident Twinning Adviser Assistant, Translator