**The instruction manual of the demoLuminexPipeline**

# The demoLuminexPipeline instruction manual

Ncité Lima Da Camara

**Created:** 23 May 2022

**Updated:** 20 October 2022

After installing the R package demoLuminexPipeline from GitHub by following the instruction in step 1, download or clone the "demoLuminexPipeline\_code" repository from GitHub using the link provided in step 2 and save the folder (i.e., "code" folder) with all the files in your personal "code" directory that you create. The directory that you create for this folder is important since all the files that will be used and generated will be saved in the directory that you created. Lastly, please download and save the data sets from OneDrive (i.e., the link has been emailed to the analyst’s SUN email address) in the "data" directory that you need to create in the "code" directory (i.e., "code" folder) that you created.

# Table of content

[The demoLuminexPipeline instruction manual 1](#_Toc115855517)

[Table of content 2](#_Toc115855518)

[List of figures 3](#_Toc115855519)

[List of tables 3](#_Toc115855520)

[1. Install demoLuminexPipeline R package from GitHub: 4](#_Toc115855521)

[2. Download or clone the "demoLuminexPipeline\_code" repository from GitHub using the link provided: 5](#_Toc115855522)

[3. Download and save the data sets from OneDrive 5](#_Toc115855523)

[4. Please do the following steps: 6](#_Toc115855524)

[4.1. Before running the demoLuminexPipeline. Please make sure that the following files are present: 6](#_Toc115855525)

[4.2. Run the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio. 8](#_Toc115855526)

[4.3. After running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio. 9](#_Toc115855527)

[4.4. Please send the two HTML report files per investigation to me using the following email address ndc@sun.ac.za, namely: 10](#_Toc115855528)

[4.5. Please note the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, needs to be run three times, once for every investigation (i.e., R01AI128765, nexgen, and mph\_hiv). 10](#_Toc115855529)

[Appendices 12](#_Toc115855530)

[Appendix AA 12](#_Toc115855531)

# List of figures

[**Figure 1.1. Package update notification message. Please select "None".** 4](file:///I:\PhD_2018_2020\R_proj\TestRun20220915_46\code\demoLuminexPipelineInstructionManual_20220528_NL_08.docx#_Toc115855496)

[**Figure 2.1. Image of the "demoLuminexPipeline\_code" repository from GitHub.** 5](#_Toc115855497)

[**Figure 4.1. The "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, loaded in RStudio.** 8](file:///I:\PhD_2018_2020\R_proj\TestRun20220915_46\code\demoLuminexPipelineInstructionManual_20220528_NL_08.docx#_Toc115855498)

[**Figure 4.2. Running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio.** 9](#_Toc115855499)

[**Figure 4.3. The demoLuminexPipeline during a run. A green "Run All" rectangle bar with a red square bar will appear. Please leave to complete the run.** 9](#_Toc115855500)

[**Figure 4.4. After running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio.** 10](file:///I:\PhD_2018_2020\R_proj\TestRun20220915_46\code\demoLuminexPipelineInstructionManual_20220528_NL_08.docx#_Toc115855501)

[**Figure 4.5. Please change the investigation and the directory in the "config\_NL\_01.R" file before running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, (red arrows).** 11](file:///I:\PhD_2018_2020\R_proj\TestRun20220915_46\code\demoLuminexPipelineInstructionManual_20220528_NL_08.docx#_Toc115855502)

# List of tables

[**Table AA.1. Description of the required files.** 12](#_Toc115855508)

**Steps:**

## Install demoLuminexPipeline R package from GitHub:

Please note this is a private repository (repo) therefore the following steps are required

(i.e., run in the R Console):

**#Set config**

usethis::use\_git\_config(user.name = "YourName", user.email = "your@mail.com")

**#Go to github page to generate token**

usethis::create\_github\_token()

**#Paste the personal access token (PAT) into pop-up that follows:**

credentials::set\_github\_pat(504631559)

**#Now remotes::install\_github() will work**

remotes::install\_github("Ncite/demoLuminexPipeline/demoLuminexPipeline")

A computer screen capture

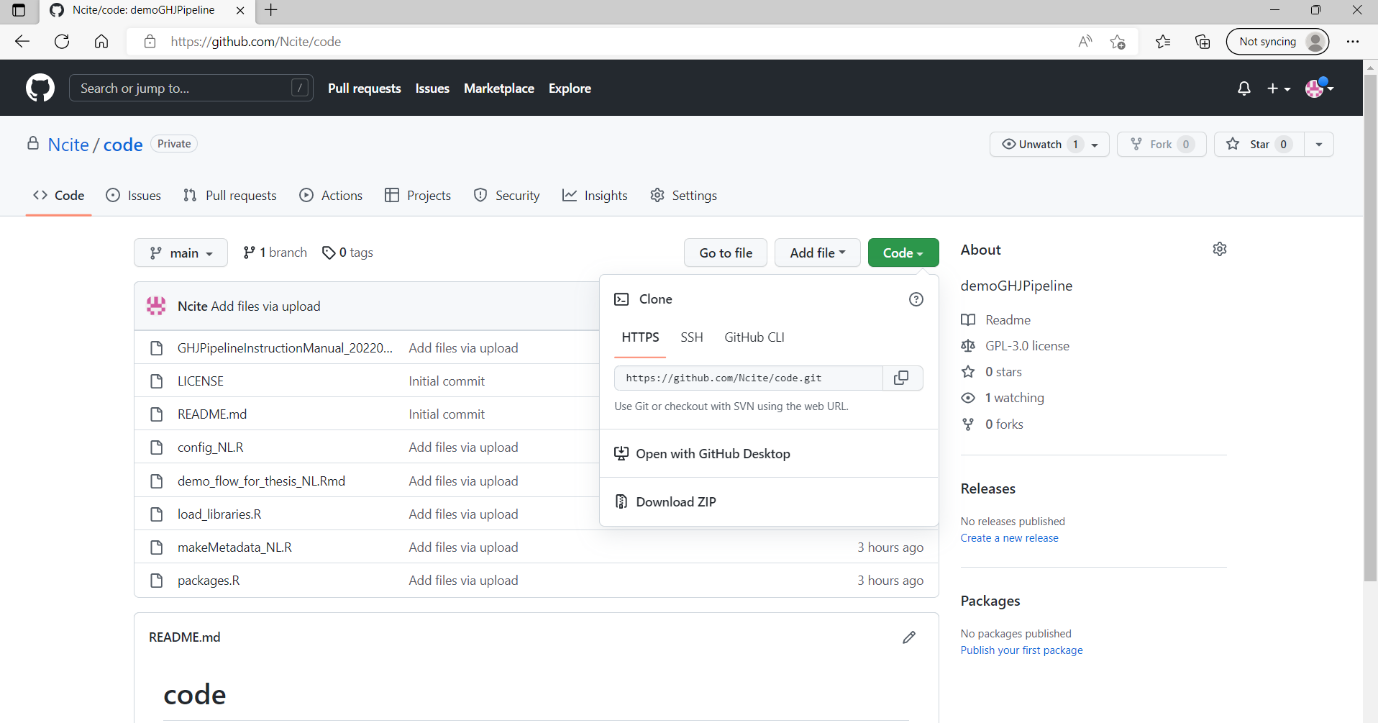
Description automatically generated with medium confidenceIf prompted with "These packages have more recent versions available. It is recommended to update all of them. Which would you like to update?" message, please select "None" (i.e., option 3) (see **Figure 1.1**).

**Figure 1.1. Package update notification message. Please select "None".**

## Download or clone the "demoLuminexPipeline\_code" repository from GitHub using the link provided:

[https://github.com/Ncite/demoLuminexPipeline\_code](https://github.com/Ncite/code)

Create a **"code" directory** on a personal machine. Download or clone the "demoLuminexPipeline\_code" repository from GitHub using the link provided. In the GitHub repository under "Code" (i.e., green button) there is a drop-down menu, and select the "Download ZIP" option (**Figure 2.1**). Please change the name of the downloaded unzipped folder from "code-main" to "code". Please save the "code" folder in the desired location (e.g., a R project you created saved with a relevant name such as "TestRun<Date>\_<Number>"), since all the files that will be used and generated will be saved in this directory.



**Figure 2.1. Image of the "demoLuminexPipeline\_code" repository from GitHub.** Instruction how to download the "demoLuminexPipeline\_code" repository from GitHub using "Code" (i.e., green button) and selecting "Download ZIP".

## Download and save the data sets from OneDrive

Download and save the data sets from OneDrive (i.e., link have been emailed to the analyst’s SUN email address) in the "data" directory that you need to create in the "code" directory (i.e., "code" folder) that you created.

## Please do the following steps:

### Before running the demoLuminexPipeline. Please make sure that the following files are present:

#### "load\_libraries.R"

#### "config\_NL\_01.R"

#### "makeMetadata\_NL\_01.R"

#### The data is present in the "data" folder and that the directory name is the same as the investigation name used. The trial data sets are in separate folders (i.e., R01AI128765, nexgen, and mph\_hiv) in the "data" folder.

#### The "demo\_flow\_for\_thesis\_NL.Rmd" is in the current working directory ("code/") in R.

Very important!

#### Please add user initials (i.e., analyst initials) on line 28 (e.g., user <- "NL") in the "config\_NL\_01.R" file.

#### Please set/activate investigation by removing the hash (i.e., "#") before "investigation" on lines 33 to 35. Alternatively, change the "investigation" by adding a hash (i.e., "#") before "investigation" to deactivate the investigation. Please note, that the **"investigation"** (i.e., investigation name) that you set by removing the "#" should match the set/activated (i.e., by removing the "#") data path you added in step 1.9 (i.e., see below). In addition, please make sure that the investigation name should match the directory name in which the data files are stored.

#### Please add your file paths to the data files that you downloaded from OneDrive in the "config\_NL\_01.R" file on lines 74 to 76. Your file path is the path to the "code/data" directory that you created at the start. Please see examples on lines 68 to 70 for the three data sets used and add a hash (i.e., "#") on line 68 before "directory\_data". Please also add "#" before 2 of the three data sets file path (i.e., your file paths that you added in lines 74 to 76) and remove the "#" in front of "directory\_data" containing the file path to the data set (i.e., data files) that correspond to the "investigation" (i.e., investigation name) that you set/activated (i.e., by removing the "#") in step 4.1.7 (i.e., above).

#### The file path that does not have a "#" sign before "investigation" and "directory\_data" will be used during the demoLuminexPipeline. Please change the investigation and directory before running the "demo\_flow\_for\_thesis\_NL.Rmd" file (i.e., the demoLuminexPipeline). This is done by loading the "config\_NL\_01.R" file in RStudio and removing the "#" sign before i) the "investigation", i.e., lines 33 to 35 depending on the investigation name and 2) the "directory\_data", i.e., lines 74 to 76 depending on the investigation set by removing one of the "#" on lines 33 to 35 and lines 74 to 76 (see red arrows in **Figure 4.5**). Please note that "investigation" and "directory\_data" should match the investigation (i.e., R01AI128765, nexgen, and mph\_hiv) that needs to be pre-processed. Only one investigation can be run in the demoLuminexPipeline at a time. See example **Figure 4.5**.

**Importantly**, when changing investigations please add the "#" sign to the investigations and the directories that you do not want to run. Alternatively, remove the "#" sign of the investigation and directory that you want to run (**Figure 4.5**).

#### Please note, that the demoLuminexPipeline will be run 3 times, once for every data set (i.e., R01AI128765, nexgen, and mph\_hiv). Thus, continue with step 4.1.8 until all the investigations have been ran. Please note that six HTML files will be sent to me (i.e., A "QCreport\_summaryList.html" from each investigation and a "Testreport\_summaryList.html" from each investigation).

#### A screenshot of a computer Description automatically generated with medium confidencePlease open the "demo\_flow\_for\_thesis\_NL.Rmd" file (i.e., thedemoLuminexPipeline) in RStudio (see **Figure 4.1**).

**Figure 4.1. The "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, loaded in RStudio.**

### Run the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio.

By clicking the down-facing arrow in the "Run" tab and selecting "Run All" (see **Figure 4.2** below). The demoLuminexPipeline will start the data pre-processing run. Please leave to complete the run, a green "Run All" rectangle bar with a red square bar will appear (see **Figure 4.3** below).

Graphical user interface, application, Word

Description automatically generated

**Figure 4.2. Running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio.**

Graphical user interface, application, Word

Description automatically generated

**Figure 4.3. The demoLuminexPipeline during a run. A green "Run All" rectangle bar with a red square bar will appear. Please leave to complete the run.**

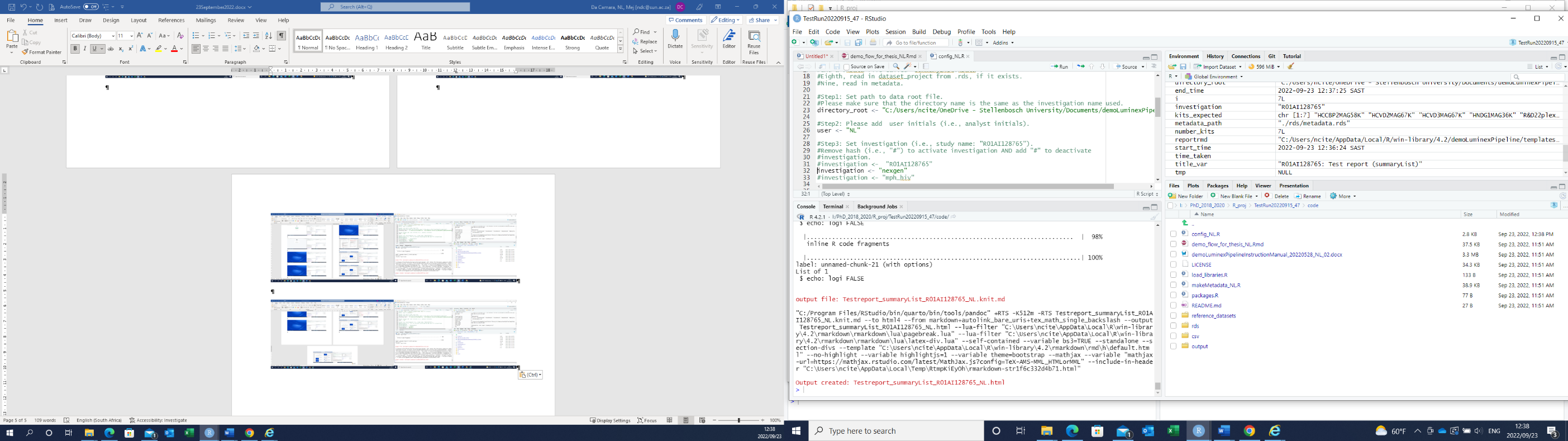
### After running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio.

The pre-processed Luminex data files are saved as .rds files in the rds folder and the two HTML report files generated per investigation are saved in the current working directory in the "output" folder (see red arrows in **Figure 4.4**), namely:

#### The pre-processed Luminex data files are saved as .rds files in the rds folder.

#### The "QCreport\_summaryList.html" and the "Testreport\_summaryList.html" per investigation are saved in the "output" folder.

**Figure 4.4. After running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, in RStudio.** The pre-processed Luminex data files are saved as .rds files in the rds folder and the two HTML report files, namely: 1. The "QCreport\_summaryList.html" and 2. The "Testreport\_summaryList.html" per investigation are saved in the "output" folder.



### Please send the two HTML report files per investigation to me using the following email address ndc@sun.ac.za, namely:

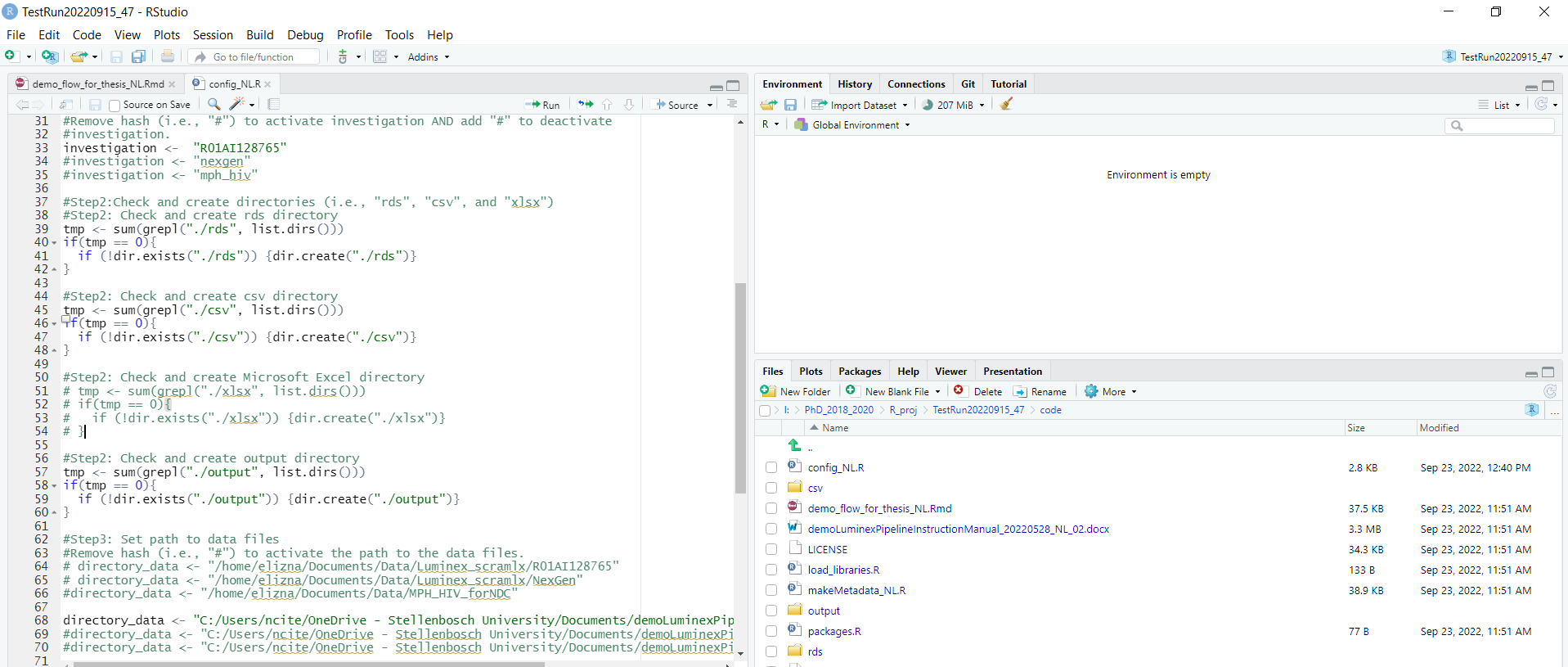
#### The "QCreport\_summaryList.html"

#### The "Testreport\_summaryList.html"

### Please note the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, needs to be run three times, once for every investigation (i.e., R01AI128765, nexgen, and mph\_hiv).

#### Please change the "investigation" and the "directory\_data" before running the "demo\_flow\_for\_thesis\_NL.Rmd" file (i.e., the demoLuminexPipeline). This is done by loading the "config\_NL\_01.R" file in RStudio and removing the "#" sign before the "investigation" and the "directory\_data", i.e., lines 33 to 35 and lines 74 to 76 depending on the investigation that you want to run (see red arrows in **Figure 4.5**). Please note that "investigation" and "directory\_data" names should match the investigation (i.e., R01AI128765, nexgen, and mph\_hiv) that needs to be pre-processed.

Importantly, when changing investigations please add the "#" sign to the investigations and the directories that you do not want to run. Alternatively, remove the "#" sign of the investigation and the directory that you want to run (see **Figure 4.5**). Then, continue with steps 4.2 until all the investigations have been ran. Please note that six HTML files will be sent to me (i.e., A "QCreport\_summaryList.html" and a "Testreport\_summaryList.html" from each investigation).



**Figure 4.5. Please change the investigation and the directory in the "config\_NL\_01.R" file before running the "demo\_flow\_for\_thesis\_NL.Rmd" file, i.e., the demoLuminexPipeline, (red arrows).**

Thank you!

Enkosi!

Dankie!

# Appendices

# Appendix AA

**Table AA.1. Description of the required files.**

|  |  |
| --- | --- |
| **Required files** | **Description** |
| "config\_NL\_01.R" | The "config\_NL\_01.R" is the  configuration file in which the investigation, the data set path, and the metadata path is set, as well as configure the output folders prior to the demoLuminexPipeline run.  Set the "investigation" and "directory\_data" before running the demoLuminexPipeline. |
| "data" folder | The "data" folder contains the raw Luminex data for each investigation in separate folders (directories). The directory name should be the same name and style (lower case with underscores as separators) as the investigation name used. |
| "demo\_flow\_for\_thesis\_NL.Rmd" | The "demo\_flow\_for\_thesis\_NL.Rmd" file is the actual Luminex data pre-processing pipeline. The R function code are structured and ordered for standardized automated data pre-processing of Luminex data.  Important, only one investigation can be run in the demoLuminexPipeline at a time. See example **Figure 4.5**. |
| "load\_libraries.R" | List of the required R libraries that need to be loaded in R. |
| "makeMetadata\_NL\_01.R" | List all the elements and the expected data from the study design of an investigation (study) prior to receiving the raw Luminex data from the laboratory. |

# Appendix AB

**Table 2. List of duplicate** **"simple" synonym analyte names for a list of analytes (10).** A list of analyte names that have duplicate "simple" synonym analyte names and the analyte name that we choose to keep the duplicate synonym analyte name in the "simple" variable of the analyte reference table. The other analyte’s duplicate synonym analyte name was removed from the analyte reference table "simple" variable.

|  |  |  |
| --- | --- | --- |
| **simple** | **consensus** | **uniqueid\_nf** |
| il1beta | il1b | 159 |
| tpo | tpo | 363 |
| factori | fibrinogen | 511 |
| ehk | krt1 | 607 |
| ulbp2 | ulbp2 | 464 |
| gcp2 | cxcl6 | 312 |
| psa | pros1 | 272 |
| c5a | c5 | 45 |
| ssa | ssaro52 | 550 |
| sm | sm | 536 |