Automated Descriptive Epidemiology Report Production

The following is a step-by-step guide describing how to produce automated descriptive epidemiology reports to investigate weekly exceedances.

Investigating exceedances consists on the following steps:

1. Review the weekly exceedance report
2. Review the weekly exceedance summary
3. Extract data from SGSS
4. Run the automated descriptive epidemiology R script

In order to process the data you will need to have:

* R installed (<https://cran.r-project.org/bin/windows/base/>)
* RStudio installed (<https://www.rstudio.com/products/rstudio/download/>)
* Have the following packages installed in R: ‘ggplot2’, ‘reshape2’, ‘ISOweek’, ‘pander’, ‘rmarkdown’

(these can be install by typing install.packages(“insert\_package\_name”) into the R console e.g. install.packages(“ggplot2’”))

# Review the weekly exceedance report

## The exceedance report is generated by the Statistics, Modelling and … department and forwarded to FES teams every Monday. This can be accessed at the following intranet page:

<http://phenet.phe.gov.uk/Resources/laboratory-reporting-surveillance/LabBase/Pages/SGSS-exceedance-reports.aspx>

# Review the weekly exceedance summary

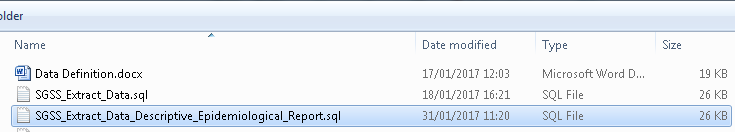
## The exceedance summary report is generated by the duty information manager shortly after the exceedance report is produced each Monday. The exceedance summary report will be forwarded to all FES South West Epidemiologists.

# Extract data from SGSS

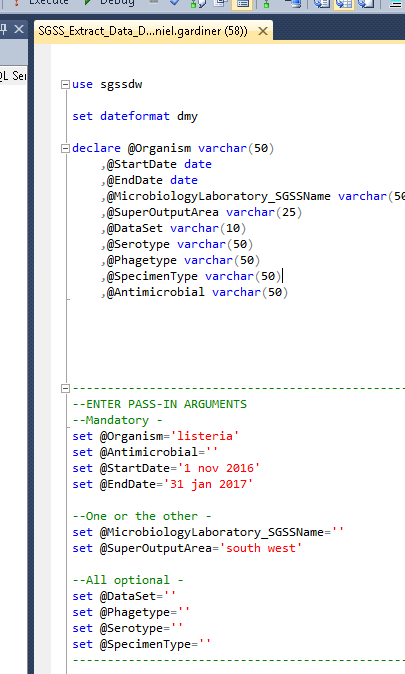
## Go to the following folder:

U:\Informatics\Statistical Analysis\SGSS\EXTRACTION AND DEDUPLICATION SCRIPT

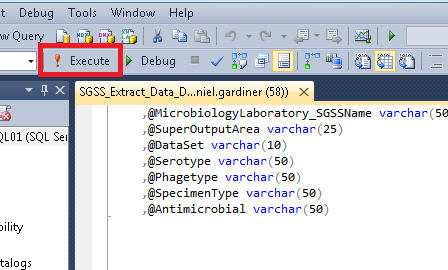
## Open the file titled **SGSS\_Extract\_Data\_Descriptive\_Epidemiological\_Report.sql**



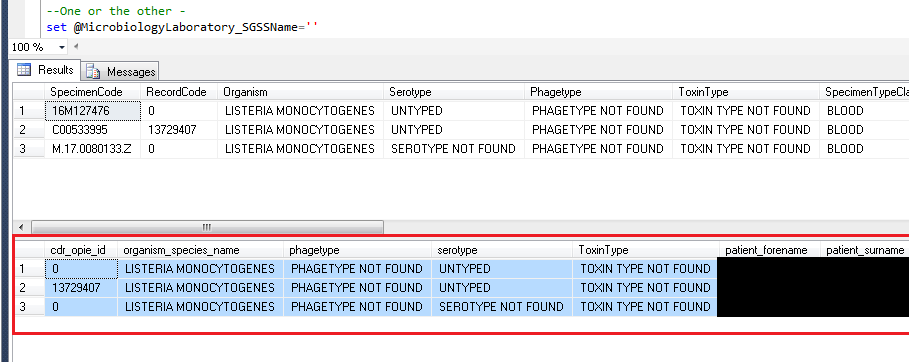
## This should look something like this; select the appropriate pass-in arguments (please see the [User guide](file:///U:\Informatics\Statistical%20Analysis\SGSS\EXTRACTION%20AND%20DEDUPLICATION%20SCRIPT\User%20Guide.docx) for details)



## Execute the query



## Then copy and paste the bottom table into a .csv file



## Save the .csv in the following folder:

U:\Core Services\Field Epidemiology Team Bristol\Incidents and Outbreaks\Outbreaks\EXCEEDANCE\Automated descriptive epidemiology\Data

# Run the automated descriptive epidemiology R script

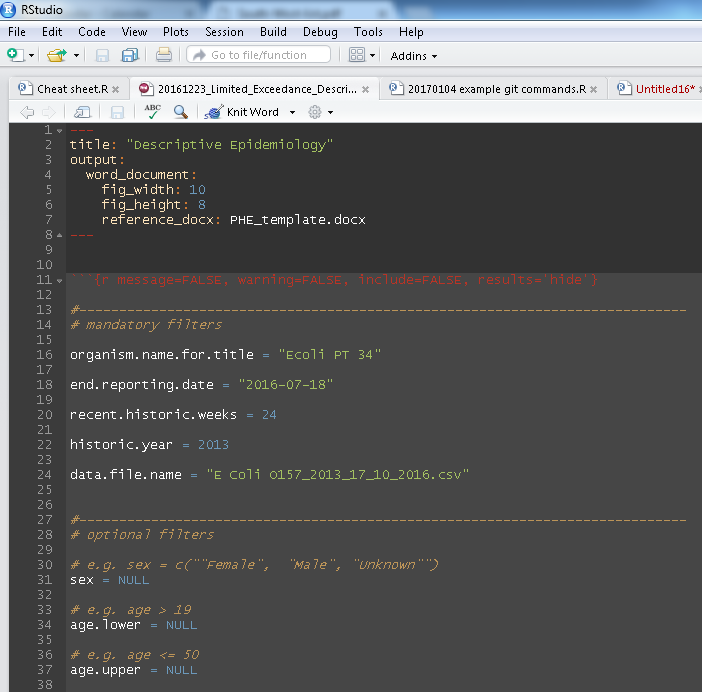
## Go to the following folder:

U:\Core Services\Field Epidemiology Team Bristol\Incidents and Outbreaks\Outbreaks\EXCEEDANCE\Automated descriptive epidemiology\R scripts

## Open the file titled **20170202\_MASTER\_Exceedance\_Descriptive\_Epidemiology.Rmd**

## 

The file should look something like this



There are three sections the user can enter parameters into (1) mandatory parameters, (2) optional filter parameters and (3) optional figure aesthetic parameters

## Information must be entered into all mandatory parameters, the user must specify each of the following:

|  |  |
| --- | --- |
| **Parameter** | **Description** |
| organism.name.for.title | Organism name to be used in the report title |
| end.reporting.date | The last date the report is to be run until (i.e. the maximum specimen date) |
| recent.historic.weeks | The number of recent weeks to be reported on |
| historic.year | The year from which to report on |
| data.file.name | The name of the .csv file containing the raw linelist data |

## The user can then choose to restrict the data using any of the parameters within the optional filter parameters section, examples below

Suppose we want to only include cases over the age of 18, we would set:

* age.lower >= 18
* we would leave all other parameters set to NULL

Suppose we want to only include cases over the age of 18 and resident in Exeter local authority, we would set:

* age.lower >= 18
* local.authority = c(“Exeter”)
* we would leave all other parameters set to NULL

Suppose we want to only include cases over the age of 18 and resident in Exeter local authority or Cornwall local authority, we would set:

* age.lower >= 18
* local.authority = c(“Exeter”, “Cornwall”)
* we would leave all other parameters set to NULL

Suppose we want to only include cases over the age of 18, resident in in Exeter local authority or Cornwall local authority and are not known to have reported foreign travel, we would set:

* age.lower >= 18
* local.authority = c(“Exeter”, “Cornwall”)
* travel = c(“No”, “Unknown”)
* we would leave all other parameters set to NULL

## The user can then choose to change the figure aesthetics using any of the parameters in the optional figure aesthetic parameters, parameter descriptions in below

|  |  |
| --- | --- |
| **Parameter** | **Description** |
| day.label.break | Number of x-axis breaks where specimen date is used |
| week.label.break | Number of x-axis breaks where specimen week is used |
| month.label.break | Number of x-axis breaks where specimen month is used |
| recent.tile.plot.text | Specifies whether numbers should be overlayed on recent tile plots |
| historic.tile.plot.text | Specifies whether numbers should be overlayed on historic tile plots |
|  |  |

## The aesthetic parameters are necessary as the time period the report is generated over can vary and is set by the user using the recent.historic.weeks and historic.year parameters in the mandatory parameters section.

## Below is an example of Figure 1where the recent.historic.weeks parameter has been set to 10 and the day.label.break parameter has been set to 0, clearly the x-axis is hard to read

## Below is an example of Figure 1where the recent.historic.weeks parameter has been set to 10 and the day.label.break parameter has been set to 1, now the the x-axis is much easier to read