# Strategy for CyToF analysis

## Clean the data before further analysis

Follow the script “clean up gating alle filer panel 1 mars 2022.R” where each step is described in “introduction-to-clean-up-gating.docx”. Both files can be found on github <https://github.com/folkehelseinstituttet/cytof>. The file “clean up gating alle filer panel 1 mars 2022.R” must be individually adapted to your dataset, where the amount of gating is decided based on lower\_gate\_percent and upper\_gate\_percent. With higher percentages the gates will be tighter, how tight you want your gate might depend on your question, you might lose something important by having to tight gates, but you are more likely to get rid of duplicates by having tight gates. The script saves files about each gating step to predefined folders. These plots should be inspected to make sure that you are happy with the gates. If not, the script should be rerun either with tighter gates for all files or with forced gates on some files. Continue adjusting the percent gating until you are happy with the clean-up gates.

## Plot the distribution of each channel after clean-up

The next step is to plot your clean data. Plot all signals against time using the file “FigSignalPerChannelOfCleanData.R”. This script should be adjusted to your dataset by updating folders where the data to used are stored and folders where you want your results. You also must adjust so that the channel names are correct for your dataset. When the script has run, inspect all channels for all files to see if any of your cytof files or channels has failed. Decide if al data can be used for further analysis.

## Positive - Negative cells (or high-low cells)

The file marker\_gating\_panel\_1\_Immune\_response\_aga\_tighterGates.R makes gates for all or only selected markers. If further analysis should be done on for instance only B cells, CD3, CD45 and CD4, has to be pregated for all cells so that cells that are positive for all three markers could be found and selected for further analysis. Explanation of different gating strategies can be found in Gating.docx.

## FlowSOM analysis

runFlowSOM\_analyse.R source the script FlowSOM\_analyse.R for different seeds and different selected cells to analysis. Update how many ks that you want to use for your analysis, use a vector with some different ks. Update where the files are read from, which files to use and where results should be saved. Make the folder where the result should be saved. All this is described in “introduction to FlowSOM analysis.docx”. This analysis cluster the cells into predefined number of clusters, count how many cells that are in each cluster for each file and find the median, 10 % quantile and 90 % quantile of the signal in each cluster. It also makes heatmap of the median signals. The file runFlowSOM\_analyse.R also source a file that read the data, i.e. readDataToAnalysePanel1TighterCleanUp.R, this file is described in introduction\_to\_readDataToAnalyse.docx.

## Results Regression of clusters

The file runResults\_Regression\_Clusters.R source the file resultater\_Regression\_clusters.Rmd. When running this a word document with the result of regression of explanatory variables on the counts of cells in each cluster is done. In this regression age, sex and status is used as explanatory variables. The resultater file must be updated based on your research question and variables that you want to analyse for.