# Discussion 2017-10-25

**Data Analysis Report**

Data Analysis report was good. It was written in a very objective manner. It lacked the writer’s opinion a bit in deciding which tool is best or which tool is best used when. For the next month that will be the main focus.

**Framework guidelines**

For the known and discussed tools, add hypotheses and guidelines when these tools should be used. Two aspects are important for such guidelines.

* What does the data consist of? Data sets can be very different of each other. Several details can be important. Examples are: The size of the data set. The number of features present in the data set, the feature type (Categorical, numerical) and the presence of unknown values.
* What should be researched for with the data? Data sets are usually created with the idea of researching a specific topic. The desired end product creates questions that can be answered better by using a certain technique using a certain tool. Example questions are: Should correlations be found between several variables? Do we need to make clusters for to define several risk groups?

To create those guidelines the tools themselves should be tested, as well. Their advantages and disadvantages can best be found when using them.

**Example test data set**

An already finished project used a data set from the hospital and performed biomedical analysis on it. The way the project was handled is one way, but maybe from a framework – selector another approach would be better. To put the above guidelines to the test, one or more alternative approaches should be made for the project. The only aspects of the data that are known are within the project report, so the challenge is to create good and efficient task descriptions to analyse the biomedical data.