**Hand-on day 2**

* Load the data “covid\_analytic\_clinical\_data.csv” (read data information at “covid\_analytic\_clinical\_data.doc”)
* Check for the 10 first lines
* Get a summary of each variable
* Plot a histogram of proportion of mortality and co-morbidities (i.e. Diabetes, Cardiovascular diseases, Cerebrovascular diseases, Co-morbidity, Obesity, Liver diseases, COPD, and Cancer). Do you have any comments?
* Plot a scatterplot of each co-morbidity (x-axis) with mortality outcome (y-axis). Please also add a linear regression trend line. Do you have any comment?
* Fitting each co-morbidity on the mortality outcome, by using linear regression.
* Which co-morbidity have a highest impact on the mortality outcome? Please discuss.
* Which co-morbidity have a highest significant on the mortality outcome? Please discuss.
* Please using number of patients in each study as a weight on the linear regression, lm(…, weight = N). How the results different from previous? Please discuss.
* Please perform a multiple regression using

lm(Mortality~Diabetes+Cardiovas+Liver+COPD+Cancer, weights = N)

How do you interpret the result? Please discuss.