

AdvanceR

Assignment sets for advanced R users

1. Upload "Pollution.xls" data, calculate the correlation between SO2 level and every other measurements, plot and save the figures. Write a function and use lapply instead of for loop.
2. Upload "dist_length.txt" and "size-assembly.txt" as input data. Plot "asbsize", "N50", "Ncont", "Maxsize", and "Minsize" against "cutoff" for all different number of cells and input coverage (calculated by "base_to_assembly/(1.2*10⁹)"). An example plot is provided.
3. Upload "Preprocessing_Summary.txt" as input data. This is a table created for many possible preprocessing metrics on the RNASeq dataset we are using in our workshop. Some of these preprocessing metrics are correlated one to another. Remove the correlated metrics, carry out PCA analysis on the reduced table and then plot the first two PCs. An example plot is provided.