Note that pre-processing and processing of the data can be resource intensive (i.e., in terms of RAM requirements). As such it is recommended to use gc() where appropriate, and also increase memory limit in R if necessary.

Cleaning the R environment after each step of the (pre)processing should also help with memory requirements.