Dear Leon,

We are group 13 and these are the prospective data sets (1) we are considering working with, in our final project. As you can see in the attached references (2), this paper is about comparing single cell gene expression between IPF/COPD patients and normal subjects. This is an article advancing the field’s understanding on diseases that have elevated chronic respiratory diseases to the 3rd cause of death in the world and we would like to try making it more accessible – within our capacity and project management planning.

We would like to perform a single-cell transcriptomic analysis by down-sampling the data set to around ~10 people in total (by including 2-3 representative participants per cohort) and use Seurat (3) - which is a tidyverse compatible R package - as the tool to generate insights from the data. The platform is the 10X Genomics/Illumina HiSeq 4000.

The data lay as individual .rds patient data that collectively get integrated during the analysis. We have managed to load the data in Rstudio and it seems that they require tidying, i.e. we believe we comply with the data wrangling requirements of the project description. What’s more, we consider building a Shiny app (if time and limitations permit) presenting our analysis with the purpose of making it more accessible to clinical audiences - that might be interested in interactive visualizations of these data.

There is a Shiny app provided by the authors (4) but we will build our own based on our analysis and we are also taking the opportunity to mention that there is no GitHub repo provided by the authors of the paper neither any code was found associated with the project in any place we looked up (nor for the Shiny app). Therefore, whatever we either re-create or build, it will be of course of our own making, following the instructions of the paper and relevant tutorials online (5).

Kind regards,

Group 13

**References:**

1. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136831>
2. <https://advances.sciencemag.org/content/6/28/eaba1983>
3. <https://github.com/satijalab/seurat>
4. <http://www.ipfcellatlas.com/>
5. <https://satijalab.org/seurat/articles/get_started.html>