For homework

Note: The cell ranger count you executed in class should take around 20minutes. You can proceed with the homework after you successfully run both samples.

1. Aggregate Agg\_EthanEric\_Controls.

* Use the text editor of your choice to make **csv file** with your library id and molecule\_h5 file
* Template

library\_id, molecule\_h5

<the --id you give while running cell ranger count>, /path\_to/molecule\_info.h5

<the --id you give while running cell ranger count>, /path\_to/molecule\_info.h5

* From sbatch file, module load cell ranger/3.0.1

cellranger aggr --id=any\_name\_you\_prefer --csv=/path\_to/csv\_file –-normalize=mapped

2. Answer the following questions with the information from the summary file.

* What is the mean reads per cell after aggregation? Why does this number approximately equal to the mean reads in dataset Eric\_Control but not dataset Ethan\_Control?
* Change the clustering type to k-means with k=2. Are the cells clustered by their datasets? Increase the value of k and observe how the clusters change. Does each cluster contain a similar number of cells from these two datasets?

3. (Advanced) Try install Seurat in R(https://satijalab.org/seurat/articles/install.html).

Transfer the filtered\_feature\_bc\_matrix/ folder to your local computer.

Follow <https://satijalab.org/seurat/articles/pbmc3k_tutorial.html> to run QC on our data.