

# Cho Lab Technical Exam

2021-03-23

## Instructions

- This is an open-book exam that should take a few hours to complete.
- Retain all code and notation for evaluation.
- Organizing content in formats like Rmarkdown or a Jupyter notebook is preferred but not required.
- No question has a single correct method. Even if you aren't able to arrive at an answer, still expand on your thought process.
- All tools, languages, and resources are fair game.
- **Upload all files for your submission here** with titles that include your name (e.g. `chris_tastad_output.pdf`)

## Questions

1. **Print the values of an integer vector 1:100, but**
  - print “Cho” if a value is divisible by 3
  - print “Lab” if a value is divisible by 7
  - print “ChoLab” if a value is divisible by both 3 and 7
2. **Consider the unknown sequencing data given here.**
  - perform a QC process on this data, and aggregate your results into a single output
    - a suggested approach would be to use a tool combo like fastQC/MultiQC
  - provide a detailed assessment of the nature and quality of this data given your QC output
  - speculate on the type of sequencing used to generate this data
3. **Process and cluster the tutorial pbmc3k single cell dataset using the parameters below:**
  - use a single cell analysis workflow of your choice (Seurat, Scanpy, etc)
  - include any figures, outputs, or interpretation you feel are relevant in QC and preprocessing
  - provide the 5 most differentially expressed markers from cluster 2

*# clustering and filtering params*

```
dimensions = 10
resolution = 1.2
nFeature_RNA > 200
nFeature_RNA < 2000
percent.mt < 10
```