

## Cbb752b23 Quiz #1 Study Guide

There are **some topics** from the lectures that we **drilled down** into detail on. Here is a list of them that you should know for the quiz:

- How to do a dynamic program sequence alignment
- How to compute sequence variability (with entropy)
- How to use simple hash tables & indexes to speed up alignment
- Table normalization + insert/delete/update/select statements
- How to construct a simple decision tree from a set of features
- Definitions of accuracy, sensitivity, specificity, PPV, FPR, and FDR
- How to do simple clustering (via K means or agglomerative approaches)
- Network quantities (clustering coefficient & path length)
- What is the degree distribution of a scale free network
- How to interpret the SVD of a data matrix -- e.g. what is a biplot
- Principle of SMV
- How to build up a profile matrix (e.g. using a simple EM/psi-blast approach)
- What are the major challenges for single-cell sequencing analysis? Why do these challenges occur in single-cell sequencing analysis?

For **all other topics**, you should understand at a **high level**, i.e., the contents of the lecture.

Quiz #1 will cover everything from the 1st half of class (1/18 to 2/22).

You can find all slides on the class website: <http://cbb752b23.gersteinlab.org/syllabus>, and all recorded lectures on canvas->media library. Examples of previous quizzes may also be helpful to get an idea of what may be on this year's: <http://cbb752b23.gersteinlab.org/quiz>