**Software Developer At-Home Interview Task: Subgroup Passing Rates**

Given the sample input files provided (demographics.txt, schools.txt, and scores.txt) and the requirements below:

\* Write code to fulfill the requirements (you may use any language you prefer). For any files you submit, ensure that each has a comment on top, with your name and the language the code is written in.

\* Add documentation explaining what you did and why you did it including a list of business rules, in any form you prefer.

\*If something is ambiguous or otherwise unclear, make a decision about how to handle it and add a note to your response documentation explaining what was ambiguous, why, and why you chose the option that you did.

\* All input files have a header row, are tab-separated, and have a newline at the end.

\* a student is receiving Special Ed if his/her Special\_Ed flag is a 'Y'

\* a student is ELL if his/her ELL flag is an 'A', 'E', 'I', 'O', or 'U'

We will evaluate the task in terms of:

1. Code logic and readability – we work as a team so our code needs to be understandable so others can maintain
2. Generalizability – this code should be adaptable to data and situations similar to the one presented but not identical
3. Documentation – assumptions and decisions of how to handle anything that is unclear should be explained in a separate document
4. Functionality – the output should be correct (given whatever assumptions you have noted)

Requirements:

A) Find the pass rates as follows: ("pass rate" = percent of students that receive at least a 65.)

A1) Each of these should exist for each school (i.e. one set of records for school A, one set of records for school B)

A2) Each of these should exist for Math and separately for ELA

A3) Each of these should exist for the following distinct populations:

A3i) all students

A3ii) ELL students

A3iii) special ed recipients

A3iv) both ELL and special ed recipients

(This should result in 2 schools \* 2 subjects \* 4 populations = up to 16 resulting numbers)

B) Produce a flat student-level dataset (as a tab-separated file) - this should have one record per student with all of that student's data listed. The file should be recoded as follows:

B1) any student who is ELL should have ELL = 1, otherwise ELL = 0

B2) any student who is receiving Special Ed should have Special\_Ed = 1, otherwise Special\_Ed = 0

B3) Any student who has passed a particular exam should have a 1 for that exam. Any student who has failed a particular exam should have a 0 for that exam. If the student did not take an exam, the value should be left empty

Example: (the order and exact names of the columns are irrelevant)

ID | School | score\_Math | score\_ELA | pass\_Math | pass\_ELA | ELL | Special\_Ed

55 | C | 85 | 55 | 1 | 0 | 0 | 1