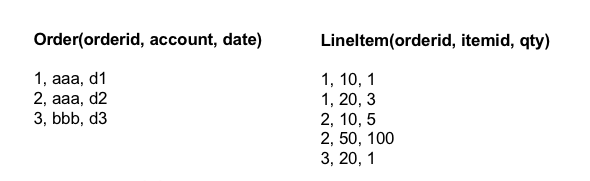
Quiz #2: MapReduce Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ID: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q1. Where does the MapReduce program store 1. Input data, 2. Intermediate files, and 3. Output data? (1 pt) Why do we care about where they store the data? (1 pt)

Q2. Write the Map and Reduce tasks and their output for joining these two tables: (2pts)



Q3. Write a MapReduce program that multiplies two matrices A and B in **two** stages. You can assume that the matrices are provided to you in a file in a sparse matrix format. Each line of the file represents an element in a matrix. For example, a line: [‘A’, 0, 0, 1] indicates that A[0, 0] = 1. You may assume that both matrices are 5 x 5. (3 pts)

Q4. What is reducer size (1 pt) and replication rate (1 pt) and why do we care about the reducer size (1 pt)?