1. We want to apply MapReduce to solve a problem of database joins with computation of summary statistics. Assume that we have two databases, one with personal information and another with income by year. Both databases are indexed by social security number (SSN).
    a. Describe an architecture to compute average income in each city in 2007.
    b. State the MapReduce pseudo-code for solving this task.

    SS-DB 1: (SSN, {Personal Information})
    123456:(John Smith;Sunnyvale, CA)
    123457:(Jane Brown;Mountain View, CA)
    123458:(Tom Little;Mountain View, CA)

    SS-DB 2: (SSN, {year, income})
    123456:(2007,$70000),(2006,$65000),(2005,$6000),...
    123457:(2007,$72000),(2006,$70000),(2005,$6000),...
    123458:(2007,$80000),(2006,$85000),(2005,$7500),...

    Note: **Both inputs sorted by SSN**