Map Reduce Pop Quiz

1. Given a log with lines of the form “<user>,<gender>,<product>” write a MapReduce query that computes the number of products purchased by each female user. Provide the map and reduce functions. (Pseudocode ok.)

2. What is the equivalent relational query? Assume an input relation of type user × gender × product. (use SQL or relational algebra.)

3. For any deterministic map() and reduce() functions, is MapReduce guaranteed to provide the same answer for a key on every run? Explain.