State if the following statements are TRUE (T) or FALSE (F)

1. Consider a database schema with the following two relations:

   \[ S(eid, \text{dob}, \text{project}, \text{hours}) \]
   \[ R(eid, \text{name}, \text{sex}) \]

   where \( eid \) is the employee id, \( \text{dob} \) is the date-of-birth, \( \text{project} \) is the project that the employee participates in, \( \text{hours} \) is the number of hours the employee spent on the project, \( \text{name} \) is the employee name, and \( \text{sex} \) is the sex of the employee. The key attributes are underlined (i.e., \( (eid, \text{dob}) \) is a composite key of \( S \), and \( eid \) is the key of \( R \)). Moreover, \( \text{sex} \) can only be M (male) or F (female). The following predicates appear in the majority of queries:

   \[ \text{hours} \leq 5, \text{hours} \leq 15, \text{sex} = M, \text{sex} = F \]

   We would like to fragment \( R \) and \( S \) using the techniques we learn in the lecture.

   a) \( S1 = \sigma_{\text{hours} \leq 15} S \) is one fragment for \( S \). [ F ]

   b) There are in total 2 fragments of \( S \) and 2 fragments of \( R \). [ F ]

   c) Relation \( R \) can be further fragmented using derived horizontal fragmentation based on the fragmentations of \( S \). [ F ]

   d) The vertical fragmentation of \( S \) into \( S1(eid, \text{project}) \) and \( S2(eid, \text{dob}, \text{hours}) \) is desirable if queries frequently access \((eid, \text{dob}, \text{hours})\). [ F ]

2. In query optimization, an exhaustive strategy with pruning heuristics (e.g., avoid cross products) will always produce a better plan than one with across products. [ F ]

3. \( R1 \text{ SEMIJOIN } R2 = R1 \text{ SEMIJOIN } (R2 \text{ SEMIJOIN } R1) \) [ T ]

4. \( R1 \text{ JOIN } R2 \text{ JOIN } R3 = R2 \text{ JOIN } R3 \text{ SEMIJOIN } (R1 \text{ JOIN } R2) \) [ F ]

5. In query optimization, the number of tuples per relation can influence the search space (i.e., the number of query plans). [ F ]
Consider the following hybrid fragmentation:

\[
\begin{align*}
\text{EMP1} &= \sigma_{\text{ENO} \leq "4" } (\Pi_{\text{ENO}, \text{ENAME}} (\text{EMP})) \\
\text{EMP2} &= \sigma_{\text{ENO} > "4" } (\Pi_{\text{ENO}, \text{ENAME}} (\text{EMP})) \\
\text{EMP3} &= \Pi_{\text{ENO}, \text{TITLE}} (\text{EMP})
\end{align*}
\]

Consider the query: \textbf{SELECT TITLE FROM EMP WHERE ENO > “5”}.

a) The optimal plan involves EMP2 and EMP3 only. \[ F \]

b) The optimal plan pushes down selection on ENO to EMP2 and pushes down projection to EMP2 (to prune away ENAME). \[ F \]