The assignment is to be turned in before Midnight (by 11:59pm) on January 25th. You should turn in the solutions to this assignment as a PDF file through Canvas. The solutions should be produced using editing software programs, such as LaTeX or Word, otherwise they will not be graded. The assignment should be done in groups of two students. Each group must submit only one file that contains the full name, OSU email, and ONID of every member of the group.

1: Incomplete Information (3.5 points)

(a) Find a SQL that computes set difference and returns false positive over a database with null values. You may show the correctness of your answer using an example. (0.5 point)

(b) Find a SQL construct other than subquery, *in*, and *not exists* that if it appears in a SQL queries, the query may return false positive or negative over a database with null values. You may show the correctness of your answer using an example. (1 point)

(c) Find a SQL query that returns false negatives over databases with null values. Then, propose a transformation that translates this query to a SQL query that does not return any false positive and retains all true positives delivered by the original query over databases with null values. You may also find a SQL query that returns false positives over databases with null values and transform it to a query that does not return any false positive and retains all true positives delivered by the original query over a database with null values. (1.5 points)

(d) Find the largest subset of SQL queries that does not return any false positives over a database with null values. (0.5 points)

3: Constraint Inference (0.5 point)

(a) Given that X, Y, W, Z are attributes in a relation, using the Armstrong's axioms, prove that if we have $X \to Y$ and $YW \to Z$, then $XW \to Z$. (0.25 point)

(b) Given that X, Y, Z are attributes in a relation, using the Armstrong's axioms, prove that if we have $X \to Y$ and $X \to Z$, then $X \to YZ$. (0.25 point)

2: Schema Decomposition (2.5 points)

Consider the relation schema R with attributes A, B, C, and D and the following functional dependencies: AB→C, AC→B, B→D, BC→A.

(a) List all keys for R. (0.5 point)

(b) Is R in BCNF? If it is not, decompose it into a collection of BCNF relations. (0.5 point)

(c) Is R in 3NF? If it is not, convert it into a collection of 3NF relations. (0.5 point)

(d) Prove that, if relation R has only one key, it is in BCNF if and only if it is in 3NF. (1 point)

3: Information preservation (0.5 point)

(a) Suppose you are given a relation R(A,B,C,D) with functional dependencies B→C and D→A. State whether the decomposition of R to S_1(B,C) and S_2(A,D) is lossless or dependency preserving and briefly explain why or why not. (0.5 point)