

Name: _____

This midterm is **open book, open notes**, but there can be no sharing of any material, and no communication except with me. You can use the Internet, but only as a library. If you are not sure if something is allowed, check with me.

Some questions in this midterm use the database schema and sample instance depicted below. [The following details are identical to those in quiz 1.] The **Boxes** table records information about boxes used by a packing store. For each type of box, it stores the box's name and supplier, the maximum weight of its contents (in pounds), its price in dollars, its dimensions (width, depth, and height) in inches, and the number available (inventory). The **Products** table records information about products stocked by the packing store for its customers. For each product, it stores the product name, customer name, weight (in pounds), and dimensions (as in **Boxes**). The type of each attribute appears directly below its name. *Primary key attributes* are underlined.

When asked for queries, you must provide answers that work for all possible database instances, not just the example instance depicted below. For brevity, relational algebra expressions abbreviate the schemas of the **Boxes** and **Products** tables as $B(N, S, L, P, W, D, H, I)$ and $P(N, C, L, W, D, H)$, respectively.

The following tables are repeated on the last page of the midterm. You may detach that page and use it for reference. There is no need to reattach it.

Boxes

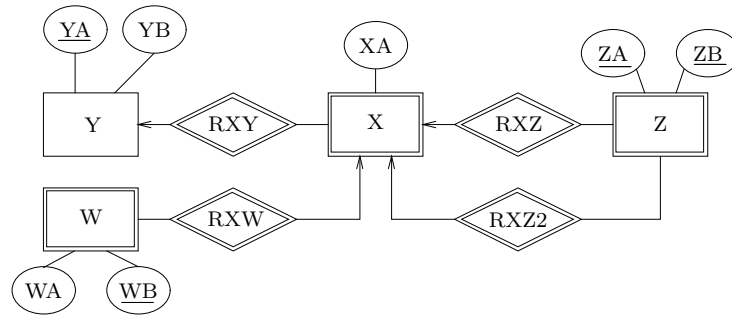
<u>name</u>	<u>supplier</u>	load	price	width	depth	height	num
varchar(20)	varchar(20)	float	decimal(6,2)	float	float	float	integer
small cube	box-o-rama	10.5	2.99	24	24	24	100
large cube	box-o-rama	20.0	4.99	50	50	50	42
small cube	box world	1.0	1.99	8	8	8	22
economy pack	box world	1.5	0.99	9	12	0.25	9

Products

<u>pname</u>	<u>customer</u>	weight	width	depth	height
varchar(20)	varchar(20)	float	float	float	float
PDR	Alice Armstrong	5	8	10	6
bookshelf speaker	Bob Beerl	15	24	13	16

- (1 pt.) Write your name in the space provided above.

2. (5 pts.) What is the primary key of X in the following E-R diagram?



3. (5 pts.) Using the method discussed in class, generate the relational schema for the entity sets X and W for the E-R diagram of Question 2. Underline primary-key attributes.

4. (5 pts.) Suppose an E-R diagram D has five double diamonds. For each of the following statements, indicate whether the statement is definitely true, definitely false, or neither (could be true or false) by writing **True**, **False**, or **Neither**, respectively, in the provided spaces.

- (a) _____ There are exactly four double boxes in D .
- (b) _____ There are exactly five double boxes in D .
- (c) _____ There are exactly six double boxes in D .
- (d) _____ There are no single boxes in D .
- (e) _____ There is exactly one single box in D .
- (f) _____ There are exactly five single boxes in D .

5. (5 pts.) Suppose the following is a valid instance:

BirdCounts

A	B	C	D	E	F
1	1	1	1	1	1
1	1	1	1	1	2
1	1	1	1	2	2
2	2	2	1	1	1

Some of the following functional dependencies are **definitely false** (i.e., they definitely do not hold in BirdCounts). Circle the dependencies that are definitely false:

$$A \rightarrow B$$

$$B \rightarrow A$$

$$F \rightarrow A$$

$$E \rightarrow F$$

$$F \rightarrow E$$

$$ABCD \rightarrow EF$$

$$EF \rightarrow ABCD$$

$$ABCDE \rightarrow F$$

$$ABC \rightarrow DEF$$

$$C \rightarrow D$$

6. (4 pts.) Consider a relation $R(A, B, C, D, E, F)$ with the following basis set of functional dependencies:

$$A \rightarrow BCF$$

$$B \rightarrow C$$

$$DE \rightarrow A$$

$$CF \rightarrow E$$

Compute $\{A, C\}^+$.

7. (5 pts.) List **all** the keys of the relation R of Question 6. (A copy of R 's schema appears on the scratch page.)

8. (5 pts.) Is R in BCNF? Justify your answer.

9. (5 pts.) Consider a relation $S(A, B, C, D, E)$. If $\{AC \rightarrow E\}$ is a basis set of S 's functional dependencies, list **all** possible *lossless decompositions* of S .

10. (5 pts.) The *price per unit volume* (*PPUV*) for a box is defined as the ratio of the box's price (in dollars) to the box's volume (in cubic inches). Write a SQL statement that creates a view `BoxesPPUV(name, supplier, load, ppuv)` consisting of tuples of the form (n, s, l, x) , where x and l are, respectively, the PPUV and the load for the box named n from supplier s .
11. (5 pts.) We wish to compute a bar chart that indicates how PPUV (see Question 10) varies with box loads. We will use 10-pound-wide ranges for loads in the bar chart: $[0, 10)$, $[10, 20)$, $[20, 30)$, \dots . Write a SQL query that produces a list of tuples of the form (r, p) where p is the **average** PPUV of boxes with loads in the range $[r, r + 10)$. The list should be sorted in ascending order of r . Skip empty ranges (bar chart buckets with no entries). [Hints: You may use views, including the one defined in Question 10. The function `floor(x)`, which returns $\lfloor x \rfloor$, may be useful.]

12. (5 pts.) Consider the bar chart of Question 11. Write a SQL query to find the bar chart ranges (lower ends of the intervals) that have the maximum values (average PPUV values). The desired output is a list of tuples of the form (r) sorted in ascending order of r .

Scratch page

Material here will not be graded. You may detach and discard this page.

Boxes

<u>name</u>	<u>supplier</u>	load	price	width	depth	height	num
varchar(20)	varchar(20)	float	decimal(6,2)	float	float	float	integer
small cube	box-o-rama	10.5	2.99	24	24	24	100
large cube	box-o-rama	20.0	4.99	50	50	50	42
small cube	box world	1.0	1.99	8	8	8	22
economy pack	box world	1.5	0.99	9	12	0.25	9

Products

<u>pname</u>	<u>customer</u>	weight	width	depth	height
varchar(20)	varchar(20)	float	float	float	float
PDR	Alice Armstrong	5	8	10	6
bookshelf speaker	Bob Beerl	15	24	13	16

The following schema is for the relation in Question 6:

$$R(A, B, C, D, E, F)$$
$$A \rightarrow BCF$$
$$B \rightarrow C$$
$$DE \rightarrow A$$
$$CF \rightarrow E$$