

Name: _____

Please note that much of this assignment is devoted to getting familiar with the class accounts and submission procedure; thus, it is much simpler than the ones that will follow. You should submit (1) a hard-copy of this assignment with your answers filled in and (2) an electronic file as detailed below. You are welcome (and encouraged) to use any resources (e.g., Web sites) to help you with your work. However, *all such help must be clearly noted* in your submissions. Further, no matter what you use, *you must be able to explain* how and why it works.

For this homework, questions marked with \star are required for COS 580 students only. A \star mark on a question in this homework does *not* mean a similar question is not applicable to COS 480 students in future homeworks or exams! Therefore, COS 480 students are encouraged to answer the \star questions for extra credit on this homework as well.

COS 480 students earn points as indicated for each question below. The non- \star questions sum to 100, the full score. Any points earned on \star questions (potentially another 50 points) are simply added to the score. If the resulting score is greater than 100, the remaining points are extra credit. COS 580 students must answer all questions to yield a raw score (maximum 150) that will be scaled by $2/3$ to obtain the score (maximum 100) on the homework. (COS 580 students who wish to work for extra credit should contact me for a suitable assignment.)

1. (1 pt.) Write your name in the space provided above.
2. (2 pts.) Read the material on the class Web site. Sign your name here to indicate that you have read this material: _____
3. (2 pts.) Change the passwords on your Gandalf (Unix) and PostgreSQL accounts. Fill in the following information. (Change your passwords and note *yes* in the last column.)

Account	User Name	Password Changed?
Gandalf	_____	_____
PostgreSQL	_____	_____

4. (9 pts.) Indicate the result of evaluating each of the following relational algebra expressions on the database instance depicted below. (The two title rows of each table indicate the attribute names and types.)

Rooms			Facilities			
building	floor	room	building	room	chairs	screens
varchar(20)	integer	integer	varchar(20)	integer	integer	integer
Neville	1	110	Neville	210	25	0
Neville	2	210	Neville	227	35	1
Neville	2	227	Corbett	101	2	0
Corbett	1	101				
Corbett	2	201				

For notational convenience in relational algebra, we shall refer to the Rooms(*building*, *floor*, *room*) and Facilities(*building*, *room*, *chairs*, *screens*) relations as $R(B, F, R)$ and $F(B, R, C, S)$, respectively.

(a) $\pi_{BR}\sigma_{F>1}R$

(b) $\pi_{BR}\sigma_{F>1}R \cup (\sigma_{R\geq 200}\pi_{BR}R \cap \pi_{BR}\sigma_{S>0}F)$

(c) $(R \cup R) - \pi_{BFR}(R \bowtie \pi_C F \bowtie \sigma_{E<5} F)$

5. (9 pts.) Provide relational algebra queries that return the information requested in each part below. Reminder: As in all query-writing questions, your answers should work on *all* database instances, not only the one provided as an example above.

(a) The room numbers of rooms on the second floor of Neville.

- (b) The rooms (building names and room numbers) that have at least 25 chairs and 1 screen each.
- (c) Rooms (building names and room numbers) whose facilities are unknown (i.e., rooms with no corresponding information in the Facilities relation).
6. (9 pts.) Indicate the result of evaluating each of the following SQL queries on the database instance of Question 4.
- (a) `select building from Rooms;`
- (b) `select distinct building from Facilities where chairs > 1;`
- (c) `(select building from Rooms where floor > 1)
intersect
(select building from Facilities where chairs > 20);`
7. (9 pts.) For each part below, write a SQL query that is equivalent to the query in the corresponding part of Question 5.

(a)

(b)

(c)

8. (10 pts.) Write SQL statements to create the database instance depicted in Question 4. You may assume integer types for all attributes. In addition to creating the necessary tables, your answer must populate them with data as indicated in Question 4.

9. (10 pts.) Read about the *script*, *col*, *tar*, *gzip*, *top*, *ps*, and *kill* Unix utilities by using the *man* command on Gandalf. (Type `man script`, for example.) Answer the following in brief:

(a) What does *script*'s `-a` option do? Provide an example of its use.

(b) Indicate how we may use *col* to remove backspace-generating control-characters from a file.

(c) Indicate how we may use *tar* to produce a listing of the contents of (i.e., file names within) a tar archive file `mystery.tar`.

(d) Explain the difference between the files `a1.tar` and `a2.tar` resulting from the following actions:

```
cd /tmp; mkdir d; echo foo > d/f1; echo bar > d/f2
tar cf a1.tar d
cd d; tar cf a2.tar *
```

(e) Suppose you notice that the CPU usage on a machine is very high. Indicate how you may determine the process responsible for that usage and the owner of that process. Assuming that the owner is you, indicate how you would terminate the process without access to that process's user interface, if any.

10. (10 pts.) Execute the statements in your answer to Question 8 in PostgreSQL using the *psql* program. Using Question 9 as a hint, capture your *psql* session in a *plain-text ASCII* file called `dbgen-log.txt`.

For this and the next question, please check carefully that your file is a plain-text ASCII file. (For example, check whether your file can be viewed properly using the command `more dbgen-log.txt`.)

Indicate the size of your file, in bytes: _____

11. (10 pts.) Execute the statements in your answer to Question 7 in *psql*. As in Question 10, capture your *psql* session in a *plain-text ASCII* file called `query-log.txt`.

Indicate the size of your file, in bytes: _____

12. (10 pts.) You should generate a gzip-compressed tar archive file called `M-hw01-N.tgz`, replacing *M* with your last-name and replacing *N* with an arbitrary 4-digit integer

16. (5 pts.) ★ Define an *antijoin* from first principles (without using any other relational algebra expressions).

17. (10 pts.) ★ Given a relational schema with $S(I, J, K)$ and $T(I, J, L)$ indicate whether the expressions $\pi_{IJ}S - \pi_{IJ}(S \bowtie T)$ and $\pi_{IJ}S - \pi_{IJ}T$ are equivalent. If so, prove the equivalence. If not, provide a counterexample.

18. (15 pts.) ★ For each relational algebra expression in Question 4, provide an equivalent SQL query in the corresponding part below. Although you do not need to (but may) provide a formal proof of equivalence, you must explain why your answer is correct.

(a)

(b)

(c)

19. (15 pts.) ★ Prove or disprove each of the following for $S(I, J, K)$ and $T(L, M, N)$, with all attributes of integer type:

(a) $S \cap T \equiv S - (S - T)$

$$(b) \pi_{IJ} \sigma_{I < J} S \equiv \sigma_{I < J} \pi_{IJ} S$$

$$(c) \sigma_{I=K}(S \times T) \equiv \sigma_{I=K} S$$

