MCDA 5540 – Assignment 1

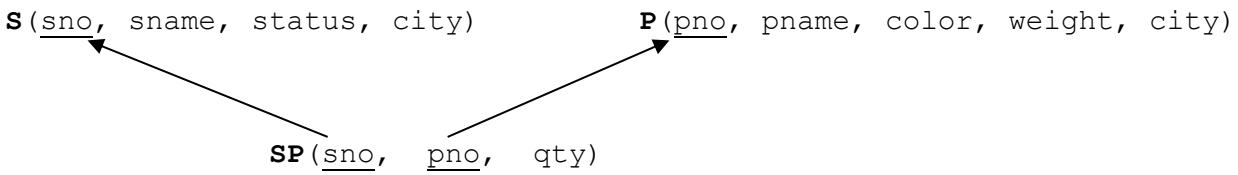
==========

Date assigned: Sep. 17 Time due: Sep. 30 at 11:59pm

(***how*** to submit: see comments further ***below***)

PROBLEM 1 (25 marks)

Consider the following relational database schema (suppliers-parts):



Prepare a text file question1.sql containing SQL expressions for each of the following queries (Also, attach sample screenshot of all queries in question1.pdf – DO NOT include any extra files such as output file or docx. You don't require to give any explanation):

(a) Get the names and locations of the suppliers who have shipped part with pno = 3.
(b) Get the part numbers and names of parts that have been shipped by suppliers located in Paris with status at least 20.
(c) For each part, show the part number, name, and the number of suppliers who have supplied the part.
(d) For each London supplier who has shipped at least 1000 parts, show the name of the supplier and the total number of parts he/she has shipped.
(e) Get the names and cities of the suppliers who have supplied all parts that weigh less than 4 grams.

PROBLEM 2 (25 marks)

Consider the University database schema in the last page of this assignment. Write relational algebra expressions for each of the following queries. Write your answers in question2.txt file (DO NOT include any extra files).

(a) Get the names of courses in the CS department.
(b) Get the names of students who got an 'A' in CS3380.
(c) Get the instructors who have taught CS1310 and CS3380.
(d) Get the instructors who have taught all courses of the CS department.

PROBLEM 3 (50 marks)

Create 6 tables as shown below and load data into the tables from the given data .csv files. (Note: Choose appropriate data types and size for the columns). Write sql queries (prepare question3.sql) for the following questions (Also, attach sample screenshot of output in question3.pdf – DO NOT include any extra files such as output file or docx. You don't require to give any explanation):

Database/ Schema Name: samplecompany

Table Name	.csv file	Description
customer	customer.csv	It contains company's customer information
order_fact	order_fact.csv	It contains order information
product_dim	product_dim.csv	It contains information related to products provided by suppliers
employee_payroll	employee_payroll.csv	It contains payroll information for all company employees
employee_addresses	employee_addresses.csv	It contains employee's address
staff	staff.csv	It contains confidential employee data for both existing and past employees

(a) **Goal:** Summarizing Data in Groups
Table: customer
Query: Write a query that displays the following statistics for each country:

- Total number of customers
- Total number of male customers
- Total number of female customers
- Percent of all customers that are male (Percent Male).

Display the result by value of Percent Male so that the country with the lowest value is listed first, with the remaining countries following in ascending order.

(b) **Goal:** Summarizing Data in Groups
Table: product_dim, order_fact
Query: Create a result by combining two tables.

- Include columns Product_ID, Product_Name from product_dim table.
- Include a column with the label Total Sold. Use a summary function to create this column, which displays the quantity sold for each product.
- Specifies the tables product_dim, with the alias p and order_fact with the alias o.
- Join the tables by matching the values of the appropriate columns in each table.
- Groups the results by Product_ID from product_dim table and Product_Name.
- Orders the rows so that products with the highest number sold appear at the top of the report and then by Product_Name.

Note: DO NOT use nested queries

(c) **Goal:** Create a result with a self-join.
Table: employee_addresses, staff
Query: Display result of all trainees and workers at company. For each trainee or temporary worker, the report should include the employee ID, name and job title, and manager ID and name. The report should be ordered by Employee_ID.

(d) **Goal:** LEAD and LAG functions
Table: employee_payroll
Query: 1) Calculate the difference between the salary of the current row and the previous row. 2) Calculate the difference between the salary of current row and the following row.

IMPORTANT COMMENTS

1. Late submissions will not be accepted (not even a single minute!).
2. How to submit: Make a zip file Axx.zip, where Axx is your A-number. Submit it on Brightspace. Email submission will not be considered.
3. This is an individual assignment. While you are permitted to discuss the problems, you are not supposed to take notes during such discussions. The submitted solutions must consist entirely of your individual work.

STUDENT

Name	Student_number	Class	Major
Smith	17	1	CS
Brown	8	2	CS

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

A database that stores student and course information.