**Databases for Data Science**

**Homework 4**

**Provide answers to the following questions as a PDF or Word file:**

**Consider the following schema (key fields are underlined):**

**Suppliers(sid: integer, sname: text, address: text)**

**Parts(pid: integer, pname: text, color: text)**

**Catalog(sid: integer, pid: integer, cost: real)**

**The catalog contains parts supplied by a supplier at a particular cost.**

**Convert the following English queries into SQL:**

**1. Find the names and addresses of suppliers who supply parts that cost less than $100.**

SELECT DISTINCT S.sname, S.address

FROM Suppliers S, Catalog C

WHERE S.sid = C.sid

AND C.cost < 100;

**2. Find the names of suppliers who supply parts that are either green or blue.**

SELECT DISTINCT S.sname

FROM Suppliers S, Catalog C, Parts P

WHERE S.sid = C.sid

AND C.pid = P.pid

AND P.color IN ('green', 'blue');

**3. Find the names of suppliers who supply both a red part and a green part (at least one of each color).**

SELECT DISTINCT S.sname

FROM Suppliers S, Catalog C, Parts P

WHERE S.sid = C.sid

AND C.pid = P.pid

AND S.sid IN (

SELECT sid

FROM Catalog C1

JOIN Parts P1 ON C1.pid = P1.pid

WHERE P1.color = 'red'

)

AND S.sid IN (

SELECT sid

FROM Catalog C2

JOIN Parts P2 ON C2.pid = P2.pid

WHERE P2.color = 'green'

);

**4. Find the names of suppliers who supply every red and blue part.**

SELECT DISTINCT S.sname

FROM Suppliers S

WHERE NOT EXISTS (

SELECT P.pid

FROM Parts P

WHERE P.color IN ('red', 'blue')

AND P.pid NOT IN (

SELECT C.pid

FROM Catalog C

WHERE S.sid = C.sid

)

);

**5. Find the sids of suppliers who supply every red part or supply every blue part (use nested query).**

SELECT DISTINCT sid

FROM Suppliers

WHERE (SELECT COUNT(DISTINCT pid) FROM Parts WHERE color = 'red') =

(SELECT COUNT(DISTINCT pid) FROM Catalog WHERE Catalog.sid = Suppliers.sid AND Catalog.pid IN (SELECT pid FROM Parts WHERE color = 'red'))

OR

(SELECT COUNT(DISTINCT pid) FROM Parts WHERE color = 'blue') =

(SELECT COUNT(DISTINCT pid) FROM Catalog WHERE Catalog.sid = Suppliers.sid AND Catalog.pid IN (SELECT pid FROM Parts WHERE color = 'blue'));

**6. Find pairs of suppliers who supply the same part.**

SELECT DISTINCT S1.sname, S2.sname

FROM Catalog C1, Catalog C2, Suppliers S1, Suppliers S2

WHERE C1.pid = C2.pid

AND C1.sid < C2.sid # Makes sure both Suppliers aren’t the same

AND C1.sid = S1.sid

AND C2.sid = S2.sid;

**7. Find the pids of parts supplied by at least two different suppliers.**

SELECT DISTINCT pid

FROM Catalog

GROUP BY pid

HAVING COUNT(DISTINCT sid) > 1;

**8. Find the supplier who supplies the most expensive part in the whole catalog.**

SELECT DISTINCT S.sname

FROM Suppliers S, Catalog C1

WHERE S.sid = C1.sid

AND C1.cost = (SELECT MAX(cost) FROM Catalog);

**9. Find all suppliers who don’t supply a red part.**

SELECT DISTINCT sname

FROM Suppliers

WHERE sid NOT IN (

SELECT sid

FROM Catalog

JOIN Parts ON Catalog.pid = Parts.pid

WHERE color = 'red'

);

**10. Find the count of all parts supplied by “Acme Suppliers” grouped by their color. Do same query but show only when the count is less than 100.**

SELECT P.color, COUNT(\*) AS part\_count

FROM Suppliers S

JOIN Catalog C ON S.sid = C.sid

JOIN Parts P ON C.pid = P.pid

WHERE S.sname = 'Acme Suppliers'

GROUP BY P.color;

SELECT P.color, COUNT(\*) AS part\_count

FROM Suppliers S

JOIN Catalog C ON S.sid = C.sid

JOIN Parts P ON C.pid = P.pid

WHERE S.sname = 'Acme Suppliers'

GROUP BY P.color

HAVING COUNT(\*) < 100;