Ace Cassidy   CPTS451   HW5   Spring'20

Question 1

a. Create a table "MySales " in the database and load the data from the provided file into that table (use PostgreSQL). Please refer to the Milestone-1 description for instructions on how to import a .csv file in PostgreSQL.

```
CREATE TABLE mysales(
pname VARCHAR,
category VARCHAR,
price INT,
discount INT,
month VARCHAR);

\copy mysales(pname,category,price,discount,month) FROM 'salesData.csv' DELIMITER ',' CSV
```

b. You find the following FDs in sales data:

  - pname, category -> price
  - month->discount

Write two SQL queries which will verify that these FDs hold on the sales data.

```
SELECT pname, category
FROM mysales
GROUP BY (pname, category)
HAVING COUNT (DISTINCT price) > 1;

SELECT month
FROM mysales
GROUP BY month
HAVING COUNT (DISTINCT discount) > 1;
```

c. Decompose the table to BCNF relations and create SQL tables for the decomposed schema. Show your work. Create primary keys and foreign keys where appropriate.Turn in the SQL commands for creating the tables

```
CREATE TABLE price (
    pname varchar,
    category varchar,
    UNIQUE (pname, category),
    price int
);

CREATE TABLE discount (
    month varchar,
    UNIQUE (month),
    discount int
);

CREATE TABLE sales (
    pname varchar,
    category varchar,
    month varchar,
    PRIMARY KEY (pname, category, month),
    FOREIGN KEY (pname, category) REFERENCES price (pname, category),
    FOREIGN KEY (month) REFERENCES discount (month)
);
```

d. Provide the INSERT statements and give the number of tuples in each decomposed table

```sql
-- 27 rows
INSERT INTO price (pname, category, price)
SELECT DISTINCT ON (pname, category)
    pname,
    category,
    price
FROM
    mysales;

-- 12 rows
INSERT INTO discount (month, discount)
SELECT DISTINCT ON (month)
    month,
    discount
FROM
    mysales;

-- 319 rows
INSERT INTO sales (pname, category, month)
SELECT
    pname,
    category,
    month
FROM
    mysales;
```

Question 2

R(A,B,C,D,E,F) and FD's A →BC, D →AF

   a. KEYS(D, E)

   b. Relation can be decomposed into R(D, E) R(D, A, B, C, F)

   c. FD's are preserved

S(A,B,C,D) and FD's ABC →D, D →A

   a. KEYS(A, B, C) OR KEYS(D, B, C)

   b. Relation can be decomposed into R(D, B, C) R(D, A)

   c. FD ABC →D may not be preserved