

MIS381N (Fall 2018) Homework1

Due on 10/08 before 12 pm

NOTE: To complete this homework, you will need to really step back and ask what problem you are trying to address and how to break the bigger problem into smaller questions to solve using SQL. Given the nature of questions, there will be subjectivity and variance in the answers.

Please submit SQL statement (well formatted) with results in a Word document. Whenever there are more than 10 records in the result, please copy and paste the first 10 records.

The data files are available on Canvas. Here are the tables you need to create; Note: PK is primary key and FK is the foreign key.

TABLE: MANAGERS (REGID is the PK)

❖ COLUMN_NAME	❖ DATA_TYPE
REGID	NUMBER
REGION	VARCHAR2 (10 BYTE)
REGMANAGER	VARCHAR2 (10 BYTE)

CONSTRAINT:

REGION can be only 'East', 'South', 'Central', 'West'.

TABLE: PRODUCTS (ProdID is the PK)

❖ COLUMN_NAME	❖ DATA_TYPE
PRODID	NUMBER
PRODNAME	VARCHAR2 (100 B...
PRODCAT	VARCHAR2 (30 BYTE)
PRODSUBCAT	VARCHAR2 (30 BYTE)
PRODCONT	VARCHAR2 (20 BYTE)
PRODUNITPRICE	NUMBER (7, 2)
PRODMARGIN	NUMBER (5, 3)

CONSTRAINTS:

PRODCAT can only be 'Technology' 'Furniture' or 'Office Supplies'

PRODCONT take on only 'Jumbo Drum', 'Medium Box', 'Jumbo Box', 'Wrap Bag', 'Large Box', 'Small Box', 'Small Pack'

TABLE: ORDERS (*OrderID* is the PK)

↕ COLUMN_NAME	↕ DATA_TYPE
ORDERID	NUMBER
STATUS	VARCHAR2 (10 BYTE)

TABLE: CUSTOMERS (*CustID* is the PK; *CustReg* is the FK on delete cascade)

↕ COLUMN_NAME	↕ DATA_TYPE
CUSTID	NUMBER
CUSTNAME	VARCHAR2 (35 BYTE)
CUSTREG	NUMBER (1, 0)
CUSTSTATE	VARCHAR2 (20 BYTE)
CUSTCITY	VARCHAR2 (20 BYTE)
CUSTZIP	NUMBER (5, 0)
CUSTSEG	VARCHAR2 (15 BYTE)

CONSTRAINT:

CUSTSEG can be only Home Office 'Corporate', 'Small Business', 'Consumer'.

TABLE: ORDERDET (*OrderID* (FK), *CustID* (FK), *ProdID* (FK) are together a PK; All FK are on delete restrict)

❖ COLUMN_NAME	❖ DATA_TYPE
ORDERID	NUMBER
CUSTID	NUMBER
PRODID	NUMBER
ORDPRIORITY	VARCHAR2(15 BYTE)
ORDDISCOUNT	NUMBER(3,2)
ORDSHIPMODE	VARCHAR2(15 BYTE)
ORDDATE	DATE
ORDSHIPDATE	DATE
ORDSHIPCOST	NUMBER(5,2)
ORDQTY	NUMBER
ORDSALES	NUMBER(7,2)

CONSTRAINTS

ORDPRIORITY can be 'Low', 'Medium', 'High', 'Critical', 'Not Specified'

ORDSHIPMODE can be 'Regular Air', 'Delivery Truck', 'Express Air'

TASKS:

Do the following and copy into Word document the DDL, DML, results, and any errors. Like in Part A, please copy and paste the first 10 rows if there are more than 10 rows in the answer.

QUESTION 1: Create the 5 tables given above. You should define primary keys, foreign keys, and other CHECK constraints. And, load the data from Excel spreadsheet.

```
CREATE TABLE MANAGERS(
    Regid Number PRIMARY KEY NOT NULL,
    REGION VARCHAR2(10) CONSTRAINT Chk_Region CHECK (REGION IN ('East',
'Central', 'West', 'South')),
    REGMANAGER VARCHAR2(10)
);

CREATE TABLE PRODUCTS(
    PRODID NUMBER PRIMARY KEY NOT NULL,
    PRODNAME VARCHAR2(100) NOT NULL,
    PRODCAT VARCHAR2(30) NOT NULL CONSTRAINT Chk_prodcats CHECK(PRODCAT
IN ('Technology', 'Furniture', 'Office Supplies')),
    PRODSUBCAT VARCHAR2(30) NOT NULL,
    PRODCONT VARCHAR2(20) NOT NULL CONSTRAINT Chk_prodcats CHECK(PRODCONT
IN ('Jumbo Drum', 'Medium Box', 'Jumbo Box', 'Wrap Bag', 'Large Box',
'Small Box', 'Small Pack')),
    PRODUNITPRICE NUMBER(7,2),
    PRODMARGIN NUMBER(5,3)
);
```

```

CREATE TABLE ORDERS(
    ORDERID NUMBER PRIMARY KEY NOT NULL,
    STATUS VARCHAR2(10)
);

CREATE TABLE CUSTOMERS(
    CUSTID NUMBER PRIMARY KEY NOT NULL,
    CUSTNAME VARCHAR2(35),
    CUSTREG NUMBER(1,0) REFERENCES MANAGERS(Regid) ON DELETE CASCADE,
    CUSTSTATE VARCHAR2(20),
    CUSTCITY VARCHAR2(20),
    CUSTZIP NUMBER(5,0),
    CUSTSEG VARCHAR2(15) CONSTRAINT Chk_custseg CHECK (CUSTSEG IN ('Home
Office', 'Corporate', 'Small Business', 'Consumer'))
);

CREATE TABLE ORDERDET(
    ORDERID NUMBER NOT NULL REFERENCES ORDERS(ORDERID),
    CUSTID NUMBER NOT NULL REFERENCES CUSTOMERS(CUSTID),
    PRODID NUMBER NOT NULL REFERENCES PRODUCTS(PRODID),
    ORDPRIORITY VARCHAR2(15) CONSTRAINT Chk_priority CHECK (ORDPRIORITY
IN ('Low', 'Medium', 'High', 'Critical', 'Not Specified')),
    ORDDISCOUNT NUMBER(3,2),
    ORDSHIPMODE VARCHAR2(15) CONSTRAINT Chk_ship CHECK (ORDSHIPMODE IN
('Regular Air', 'Delivery Truck', 'Express Air')),
    ORDDATE DATE,
    ORDSHIPDATE DATE,
    ORDSHIPCOST NUMBER(5,2),
    ORDQTY NUMBER,
    ORDSALES NUMBER(8,2),
    PRIMARY KEY(ORDERID, CUSTID, PRODID)
);

```

QUESTION 2: ORDER Cancellations

a) What fraction of the orders was cancelled?

```

SELECT ROUND((SELECT COUNT(*) FROM ORDERS WHERE STATUS = 'Returned')/
(SELECT COUNT(*)FROM ORDERS),5) AS Prop_Returned_Orders
FROM DUAL;
Answer:0.00933

```

b) What were the sales from cancelled orders?

```

SELECT SUM(ORDSALES) AS Sales_By_Returned
FROM
ORDERS A
INNER JOIN
ORDERDET B
ON A.ORDERID = B.ORDERID and status = 'Returned';
Answer: 308,455.12

```

c) Who are the top five customers in terms of cancelled orders?

```
SELECT C.CUSTID, D.CUSTNAME, C.Count_Returned
FROM
(
  (SELECT B.CUSTID, COUNT(A.ORDERID) AS Count_Returned
   FROM
    (
      ORDERS A
      INNER JOIN
      ORDERDET B
      ON A.ORDERID = B.ORDERID AND STATUS = 'Returned'
    )
   GROUP BY CUSTID
   ORDER BY COUNT(A.ORDERID) DESC ) C
  INNER JOIN
  CUSTOMERS D
  ON C.CUSTID = D.CUSTID
)
WHERE ROWNUM < 6;
```

Answer:

	CUSTID	CUSTNAME	COUNT_RETURNED
1	1228	Hazel Jennings	7
2	1314	Keith Marsh	5
3	699	Jenny Gold	5
4	3075	Gordon Brandt	4
5	1106	Maxine Collier Grady	4

QUESTION 3: CUSTOMER related:

a) Who are the top 10 customers in terms of revenues generated?

```
SELECT C.CUSTID, D.CUSTNAME, C.TOT_REVENUE
FROM
(
  (SELECT A.CUSTID, SUM(A.ORDSALES) AS Tot_Revenue
   FROM
    (
      ORDERDET A
      INNER JOIN
      ORDERS B
      ON A.ORDERID = B.ORDERID
    )
   WHERE STATUS IS null
   GROUP BY A.CUSTID
   ORDER BY SUM(A.ORDSALES) DESC) C
  INNER JOIN
  CUSTOMERS D
  ON C.CUSTID = D.CUSTID
)
WHERE ROWNUM < 11;
```

Answer:

	CUSTID	CUSTNAME	TOT_REVENUE
1	3075	Gordon Brandt	121500.21
2	308	Glen Caldwell	83443.02
3	2571	Rosemary O'Brien	80835.09
4	553	Kristine Connolly	79860.51
5	1733	Nina Horne Kelly	77805.35
6	640	Neal Wolfe	62431.66
7	2491	Sean N Boyer	55241.63
8	1999	Priscilla Kane	54725.99
9	68	Scott Bunn	54091.64
10	2756	Ruth McConnell Young	52884.35

b) Are there customers who buy mostly some categories of products and there is a potential for them to buy other product categories?

```

SELECT C.CUSTID, D.CUSTNAME, D.CUSTSEG, COUNT(DISTINCT C.PRODCAT) AS
Categories_Purchased, SUM(C.ORDQTY) as Qty_Purchased
FROM
((SELECT *
FROM(
ORDERDET A
INNER JOIN
PRODUCTS B
ON A.PRODID = B.PRODID
)) C
INNER JOIN
CUSTOMERS D
ON C.CUSTID = D.CUSTID
)
GROUP BY C.CUSTID, D.CUSTNAME, D.CUSTSEG
HAVING SUM(C.ORDQTY) > 80 AND COUNT(DISTINCT C.PRODCAT) <>3 AND CUSTSEG IN
('Corporate', 'Small Business')
ORDER BY Categories_Purchased ASC, Qty_Purchased DESC;

```

Answer:

	CUSTID	CUSTNAME	CUSTSEG	CATEGORIES_PURCHASED	QTY_PURCHASED
1	2980	Joanna Kenney	Corporate	1	124
2	746	Lindsay York	Corporate	1	114
3	2889	Kimberly Melvin	Corporate	1	113
4	3399	Marvin Reid	Small Business	1	85
5	1959	Bonnie Matthews Rowland	Corporate	2	599
6	2498	Arlene Long	Corporate	2	505
7	607	Clara Hauser	Corporate	2	432
8	1810	Denise Albright	Corporate	2	317
9	1280	Harold Albright	Corporate	2	310
10	2618	Amy Hamrick Melvin	Corporate	2	291
11	1313	Katherine Knight	Corporate	2	252
12	898	Harriet Hodges	Small Business	2	230
13	1733	Nina Horne Kelly	Small Business	2	223

This query generated a table showing customers who have purchased from more than one category. Using 'Small Business' and 'Corporate' as examples, there does appear to be potential for other category purchases.

QUESTION 4: There are differences in the actual (theoretical) price ((unit price * number of units*(1-discount) + shipping cost) and the actual sales for all products. There are some discounts and shipping costs. Yet, there are discrepancies in the theoretical sales and actual sales.

a) How much more or less are the actual sales value compared to the theoretical sales value?

```
SELECT E.ORDERID, E.ORDDISCOUNT, E.ORDSHIPCOST, E.ORDQTY, E.ORDSALES,
E.PRODUNITPRICE, E.PRODMARGIN, F.REGION, F.REGMANAGER,
(E.PRODUNITPRICE*E.ORDQTY*(1-E.ORDDISCOUNT) + E.ORDSHIPCOST) AS
Theoretical_Price,
(E.ORDSALES - (E.PRODUNITPRICE*E.ORDQTY*(1-E.ORDDISCOUNT) +
E.ORDSHIPCOST)) AS Difference
FROM(
  (SELECT *
  FROM(
    (SELECT * FROM(
      ORDERDET A
      INNER JOIN
      PRODUCTS B
      ON A.PRODID = B.PRODID
    )) C
    INNER JOIN
    CUSTOMERS D
    ON C.CUSTID = D.CUSTID
  )
  ) E
  INNER JOIN
  MANAGERS F
  ON F.REGID = E.CUSTREG
);
```

Answer:

	ORDERID	ORDDISCOUNT	ORDSHIPCOST	ORDQTY	ORDSALES	PRODUNITPRICE	PRODMARGIN	REGION	REGMANAGER	THEORETICAL_PRICE	DIFFERENCE
1	19429	0.05	51.94	4	480.48	124.49	0.63	Central	Chris	525.002	-44.522
2	19556	0.08	4.23	65	6013.69	115.99	0.56	West	William	6940.432	-926.742
3	19556	0.07	30	83	10671.2	130.98	0.78	West	William	10140.3462	530.8538
4	19556	0.09	4.86	41	2109.51	55.98	0.36	West	William	2093.4738	16.0362
5	19812	0.01	45.7	3	1358.78	550.98	0.71	West	William	1682.1106	-323.3306
6	19841	0.08	4.2	13	2092.05	195.99	0.6	East	Erin	2348.2404	-256.1904
7	19841	0.06	8.08	86	9375.35	125.99	0.57	East	Erin	10193.1116	-817.7616
8	19841	0.03	5.81	31	388.52	11.97	0.6	East	Erin	365.7479	22.7721
9	19936	0.06	1.49	68	131.53	1.88	0.37	Central	Chris	121.6596	9.8704
10	19936	0.09	1.99	148	4025.6	29.89	0.5	Central	Chris	4027.5752	-1.9752

The difference column in the difference between the actual and theoretical sales for each Order. The overall difference is -21,791, so the actual sales is less than the theoretical.

b) Are certain managers generally pricing more or less than theoretical sales? Analyze the differences based on the regions/managers.

```

SELECT *
FROM(
    (SELECT REGMANAGER, REGION, PRODCAT, SUM(DIFFERENCE) as Overall_Gap
    FROM TEMP1
    GROUP BY REGMANAGER, REGION, PRODCAT)
    PIVOT
    (
        SUM(Overall_Gap)
        FOR PRODCAT
        IN ('Technology', 'Furniture', 'Office Supplies')
    )
);

```

Answer:

REGMANAGER	REGION	'Technology'	'Furniture'	'Office Supplies'
Sam	South	-13818.8453	838.0148	9680.6284
Chris	Central	-34990.3577	13080.471	18222.441
Erin	East	-29224.3436	7662.5335	19040.6171
William	West	-32384.3371	6194.8618	13907.0595

From the above table it appears that Technology is given at steep discounts by most managers, which should be investigated. While, furniture and office supplies are sold for profits.