Assignment 2

Due: September 26 (Submit through Canvas by 5pm)
Type: Individual (OK to ask others for simple clarification question)

WARNING: There are two parts to this assignment – Part A and Part B. The assignment is not something that be done in one sitting or the day before the deadline. You 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.

PART A – Coffee Sales

The following questions are based on the Coffee store sales data. Answer the following queries and you must use SQL to extract data and not eyeball some data to answer the questions.

A. Just for starters - SQL questions:

1. In each state, find the area codes with sales 10% higher than average sales of all area codes within that state for the year 2013 (that is, where they are 110% or more of average for the state).
   Solution – 58 rows

```
SELECT Z.areaid,Z.stateid
from
(select sum(f.actsales) act_sales,a.areaid,s.stateid
from areacode a, factcoffee f, states s
where a.areaid = f.areaid and a.stateid = s.stateid and  extract(year from f.factdate) = 2013
group by a.areaid,s.stateid
) Z
JOIN
(select stateid,avg(X.sum_sales) avg_sales from
(select a.areaid,sum(f.actsales) sum_sales
from areacode a,factcoffee f,states s
where a.areaid = f.areaid and a.stateid =s.stateid and extract(year from f.factdate) = 2013
group by a.areaid
) X,
areacode a
where X.areaid = a.areaid
group by stateid
) Y


ON Z.stateid = Y.stateid
```

WHERE Z.act_sales > (1.10 * Y.avg_sales)

	AREAID	STATEID
1	203	1003
2	206	1019
3	209	1001
4	213	1001
5	234	1014
6	239	1004
7	254	1017
8	262	1020
9	281	1017
10	310	1001

2. Find the products with profit margins as percentage of sales (profits/sales) of at least 15%. Display the results in descending order of total actual sales. Round the percentage to two digits using ROUND(….,2) function.

Solution – 12 rows

```
select pc.prodname, sum(fc.actsales) as ActualSales,
        round(sum(fc.actprofit)/sum(fc.actsales),2)
from prodcoffee pc join factcoffee fc
on pc.productid = fc.productid
having round(sum(fc.actprofit)/sum(fc.actsales),2) > 0.15
group by pc.prodname
order by ActualSales desc;
```

	PRODNAME	ACTUALSALES	ROUND(SUM(FC.ACTPROFIT)/SUM(FC.ACTSALES),2)
1	Colombian	128311	0.43
2	Lemon	95926	0.31
3	Caffe Mocha	84904	0.21
4	Decaf Espresso	78162	0.38
5	Chamomile	75578	0.36
6	Darjeeling	73151	0.4
7	Earl Grey	66772	0.36
8	Decaf Irish Cream	62248	0.22
9	Caffe Latte	35899	0.32
10	Mint	35710	0.17

3. Find AreaIDs where the total profits from leaves in 2012 are 1.2 times greater than that from beans.

Solution – 43 rows

```
select * from
(select areaid, sum(fc.actprofit) as tp1
from prodcoffee pc join factcoffee fc
on pc.productid = fc.productid
where pc.prodline='Leaves' and extract(year from fc.factdate) = '2012'
group by fc.areaid) ProfitLeaves
join
(select areaid, sum(fc.actprofit) as tp2
from prodcoffee pc join factcoffee fc
on pc.productid = fc.productid
```

```
where pc.prodline='Beans' and extract(year from fc.factdate) = '2012'
group by fc.areaid) ProfitBeans
on
ProfitLeaves.areaid = ProfitBeans.areaid
where tp1 > (1.2 * tp2);
```

	AREAID	TP1	AREAID_1	TP2
1	619	147	619	-56
2	337	295	337	233
3	740	392	740	285
4	209	421	209	295
5	330	227	330	51
6	541	785	541	511
7	712	1818	712	241
8	607	563	607	-208
9	518	889	518	-284
10	213	598	213	423

B. DECLINING PROFITS:

1. Which are the top 5 area codes with declining profits and how much did the profits decline for these 5 area codes?

Solution –

```
select P2012.areaid, (P2013.tp2-P2012.tp1) as Diff from
(select areaid, sum(fc.actprofit) as tp1
from prodcoffee pc inner join factcoffee fc
on pc.productid = fc.productid
where extract(year from fc.factdate) = '2012'
group by fc.areaid) P2012
join
(select areaid, sum(fc.actprofit) as tp2
from prodcoffee pc inner join factcoffee fc
on pc.productid = fc.productid
where extract(year from fc.factdate) = '2013'
group by fc.areaid) P2013
on P2012.areaid = P2013.areaid
order by Diff
fetch first 5 rows only
```

	AREAID	DIFF
1	845	-2082
2	508	-1131
3	626	-1064
4	712	-1039
5	631	-810

2. Among the five profit-declining area codes, are the profits consistently declining for all products? If not, identify the products for which they had significantly higher profit decline.
01-JUL-12

```
WITH prof_dec AS (
SELECT areaid, Year1, Year2,
  ROUND((Year2 - Year1) / Year2 * 100, 2)  AS "%_profit_change"
```

```
FROM (
  SELECT * FROM (
    SELECT areaid, EXTRACT(YEAR FROM factdate) AS year, sum(actprofit) AS profits
    FROM factcoffee
    GROUP BY areaid, EXTRACT(YEAR FROM factdate))
  PIVOT (
    SUM(profits)
    FOR year IN (2012 AS Year1, 2013 AS Year2)
  )
)
ORDER BY "%_profit_change" ASC
FETCH FIRST 5 ROWS ONLY
)
SELECT productid, prodname, Year1, Year2,
  ROUND((Year2 - Year1) / Year2 * 100, 2)  AS "%_profit_change"
FROM (
  SELECT * FROM (
    SELECT fc.productid, pc.prodname,
      EXTRACT(YEAR FROM factdate) AS year, SUM(actprofit) AS profits
    FROM factcoffee fc, prof_dec pd, prodcoffee pc
    WHERE fc.areaid = pd.areaid AND fc.productid = pc.productid
    GROUP BY fc.productid, pc.prodname, EXTRACT(YEAR FROM factdate)
  ) PIVOT (
    SUM(profits)
    FOR year in (2012 AS Year1, 2013 AS Year2)
  )
)
WHERE Year2 - Year1 < 0
ORDER BY "%_profit_change" ASC;
```

	PRODUCTID	PRODNAME	YEAR1	YEAR2	%_profit_change
1	8	Chamomile	207	75	-176
2	6	Decaf Espresso	320	119	-168.91
3	3	Decaf Irish Cream	96	39	-146.15
4	9	Lemon	91	53	-71.7
5	5	Caffe Mocha	297	184	-61.41
6	7	Regular Espresso	301	277	-8.66

C. BUDGETED Numbers:

1. All the budgeted numbers are expected targets for 2012 and 2013. Identify the top 5 states for the year 2012 that have substantially higher actual numbers relative to budgeted numbers for profits and sales.

   Solution – 50 rows

```
select (s.statename), sum(fc.actprofit) - sum(fc.budprofit) as d1, sum(fc.actsales) -
sum(fc.budsales) as d2
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
```

```
where extract(year from fc.factdate) = '2012'
group by s.statename
order by d1 desc, d2 desc
fetch first 5 rows only;
```

	STATENAME	D1	D2
1	Iowa	-109	3298
2	Massachusetts	-588	1157
3	Louisiana	-858	746
4	Connecticut	-867	1491
5	Florida	-871	2041

2. Identify area codes within these 5 states that beat budgeted sales and profits significantly (You need to define what significant means here).

```
WITH beat_states AS (SELECT s.stateid, s.statename,
  SUM(fc.budprofit) AS budgeted_profits, SUM(fc.actprofit) AS actual_profits,
  ROUND(SUM(fc.actprofit) / SUM(fc.budprofit) * 100, 2) AS "%_budgeted_profit_met",
  SUM(fc.budsales) AS budgeted_sales, SUM(fc.actsales) AS actual_sales,
  ROUND(sum(fc.actsales) / SUM(fc.budsales) * 100, 2) AS "%_budgeted_sales_met"
FROM states s, factcoffee fc, areacode ac
WHERE fc.areaid = ac.areaid AND s.stateid = ac.stateid
  AND EXTRACT(YEAR FROM factdate) = 2012
GROUP BY s.stateid, s.statename
HAVING ROUND(sum(fc.actsales) / SUM(fc.budsales), 2) >= 1.1 OR
  ROUND(sum(fc.actprofit) / SUM(fc.budprofit), 2) >= 1.1
)
SELECT fc.areaid,
  SUM(fc.budprofit) AS budgeted_profits, SUM(fc.actprofit) AS actual_profits,
  ROUND(SUM(fc.actprofit) / SUM(fc.budprofit) * 100, 2) AS "%_budgeted_profit_met",
  SUM(fc.budsales) AS budgeted_sales, SUM(fc.actsales) AS actual_sales,
  ROUND(sum(fc.actsales) / SUM(fc.budsales) * 100, 2) AS "%_budgeted_sales_met"
FROM beat_states bs, factcoffee fc, areacode ac
WHERE bs.stateid = ac.stateid AND fc.areaid = ac.areaid
  AND EXTRACT(YEAR FROM factdate) = 2012
GROUP BY fc.areaid
HAVING ROUND(sum(fc.actsales) / SUM(fc.budsales), 2) >= 1.1
  OR ROUND(sum(fc.actprofit) / SUM(fc.budprofit), 2) >= 1;
```

	AREAID	BUDGETED_PROFITS	ACTUAL_PROFITS	%_budgeted_profit_met	BUDGETED_SALES	ACTUAL_SALES	%_budgeted_sales_met
1	407	350	297	84.86	900	1027	114.11
2	475	1170	939	80.26	3070	3474	113.16
3	712	2100	2059	98.05	5160	5831	113
4	541	1570	1296	82.55	4650	5317	114.34
5	253	1230	945	76.83	3760	4230	112.5
6	360	730	587	80.41	2160	2562	118.61
7	561	420	402	95.71	1070	1257	117.48
8	702	2340	1915	81.84	11430	14970	130.97
9	959	750	612	81.6	1990	2313	116.23
10	239	310	199	64.19	930	1042	112.04

D. PRODUCT related:

1. In each market, which products have the greatest increase in profits?

<u>Solution</u> –

Market	Product	Diff
Central	Caffe Mocha	2704
Central	Chamomile	2677
East	Colombian	4998
East	Regular Expresso	1881
South	Colombian	1615
South	Decaf Expresso	1088
West	Lemon	2413
West	Decaf Expresso	2276

```
select m13.m, m13.pn, max(m13.p13 - m12.p12) as diff from
(select s.statemkt as m, (sum(fc.actprofit)) as p13, pc.prodname as pn
from states s join areacode a
on s.stateid = a.stateid
join factcoffee fc
on a.areaid = fc.areaid
join prodcoffee pc
on pc.productid = fc.productid
where extract(year from fc.factdate) = '2013'
group by s.statemkt, pc.prodname
order by s.statemkt, pc.prodname, (sum(fc.actprofit)) desc) m13
join
(select s.statemkt as m, (sum(fc.actprofit)) as p12, pc.prodname as pn
from states s join areacode a
on s.stateid = a.stateid
join factcoffee fc
on a.areaid = fc.areaid
join prodcoffee pc
on pc.productid = fc.productid
where extract(year from fc.factdate) = '2012'
group by s.statemkt, pc.prodname
order by s.statemkt, pc.prodname, (sum(fc.actprofit)) desc) m12
on m13.pn = m12.pn and m13.m = m12.m
group by m13.m, m13.pn
order by m13.m, diff desc;
```

2. In each market, which **product types** have greatest increase in sales?

Solution –

Market	Product Type
Central	Tea
Central	Coffee
East	Coffee
East	Expresso
South	Expresso
South	Coffee

West	Tea
West	Expresso

```
select m13.m, m13.pn, max(m13.p13 - m12.p12) as diff from
(select s.statemkt as m, (sum(fc.actsales)) as p13, pc.prodtype as pn
from states s join areacode a
on s.stateid = a.stateid
join factcoffee fc
on a.areaid = fc.areaid
join prodcoffee pc
on pc.productid = fc.productid
where extract(year from fc.factdate) = '2013'
group by s.statemkt, pc.prodtype
order by s.statemkt, pc.prodtype, (sum(fc.actsales)) desc) m13
join
(select s.statemkt as m, (sum(fc.actsales)) as p12, pc.prodtype as pn
from states s join areacode a
on s.stateid = a.stateid
join factcoffee fc
on a.areaid = fc.areaid
join prodcoffee pc
on pc.productid = fc.productid
where extract(year from fc.factdate) = '2012'
group by s.statemkt, pc.prodtype
order by s.statemkt, pc.prodname, (sum(fc.actsales)) desc) m12
on m13.pn = m12.pn and m13.m = m12.m
group by m13.m, m13.pn
order by m13.m, diff desc;
```

3. Have all products within the product types show similar behavior, or some products within a product type have greatest increase in sales?

```
WITH prods AS (
  SELECT prodtype
  FROM (
    SELECT statemkt, prodtype, Year1, Year2,
      (Year2 - Year1) / Year2 * 100 AS "%_change_in_sales",
      ROW_NUMBER() OVER (PARTITION BY statemkt ORDER BY statemkt) AS rank
    FROM (
      SELECT statemkt, prodtype, EXTRACT(YEAR FROM factdate) AS year,
        SUM(actsales) AS sum_actsales
      FROM prodcoffee pc, factcoffee fc, states s, areacode ac
      WHERE pc.productid = fc.productid AND s.stateid = ac.stateid AND fc.areaid =
ac.areaid
      GROUP BY statemkt, prodtype, EXTRACT(YEAR FROM factdate)
    ) PIVOT (
      SUM(sum_actsales)
      FOR YEAR IN (2012 AS Year1, 2013 AS Year2)
```

```
    )
    ORDER BY statemkt ASC, "%_change_in_sales" DESC
  )
  WHERE rank = 1
)
SELECT productid, prodname, prodtype,
  (Year2 - Year1) / Year2 * 100 AS "%_change_in_sales"
FROM (
  SELECT * FROM (
    SELECT pc.productid, prodname, pc.prodtype, EXTRACT(YEAR FROM factdate) AS Year,
      SUM(actsales) AS actsales
    FROM prodcoffee pc, prods p, factcoffee fc
    WHERE pc.prodtype = p.prodtype AND fc.productid = pc.productid
    GROUP BY pc.productid, prodname, pc.prodtype, EXTRACT(YEAR FROM factdate)
  ) PIVOT (
    SUM(actsales)
    FOR YEAR IN (2012 AS Year1, 2013 AS Year2)
  )
)
ORDER BY prodtype, "%_change_in_sales" DESC;
```

	PRODUCTID	PRODNAME	PRODTYPE	%_change_in_sales
1	7	Regular Espresso	Espresso	4.32375213744809054637244524061558505008
2	4	Caffe Latte	Espresso	4.25845147219193020719738276990185387132
3	5	Caffe Mocha	Espresso	4.21979846427006710171328429450965019485
4	6	Decaf Espresso	Espresso	4.21321577075296828816191573568458494063
5	8	Chamomile	Herbal Tea	4.24771426351368851822114014866999922298
6	10	Mint	Herbal Tea	4.11409764125068568294020844761382336807
7	9	Lemon	Herbal Tea	4.11272207473963651215029609965284868287

E. MARKETING EXPENSES (LOWEST):

1. Which top 5 states have the lowest market expenses as a percentage of their sales?

Solution

Massachusetts
Texas
Illinois
Iowa
Colarado

```
select (s.statename), sum(fc.actmarkcost)/sum(fc.actsales) as d
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
group by s.statename
order by d
fetch first 5 rows only;
```

2. Do the above 5 states also have the highest profits as a percentage of sales?

Solution - Yes

```
select (s.statename), sum(fc.actprofit)/sum(fc.actsales) as d
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
group by s.statename
order by d desc
fetch first 5 rows only;
```

3. Are there any particular product(s) within these markets with the least marketing expenses?
   Solution

```
select pc.prodname, sum(fc.actmarkcost) as s
from prodcoffee pc
join
factcoffee fc
on pc.productid = fc.productid
join
areacode a
on a.areaid = fc.areaid
join states s
on s.stateid = a.stateid
where s.statename in (select (s.statename)
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
group by s.statename
order by sum(fc.actprofit)/sum(fc.actsales) desc
fetch first 5 rows only)
group by pc.prodname
order by s;
```

	PRODNAME	S
1	Caffe Latte	504
2	Green Tea	586
3	Amaretto	842
4	Regular Espresso	1052
5	Mint	1102
6	Decaf Irish Cream	2290
7	Decaf Espresso	2652
8	Darjeeling	2796
9	Earl Grey	3304
10	Lemon	3342

F. MARKETING EXPENSES (highest):
   1. Which 5 states have the highest marketing expenses as a percentage of sales?
      Are these marketing expenses justified? (Note: you need to think how you will justify high marketing expenses)?

      Solution

	STATENAME	D
1	Nevada	0.20040226732492228926677637593710001828 49
2	Wisconsin	0.19813118025945749795881339018416039190 78
3	New Mexico	0.19141706519003272086584445003775484520 51
4	Washington	0.18607757513485743642435139994862573850 5
5	New York	0.18410207192457517077852424772765765257 16

```
select (s.statename), sum(fc.actmarkcost)/sum(fc.actsales) as d
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
group by s.statename
order by d desc
fetch first 5 rows only;
```

2. In each of these 5 states, do any area codes spend too much on marketing expenses relative to others?

Solution

1	775	6273	4346
2	702	5783	6270
3	505	3042	799
4	206	2741	3823
5	262	1962	2599
6	715	1935	2582
7	845	1355	-36
8	360	1262	1829
9	253	1258	2040
10	518	1169	2421

```
select fc.areaid, sum(fc.actmarkcost) as d, sum(fc.actprofit) as e
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
where s.statename in(
select (s.statename)
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
group by s.statename
order by sum(fc.actmarkcost)/sum(fc.actsales) desc
fetch first 5 rows only)
group by fc.areaid
order by d desc;
```

G. STRATEGY:

1. You are in a high-level strategy meeting to discuss how to improve performance. This may involve shutting down stores in losing area codes and/or expanding in very profitable/high growth area. Evaluate the data and recommend which stores to close and where?

Solution

```
select fc.areaid, sum(fc.actmarkcost) as d, sum(fc.actprofit) as e
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
where s.statename in(
select (s.statename)
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
group by s.statename
order by sum(fc.actmarkcost)/sum(fc.actsales) desc
fetch first 5 rows only)
group by fc.areaid
order by d desc;
```

	AREAID	D	E
1	775	6273	4346
2	702	5783	6270
3	505	3042	799
4	206	2741	3823
5	262	1962	2599
6	715	1935	2582
7	845	1355	-36
8	360	1262	1829
9	253	1258	2040
10	518	1169	2421
11	718	1119	2346
12	914	1093	-154
13	516	1068	2972
14	509	996	1696
15	425	987	2017
16	608	959	943
17	716	953	1721
18	585	932	1269
19	920	861	1310
20	646	848	1847
21	347	846	634
22	212	844	609
23	414	835	1268
24	631	824	2308
25	315	742	1263
26	607	691	797
27	917	560	2099

The stores to shut down is one in areaid is 845. It has a high marketing cost and low profit.

2. Where should the firm focus on expanding?

```
select fc.areaid, sum(fc.actmarkcost) as d, sum(fc.actprofit) as e
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
where s.statename in(
select (s.statename)
from factcoffee fc join areacode a
on fc.areaid = a.areaid
join states s
on s.stateid = a.stateid
group by s.statename
```

order by sum(fc.actmarkcost)/sum(fc.actsales) desc
fetch first 5 rows only)
group by fc.areaid
order by d desc;
The stores to expand is one in areaid is 917. It has a low marketing cost and high profit.

PART B: Office Product

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,
region varchar2(10),
regmanager varchar2(10),
constraint ch_region check (region in ('East','South','Central','West')))

create table PRODUCTS (
prodid number primary key,
prodname varchar2(100),
prodcat varchar2(30) CONSTRAINT ch_prodcat check (prodcat in ('Technology','Furniture','Office
Supplies')),
prodsubcat varchar2(30),
prodcont varchar2(20)CONSTRAINT ch_prodcont 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,
status varchar2(10))

create table customers(
custid number primary key,
custname varchar2(35),
custreg number(1,0),
custstate varchar2(20),
custcity varchar2(20),
custzip number(5,0),
custseg varchar2(15) CONSTRAINT ch_custseg check ( custseg in ('Corporate','Small
Business','Consumer','Home Office')))

create table orderdet(
orderid number references orders(orderid),
custid number references customers(custid),
prodid number references products(prodid),
ordpriority varchar2(15),
orddiscount number(3,2),
ordshipmode varchar2(15),
orddate date,
ordshipdate date,
ordshipcost number(5,2),
ordqty number,
ordsales number(10,2),
constraint ch_ordpriority check (ordpriority in ('Low','Medium','High','Critical','Not Specified')),
constraint ch_ordshipmode check (ordshipmode in('Regular Air','Delivery Truck','Express Air')))
```

QUESTION 2: ORDER Cancellations

a) What fraction of the orders was cancelled?
Solution

```
select a.returned/b.total from (select count(status) as returned from orders where status =
'Returned') a, (select count(orderid) as total from orders) b;
```

0.0093

b) What were the sales from cancelled orders?
Solution

```
select sum(ordsales) from orderdet o where o.orderid in (select orderid from orders where
status = 'Returned');
```

30855.12

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

Solution

```
select c.custid,sum(od.ordsales)
from orderdet od,orders o ,customers c
where od.orderid = o.orderid
and od.custid = c.custid
and o.status = 'Returned'
group by c.custid
order by sum(od.ordsales) desc
fetch first 5 rows only
```

	CUSTID	SUM(OD.ORDSALES)
1	2107	56387.55
2	349	40511.19
3	2670	28779.13
4	2867	24362.25
5	2403	21450.02

QUESTION 3: CUSTOMER related:

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

Solution

1	3075	121500.21
2	308	83443.02
3	2571	80835.09
4	553	79860.51
5	1733	77805.35
6	640	62431.66
7	2491	55241.63
8	1999	54725.99
9	68	54091.64
10	2756	52884.35

```
select c.custid, sum(o.ordsales) as total
from customers c join orderdet o
on c.custid = o.custid
join orders ord
on ord.orderid = o.orderid
where ord.status is null
group by c.custid
order by total desc
fetch first 10 rows only;
```

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

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?

Solution

```
select X.prodid,X.prodname,sum(X.ordsales) totordersales,sum(X.theorysales)
totaltheorysales,(sum(X.theorysales) - sum(X.ordsales)) diff_sale_values
FROM
(
  select orderdet.*, products.prodname,((products.produnitprice * orderdet.ordqty * (1-
orderdet.orddiscount))+orderdet.ordshipcost ) as theorysales
  from orderdet,products
  where orderdet.prodid = products.prodid
) X
GROUP BY X.prodid,X.prodname
ORDER BY X.prodid;
```

	PRODID	PRODNAME	TOTORDERSALES	TOTALTHEORYSALES	DIFF_SALE_VALUES
1	1	While you Were Out Message Book, One Form per Page	259.31	250.5411	-8.7689
2	2	#10- 4 1/8" x 9 1/2" Recycled Envelopes	1879.01	1887.8912	8.8812
3	3	#10- 4 1/8" x 9 1/2" Security-Tint Envelopes	897.31	857.6468	-39.6632
4	4	#10 Self-Seal White Envelopes	825.61	813.7509	-11.8591
5	5	#10 White Business Envelopes,4 1/8 x 9 1/2	2846.84	2724.2587	-122.5813
6	6	#10-4 1/8" x 9 1/2" Premium Diagonal Seam Envelopes	2502.8	2312.6494	-190.1506
7	7	#6 3/4 Gummed Flap White Envelopes	620.23	593.826	-26.404
8	8	*Staples* Highlighting Markers	567.81	546.438	-21.372
9	9	*Staples* Letter Opener	245.77	239.4248	-6.3452
10	10	*Staples* Packaging Labels	11.71	12.1354	0.4254

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

Solution

```
                SELECT X.custreg
    ,x.REGMANAGER
    ,SUM(X.ordsales) totordsales
    ,SUM(X.theorysales) tottheorysales
    ,(SUM(X.theorysales) -SUM(X.ordsales) ) diff_theory_minus_order
FROM
(
select orderdet.orderid
    ,orderdet.custid
    ,customers.custname
    ,customers.custreg
    ,managers.REGMANAGER
    ,orderdet.prodid
    ,((products.produnitprice*orderdet.ordqty*(1-
orderdet.orddiscount))+orderdet.ordshipcost ) theorysales
    ,orderdet.ordsales
from orderdet
  ,products
  ,customers
  ,managers
where orderdet.prodid = products.prodid
and orderdet.custid = customers.custid
and customers.custreg = managers.regid
ORDER BY customers.custreg
```

```
) X
GROUP BY X.custreg,X.REGMANAGER
ORDER BY diff_theory_minus_order
;
```

	CUSTREG	REGMANAGER	TOTORDSALES	TOTTHEORYSALES	DIFF_THEORY_MINUS_ORDER
1	2	Erin	2366343.86	2368865.053	2521.193
2	3	Sam	1568582.35	1571882.5521	3300.2021
3	1	Chris	2493762.8	2497450.2457	3687.4457
4	4	William	2360868.33	2373150.7458	12282.4158

QUESTION 5: these are product related questions:

a) Products have numbers within its name. Identify the product names with digits in their name. (hint: use REGEXP_LIKE)

```
select distinct(prodname) from products where regexp_like(prodname, '[0-9]');
```

1	Accessory8
2	Acco 6 Outlet Guardian Premium Surge Suppressor
3	Acco Pressboard Covers with Storage Hooks, 14 7/8" x 11", Dark Blue
4	Acco PRESSTEX® Data Binder with Storage Hooks, Dark Blue, 14 7/8" X 11"
5	Adesso Programmable 142-Key Keyboard
6	AT&T Black Trimline Phone, Model 210
7	Avery 05222 Permanent Self-Adhesive File Folder Labels for Typewriters, on Rolls, White, 250/Roll
8	Avery 491
9	Canon imageCLASS 2200 Advanced Copier
10	CF 688

b) Which are the top 5 selling products during the year 2011?

```
select p.prodname, count(o.prodid) as ct
from products p join orderdet o
on p.prodid = o.prodid
join orders ord
on ord.orderid = o.orderid
where extract(year from o.orddate) = '2011' and ord.status is null
group by p.prodname
order by ct desc;
```

1	O'Sullivan 3-Shelf Heavy-Duty Bookcases	10
2	US Robotics 56K V.92 External Faxmodem	9
3	Fellowes Basic 104-Key Keyboard, Platinum	9
4	Bevis 36 x 72 Conference Tables	9
5	Belkin ErgoBoard™ Keyboard	8
6	Belkin Premiere Surge Master II 8-outlet surge protector	8
7	Xerox 1920	8
8	BoxOffice By Design Rectangular and Half-Moon Meeting Room Tables	7
9	Office Impressions Heavy Duty Welded Shelving & Multimedia Storage Drawers	7
10	Canon PC940 Copier	7

c) Which are the top 10 products with greatest total profit margin? (i.e., sales*margin).

```
select p.prodname, sum(o.ordsales * p.prodmargin) as tp
from products p join orderdet o
on p.prodid = o.prodid
join orders ord
on ord.orderid = o.orderid
where ord.status is null
group by p.prodname
order by tp desc;
```

1	Riverside Palais Royal Lawyers Bookcase, Royale Cherry Finish	117920.993
2	Global Troy™ Executive Leather Low-Back Tilter	116415.384
3	Bretford CR4500 Series Slim Rectangular Table	54207.7652
4	Canon PC1080F Personal Copier	51466.385
5	BoxOffice By Design Rectangular and Half-Moon Meeting Room Tables	48811.1858
6	Bretford CR8500 Series Meeting Room Furniture	47996.17
7	Lexmark 4227 Plus Dot Matrix Printer	45328.3435
8	Hon Multipurpose Stacking Arm Chairs	44422.7107
9	Adesso Programmable 142-Key Keyboard	42275.5311
10	Canon imageCLASS 2200 Advanced Copier	39848.1601

d) Identify the worst five products in terms of sales?

```
select p.prodname, sum(o.ordsales) as ct
from products p join orderdet o
on p.prodid = o.prodid
join orders ord
on ord.orderid = o.orderid
where ord.status is null
group by p.prodname
order by ct;
```

1	Alliance Rubber Bands	7.43
2	*Staples* Packaging Labels	11.71
3	Blackstonian Pencils	13.18
4	Avery 482	16.67
5	Sony IBM Color Diskettes, 25/Pack	18.17
6	Colorific® Watercolor Pencils	20.39
7	Avery 516	35.01
8	Computer Printout Index Tabs	36.16
9	DAX Charcoal/Nickel-Tone Document Frame, 5 x 7	39.63
10	Accohide Poly Flexible Ring Binders	40.6