Mobile Manufacturer Data Analysis

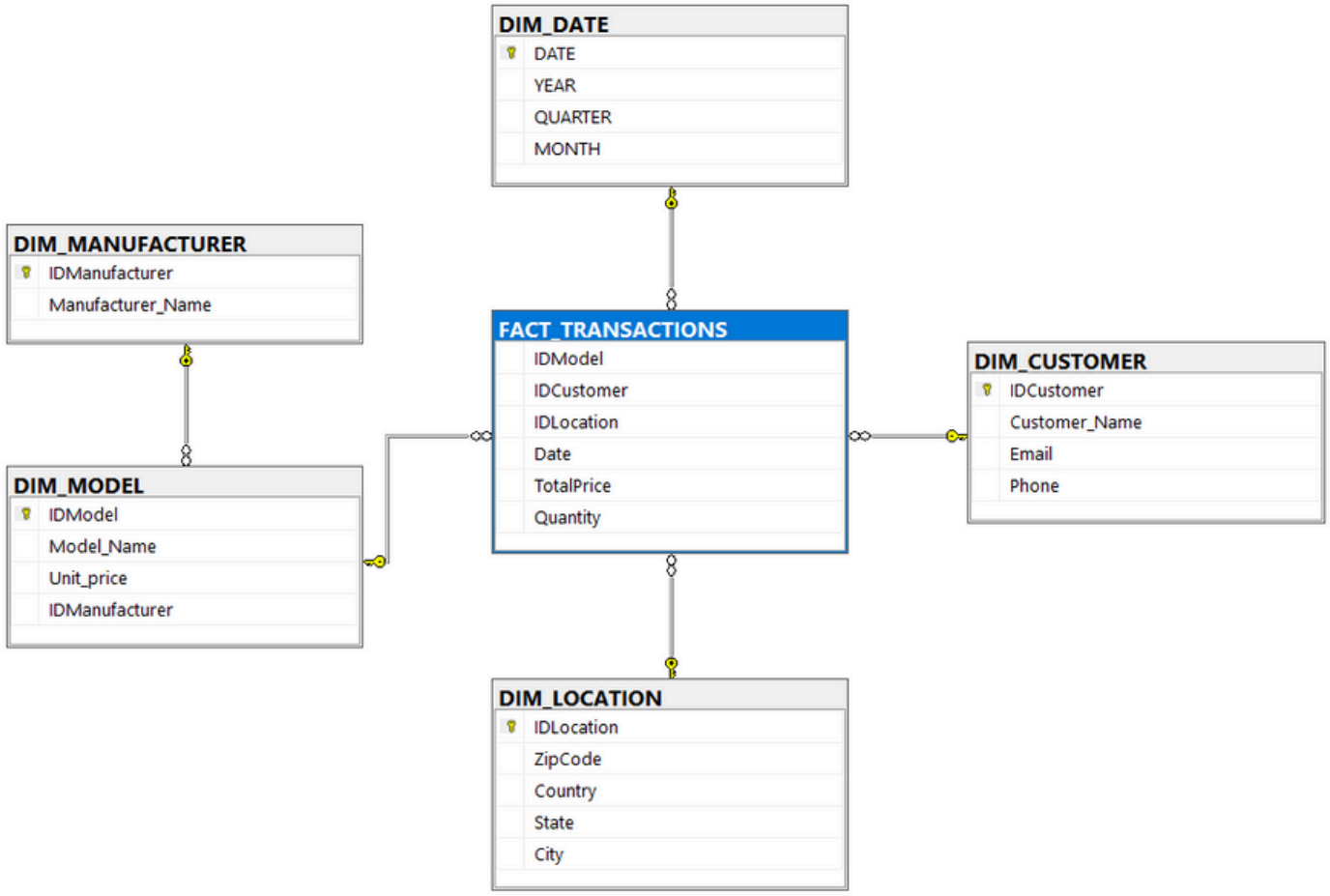
SQL SERVER, SSMS

DATA AVAILABILITY

The database “db_mobilemanufacturer” contains details on cell phone sales or transactions. Detailes stored are: Dim_manufacturer, Dim_model, Dim_customer, Dim_Location and Fact_Transactions.

The first four store entries for the respective elements and Fact_Transactions stores all the information about sales of specific cellphones.

Assume that you do not have access to the data. Hence, pls create a schema based on the representation below to work on the case study



Skills

SQL querying, data modeling, data cleaning, sales trend analysis, business insights generation.

Key Analysis

- Identifying states with active customers since 2005
- Top-performing states for Samsung device sales
- Cheapest cellphone models
- Year-wise performance of top customers and manufacturers
- Comparative manufacturer sales across 2009–2010

1. List all the states in which we have customers who have bought cellphones from 2005 till today.

```
SELECT DISTINCT L.State
FROM FACT_TRANSACTIONS AS F
INNER JOIN DIM_LOCATION AS L ON F.IDLocation = L.IDLocation
WHERE YEAR(F.Date) > 2004
```

	State
1	Arizona
2	California
3	Delhi
4	Haryana
5	Karnataka
6	Maharashtra
7	Maryland

2. Which state in the US is buying the most Samsung cellphones?

```
SELECT TOP 1 L.State, COUNT(*) AS TOTAL_PURCHASES
FROM FACT_TRANSACTIONS AS F
INNER JOIN DIM_LOCATION AS L ON F.IDLocation = L.IDLocation
INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
INNER JOIN DIM_MANUFACTURER AS MF ON M.IDManufacturer = MF.IDManufacturer
WHERE L.Country = 'US' AND MF.Manufacturer_Name = 'Samsung'
GROUP BY L.State
ORDER BY TOTAL_PURCHASES DESC
```

	State	TOTAL_PURCHASES
1	Arizona	18

3. Show the number of transactions for each model per zip code per state.

```
SELECT M.Model_Name, L.ZipCode, L.State, COUNT(*) AS TOTAL_TRANSACTIONS
FROM FACT_TRANSACTIONS AS F
INNER JOIN DIM_LOCATION AS L ON F.IDLocation = L.IDLocation
INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
GROUP BY M.Model_Name, L.ZipCode, L.State
```

	Model_Name	ZipCode	State	TOTAL_TRANSACTIONS
1	3210	21648	Maryland	2
2	3210	85086	Arizona	1
3	3210	85117	Arizona	2
4	3210	94005	California	1
5	3210	110004	Delhi	1
6	3210	122002	Haryana	1
7	3210	400006	Maharashtra	2
8	3210	530068	Karnataka	2
9	3310 (3330)	21163	Maryland	1
10	3310 (3330)	21648	Maryland	1
11	3310 (3330)	85086	Arizona	2
12	3310 (3330)	85117	Arizona	3

4. Show the cheapest cellphone.

```
SELECT TOP 1 M.IDModel ,M.Model_Name, M.Unit_price
FROM DIM_MODEL  AS M
ORDER BY M.Unit_price ASC
```

	IDModel	Model_Name	Unit_price
1	112	3210	14.00

5. Find out the average price for each model in the top 5 manufacturers in terms of sales quantity and order by average price

```
WITH TOP_5_MANUFACTURERS
AS(
  SELECT TOP 5 MF.IDManufacturer, SUM(F.Quantity) AS SALES_QTY, AVG(F.TotalPrice)
AS AVG_PRICE
  FROM FACT_TRANSACTIONS AS F
  INNER JOIN DIM_MODEL AS M
  ON F.IDModel = M.IDModel
  INNER JOIN DIM_MANUFACTURER AS MF
  ON M.IDManufacturer = MF.IDManufacturer
  GROUP BY MF.IDManufacturer
  ORDER BY SALES_QTY DESC,AVG_PRICE ASC
)
SELECT MF.Manufacturer_Name, M.Model_Name, AVG(F.TotalPrice) AS AVG_PRICE
FROM FACT_TRANSACTIONS AS F
INNER JOIN DIM_MODEL AS M
ON F.IDModel = M.IDModel
INNER JOIN DIM_MANUFACTURER AS MF ON M.IDManufacturer = MF.IDManufacturer
WHERE M.IDManufacturer IN (SELECT IDManufacturer FROM TOP_5_MANUFACTURERS)
GROUP BY MF.Manufacturer_Name, M.Model_Name
```

	Manufacturer_Name	Model_Name	AVG_PRICE
1	Nokia	3210	18.50
2	Nokia	3310 (3330)	31.4545
3	Nokia	5230	34.0909
4	Nokia	6010 (6020/6030)	58.50
5	Nokia	6230 (6233)	55.75
6	Nokia	6600	60.00
7	Motorola	C139	185.10
8	Motorola	C200	151.0833
9	Motorola	Droid Bionic	175.4444
10	Samsung	E1100	178.1818
11	Samsung	E250	170.6666
12	Samsung	Galaxy Note II	219.20
13	Samsung	Galaxy S	198.2222
14	Samsung	Galaxy S4	343.60
15	Samsung	Galaxy S5	394.00
16	Samsung	Galaxy S7	478.00
17	Samsung	Galaxy S8	665.00
18	Apple	iPhone 4	597.9523
19	Apple	iPhone 4S	535.1428
20	Apple	iPhone 5	524.7142
21	Apple	iPhone 6	503.5882
22	Apple	iPhone 7	633.7142
23	Motorola	Motorola Z	285.909
24	One Plus	OnePlus 2	192.6666
25	One Plus	OnePlus 5	320.00
26	One Plus	OnePlus 5T	407.1818
27	One Plus	OnePlus 6T	588.8181
28	One Plus	OnePlus X	184.2307
29	Motorola	RAZR V3	91.3636

6. List the names of the customers and the average amount spent in 2009, where the average is higher than 500.

```
SELECT C.Customer_Name, AVG(F.TotalPrice) AS AVERAGE_AMOUNT
FROM FACT_TRANSACTIONS AS F
INNER JOIN DIM_CUSTOMER AS C
ON F.IDCustomer = C.IDCustomer
WHERE YEAR(F.Date) = 2009
GROUP BY C.Customer_Name
HAVING AVG(F.TotalPrice) > 500
```

	Customer_Name	AVERAGE_AMOUNT
1	Celeste Korando	613.00
2	Danica Bruschke	760.00
3	Laurel Reitler	1528.00
4	Lettie Isenhower	870.00
5	Moon Parlato	823.50
6	Shawna Palaspas	569.00

7. List if there is any model that was in the top 5 in terms of quantity, simultaneously in 2008, 2009 and 2019

```
SELECT T.Model_Name
FROM(   SELECT TOP 5 M.Model_Name,SUM(F.Quantity) AS TOTAL_QTY
  FROM FACT_TRANSACTIONS AS F
  INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
  INNER JOIN DIM_DATE AS D ON F.Date = D.DATE
  WHERE D.YEAR = 2008
  GROUP BY M.Model_Name
  ORDER BY TOTAL_QTY DESC
 UNION ALL
  SELECT TOP 5 M.Model_Name,SUM(F.Quantity) AS TOTAL_QTY
  FROM FACT_TRANSACTIONS AS F
  INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
  INNER JOIN DIM_DATE AS D ON F.Date = D.DATE
  WHERE D.YEAR = 2009
  GROUP BY M.Model_Name
  ORDER BY TOTAL_QTY DESC
 UNION ALL
  SELECT TOP 5 M.Model_Name,SUM(F.Quantity) AS TOTAL_QTY
  FROM FACT_TRANSACTIONS AS F
  INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
  INNER JOIN DIM_DATE AS D ON F.Date = D.DATE
  WHERE D.YEAR = 2010
  GROUP BY M.Model_Name
  ORDER BY TOTAL_QTY DESC
) AS T
GROUP BY T.Model_Name
HAVING COUNT(*) = 3
```

	Model_Name
1	Droid Bionic

8. Show the manufacturers with the 2nd top sales in the year 2009 and the manufacturer with the second top sales in year 2010

```
SELECT T.Manufacturer_Name,T.YEAR_,T.TOTAL_SALES
FROM ( SELECT MF.Manufacturer_Name,YEAR(F.Date) AS YEAR_ ,SUM(F.TotalPrice) AS TOTAL_SALES
  FROM FACT_TRANSACTIONS AS F
  INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
  INNER JOIN DIM_MANUFACTURER AS MF ON M.IDManufacturer = MF.IDManufacturer
  WHERE YEAR(F.Date) = 2009
  GROUP BY MF.Manufacturer_Name,YEAR(F.Date)
  ORDER BY TOTAL_SALES DESC
  OFFSET 1 ROWS
  FETCH NEXT 1 ROWS ONLY
 UNION ALL
  SELECT MF.Manufacturer_Name, YEAR(F.Date) AS YEAR_,SUM(F.TotalPrice) AS TOTAL_SALES
  FROM FACT_TRANSACTIONS AS F
  INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
  INNER JOIN DIM_MANUFACTURER AS MF ON M.IDManufacturer = MF.IDManufacturer
  WHERE YEAR(F.Date) = 2010
  GROUP BY MF.Manufacturer_Name,YEAR(F.Date)
  ORDER BY TOTAL_SALES DESC
  OFFSET 1 ROWS
  FETCH NEXT 1 ROWS ONLY
) AS T
```

	Manufacturer_Name	YEAR_	TOTAL_SALES
1	Samsung	2009	3357.00
2	Apple	2010	3242.00

9. Show the manufacturers who sold cellphones in 2010 but did not in 2009

```
SELECT MF.IDManufacturer, MF.Manufacturer_Name
FROM FACT_TRANSACTIONS AS F
INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
INNER JOIN DIM_MANUFACTURER AS MF ON M.IDManufacturer = MF.IDManufacturer
WHERE YEAR(F.Date) = 2010
EXCEPT
SELECT MF.IDManufacturer, MF.Manufacturer_Name
FROM FACT_TRANSACTIONS AS F
INNER JOIN DIM_MODEL AS M ON F.IDModel = M.IDModel
INNER JOIN DIM_MANUFACTURER AS MF ON M.IDManufacturer = MF.IDManufacturer
WHERE YEAR(F.Date) = 2009
```

	IDManufacturer	Manufacturer_Name
1	16	HTC

10. Find top 10 customers and their average spend, average quantity by each year. Also find the percentage of change in their spend.

```
WITH TOP10CUSTOMERS
AS(
 SELECT TOP 10 F.IDCustomer
 FROM FACT_TRANSACTIONS AS F
 GROUP BY F.IDCustomer
 ORDER BY SUM(F.TotalPrice) DESC
),
CUSTOMER_YEARLY_SPEND
AS(
 SELECT F.IDCustomer,
 YEAR(F.Date) AS YEAR_,
 AVG(F.TotalPrice) AS AVG_SPEND,
 AVG(F.Quantity) AS AVG_QTY
 FROM FACT_TRANSACTIONS AS F
 INNER JOIN DIM_CUSTOMER AS C ON F.IDCustomer = C.IDCustomer
 WHERE F.IDCustomer IN (SELECT IDCustomer FROM TOP10CUSTOMERS)
 GROUP BY F.IDCustomer, YEAR(F.Date)
)
SELECT *,
((C.AVG_SPEND - LAG(C.AVG_SPEND,1) OVER(PARTITION BY C.IDCUSTOMER ORDER BY C.YEAR_))/
LAG(C.AVG_SPEND,1) OVER(PARTITION BY C.IDCUSTOMER ORDER BY C.YEAR_))*100 AS
PERCENTAGE_CHANGE
FROM CUSTOMER_YEARLY_SPEND AS C
```

	IDCustomer	YEAR_	AVG_SPEND	AVG_QTY	PERCENTAGE_CHANGE
1	10003	2005	319.00	1	NULL
2	10003	2006	474.00	1	48.58
3	10003	2007	1106.00	2	133.33
4	10003	2008	322.00	1	-70.88
5	10003	2009	442.00	1	37.26
6	10003	2010	435.00	1	-1.58
7	10006	2003	667.00	1	NULL
8	10006	2004	256.50	1	-61.54
9	10006	2009	823.50	2	221.05
10	10006	2010	226.50	1	-72.49
11	10007	2003	410.00	1	NULL
12	10007	2005	149.00	1	-63.65
13	10007	2006	226.00	1	51.67
14	10007	2007	288.00	1	27.43
15	10007	2008	557.00	1	93.40
16	10007	2009	1528.00	4	174.32
17	10007	2010	35.00	1	-97.70
18	10014	2004	56.00	1	NULL
19	10014	2005	338.00	2	503.57
20	10014	2006	406.00	2	20.11
21	10014	2007	920.00	2	126.60
22	10014	2008	255.00	1	-72.28
23	10022	2003	284.00	1	NULL
24	10022	2004	503.00	1	77.11
25	10022	2005	202.00	1	-59.84
26	10022	2010	390.00	1	93.06
27	10030	2004	665.00	1	NULL
28	10030	2006	415.00	1	-37.59
29	10030	2008	347.00	1	-16.38
30	10030	2010	476.00	1	37.17
31	10036	2003	283.00	1	NULL
32	10036	2004	461.00	1	62.89
33	10036	2005	31.00	1	-93.27
34	10036	2006	401.6666	1	1195.69
35	10036	2009	613.00	1	52.61
36	10036	2010	396.00	1	-35.39