U2M1.LW.Core SQL

Mastykina Elizaveta

2.1. Task 01: CREATE Storage Objects

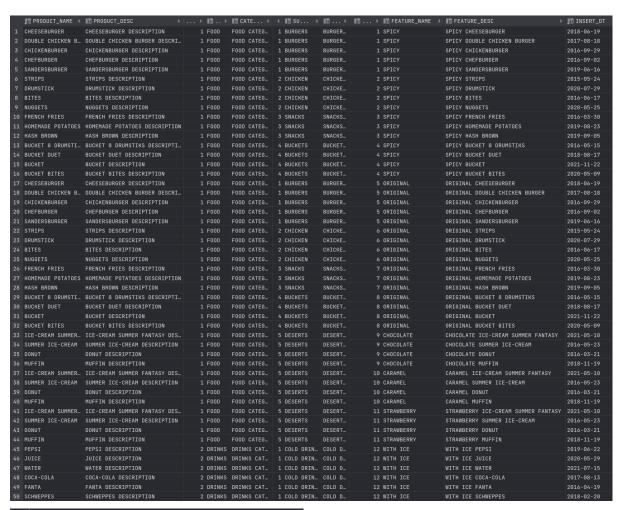
```
/* Table: SA_COUPONS_DATA
CREATE TABLE SA_COUPONS_DATA (
                                 VARCHAR2(40),
    COUPON_NAME
   COUPON_DESC
                                 VARCHAR2(40),
   MEDIA_TYPE_ID
                                 NUMBER(*,0),
                                 VARCHAR2(40),
                                 NUMBER(4,2),
    DISCOUNT_PERCENTAGE_AMOUNT
   VALID_FROM
                                 DATE,
                                 DATE,
    INSERT_DT
                                  DATE
TABLESPACE TS_SA_COUPONS_DATA_01;
```

2.2. Task 02: Generate Test Data in Storage Layers

Data generation for SA PRODUCTS DATA:

Data about products, categories, subcategories was generated by python utility (code is stored on github) in csv-format.

To add all possible combinations of products and features I create CTE statement and LEFT .IOIN





Data generation for SA_RESTAURANTS_DATA:

	■ PHONE	‡	₽ POSTAL_CODE ÷	.⊞ ADDF	RESS	\$. INSERT_DT
1	+97214054721		4620057438	STREET	CTL 96		2021-11-23
2	+47248504833		7127801055	STREET	PFDWILFLC 39		2020-03-21
3	+93016486919		3016490712	STREET	AS 59		2020-04-24
4	+22626061500		1467470147	STREET	ZFRDIQFNNOFWZ 79		2019-04-05
5	+56323871181		5607676048	STREET	SVUTL 5		2017-12-29
6	+88612743997		7311768084	STREET	PAWJXHSO 79		2015-08-02
7	+84859694689		9723132053	STREET	QZJNNTPHKLO 89		2018-07-18
8	+93327981822		4545017060	STREET	JQY 66		2015-08-10
9	+90387225703		4667783098	STREET	XLUWLMGXRFIH 20		2019-05-25
10	+75417893376		3657247962	STREET	BMN 79		2020-06-02

Data generation for SA_COUPONS_DATA:

```
INSERT INTO SA_COUPONS_DATA(COUPON_NAME,DISCOUNT_PERCENTAGE_AMOUNT,VALID_FROM,INSERT_DT, MEDIA_TYPE_ID, MEDIA_TYPE)
CTE_COUPON_NAME AS(
    SELECT 'COUPON ' || TO_CHAR(TRUNC(dbms_random.value(1, 10)*10000)) AS NAME
           ,TRUNC(dbms_random.value(1, 20)) AS DISCONT
           ,TO_CHAR(TRUNC(dbms_random.value(1, 5))) AS MEDIA_ID
                       DBMS_RANDOM.VALUE(TO_CHAR(DATE '2015-01-01', 'J')
           )AS INSERTION_DT
                       DBMS_RANDOM.VALUE(TO_CHAR(DATE '2015-01-01', 'J')
           )AS VALID_FROM_DT
    FROM DUAL
    CONNECT BY LEVEL <= 1000
CTE_MEDIA_TYPE AS (
    SELECT 2 AS ID, 'OUTDOOR' AS TYPE FROM DUAL UNION ALL
    SELECT 4 AS ID, 'PRINT' AS TYPE FROM DUAL
SELECT NAME, DISCONT, VALID_FROM_DT, INSERTION_DT, MEDIA_ID,TYPE FROM CTE_COUPON_NAME
    INNER JOIN CTE_MEDIA_TYPE ON MEDIA_ID=CTE_MEDIA_TYPE.ID;
UPDATE SA_COUPONS_DATA
UPDATE SA_COUPONS_DATA
SELECT COUNT(*) FROM SA_COUPONS_DATA
```

III COUPON ▲ :	L III COUPON_DESC ÷	■ MEDIA_TYPE_ID : ■ MEDIA_TYPE :	■ DISCOUNT_PERCENTAGE_AMOUNT ÷ ■ VALID_FROM	÷ ■■ VALID_TO ÷	: ■ INSERT_DT
COUPON 10039	COUPON 10039 OUTDOOR 13.05.20 - 13.08.20	2 OUTDOOR	4.00 2020-05-13	2020-08-13	2021-06-17
COUPON 10075	COUPON 10075 MOBILE 04.02.16 - 04.05.16	3 MOBILE	3.00 2016-02-04	2016-05-04	2017-12-08
COUPON 10169	COUPON 10169 ONLINE 07.04.20 - 07.07.20	1 ONLINE	3.00 2020-04-07	2020-07-07	2016-01-22
COUPON 10308	COUPON 10308 PRINT 13.04.21 - 13.07.21	4 PRINT	16.00 2021-04-13	2021-07-13	2020-06-17
COUPON 10421	COUPON 10421 MOBILE 12.03.15 - 12.06.15	3 MOBILE	16.00 2015-03-12	2015-06-12	2019-08-09
COUPON 10427	COUPON 10427 ONLINE 21.09.21 - 21.12.21	1 ONLINE	8.00 2021-09-21	2021-12-21	2021-10-07
COUPON 10430	COUPON 10430 OUTDOOR 19.01.15 - 19.04.15	2 OUTDOOR	19.00 2015-01-19	2015-04-19	2020-09-02
COUPON 10776	COUPON 10776 ONLINE 25.04.18 - 25.07.18	1 ONLINE	11.00 2018-04-25	2018-07-25	2020-05-05
COUPON 10839	COUPON 10839 OUTDOOR 02.03.21 - 02.06.21	2 OUTDOOR	17.00 2021-03-02	2021-06-02	2020-11-25
COUPON 10923	COUPON 10923 MOBILE 15.12.19 - 15.03.20	3 MOBILE	11.00 2019-12-15	2020-03-15	2020-01-11
COUPON 10955	COUPON 10955 PRINT 03.07.15 - 03.10.15	4 PRINT	6.00 2015-07-03	2015-10-03	2021-03-08
COUPON 10986	COUPON 10986 ONLINE 16.04.19 - 16.07.19	1 ONLINE	1.00 2019-04-16	2019-07-16	2017-05-18
COUPON 11174	COUPON 11174 OUTDOOR 01.03.17 - 01.06.17	2 OUTDOOR	9.00 2017-03-01	2017-06-01	2020-06-22
COUPON 11340	COUPON 11340 OUTDOOR 12.07.19 - 12.10.19	2 OUTDOOR	4.00 2019-07-12	2019-10-12	2020-05-26
COUPON 11426	COUPON 11426 PRINT 02.04.15 - 02.07.15	4 PRINT	15.00 2015-04-02	2015-07-02	2015-05-27
COUPON 11456	COUPON 11456 OUTDOOR 01.02.15 - 01.05.15	2 OUTDOOR	2.00 2015-02-01	2015-05-01	2016-03-15
COUPON 11471	COUPON 11471 OUTDOOR 07.06.18 - 07.09.18	2 OUTDOOR	17.00 2018-06-07	2018-09-07	2018-02-08
COUPON 11524	COUPON 11524 OUTDOOR 26.09.16 - 26.12.16	2 OUTDOOR	9.00 2016-09-26	2016-12-26	2018-05-05
COUPON 11559	COUPON 11559 PRINT 03.11.21 - 03.02.22	4 PRINT	2.00 2021-11-03	2022-02-03	2020-08-01
COUPON 11615	COUPON 11615 MOBILE 12.11.17 - 12.02.18	3 MOBILE	9.00 2017-11-12	2018-02-12	2021-03-25

Data generation for SA_GEO_DATA:

Was loaded from the source csv file.

	■ COUNTRY_NAME ÷	■ REGION_NAME \$	■ CITY_NAME ÷	I INSER
1	Argentina	Tucumán	San Miguel de Tucumán	2021-12-30
2	Belgium	Flanders	Nijlen	2021-12-30
3	Australia	New South Wales	Cessnock	2021-12-30
4	Belgium	Wallonia	Visé	2021-12-28
5	Czech Republic	Ústecký	Litoměřice	2021-12-28
6	China	Shandong Sheng	Gaomi	2021-12-26
7	Chile	Maule	Molina	2021-12-25
8	Belgium	Wallonia	Saint-Ghislain	2021-12-25
9	Australia	New South Wales	Caringbah	2021-12-25
10	China	Hubei	Huangmei	2021-12-25
11	China	Yunnan	Haikou	2021-12-24
12	Chile	Aisén	Puerto Aisén	2021-12-23
13	Switzerland	Vaud	Nyon	2021-12-21
14	Colombia	Antioquia	Chigorodó	2021-12-21
15	Germany	North Rhine-Westphalia	Wipperfürth	2021-12-21
16	China	Hunan	Leiyang	2021-12-20
17	Australia	Victoria	Langwarrin	2021-12-18
18	Australia	New South Wales	Wollongong	2021-12-17
19	Belgium	Flanders	Lommel	2021-12-16
20	China	Heilongjiang Sheng	Qinggang	2021-12-15

Data generation for SA_PERIODS_DATA:

```
/* Table: SA_PERIODS_DATA
INSERT INTO SA_PERIODS_DATA (PERIOD_NAME, BEGIN_DT, INSERT_DT)
SELECT 'PERIOD ' || TO_CHAR(ROWNUM),
      TO_DATE(
             TRUNC(
                     DBMS_RANDOM. VALUE (TO_CHAR (DATE '2015-01-01', 'J')
                        , TO_CHAR(DATE '2022-01-01', 'J')
                 ), 'J'
          ),
      TO_DATE(
             TRUNC(
                     DBMS_RANDOM. VALUE (TO_CHAR (DATE '2015-01-01', 'J')
                        , TO_CHAR(DATE '2022-01-01', 'J')
FROM DUAL
CONNECT BY LEVEL <= 100;
UPDATE SA_PERIODS_DATA
SET END_DT = ADD_MONTHS(BEGIN_DT, 12),
   PERIOD_DESC = PERIOD_NAME || ' DESC';
COMMIT;
```

	■ PERIOD_NAME \$	■ PERIOD_DESC ÷	■ BEGIN_DT ÷	■ END_DT ÷	I INSERT_DT
1	PERIOD 1	PERIOD 1 DESC	2021-01-12	2022-01-12	2015-08-18
2	PERIOD 2	PERIOD 2 DESC	2019-02-06	2020-02-06	2016-02-24
3	PERIOD 3	PERIOD 3 DESC	2017-01-27	2018-01-27	2017-05-29
4	PERIOD 4	PERIOD 4 DESC	2021-10-02	2022-10-02	2021-08-20
5	PERIOD 5	PERIOD 5 DESC	2018-12-07	2019-12-07	2019-04-13
6	PERIOD 6	PERIOD 6 DESC	2021-02-08	2022-02-08	2020-09-12
7	PERIOD 7	PERIOD 7 DESC	2020-12-27	2021-12-27	2015-06-14
8	PERIOD 8	PERIOD 8 DESC	2021-08-18	2022-08-18	2017-11-10
9	PERIOD 9	PERIOD 9 DESC	2018-10-10	2019-10-10	2017-05-28
10	PERIOD 10	PERIOD 10 DESC	2017-11-23	2018-11-23	2016-06-16

2.3. Task 03: Create Group by Plan

Transactions table structure:

```
CREATE TABLE TRANSACTIONS(
    PRODUCT_NAME
                      VARCHAR2(40 BYTE) ,
                      VARCHAR2(40 BYTE),
    PRODUCT_DESC
    CATEGORY_ID
                      NUMBER(*, 0),
                      VARCHAR2(40 BYTE),
    CATEGORY_NAME
                      VARCHAR2(40 BYTE),
    CATEGORY_DESC
                      NUMBER(*, 0),
    SUBCATEGORY_ID
   SUBCATEGORY_NAME
                      VARCHAR2(40 BYTE),
   SUBCATEGORY_DESC
                      VARCHAR2(40 BYTE),
                      NUMBER(*, 0),
   FEATURE_ID
                      VARCHAR2(40 BYTE),
    FEATURE_NAME
    FEATURE_DESC
                      VARCHAR2(40 BYTE),
    COST_DOLLAR_AMOUNT NUMBER(5,3),
    PHONE
                      VARCHAR2(40 BYTE),
    POSTAL_CODE
                      NUMBER(10,0) ,
                      VARCHAR2(40 BYTE),
    ADDRESS
    COUPON_NAME
                                VARCHAR2(40),
                                VARCHAR2(40),
    COUPON_DESC
    MEDIA_TYPE_ID
                                NUMBER(*,0) ,
                                VARCHAR2(40),
    MEDIA_TYPE
    DISCOUNT_PERCENTAGE_AMOUNT
                                NUMBER(4,2),
   VALID_FROM
                                DATE ,
   VALID_TO
                                DATE ,
    COUNTRY_NAME
                        CHAR(40),
                        CHAR(40),
    REGION_NAME
                        CHAR(40),
    CITY_NAME
    PERIOD_NAME VARCHAR2(20),
    PERIOD_DESC VARCHAR2(40),
   BEGIN_DT DATE DEFAULT TO_DATE('01/01/1990', 'DD/MM/YYYY'),
              DATE DEFAULT TO_DATE('31/12/9999', 'DD/MM/YYYY'),
   END_DT
    INSERT_DT DATE
```

Population of the table with data:

```
CTE_PRODUCTS_COUPONS_RESTAURANTS AS (
       TRUNC(DBMS_RANDOM.VALUE(1,3648)) AS GEO_ID,
        TRUNC(DBMS_RANDOM.VALUE(1,100)) AS PERIOD_ID
FROM SA_PRODUCTS_DATA P CROSS JOIN SA_COUPONS_DATA C, SA_RESTAURANTS_DATA R ORDER BY dbms_random.value
FETCH FIRST 300000 ROWS ONLY
CTE_GEO AS (
    SELECT COUNTRY_NAME, REGION_NAME, CITY_NAME, ROWNUM AS GEO_ID FROM SA_GEO_DATA
CTE_PERIODS AS (
    SELECT PERIOD_NAME, PERIOD_DESC, BEGIN_DT, END_DT, ROWNUM AS PERIOD_ID FROM SA_PERIODS_DATA
   FROM CTE_PRODUCTS_COUPONS_RESTAURANTS
   INNER JOIN CTE_GEO ON CTE_GEO.GEO_ID = CTE_PRODUCTS_COUPONS_RESTAURANTS.GEO_ID
    INNER JOIN CTE_PERIODS ON CTE_PERIODS.PERIOD_ID = CTE_PRODUCTS_COUPONS_RESTAURANTS.PERIOD_ID;
COMMIT;
```

Examples of using group by plan:

PRODUCT_NAME, COUNTRY_NAME, COUNT(*) AS QUANTITY, SUM(COST_DOLLAR_AMOUNT) AS TOATAL_SALE_DOLLAR_AMOUNT FROM TRANSACTIONS GROUP BY PRODUCT_NAME, COUNTRY_NAME ORDER BY TOATAL_SALE_DOLLAR_AMOUNT DESC;						
.	hutput Result 9 ×			csv ×		
`		■ COUNTRY_NAME \$	■ QUANTITY ≎	■ TOATAL_SALE_DOLLAR_AMOUNT ≎		
1	NUGGETS	China	2071	23257.33		
2	STRIPS	China	2069	23234.87		
3	DRUMSTICK	China	2059	23122.57		
4	CHICKENBURGER	China	2137	22652.2		
5	CHEESEBURGER	China	2122	22493.2		
6	BITES	China	1998	22437.54		
7	SANDERSBURGER	China	2069	21931.4		
8	CHEFBURGER	China	2067	21910.2		
9	DOUBLE CHICKEN BURGER	China	2042	21645.2		
10	NUGGETS	Germany	1925	21617.75		
11	DRUMSTICK	Germany	1923	21595.29		
12	BUCKET 8 DRUMSTIKS	China	2083	21392.41		
13	BUCKET	China	2076	21320.52		
14	BUCKET BITES	China	2071	21269.17		
15	DOUBLE CHICKEN BURGER	Germany	2000	21200		
16	STRIPS	Germany	1884	21157.32		
17	BUCKET DUET	China	2023	20776.21		
18	BITES	Germany	1846	20730.58		
19	CHEFBURGER	Germany	1914	20288.4		
20	BUCKET	Germany	1959	20118.93		
21	SANDERSBURGER	Germany	1888	20012.8		
22	CHICKENBURGER	Germany	1885	19981		
23	CHEESEBURGER	Germany	1843	19535.8		
24	BUCKET 8 DRUMSTIKS	Germany	1883	19338.41		
25	BUCKET DUET	Germany	1872	19225.44		
26	BUCKET BITES	Germany	1863	19133.01		
27	HASH BROWN	China	2078	12260.2		
28	HOMEMADE POTATOES	China	2039	12030.1		
29	FRENCH FRIES	China	2024	11941.6		
30	FRENCH FRIES	Germany	1892	11162.8		