Anton Slizh's

U2M11.LW.ETL Overview - Advanced Refresh Scenarios

Task 1

2.1. Task 01: Loading to SAL Layer Data

Task 1 is the same as Task 2 at Lab 10. Check previous lab.

Task 2

3.1. Task 02: Prepare Report Layout

The Main Task is to create Ad Hoc SQL for Report Layout Monthly that was developed on LabWork 2 (Use STAR schema objects for source of Data).

Task Results:

Create required objects:

- Put objects script to Git.
- Prepare Document with Screenshot of Data

First report represents the monthly sales and revenue for each store

```
-- Monthly sales and revenue for each country and store
SELECT TRUNC(sl.date_id, 'MM') AS month,
       DECODE(GROUPING(st.country), 1, 'All countries', st.country) AS country,
       DECODE(GROUPING(st.address), 1, 'All stores', st.address) AS store_address,
       COUNT(sl.amount) AS amount,
       SUM(sl.sum) AS revenue
FROM sal_data.fct_sales sl
LEFT JOIN sal_data.dim_stores st
ON sl.store_id = st.store_id
GROUP BY TRUNC(sl.date_id, 'MM'), GROUPING SETS(
    (st.country, st.address),
    (st.country),
    (TRUNC(sl.date_id, 'MM')))
ORDER BY month, st.country, SUM(sl.sum) DESC;
```

	MONTH	COUNTRY	STORE_ADDRESS	AMOUNT	REVENUE
1	01.01.21	Belarus	All stores	12347	52586,63
2	01.01.21	Belarus	Minks, Byady, 3	1074	4620,42
3	01.01.21	Belarus	Brest, Kalinouskaga, 47	1086	4574,79
4	01.01.21	Belarus	Mogilev, Kalinouskaga, 17	1061	4512,27
5	01.01.21	Belarus	Minks, Zaporozhskaya, 73	1030	4419,19
6	01.01.21	Belarus	Gomel, Kalinouskaga, 6	1033	4404,31
7	01.01.21	Belarus	Minks, Kalinouskaga, 33	1022	4397,41
8	01.01.21	Belarus	Grodno, Savetskaya, 2	1019	4344,22
9	01.01.21	Belarus	Novogrudok, Kalinouskaga, 1	1019	4338,98
10	01.01.21	Belarus	Grodno, Kirova, 13	998	4317,5
11	01.01.21	Belarus	Grodno, Kalinouskaga, 7	1014	4304,8
12	01.01.21	Belarus	Vitebsk, Kalinouskaga, 9	1013	4204,48
13	01.01.21	Belarus	Slonim, Skaryny, 18	978	4148,26
14	01.01.21	Kazakhstan	Astana, Kalinouskaga, 90	953	3937,57
15	01.01.21	Kazakhstan	All stores	953	3937,57
16	01.01.21	Russian Federation	All stores	2013	8648,34
17	01.01.21	Russian Federation	Moscow, Kalinouskaga, 12	1000	4373,55
18	01.01.21	Russian Federation	Saint Petersburg, Kalinouskaga, 90	1013	4274,79
19	01.01.21	Ukraine	Kyiv, Kalinouskaga, 11	1002	4284,77
20	01.01.21	Ukraine	All stores	1002	4284,77
21	01.01.21	All countries	All stores	16315	69457,31
22	01.02.21	Belarus	All stores	10889	46471,57
23	01.02.21	Belarus	Grodno, Kalinouskaga, 7	960	4115,58
24	01.02.21	Belarus	Vitebsk, Kalinouskaga, 9	943	4052,89
25	01.02.21	Belarus	Novogrudok, Kalinouskaga, 1	919	3951,99
26	01.02.21	Belarus	Minks, Byady, 3	918	3943,77

Second report represents the most popular product brands

```
17  -- Monthly most popular product brands
18  SELECT TRUNC(s.date_id, 'MM') AS month,
19         DECODE(GROUPING(p.brand), 1, 'All brands', p.brand) AS product_brand,
20         COUNT(s.amount) AS amount,
21         SUM(s.sum) AS revenue
22  FROM sal_data.fct_sales s
23  LEFT JOIN sal_data.dim_products_scd p
24  ON s.product_id = p.product_id
25  GROUP BY TRUNC(s.date_id, 'MM'), ROLLUP(p.brand)
26  ORDER BY month, SUM(s.sum) DESC;
27
```

	MONTH	PRODUCT_BRAND	AMOUNT	REVENUE
1	01.01.21	All brands	16315	69457,31
2	01.01.21	Lidski	4925	18889,4
3	01.01.21	Hatni	4036	16529,67
4	01.01.21	Zaporozhski	2435	15644,7
5	01.01.21	Alivarski	2481	10597,74
6	01.01.21	Ruski	2438	7795,8
7	01.02.21	All brands	14558	61932,21
8	01.02.21	Lidski	4386	16669,35
9	01.02.21	Hatni	3573	14472,75
10	01.02.21	Zaporozhski	2197	14239
11	01.02.21	Alivarski	2224	9551,51
12	01.02.21	Ruski	2178	6999,6
13	01.03.21	All brands	16248	69554,6
14	01.03.21	Lidski	4888	18588,25
15	01.03.21	Hatni	4026	16488,96
16	01.03.21	Zaporozhski	2369	15567,4
17	01.03.21	Alivarski	2449	10848,09
18	01.03.21	Ruski	2516	8061,9
19	01.04.21	All brands	15514	66361,39
20	01.04.21	Lidski	4674	17816,2
21	01.04.21	Hatni	3869	15770,07
22	01.04.21	Zaporozhski	2313	15261
23	01.04.21	Alivarski	2280	10016,02
24	01.04.21	Ruski	2378	7498,1
25	01.05.21	All brands	16292	69154,92

The third report represents the most popular products

```
28  -- Monthly most popular products
29  SELECT TRUNC(s.date_id, 'MM') AS month,
30         DECODE(GROUPING(p.description), 1, 'All products', p.description) AS product_name,
31         COUNT(s.amount) AS amount,
32         SUM(s.sum) AS revenue
33  FROM sal_data.fct_sales s
34  LEFT JOIN sal_data.dim_products_scd p
35  ON s.product_id = p.product_id
36  GROUP BY GROUPING SETS(
37      (TRUNC(s.date_id, 'MM'), p.description),
38      (TRUNC(s.date_id, 'MM')))
39  ORDER BY month, SUM(s.sum) DESC;
```

	MONTH	PRODUCT_NAME	AMOUNT	REVENUE
1	01.01.21	All products	16315	69457,31
2	01.01.21	Zaporozhski Light 2 liter	209	2085
3	01.01.21	Zaporozhski Dark 2 liter	211	1899
4	01.01.21	Zaporozhski Classic 2 liter	200	1684
5	01.01.21	Hatni Dark 2 liter	234	1395
6	01.01.21	Zaporozhski Dark 1.5 liter	202	1358
7	01.01.21	Hatni Classic 2 liter	215	1308
8	01.01.21	Zaporozhski Light 1 liter	215	1269
9	01.01.21	Alivarski Dark 2 liter	216	1263
10	01.01.21	Alivarski Light 2 liter	221	1235,4
11	01.01.21	Zaporozhski Classic 1.5 liter	210	1209
12	01.01.21	Lidski Light 2 liter	196	1148
13	01.01.21	Zaporozhski Dark 0.5 liter	198	1145,5
14	01.01.21	Zaporozhski Light 1.5 liter	198	1110
15	01.01.21	Zaporozhski Light 0.5 liter	182	1097,4
16	01.01.21	Lidski Orange 2 liter	208	1082,5
17	01.01.21	Lidski Dark 2 liter	200	1080,8
18	01.01.21	Hatni Light 2 liter	191	1061,4
19	01.01.21	Alivarski Classic 2 liter	178	1056
20	01.01.21	Hatni Cranberry 2 liter	207	1050
21	01.01.21	Lidski Classic 2 liter	209	1045
22	01.01.21	Hatni Orange 2 liter	202	1012,5
23	01.01.21	Lidski Cranberry 2 liter	190	1010

Task 3

3.2. Task 03: Compare Report Layout Performance

The Main Task is to create summarize table with comparison Performance of next Report Layout:

- Advancing Grouping (GROUP BY GROUPING SETs – LabWork 02)
- Model Clause (LabWork 05)
- Star Schema (LabWork 11)

1. Advancing Grouping (0.157 s)

```
SELECT TRUNC(date_id, 'MM') AS month,
       DECODE(GROUPING(product_name), 1, 'All products', product_name) AS product_name,
       SUM(amount) AS amount,
       SUM(amount * price) AS revenue
FROM sa_customers.sa_sale_data s
JOIN sa_products.sa_product_data p
ON s.sku_num = p.sku_num
GROUP BY GROUPING SETS(
    (TRUNC(date_id, 'MM'), product_name),
    (TRUNC(date_id, 'MM')))
ORDER BY month, SUM(price) DESC;
```

```
-------------------------------------------------------------------------------------------------
| Id  | Operation             | Name            | Rows  | Bytes |TempSpc| Cost (%CPU)| Time     |
-------------------------------------------------------------------------------------------------
|   0 | SELECT STATEMENT      |                 | 32471 |  2124K|       |  3178   (1)| 00:00:01 |
|   1 |  SORT ORDER BY        |                 | 32471 |  2124K|    22M|  3178   (1)| 00:00:01 |
|   2 |   HASH GROUP BY ROLLUP|                 | 32471 |  2124K|    22M|  3178   (1)| 00:00:01 |
|*  3 |    HASH JOIN          |                 |   300K|    19M|       |   768   (1)| 00:00:01 |
|   4 |     TABLE ACCESS FULL | SA_PRODUCT_DATA |    80 |  1440 |       |     4   (0)| 00:00:01 |
|   5 |     TABLE ACCESS FULL | SA_SALE_DATA    |   300K|    14M|       |   764   (1)| 00:00:01 |
-------------------------------------------------------------------------------------------------
```

SQL | Fetched 50 rows in 0,157 seconds

	MONTH	PRODUCT_NAME	AMOUNT	REVENUE
1	01.01.21	All products	32599	69457,31
2	01.01.21	Zaporozhski Light 2 liter	417	2085
3	01.01.21	Zaporozhski Dark 2 liter	422	1899
4	01.01.21	Zaporozhski Classic 2 liter	421	1684
5	01.01.21	Zaporozhski Dark 1.5 liter	388	1358
6	01.01.21	Hatni Dark 2 liter	465	1395
7	01.01.21	Alivarski Dark 2 liter	421	1263
8	01.01.21	Zaporozhski Light 1 liter	423	1269
9	01.01.21	Hatni Classic 2 liter	436	1308
10	01.01.21	Alivarski Light 2 liter	426	1235,4
11	01.01.21	Zaporozhski Classic 1.5 liter	403	1209
12	01.01.21	Zaporozhski Light 1.5 liter	370	1110
13	01.01.21	Zaporozhski Dark 0.5 liter	395	1145,5
14	01.01.21	Zaporozhski Light 0.5 liter	354	1097,4
15	01.01.21	Lidski Dark 2 liter	386	1080,8
16	01.01.21	Hatni Light 2 liter	366	1061,4
17	01.01.21	Lidski Light 2 liter	410	1148
18	01.01.21	Alivarski Classic 2 liter	352	1056

SQL History

SQL	Connection	TimeStamp	Type	Executed	Duration(seconds)
SELECT TRUNC(date_id, 'MM') AS month, DECODE(GROUPING(prod...	HomeConne...	21.08.22 16...	SQL	8	0.157

2. Model Clause (0.143 s)

```
--------------------------
WITH sales_by_month
AS
(
    SELECT TRUNC(date_id, 'MM') AS month,
           product_name AS product,
           SUM(amount * price) AS revenue,
           sum(amount) AS amount
            FROM sa_customers.sa_sale_data s
           JOIN sa_products.sa_product_data p
        ON s.sku_num = p.sku_num
        GROUP BY TRUNC(date_id, 'MM'), product_name
)

SELECT DISTINCT month, product, amount, revenue
FROM sales_by_month

MODEL
    PARTITION BY (month)
    DIMENSION BY (product)
    MEASURES (revenue, amount)
    RULES
        (
            revenue['All products'] = SUM(revenue)[any],
            amount['All products'] = SUM(amount)[any]
        )
ORDER BY month, revenue DESC;
```

```
-------------------------------------------------------------------------------------------------------
| Id  | Operation                | Name             | Rows  | Bytes |TempSpc| Cost (%CPU)| Time     |
-------------------------------------------------------------------------------------------------------
|   0 | SELECT STATEMENT         |                  | 32471 |  2124K|       |  3178   (1)| 00:00:01 |
|   1 |  SORT ORDER BY           |                  | 32471 |  2124K|    22M|  3178   (1)| 00:00:01 |
|   2 |   SQL MODEL ORDERED FAST |                  | 32471 |  2124K|       |  3178   (1)| 00:00:01 |
|   3 |    HASH GROUP BY         |                  | 32471 |  2124K|    22M|  3178   (1)| 00:00:01 |
|*  4 |     HASH JOIN            |                  |   300K|    19M|       |   768   (1)| 00:00:01 |
|   5 |      TABLE ACCESS FULL   | SA_PRODUCT_DATA  |    80 |  1440 |       |     4   (0)| 00:00:01 |
|   6 |      TABLE ACCESS FULL   | SA_SALE_DATA     |   300K|    14M|       |   764   (1)| 00:00:01 |
-------------------------------------------------------------------------------------------------------
```

SQL | Fetched 50 rows in 0,143 seconds

	MONTH	PRODUCT	AMOUNT	REVENUE
1	01.01.21	All products	32599	69457,31
2	01.01.21	Zaporozhski Light 2 liter	417	2085
3	01.01.21	Zaporozhski Dark 2 liter	422	1899
4	01.01.21	Zaporozhski Classic 2 liter	421	1684
5	01.01.21	Hatni Dark 2 liter	465	1395
6	01.01.21	Zaporozhski Dark 1.5 liter	388	1358
7	01.01.21	Hatni Classic 2 liter	436	1308
8	01.01.21	Zaporozhski Light 1 liter	423	1269
9	01.01.21	Alivarski Dark 2 liter	421	1263
10	01.01.21	Alivarski Light 2 liter	426	1235,4
11	01.01.21	Zaporozhski Classic 1.5 liter	403	1209
12	01.01.21	Lidski Light 2 liter	410	1148
13	01.01.21	Zaporozhski Dark 0.5 liter	395	1145,5
14	01.01.21	Zaporozhski Light 1.5 liter	370	1110
15	01.01.21	Zaporozhski Light 0.5 liter	354	1097,4
16	01.01.21	Lidski Orange 2 liter	433	1082,5
17	01.01.21	Lidski Dark 2 liter	386	1080,8
18	01.01.21	Hatni Light 2 liter	366	1061,4
19	01.01.21	Alivarski Classic 2 liter	352	1056

SQL History

SQL	Connection	TimeStamp	Type	Executed	Duration(seconds)
WITH sales_by_monthAS(SELECT TRUNC(date_id, 'MM') AS month, ...	HomeConne...	21.08.22 16...	SQL	2	0.143

3. Using Star Scheme (0.110 s)

```
SELECT TRUNC(s.date_id, 'MM') AS month,
       DECODE(GROUPING(p.description), 1, 'All products', p.description) AS product_name,
       SUM(s.amount) AS amount,
       SUM(s.sum) AS revenue
FROM sal_data.fct_sales s
JOIN sal_data.dim_products_scd p
ON s.product_id = p.product_id
GROUP BY GROUPING SETS(
    (TRUNC(s.date_id, 'MM'), p.description),
    (TRUNC(s.date_id, 'MM')))
ORDER BY month, revenue DESC;
```

```
-------------------------------------------------------------------------------------------------
| Id  | Operation             | Name             | Rows  | Bytes |TempSpc| Cost (%CPU)| Time     |
-------------------------------------------------------------------------------------------------
|   0 | SELECT STATEMENT      |                  | 32471 |  1617K|       |  2304   (1)| 00:00:01 |
|   1 |  SORT ORDER BY        |                  | 32471 |  1617K|    18M|  2304   (1)| 00:00:01 |
|   2 |   HASH GROUP BY ROLLUP|                  | 32471 |  1617K|    18M|  2304   (1)| 00:00:01 |
|*  3 |    HASH JOIN          |                  |   299K|    14M|       |   380   (1)| 00:00:01 |
|   4 |     TABLE ACCESS FULL | DIM_PRODUCTS_SCD |    80 |  2400 |       |     4   (0)| 00:00:01 |
|   5 |     TABLE ACCESS FULL | FCT_SALES        |   299K|  6152K|       |   375   (1)| 00:00:01 |
-------------------------------------------------------------------------------------------------
```

SQL | Fetched 50 rows in 0,11 seconds

	MONTH	PRODUCT_NAME	AMOUNT	REVENUE
1	01.01.21	All products	32599	69457,31
2	01.01.21	Zaporozhski Light 2 liter	417	2085
3	01.01.21	Zaporozhski Dark 2 liter	422	1899
4	01.01.21	Zaporozhski Classic 2 liter	421	1684
5	01.01.21	Hatni Dark 2 liter	465	1395
6	01.01.21	Zaporozhski Dark 1.5 liter	388	1358
7	01.01.21	Hatni Classic 2 liter	436	1308
8	01.01.21	Zaporozhski Light 1 liter	423	1269
9	01.01.21	Alivarski Dark 2 liter	421	1263
10	01.01.21	Alivarski Light 2 liter	426	1235,4
11	01.01.21	Zaporozhski Classic 1.5 liter	403	1209
12	01.01.21	Lidski Light 2 liter	410	1148
13	01.01.21	Zaporozhski Dark 0.5 liter	395	1145,5
14	01.01.21	Zaporozhski Light 1.5 liter	370	1110
15	01.01.21	Zaporozhski Light 0.5 liter	354	1097,4
16	01.01.21	Lidski Orange 2 liter	433	1082,5
17	01.01.21	Lidski Dark 2 liter	386	1080,8
18	01.01.21	Hatni Light 2 liter	366	1061,4
19	01.01.21	Alivarski Classic 2 liter	352	1056

SQL History

SQL	Connection	TimeStamp	Type	Executed	Duration(seconds)
SELECT TRUNC(s.date_id, 'MM') AS month, DECODE(GROUPING(p.d...	HomeConne...	21.08.22 16...	SQL	2	0.11

№	Source Type	Explain Plan - Statistics	Time, Sec.
1	Advancing Grouping	Cost: 3178	0.157
2	Model Clause	Cost: 3178	0.143
3	Star Schema	Cost: 2304	0.110

In my case, using the Star Scheme objects to prepare business reports is more useful than selecting data on source data tables. The benefit of selecting on Star Scheme object is slight decrease of time and cost needed for query execution.