Data Warehouse Project in SQL

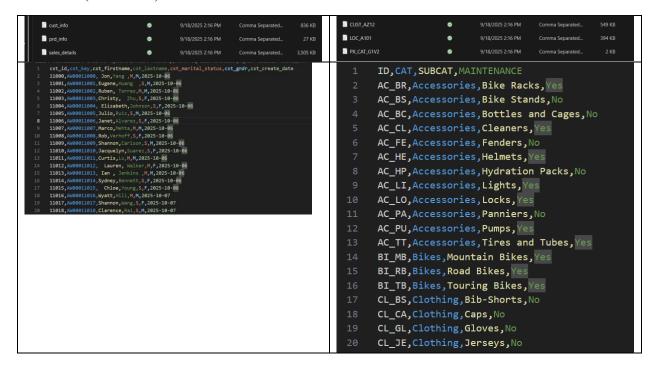
Learning Outcomes:

- Analyzing Data
- SQL Basics & Advanced (Week 1 & 2)
- Extraction Process (Data Ingestion)
- Data Architecture
- Transformation Process (Data Cleaning)
- Dimensions / Facts
- Data Modeling
- Schemas (Snow & Star Flake)
- Load Process (Data warehouse Build)

Tools & Technologies

- SQL Server
- Azure Data Studio
- SQL
- Flat Files

Datasets: (CSV Files)



```
CID, BDATE, GEN
      RF-FR-R928-58,HL Road Frame - Black- 58,,R ,2003-07-01,
RF-FR-R92R-58,HL Road Frame - Red- 58,,R ,2003-07-01,
                                                                                             NASAW00011000,1971-10-06,Male
212,AC-HE-HL-U509-R,Sport-100 Helmet- Red,12,S ,2011-07-01,2007-12-28 213,AC-HE-HL-U509-R,Sport-100 Helmet- Red,14,S ,2012-07-01,2008-12-27 214,AC-HE-HL-U509-R,Sport-100 Helmet- Red,13,S ,2013-07-01,
                                                                                             NASAW00011001,1976-05-10, Male
                                                                                            NASAW00011002,1971-02-09,Male
      HE-HL-U509, Sport-100 Helmet- Black, 12, S , 2011-07-01, 2007-12-28
HE-HL-U509, Sport-100 Helmet- Black, 14, S , 2012-07-01, 2008-12-27
                                                                                            NASAW00011003,1973-08-14, Female
      HE-HL-U509, Sport-100 Helmet- Black, 13, S., 2013-07-01, 
-SO-SO-B909-M, Mountain Bike Socks- M, 3, M, 2011-07-01, 2007-12-28
-SO-SO-B909-L, Mountain Bike Socks- L, 3, M, 2011-07-01, 2007-12-28
                                                                                            NASAW00011004,1979-08-05,Female
      -HE-HL-U509-B,Sport-100 Helmet- Blue,12,S
-HE-HL-U509-B,Sport-100 Helmet- Blue,14,S
                                          2,5,2011-07-01,2007-12-28
4,5,2012-07-01,2008-12-27
                                                                                            NASAW00011005,1976-08-01,Male
                                                                                            NASAW00011006,1976-12-02, Female
      -HE-HL-U509-8, Sport-100 Helmet- Blue, 13, S, 2013-07-01, -CA-CA-1098, AWC Logo Cap, 6, S, 2011-07-01, 2007-12-28 -CA-CA-1098, AWC Logo Cap, 5, S, 2012-07-01, 2008-12-27
                                                                                            NASAW00011007,1969-11-06,Male
                              7,S ,2013-07-01,
                                                                                            NASAW00011008,1975-07-04,Female
                 S,Long-Sleeve Logo Jersey- S,32,S ,2011-07-01,2007-12-28
                                                                                            NASAW00011009,1969-09-29,Male
                                                                                            NASAW00011010,1969-08-05, Female
                                                                                            NASAW00011011,1969-05-03,Male
                                                                                            NASAW00011012,1979-01-14, Female
                                                                                            NASAW00011013,1979-08-03,Male
                                                                                            NASAW00011014,1973-11-06,Female
                                                                                            NASAW00011015,1984-08-26,Female
                                                                                            NASAW00011016,1984-10-25,Male
                                                                                             NASAW00011017,1949-12-24, Female
                                                                                            NASAW00011018,1955-10-06,Male
                                                                                             CID, CNTRY
                                                                                            AW-00011000, Australia
                                                                                            AW-00011001, Australia
                                                                                            AW-00011002, Australia
                                                                                            AW-00011003, Australia
                                                                                            AW-00011004, Australia
                                                                                            AW-00011005, Australia
                                                                                             AW-00011006, Australia
                                                                                             AW-00011007, Australia
                                                                                            AW-00011008, Australia
                                                                                            AW-00011009, Australia
                                                                                            AW-00011010, Australia
                                                                                            AW-00011011, Australia
                                                                                             AW-00011012, US
                                                                                             AW-00011013,US
                                                                                             AW-00011014,US
                                                                                             AW-00011015,US
                                                                                             AW-00011016,US
                                                                                             AW-00011017, Australia
                                                                                             AW-00011018, Australia
```

Data Architecture:

- Bronze Layer Data Ingestion (Extraction)
- Silver Layer Data Ingestion, Data Cleaning (Extraction, Transformation)
- Gold Layer Data Integration (Transformation, Load)

Extraction (Schema Bronze): Data Ingestion

Process Used

- 1. Full Load
- 2. Batch Processing

3. Truncate & Insert

Tables Creation:

```
IF OBJECT ID('bronze.crm cust info','U') IS not NULL
 1
          drop table bronze.crm cust info;
 2
      Create table bronze.crm cust info(
 3
      cst id INT,
 4
      cst key NVARCHAR(50),
 5
      cst firstname NVARCHAR(50),
 6
 7
      cst lastname NVARCHAR(50),
 8
      cst_martial_status NVARCHAR(50),
      cst gndr NVARCHAR(50),
      cst create date DATE
10
11
      );
12
      IF OBJECT_ID('bronze.crm_prd_info' ,'U') IS not NULL
13
          drop table bronze.crm prd info;
14
      Create table bronze.crm prd info(
15
      prd id int,
16
17
      prd key NVARCHAR(50),
      prd nm NVARCHAR(50),
18
      prd cost int,
19
      prd line NVARCHAR(10),
20
      prd_start_date_Date,
21
      prd end date Date
22
23
      );
24
```

Data Ingestion:

```
TRUNCATE TABLE bronze.crm cust info
     BULK INSERT bronze.crm_cust_info
14
     from 'C:\Users\mtalha.saeed\OneDrive - Systems Limited\Internship_Work\DataWarehouse\datasets\source_crm\cust_info.csv
15
     ( FIRSTROW = 2,
    Fieldterminator = ',', tablock
17
18
20
21
     TRUNCATE TABLE bronze.crm_prd_info
23
     BULK INSERT bronze.crm_prd_info
      from 'C:\Users\mtalha.saeed\OneDrive - Systems Limited\Internship_Work\DataWarehouse\datasets\source_crm\prd_info.csv'
     with
     ( FIRSTROW = 2,
     Fieldterminator = ',',
     tablock
```

Extraction & Transformation (Schema Silver):

For Silver Layer – Metadata Column added "dwh create date" in each column.

```
32 ∨ IF OBJECT ID('silver.crm sales details' ,'U') IS not NULL
          drop table silver.crm sales details;
33
34
     Create table silver.crm sales details(
      sls order num NVARCHAR(50),
35
      sls prd key NVARCHAR(50),
36
      sls cust id int,
37
      sls order dt int,
38
39
     sls ship dt int,
     sls due dt int,
40
     sls sales int,
41
42
      sls quantity int,
      sls price int,
43
      dwh create date DATETIME2 DEFAULT GETDATE()
44
45
      );
46
47
     IF OBJECT ID('silver.erp cust az12' ,'U') is not NULL
48
     Drop table silver.erp cust az12;
49
50

∨ create table silver.erp_cust_az12(
          cid NVARCHAR(50),
51
          bdate date,
52
          gen NVARCHAR(50),
53
          dwh create date DATETIME2 DEFAULT GETDATE());
54
```

Transformation:

- Data Cleansing Removing Nulls, Removing Duplicates, Remove String Spaces
- Data Standarization Mapping Values
- Data Normalization Normalize Data Into single type
- Derived Columns Addition of New Columns from previous ones.
- Data Enrichment Refining Data

SQL Scripts for Data Transformation (Schema Silver):

• Silver.crm cust info

```
SELECT
10
11
          cst_id,
12
          cst_key,
13
          Trim(cst_firstname) as cst_firstname,
14
          Trim(cst_lastname) as cst_lastname,
15
16
          CAse when upper(trim(cst_martial_status)) = 'S' then 'Single'
17
          when upper(trim(cst_martial_status)) = 'M' then 'Married'
18
          else 'n/a'
19
          end cst_martial_status,
20
21
          CAse when upper(trim(cst_gndr)) = 'F' then 'Female'
          when upper(trim(cst_gndr)) = 'M' then 'Male'
22
23
          else 'n/a'
24
          end cst gndr,
25
          cst_create_date
26
      FROM
27
28
              SELECT
29
                  ROW_Number() OVER (PARTITION BY cst_id ORDER BY cst_create_date DESC) as flag_last
30
              FROM
31
32
                  bronze.crm_cust_info
33
34
      WHERE
35
          flag_last = 1;
```

Silver.crm prd info

```
53
         Select
54
      prd_id,
       replace(SUBSTRING(prd_key, 1, 5), '-' , '_') as cat_id,
55
56
       SUBSTRING(prd_key , 7 , Len(prd_key)) as prd_key,
57
58
       isnull(prd_cost,0) as prd_cost,
59
       case upper(trim(prd_line))
         When 'M' Then 'Mountain'
60
61
         When 'R' Then 'Road'
         When'S' Then 'Other Sales'
         When 'T' Then 'Touring'
63
      else 'n/a'
64
       end prd_line,
65
      cast (prd_start_date as date) as prd_start_date,
     cast (DATEADD(day, -1, LEAD(prd_start_date) over(PARTITION by prd_key order by prd_start_date)) as date) as prd_end_date_test
     from bronze.crm_prd_info;
```

• Silver.crm_sales_details

```
85
    select
86 sls_order_num,
87 sls_prd_key,
88 sls_cust_id,
     case when sls_order_dt = 0 OR len(sls_order_dt) != 8 then null
89
     else cast(cast(sls_order_dt as Varchar) as date)
91
    end as sls_order_dt,
92 case when sls_ship_dt = 0 OR len(sls_ship_dt) != 8 then null
93 else cast(cast(sls_ship_dt as Varchar) as date)
94 end as sls_ship_dt,
95
     case when sls due dt = 0 OR len(sls due dt) != 8 then null
     else cast(cast(sls due dt as Varchar) as date)
97
     end as sls_due_dt,
     case when sls_sales is null or sls_sales <= 0 or sls_sales != sls_quantity * ABS(sls_price)</pre>
98
         then sls_quantity * ABS(sls_price)
99
100
         else sls sales
    end as sls_sales,
101
102
     sls_quantity,
     case when sls_price is null or sls_price <= 0
103
         then sls_sales /nullif(sls_quantity,0)
104
105
    else sls_price
    end as sls_price
107
     FROM
108
    bronze.crm_sales_details;
  • Silver.erp cust az12
122
         select
         case when cid LIKE 'NAS%' then SUBSTRING(cid, 4, len(cid))
123
124
         else cid
125
       end as cid,
126
        case when bdate > getdate() then null
127
         else bdate
128
         end as bdate,
        case when upper(TRIM(GEN)) iN ('F' , 'Female') then 'Female'
129
         when upper(TRIM(GEN)) iN ('M', 'male') then 'Male'
130
         else ' n/a'
131
132
         end as gen
133
         from bronze.erp_cust_az12;
```

• Silver.erp loc a101

```
141
      select
      replace(cid, '-', '') as cid,
142
143
      CASE
          WHEN TRIM(cntry) = 'DE' THEN 'Germany'
144
          WHEN TRIM(cntry) IN ('US', 'USA') THEN 'United States'
145
          WHEN TRIM(cntry) = '' OR TRIM(cntry) IS NULL THEN 'n/a'
146
          ELSE TRIM(cntry)
147
        END AS newcntry
148
      from bronze.erp_loc_a101
149
150
 • Silver.erp_px_cat_g1v2
      select
             id,
            cat,
            subcat,
           maintenance FROM
```

bronze.erp_px_cat_g1v2;

Loads:

- 1. Data Integrations
- 2. Aggregations
- 3. Star Schema
- 4. Aggregated Table
- Creating Dimension view Customers:

```
20 ∨ IF OBJECT_ID('gold.dim_customers', 'V') IS NOT NULL
          DROP VIEW gold.dim customers;
21
22
      GO
23
      CREATE VIEW gold.dim customers AS
24

✓ SELECT

          ROW NUMBER() OVER (ORDER BY cst id) AS customer key,
25
26
          ci.cst id
                          AS customer id,
27
          ci.cst key
                         AS customer number,
          ci.cst firstname AS first name,
28
29
          ci.cst lastname AS last name,
          la.cntry
                         AS country,
30
          ci.cst martial status AS marital status,
31
          CASE
32
              WHEN ci.cst gndr != 'n/a' THEN ci.cst_gndr
33
              ELSE COALESCE(ca.gen, 'n/a')
34
                  AS gender,
35
          END
                       AS birthdate,
36
          ca.bdate
          ci.cst create date AS create date
37
38
      FROM silver.crm cust info ci

∨ LEFT JOIN silver.erp cust az12 ca

39
40
          ON ci.cst key = ca.cid

∨ LEFT JOIN silver.erp loc a101 la

          ON ci.cst key = la.cid;
42
43
      GO
```

View:

	customer_key 🗸	customer_id 🗸	customer_number 🗸	first_name 🗸	last_name 🗸	country 🗸	marital_status 🗸	gender 🗸	birthdate 🗸	create_date 🗸
1	1	11000	AW00011000	Jon	Yang	Australia	Married	Male	1971-10-06	2025-10-06
2	2	11001	AW00011001	Eugene	Huang	Australia	Single	Male	1976-05-10	2025-10-06
3	3	11002	AW00011002	Ruben	Torres	Australia	Married	Male	1971-02-09	2025-10-06
4	4	11003	AW00011003	Christy	Zhu	Australia	Single	Female	1973-08-14	2025-10-06
5	5	11004	AW00011004	Elizabeth	Johnson	Australia	Single	Female	1979-08-05	2025-10-06
6	6	11005	AW00011005	Julio	Ruiz	Australia	Single	Male	1976-08-01	2025-10-06
7	7	11006	AW00011006	Janet	Alvarez	Australia	Single	Female	1976-12-02	2025-10-06
8	8	11007	AW00011007	Marco	Mehta	Australia	Married	Male	1969-11-06	2025-10-06
9	9	11008	AW00011008	Rob	Verhoff	Australia	Single	Female	1975-07-04	2025-10-06
1	10	11009	AW00011009	Shannon	Carlson	Australia	Single	Male	1969-09-29	2025-10-06

• Creating Dimension view Products:

```
CREATE VIEW gold.dim_products AS
53
      SELECT
          ROW_NUMBER() OVER (ORDER BY pn.prd_start_date, pn.prd_key) AS product_key,
54
          pn.prd id
                          AS product id,
55
                          AS product_number,
          pn.prd key
56
                         AS product_name,
57
          pn.prd_nm
          pn.cat id
                          AS category_id,
                          AS category,
59
          pc.cat
                         AS subcategory,
60
         pc.subcat
          pc.maintenance AS maintenance,
61
         pn.prd_cost
                         AS cost,
62
         pn.prd_line
                          AS product_line,
63
         pn.prd_start_date AS start_date
64
      FROM silver.crm_prd_info pn
65
      LEFT JOIN silver.erp_px_cat_g1v2 pc
66
         ON pn.cat_id = pc.id
67
     WHERE pn.prd_end_date IS NULL;
68
69
      GO
```

View:

product_key	product_id	product_nu	product_na	category_id	category	subcategory	maintenance	cost	product_line	start_date
1	210	FR-R92B-58	HL Road Fra	CO_RF	Components	Road Frames	Yes	0	Road	2003-07-01
2	211	FR-R92R-58	HL Road Fra	CO_RF	Components	Road Frames	Yes	0	Road	2003-07-01
3	348	BK-M82B-38	Mountain-1	BI_MB	Bikes	Mountain Bi	Yes	1898	Mountain	2011-07-01
4	349	BK-M82B-42	Mountain-1	BI_MB	Bikes	Mountain Bi	Yes	1898	Mountain	2011-07-01
5	350	BK-M82B-44	Mountain-1	BI_MB	Bikes	Mountain Bi	Yes	1898	Mountain	2011-07-01

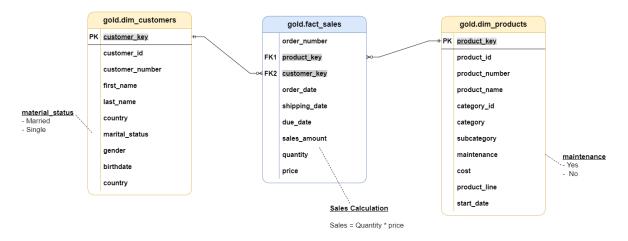
• Creating Fact view Sales:

```
CREATE VIEW gold.fact sales AS
82
      SELECT
83
          sd.sls order num AS order number,
84
          pr.product key AS product key,
85
          cu.customer key AS customer key,
86
          sd.sls order dt AS order date,
87
          sd.sls ship dt AS shipping date,
88
          sd.sls due dt AS due date,
89
          sd.sls sales AS sales amount,
90
          sd.sls_quantity AS quantity,
91
          sd.sls price AS price
92
      FROM silver.crm sales details sd
93
      LEFT JOIN gold.dim products pr
94
          ON sd.sls prd key = pr.product number
95
      LEFT JOIN gold.dim customers cu
96
          ON sd.sls cust id = cu.customer id;
97
      GO
98
00
```

View:

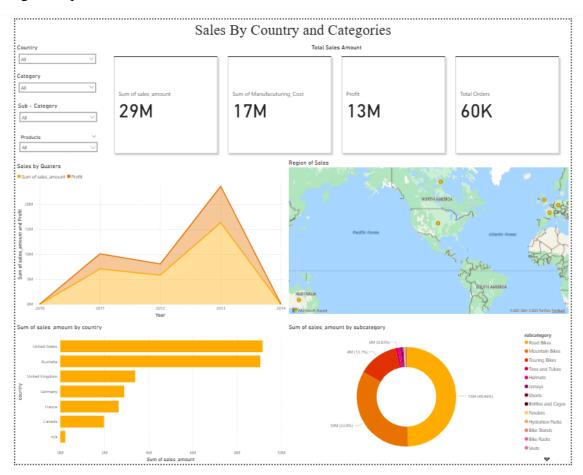
	order_number 🗸	product_key 🗸	customer_key 🗸	order_date 🗸	shipping_date 🗸	due_date 🗸	sales_amount 🗸	quantity 🗸	price 🗸
1	S043697	20	10769	2010-12-29	2011-01-05	2011-01-10	3578	1	3578
2	S043698	9	17390	2010-12-29	2011-01-05	2011-01-10	3400	1	3400
3	S043699	9	14864	2010-12-29	2011-01-05	2011-01-10	3400	1	3400
4	S043700	41	3502	2010-12-29	2011-01-05	2011-01-10	699	1	699
5	S043701	9	4	2010-12-29	2011-01-05	2011-01-10	3400	1	3400
6	S043702	16	16646	2010-12-30	2011-01-06	2011-01-11	3578	1	3578
7	S043703	20	5625	2010-12-30	2011-01-06	2011-01-11	3578	1	3578
8	S043704	6	6	2010-12-30	2011-01-06	2011-01-11	3375	1	3375
9	S043705	7	12	2010-12-30	2011-01-06	2011-01-11	3400	1	3400

Star Schema:



End of Data Warehouse.

Building BI Report



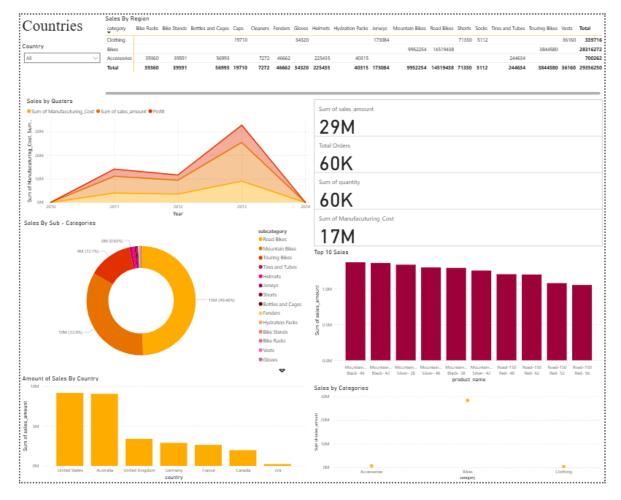
First Page of BI Report:

Charts Used:

- Slicers Filtering by All Possibilities including Countries, Categories, Sub-Categories, Products.
- Heat Map Shows Region with Sales
- Stacked Line Chart Shows Sales Amount, Profit by Year
- Horizontal Bar Chart Shows Country Sales
- Pie Chart Shows Sales Amount by Sub Categories.
- Cards Shows Number

DAX Used:

- Total Orders Count [order Num]
- Profit Sum (Sales Amount) Sum (Manufacturing Price)



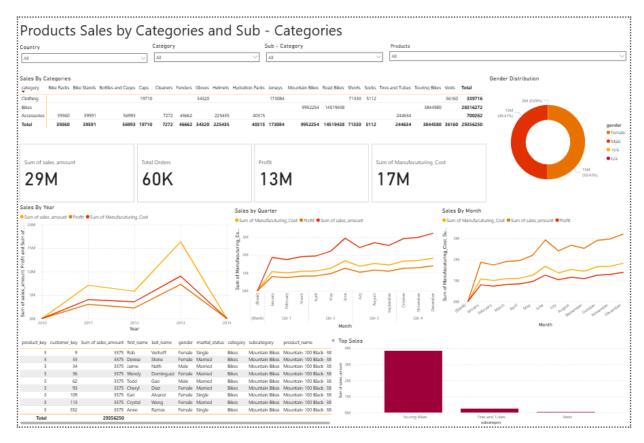
Second Page of BI Report:

Charts Used:

• Slicers - Filtering by All Possibilities including Countries

- Stacked Line Chart Shows Sales Amount, Profit by Quarters
- Bar Chart Shows Country Sales
- Pie Chart Shows Sales Amount by Sub Categories.
- Cards Shows Number
- Table Shows Sales Across all Categories and Sub Categories
- Bar Chart Top 10 Sales

Third Page of BI Report:



Charts Used:

- Slicers Filtering by All Possibilities including Countries, Categories, Sub-Categories, Products
- Stacked Line Chart Shows Sales Amount, Profit by Year, Quarter, Months
- Bar Chart Shows Top Sales
- Pie Chart Gender Sales
- Cards Shows Sales, Profit, Manufacturing Price, Total Orders
- Table Shows All the specific details of Customers

End of BI Report

By: Muhammad Talha @Muhammad Talha

Submitted To: Hassan Muhuyuddin

Dated: 29-Sep-2025