#### Solution Concept of TESLA

#### Saveli Litvin

Contents

[Overview 3](#_Toc78162800)

[Business Background 3](#_Toc78162801)

[Benefits 4](#_Toc78162802)

[Requirements 5](#_Toc78162803)

[Business Requirements 5](#_Toc78162804)

[Technical requirements 7](#_Toc78162805)

[Solution Sketch 9](#_Toc78162806)

[Source Tables Structure 9](#_Toc78162807)

[Star Scheme 11](#_Toc78162808)

[Physical Diagram 11](#_Toc78162809)

[Logical Diagram 12](#_Toc78162810)

[Snowflake Scheme 13](#_Toc78162811)

[Physical Diagram 13](#_Toc78162812)

[Logical Diagram 14](#_Toc78162813)

[Dimension Types Description 15](#_Toc78162814)

[Dimension Hierarchies 16](#_Toc78162815)

[Facts aggregation 19](#_Toc78162816)

[DataFlow Diagram 20](#_Toc78162817)

[Partitioning Facts 22](#_Toc78162818)

[Strategy of Parallel Execution 23](#_Toc78162819)

Overview

# Business Background

Automotive company Tesla was founded in 2003 by a group of engineers who wanted to prove that people didn’t need to compromise to drive electric – that electric vehicles can be better, quicker and more fun to drive than gasoline cars. Today, Tesla builds not only all-electric vehicles but also infinitely scalable clean energy generation and storage products. Tesla believes the faster the world stops relying on fossil fuels and moves towards a zero-emission future, the better.

Today’s Tesla offers cars around over the world. The idea of future policy is to take overview and control not only all cars sales, but also clients needs and dealers in different countries. As well Tesla keeps eyes on their clients and would like to perform cars as well increase sales. Tesla has only 4 Models. Each of them can have their own modification, like Tesla Model S Plaid or Tesla Model 3 Performance.

* Model S
  + Long Range: AWD, 405 miles, 0-60mph in 3.1 seconds
  + Plaid: AWD, 390 miles, 0-60mph in 1.9s
* Model 3
  + Standard Range Plus: RWD, 263 miles, 0-60mph in 5.3 seconds, partial premium interior
  + Long Range: AWD, 353 miles, 0-60mph in 4.2s, premium interior
  + Performance: AWD, 315 miles, 0-60mph in 3.1s, premium interior
* Model X
  + Long Range: AWD, 360 miles, 0-62mph in 3.8 seconds
  + Plaid: AWD, 340 miles, 0-62mph in 2.5s
* Model Y
  + Long Range: AWD, 326 miles, 0-60mph in 4.8s
  + Performance: AWD, 303 miles, 0-60mph in 3.5s

To realise patern of worldwide distribution one‘s can be a great help – DWH storage.

# Benefits

The next benefits could be received by creating OLAP solution:

* Data models are mainly used to support the design and implementation of information systems and as an information management tool for the determination of information supply and demand as well as for the structuring and documentation of operational contexts.
* Multidimensional data models were designed for the special requirements of management support systems. They should represent the data as much as possible as they are seen by the users in reality.
* The access to the data takes place in an intuitive form, which no longer requires the formulation of complex queries with special languages.
* OLAP models can be scaled within the existing system to expand the interest of analytical data processing.

Requirements

# Business Requirements

The Companies vehicle types are regularly examined as part of a financial analysis. The business key figures sales, contribution to cover and operating profit are examined in relation to each individual vehicle type and to the total quantity sold of this type. In this regard, the paragraph refers only to the dealers and does not contain any data on deliveries to end customers. Thus, the above financial data also refers only to the dealer network. Furthermore, the financial data are compared with each other, so that the shares of the contribution margin and the operating result in sales are identified can.

The vehicle types are also analysed from a market perspective. Here, the market volume of each type is considered as the relevant variable. In general, market volume differs from sales, as it refers to the end customer. For example, the market volume may be higher than the sales volume due to stock sales by dealers. Of course, it is also possible that sales are higher than market volume. In that case, the dealers could not sell all the vehicles purchased from the factory to the end customer and have to increase their stock. A differentiated view is therefore necessary. If the market volume is put in relation to the total market volume, the market share of a vehicle type can be determined. By comparing the individual market volumes of a type in different countries, the country mix of a vehicle type can also be determined

The analysis of the above deviations will allow the formation of a market supply able to meet the demand in the region. Thus, an objective Task in the implementation of the platform becomes the reduction of the final costs of vehicle storage and sales. This will ultimately have a positive impact on the company's balance sheet.

Ultimately the implementation will not only affect the current situation, but also allow for analysis and forecasting of future performance by region.

The time dimension will allow demand to be shaped on the basis of seasonal variations. Thus, the main task of time variable analysis becomes the consolidation of yearly, quarterly and monthly data in the regions.

# Technical requirements

One’s important think is TR. We will construct table to pin key points for DWH creating.

|  |  |  |
| --- | --- | --- |
| # | Action | Description |
| 1 | Availability - 24/7/365 | Because main Idea is to build DWH to be accessed from multiple countries – need to perform availability. In this stage availability is accessibility from any time zone. |
| 2 | Historical Date | To start making prognoses and predictions faster, you need to collect as much data as possible. Thus, historical data from regions can help in building this part. It is clear that if historical data is not available or available, we will collect and store it for a long time. |
| 3 | High performance | In a large data warehouse, it is necessary to create all conditions for fast data access. Thus, it is necessary to provide high-speed two-way communication. Probably using patterns to provide access to the most popular queries. |
| 4 | All the information must be protected according to the company’s security Policy | Obliviously, data will be protected according to company rules. |
| 5 | Adjust granuality | Determine the desired level of detail, balancing the business needs and the performance and cost implications |
| 6 | Scaliability | There's no question that by creating OLAP system - volume of data stored in the storage increasly fast scaling. At this way one of the biggest part is creating conditions of scaleability basis. |
| 7 | Data Backup strategy | Every week backup as well store incremental cloud backups |

Basic Implementation:

Recreating DWH will be a big helpful by collecting information from all over departments that basically differentiated by Region, Time Zone and Local storages to provide reports for future expandation and at market influence.

Solution Sketch

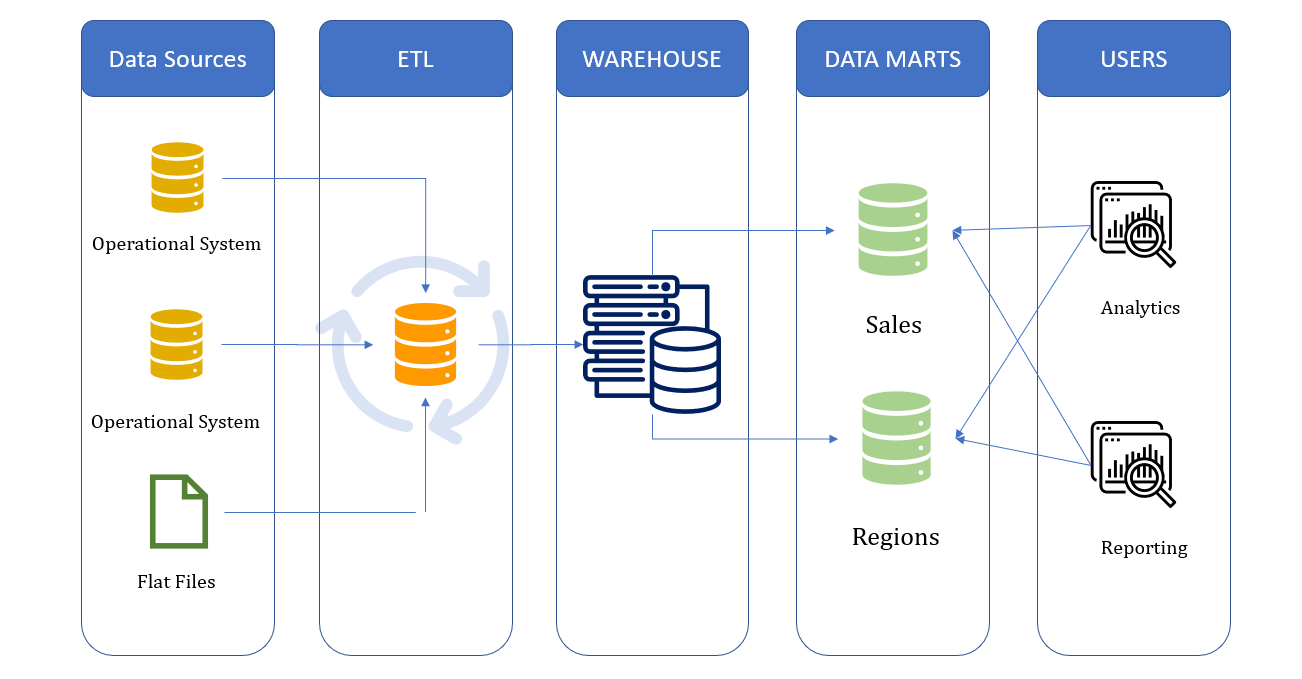
# Source Tables Structure

|  |  |
| --- | --- |
| Name | Products |
| Abbreviation | PRODUCTS |
| Unit of measurement | Unique Model and additional modifications |
| Status | Date |
| Description | Products include actual model lineup and order availability |
| Data origin | Collected from operational systems of the brands |
| Update | Of necessity |

|  |  |
| --- | --- |
| Name | Customers |
| Abbreviation | CUSTOMERS |
| Unit of measurement | One client |
| Status | Date |
| Description | Each customers include first and last name, contact details, purchase details (product, bill, city / region), date |
| Data origin | Collected from operational systems of the brands as well flat files |
| Update | Every day at 12 p.m. |

|  |  |
| --- | --- |
| Name | Regions |
| Abbreviation | GEO |
| Unit of measurement | - |
| Status | Date |
| Description | Regions table describes each for delivery available end point. |
| Data origin | Collected from operational systems of the brands as well flat files |
| Update | Each new distribution center filled in the dimension |

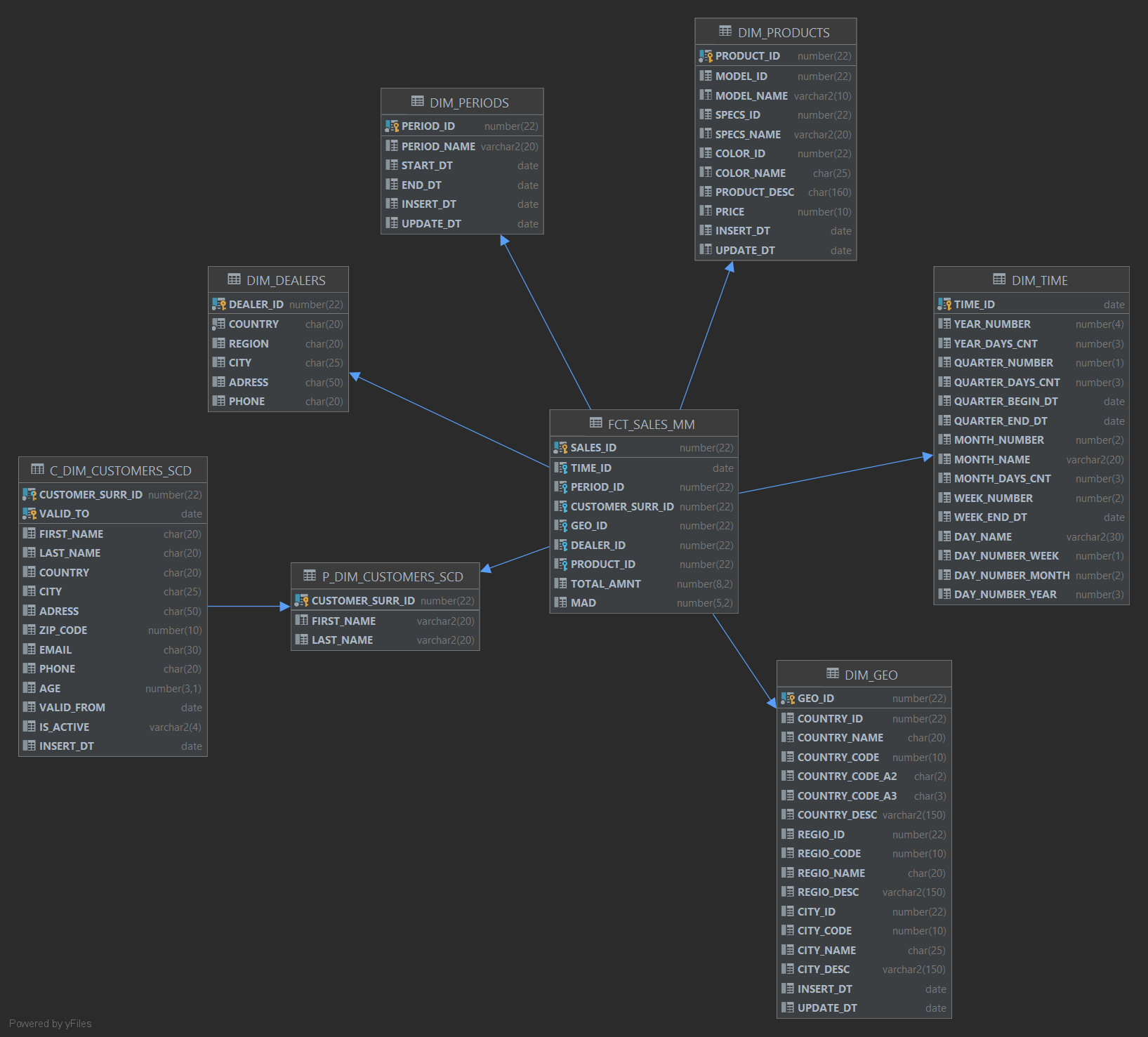
|  |  |
| --- | --- |
| Name | Dealers |
| Abbreviation | DEALERS |
| Unit of measurement | - |
| Status | Date |
| Description | Contains information about Local Offices |
| Data origin | Collected from flat files |
| Update | Up to Date |



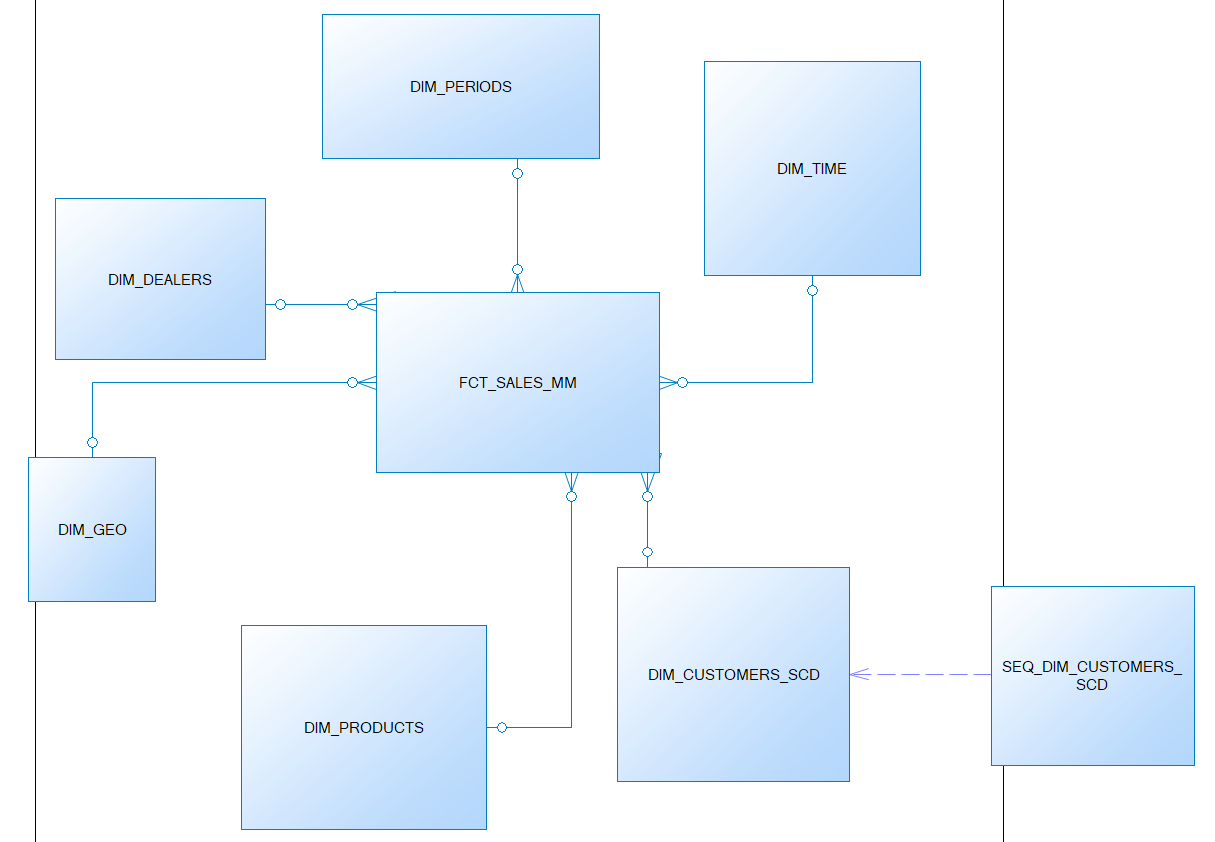
Conception Design

Star Scheme

# Physical Diagram

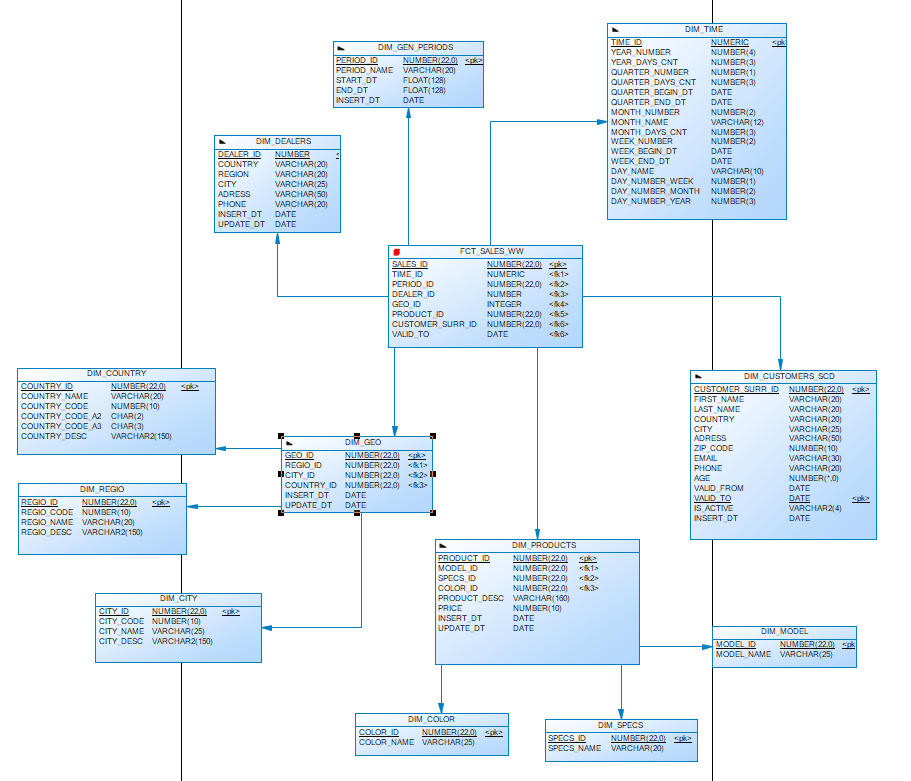


# Logical Diagram

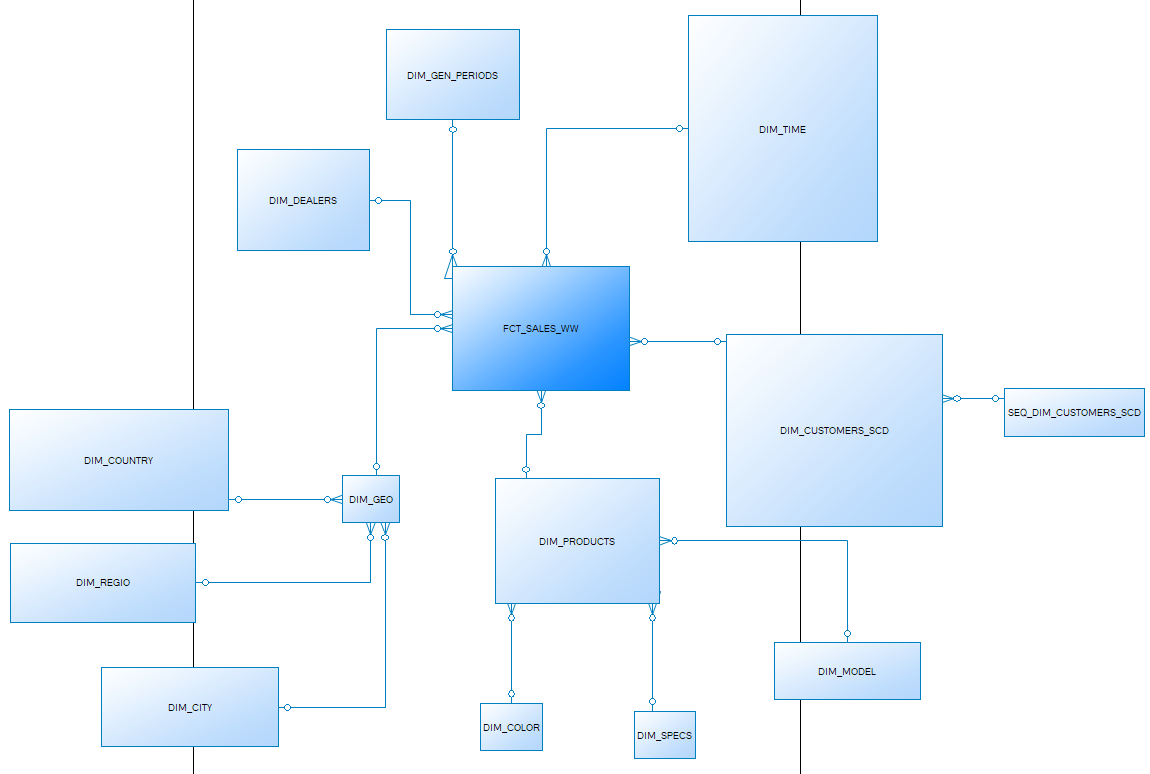


Snowflake Scheme

# Physical Diagram



# Logical Diagram



Dimension Types Description

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Type | Size | DW – Merged Dimensions | Descriptions |
| DIM \_TIME | SCD1 | BIG | DW.T\_DAYS, DW.T\_WEEKS, DW.T\_MONTHS, DW.T\_QUARTERS,  DW.T\_YEARS | A Calendar, contain wide information about days, weeks, months, quarters and years. |
| DIM\_GEN\_PERIODS | SCD1 | BIG | PERIOD\_ID,  PERIOD\_NAME,  START\_DT,  DAYS\_CNT,  END\_\_DT | Processing Dimension – Useful in Context of Data Aggregation. |
| DIM\_DEALERS | SCD1 | SMALL | DEALER\_ID,  COUNTRY,  REGION,  CITY,  ADRESS,  PHONE | This dimension contain all information about distributors. |
| DIM\_GEO | SCD1 | SMALL | DW.T\_COUNTRIES,  DW.T\_REGIONS,  DW.T\_CITIES | This one stored data about all available locations all over the world. |
| DIM\_PRODUCTS | SCD1 | SMALL | MODELS,  SPECS,  COLORS,  PRICES | Dimensions stores information about available range of car models and their adaptive specs. |
| DIM\_CUSTOMERS\_SCD | SCD2 | BIG | FIRST\_NAME,  LAST\_NAME,  COUNTRY,  CITY,  ADRESS,  ZIP\_CODE  EMAIL,  PHONE,  AGE | Dimension is stored detailed information about customers for widespread representative uses. This Table was chosed as SCD2 dimension type for rejecting data merging. |

Dimension Hierarchies

**DIM\_TIME:**

**Hierarchy DAY-WEEK-MONTH- QUARTER -YEAR**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| DAYs | DAY | Store all day at the calendar | DAY\_NUMBER\_YEAR |
| WEEKs | WEEK | Store all weeks at the calendar year | WEEK\_NUMBER |
| MONTHs | MONTH | Store all months at the calendar year | MONTH\_NUMBER |
| QUARTERs | QUARTER | Store all quarter at the calendar year | QUARTER\_NUMBER |
| YEARs | YEAR | Store all years | YEAR\_NUMBER |

**DIM\_GEO:**

**Hierarchy CITY-REGIO-COUNTRY**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| CITYs | CITY | Store all cities in the country | CITY\_ID |
| REGIOs | REGIO | Store all regios in the country | REGIO\_ID |
| COUNTRYs | COUNTRY | Store all countries | COUNTRY\_ID |

**DIM\_GEN\_PERIODS:**

**Hierarchy START\_DATE – END\_DATE – PERIOD\_NAME**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| START DATEs | START\_DT | Store start date of looking period | START\_DT |
| END DATEs | END\_DT | Store end date of looking period | END\_DT |
| PERIOD NAMEs | PERIOD\_ID | Store all periods | PERIOD\_ID |

**DIM\_PRODUCTS:**

**Hierarchy PRICE-COLOR-SPECS-MODEL**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| PRICEs | PRICE | Contain information about the prices according to models | PRICE |
| COLORs | COLOR | Contain informatin about colors according to all models | COLOR\_ID |
| SPECIFICATIONs | SPECIFICATION | Contain information about available customization to each model | SPECS\_ID |
| MODELs | MODEL | Contain information about all available models | MODEL\_ID |

**DIM\_CUSTOMERS\_SCD:**

**Hierarchy FIRST\_NAME – LAST\_NAME – COUNTRY – CITY – ZIP\_CODE – ADDITITVE\_DATA - CUSTOMER**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| FIRST NAMEs | FIRST\_NAME | Contain information about all customers first names | FIRST\_NAME |
| LAST NAMEs | LAST\_NAME | Contain information about all customers last names | LAST\_NAME |
| COUNTRYs | COUNTRY | Contain information about all countries of existing customers | COUNTRY |
| CITYs | CITY | Contain information about all cities of existing customers | CITY |
| ZIP CODEs | ZIP\_CODE | Contain information about all zip codes of existing customers | ZIP\_CODE |
| Additive data | EMAIL, PHONE, AGE | Contain additive information of existing customers. Is for subjectification while searching | EMAIL,  PHONE, AGE |
| CUSTOMERs | CUSTOMER | Contain Information about all existing customers | CUSTOMER\_SURR\_ID |

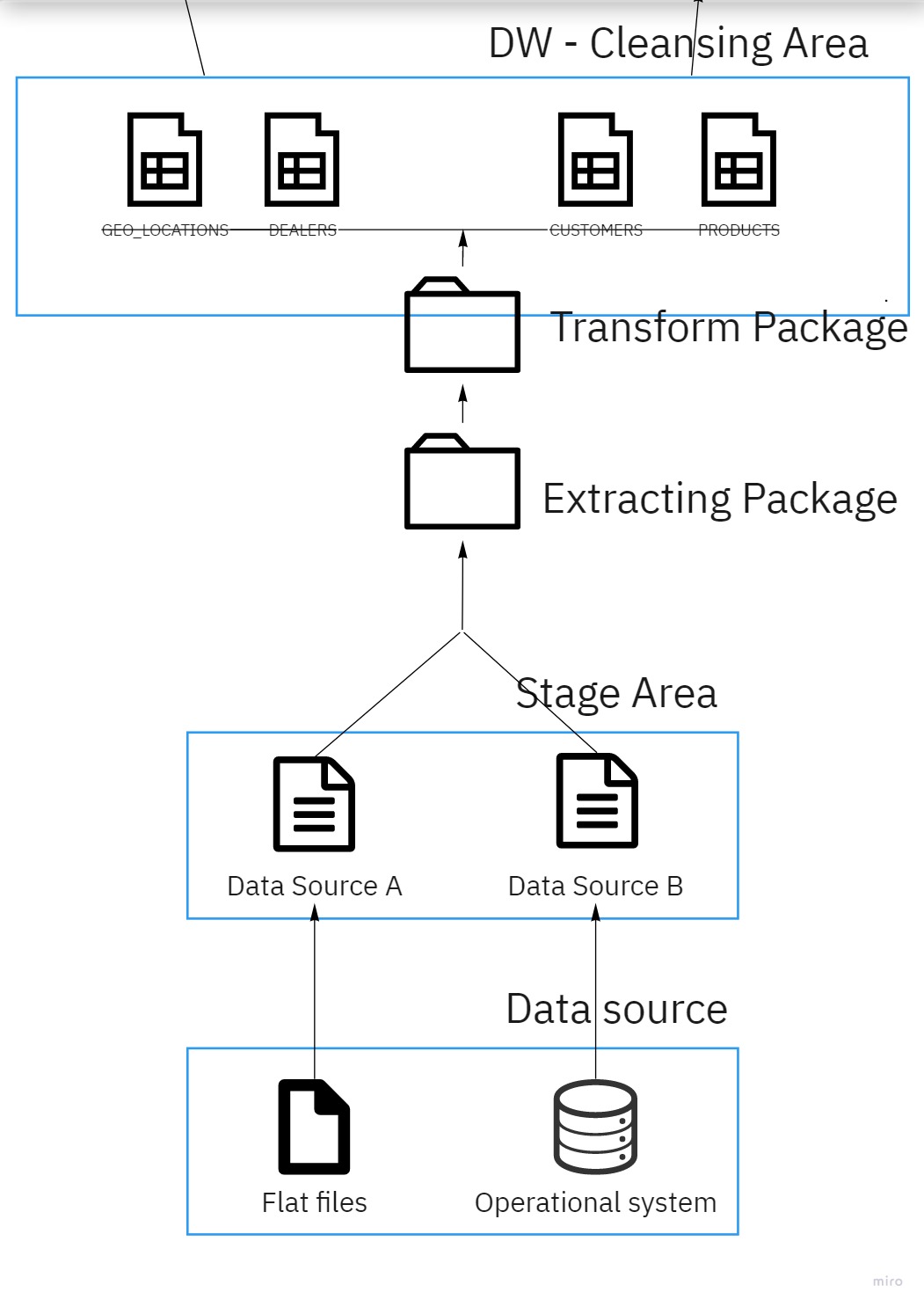
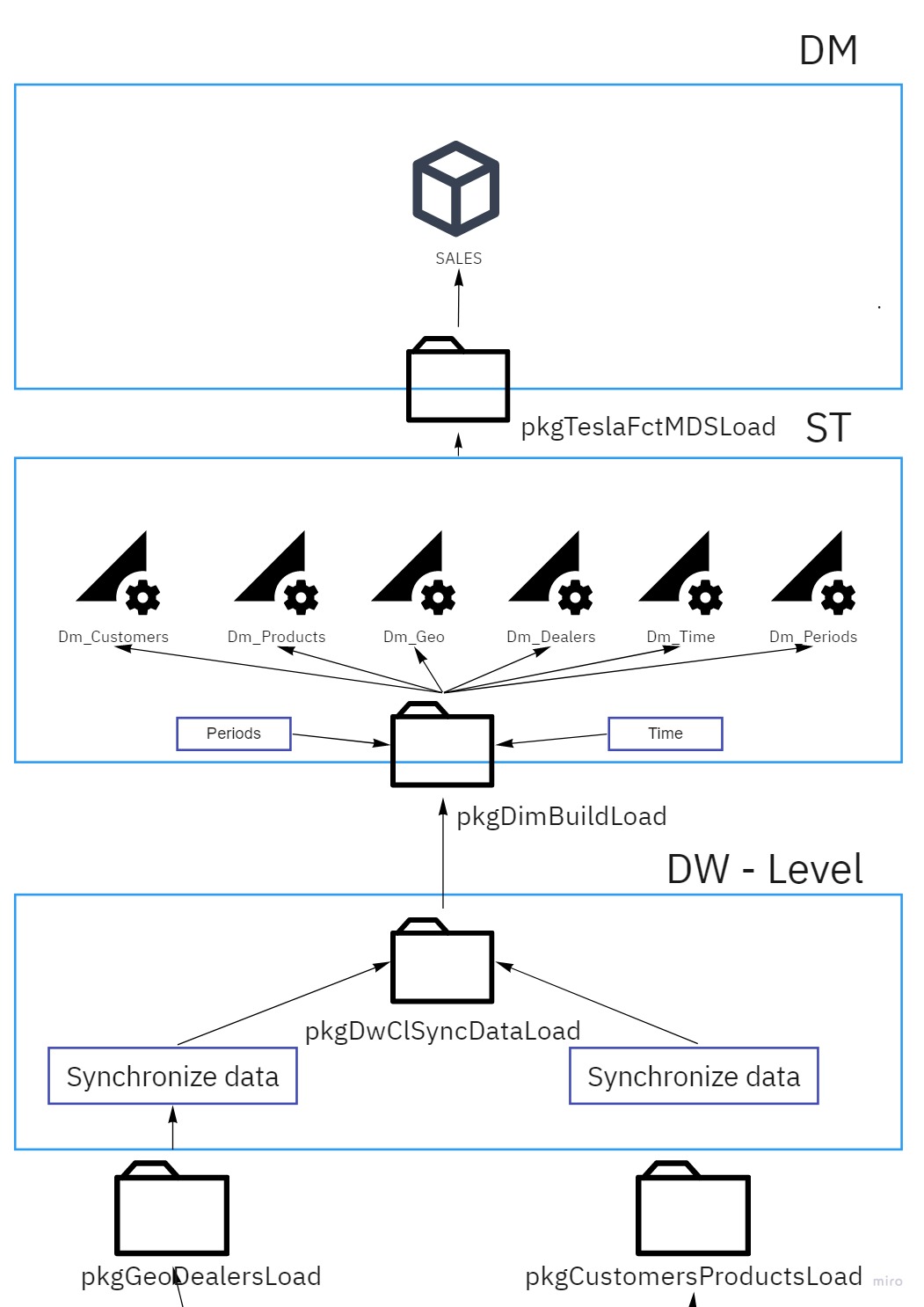
**DIM\_DEALERS:**

**Hierarchy PHONE – ADRESS – CITY – REGIO – COUNTRY – DEALERSHIPS**

|  |  |  |  |
| --- | --- | --- | --- |
| Name | LEVEL\_CODE | LEVEL\_DESC | LEVEL\_NATURAL\_KEY |
| PHONEs | PHONE | Store all phones at all dealerships | PHONE |
| ADRESSes | ADRESS | Store all local addresses at all dealerships | ADRESS |
| CITYs | CITY | Store all citys where all dealerships are located | CITY |
| REGIOs | REGIO | Store all regios at all dealerships | REGION |
| COUTNRYs | COUNTRY | Store all countries at all dealerships | COUNTRY |
| DEALERSHIPs | DEALERSHIP | Store all dealerships | DEALER\_ID |

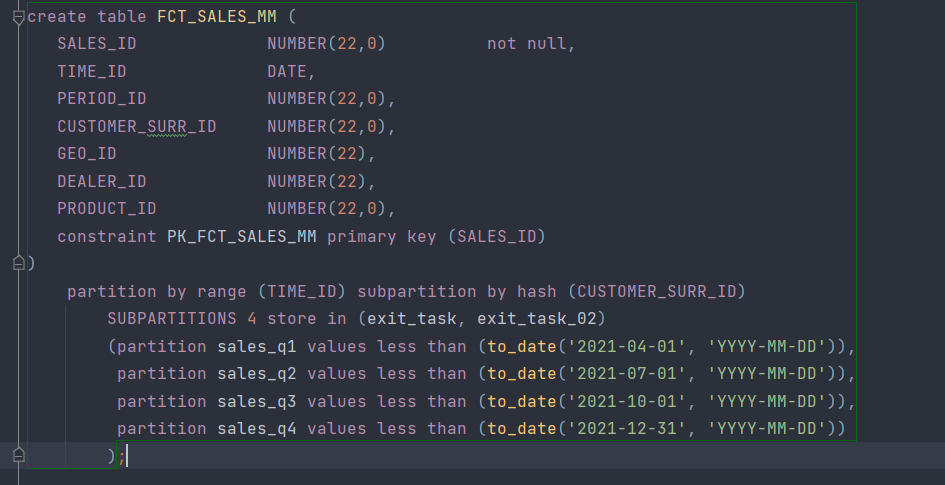
Facts aggregation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Code | Table Name | Additive | Descriptions |
| Total amount | TOTAL\_AMNT | FCT\_SALES\_MM | + | The estimated total sales in the fact table for the corresponding period. |
| Mean Absolute Deviation | MAD | FCT\_SALES\_MM | - | Mean absolute deviation from the sales norm for a certain period of time. The level of assessment of the performance of an individual entity. The assessment is based on data on the previous periods of the subject's activity. |

DataFlow Diagram

Partitioning Facts

The concept of Fact table is to create opportunity to suggest successful reporting and analyse trends as well perspectives. Keys to get it is to calculating Total Amount of Sales and Mean Absolute Deviation. Right Strategy is to create a composite range-hash partitioned table. Partitioning by Range on Time\_id where Ranging will be adoptet by quarters, cause the reporting will be over the months. And is logic to divide each 3 months at 4 sections. And subpartition by hash over the Customer\_SURR\_ID by 4.



Strategy of Parallel Execution

One way or another Concept Solution will makes report building automatization. Thus main parallel strategy will go over 2 main tables. Is by selecting actual data from customers table, this table is SCD2 and Big . And by creating new fact sales table. By selecting we will achive it by using Parallel execution with 4 threads. By creating – DDL execution with 2 threads. We will go over customers table also by inserting or merging, deleting data from with using DML parallelization, I presume that 4 threads will be enough, but could be increased in future.