

|  |
| --- |
| Business Template |
| **Logo / Image** |

Contents

[1 Business Description 3](#_Toc155614186)

[1.1 Business background 3](#_Toc155614187)

[1.2 Problems because of poor data management 3](#_Toc155614188)

[1.3 Benefits from implementing a Data Warehouse 3](#_Toc155614189)

[1.4 DATASETS DESCRIPTION 3](#_Toc155614190)

[1.5 GRAIN / DIM / FACT 4](#_Toc155614191)

[2 Business Layer 3NF 4](#_Toc155614192)

[3 Business Layer Dimensional Model 4](#_Toc155614193)

[4 Logical Scheme 4](#_Toc155614194)

[5 Data Flow 4](#_Toc155614195)

[6 Fact Table Partitioning Strategy 4](#_Toc155614196)

# 

# Business Description

## Business background

In the dynamic world of retail, our business operates at the intersection of online and offline sales channels, offering a diverse range of products to meet customer needs. The modern retail landscape is complex and competitive, requiring sophisticated data management and analytical capabilities to understand customer behavior, optimize sales strategies, and make informed business decisions.

## Problems because of poor data management

Traditional retail businesses often struggle with fragmented data sources, making it difficult to gain comprehensive insights. Without a robust data management system, retailers face significant challenges:

* Inability to track sales performance across different channels
* Limited understanding of customer preferences and buying patterns
* Difficulty in identifying cross-selling and upselling opportunities
* Inefficient inventory management
* Challenges in personalizing customer experiences

## Benefits from implementing a Data Warehouse

A comprehensive Data Warehouse solution can transform these challenges into strategic advantages:

* Unified view of sales across online and offline channels
* Detailed customer segmentation and behavior analysis
* Real-time performance tracking
* Predictive analytics for inventory and sales forecasting
* Enhanced decision-making through comprehensive data insights

## DATASETS DESCRIPTION

Online Store Dataset: online store dataset captures digital transactions with rich, multidimensional data:

* Comprehensive sales tracking
* Detailed customer information
* Product categorization
* Geographic distribution
* Digital transaction specifics (delivery method, payment types)

Offline Store Dataset: Complementing the online data, offline store dataset provides:

* In-store sales transactions
* Local market insights
* Physical store performance metrics
* Alternative customer interaction points

Unique Features:

* Two distinct yet complementary sales channels
* Comprehensive tracking of 500,000 transactions per channel
* Historical data covering two years
* Multiple dimensional perspectives (customer, product, geographic, temporal)

Hierarchical Relationships:

* Geographic Hierarchy: Country → Region → State/Province → City
* Product Hierarchy: Category → Sub-Category → Product Name
* Time Hierarchy: Year → Quarter → Month → Week → Day

Fact Table:

* Name: FCT\_SALES\_DD
* Attributes:
  + SALE\_SURR\_ID (Primary Key)
  + CUSTOMER\_SURR\_ID (FK)
  + PRODUCT\_SURR\_ID (FK)
  + ORDER\_SURR\_ID (FK)
  + EMPLOYEE\_SURR\_ID (FK)
  + GEOGRAPHIC\_SURR\_ID (FK)
  + DATE\_SURR\_ID (FK)
  + SALES\_TOT\_AMT
  + QUANTITY\_CNT
  + DISCOUNT\_PCT
  + PROFIT\_TOT\_AMT

Dimension Tables:

1. Customer Dimension (DIM\_CUSTOMERS):
   * CUSTOMER\_SURR\_ID
   * CUSTOMER\_NAME
   * CUSTOMER\_SEGMENT\_NAME
   * COUNTRY\_NAME
   * CITY\_NAME
   * POSTAL\_CODE\_NO
2. Product Dimension (DIM\_PRODUCTS):
   * PRODUCT\_SURR\_ID
   * PRODUCT\_NAME
   * PRODUCT\_CATEGORY\_NAME
   * PRODUCT\_SUBCATEGORY\_NAME
   * PRODUCT\_DESC
3. Order Dimension (DIM\_ORDERS):
   * ORDER\_SURR\_ID
   * EVENT\_DT
   * SHIP\_DT
   * SHIP\_MODE\_NAME
   * ORDER\_TYPE\_NAME
4. Geographic Dimension (DIM\_GEOGRAPHY):
   * GEOGRAPHIC\_SURR\_ID
   * COUNTRY\_NAME
   * REGION\_NAME
   * STATE\_NAME
   * CITY\_NAME
   * POSTAL\_CODE\_NO
5. Employee Dimension (DIM\_EMPLOYEES):
   * EMPLOYEE\_SURR\_ID
   * EMPLOYEE\_NAME
   * EMPLOYEE\_SURNAME
   * DEPARTMENT\_NAME
   * ROLE\_NAME

Date Dimension (DIM\_TIME\_DAY):

* DATE\_SURR\_ID
* FULL\_DT
* YEAR\_NO
* QUARTER\_NO
* MONTH\_NO
* WEEK\_NO
* DAY\_NO

## GRAIN / DIM / FACT

# Business Layer 3NF

# Business Layer Dimensional Model

# Logical Scheme

# Data Flow

# Fact Table Partitioning Strategy