Stanisław Lemański 193333
Jakub Kobyliński 193618

Developer - Data warehouse design

Business process

The business processes focus on optimizing material orders and transfers for construction projects. Key activities include tracking order costs, delivery efficiency, material loss during transfers, and monitoring historical changes in suppliers and projects to improve decision-making and cost management.

Relational Database schema

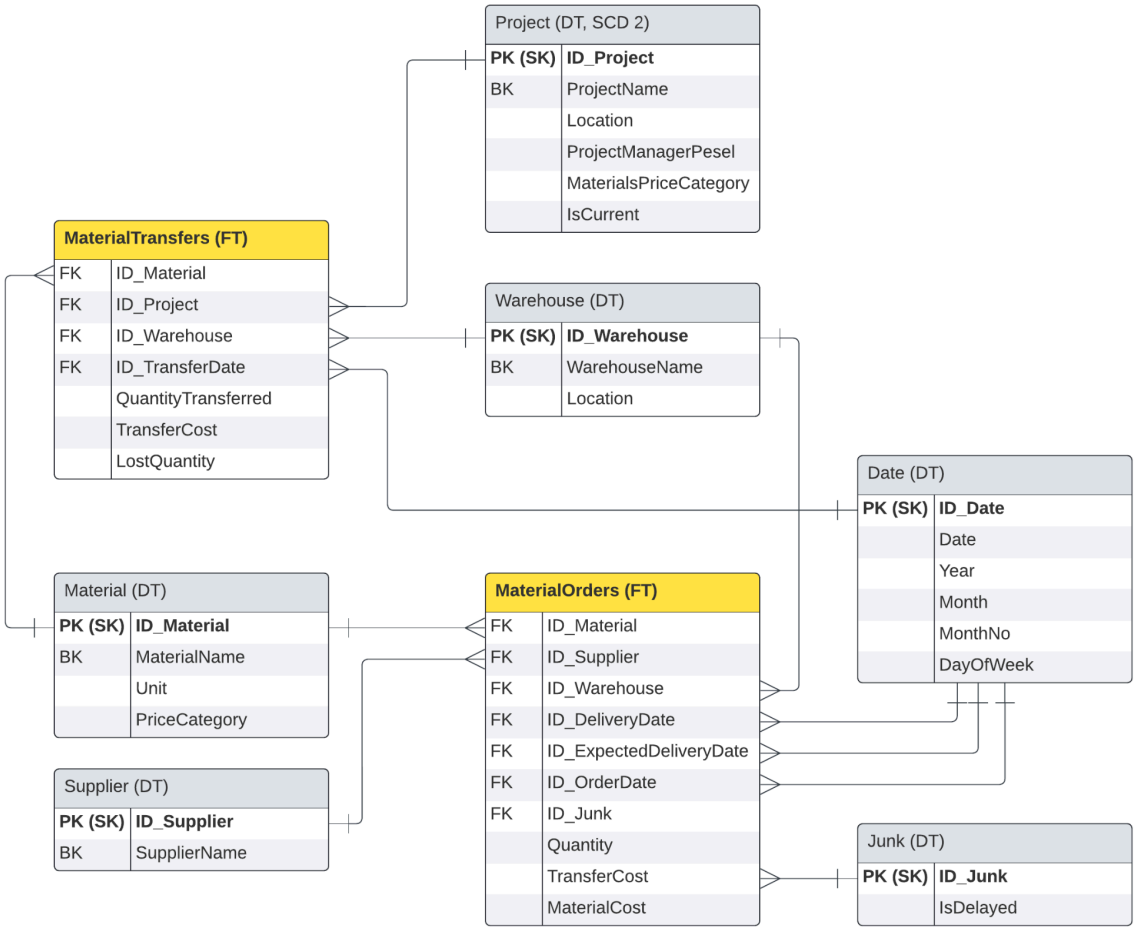

TABLE NAME	ATTRIBUTE	ATTRIBUTE TYPE	DESCRIPTION
MaterialOrders	**One tuple describes one fact of material ordering.**		

(FACT TABLE)			
	ID_Material	Numeric	FK Material Material
	ID_Supplier	Numeric	FK Supplier Supplier
	ID_OrderDate	Numeric	FK Date Order date
	ID_DeliveryDate	Numeric	FK Date Delivery date
	ID_ExpectedDeliveryDate	Numeric	FK Date Expected delivery date
	ID_Warehouse	Numeric	FK Warehouse Warehouse
	ID_Junk	Numeric	FK Junk Junk
	Quantity	Numeric	Quantity of material ordered
	TransferCost	Money	Cost associated with transferring materials
	MaterialCost	Money	Cost of material ordered (Quantity * UnitPrice)
MaterialTransfers (FACT TABLE)	**One tuple describes one fact of material transfer from warehouse to project.**		
	ID_Material	Numeric	FK Material Material
	ID_Warehouse	Numeric	FK Warehouse Warehouse
	ID_Project	Numeric	FK Project Project
	ID_TransferDate	Numeric	FK Date Transfer Date
	QuantityTransferred	Numeric	Quantity of material transferred to the project.
	TransferCost	Money	Cost of the transfer of materials.
	LostQuantity	Numeric	Quantity of material lost during transfer.
Material	**One tuple describes each type of material.**		

(DIMENSION TABLE)			
	ID_Material	Numeric	PK
	MaterialName	Varchar(40)	Material name
	Unit	Varchar(10)	Defines the material unit. Allowed values: [kg, m, m², m³, l]
	PriceCategory	Varchar(10)	Price category. Allowed values: very cheap, cheap, moderate, expensive, exclusive
Supplier (DIMENSION TABLE)	**One tuple describes each supplier.**		
	ID_Supplier	Numeric	PK
	SupplierName	Varchar(40)	Supplier name
Warehouse (DIMENSION TABLE)	**One tuple describes each warehouse.**		
	ID_Warehouse	Numeric	PK
	WarehouseName	Varchar(40)	Warehouse name
	Location	Varchar(40)	Location of the warehouse.
Project (DIMENSION TABLE)	**One tuple describes each project.**		
	ID_Project	Numeric	PK
	ProjectName	Varchar(40)	Name of the construction project.
	Location	Varchar(40)	Location of the project.
	ProjectManagerPesel	Varchar(11)	Pesel number of project manager.
	MaterialPriceCategory	String	All materials costs for the project summed up category. Allowed values: [Economical, Standard, Premium]
	IsCurrent	Boolean	1 if information is current, otherwise 0. (SCD2 implementation)
Date	**One tuple describes one day.**		

(DIMENSION TABLE)			
	ID_Date	Numeric	PK
	Date	Date	Date
	Year	4 digits	Year
	Month	Varchar(10)	Month. Allowed values: January, February, March, April, May, June, July, August, September, October, November and December.
	MonthNo	Numeric	Month's numeric value
	DayOfWeek	Varchar(10)	Day of week. Allowed values: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday
Junk (DIMENSION TABLE)	**The tuples correspond to "all" possible combinations of values for isDelayed.**		
	ID_Junk	Numeric	PK
	IsDelayed	Booloean	Indicates whether a delay occurred for deliveries on this specific day. A value of "1" means there was a delay, and "0" means there was no delay.

Dimensional model

Fact definitions

Fact 1: Order Fact: Represents each individual material order placed for storage in a warehouse before distribution to construction projects. The order includes details about the requested material, the supplier providing it, and the warehouse where it will be stored. This fact captures the order placement date, expected delivery date, whether the delivery was delayed, and all associated costs, including material and transfer expenses. This data allows for accurate tracking of material stock and storage efficiency.

Fact table: MaterialOrders

Granularity:

- a specified warehouse in a designated location, which will store materials temporarily.
- a specified supplier
- a specified order date, delivery date and expected delivery date
- a specified material with specified unit and price category
- an information if delivery was delayed

Measures and aggregation functions:
Number of order facts - COUNT(1)
Total transfer cost - SUM (TransferCost)
Total material cost - SUM (MaterialCost)
Total quantity ordered - SUM (Quantity)
Number of suppliers - DISTINCT COUNT (ID_Supplier)

Fact 2 Transfer fact: Captures every transfer of materials from a warehouse to a construction project site, enabling detailed tracking of logistics and inventory movement. Each transfer record includes the specific material being transferred, the warehouse it originates from, and the intended project destination. Additionally, the fact records the date of transfer, allowing for tracking of logistical timelines, and accounts for any material lost during transit, ensuring accurate records of material availability at project sites. This fact also tracks costs directly related to the transfer process, contributing to an understanding of the total logistical expenditure per project.

Fact table: MaterialTransfers

Granularity:

- a specified material with specified unit and price category
- a specified warehouse with specified location
- a specified project in the specified location, with the specified project manager and total material cost
- a specified transfer date

Measures and aggregation functions:
Number of transfer facts - COUNT(1)
Number of warehouses - DISTINCT COUNT (ID_Warehouse)
Total quantity transferred - SUM (QuantityTransferred)
Total transfer cost - SUM (TransferCost)
Total lost quantity - SUM (LostQuantity)

Dimensional definitions

Dimensions for Fact 1 Order fact:

DIMENSION/DIMENSION ATTRIBUTE	TABLE/COLUMN	TYPE
MATERIAL	Material	Dimension
MATERIAL NAME	Material.MaterialName	Dimension attribute
MATERIAL UNIT	Material.Unit	Dimension attribute
PRICE CATEGORY	Material.PriceCategory	Dimension attribute
SUPPLIER	Supplier	Dimension
SUPPLIER NAME	Supplier.SupplierName	Dimension attribute
WAREHOUSE	Warehouse	Dimension
WAREHOUSE LOCATION	Warehouse.Location	Dimension attribute
DELIVERY DATE	Date	Dimension
DELIVERY YEAR	Date.Year	Dimension attribute
DELIVERY MONTH	Date.Month	Dimension attribute
DELIVERY DAY OF WEEK	Date.DayOfWeek	Dimension attribute
DELIVERY DATE HIERARCHY	• Date.Year •• Date.Month ••• Date.Date	Hierarchical dimension
EXPECTED DELIVERY DATE	Date	Dimension
EXPECTED DELIVERY YEAR	Date.Year	Dimension attribute
EXPECTED DELIVERY MONTH	Date.Month	Dimension attribute
EXPECTED DELIVERY	Date.DayOfWeek	Dimension

DAY OF WEEK		attribute
ORDER DATE	Date	Dimension
ORDER YEAR	Date.Year	Dimension attribute
ORDER MONTH	Date.Month	Dimension attribute
ORDER DAY OF WEEK	Date.DayOfWeek	Dimension attribute
ORDER DATE HIERARCHY	• Date.Year •• Date.Month ••• Date.Date	Hierarchical dimension
JUNK	Junk	Dimension
IS DELAYED	Junk.IsDelayed	Dimension attribute

Dimensions for Fact 2 Transfer fact:

DIMENSION/DIMENSION ATTRIBUTE	**TABLE/COLUMN**	**TYPE**
MATERIAL	Material	Dimension
MATERIAL NAME	Material.MaterialName	Dimension attribute
MATERIAL UNIT	Material.Unit	Dimension attribute
PRICE CATEGORY	Material.PriceCategory	Dimension attribute
PROJECT	Project	Dimension
PROJECT NAME	Project.ProjectName	Dimension attribute
PROJECT LOCATION	Project.Location	Dimension attribute
PROJECT MANAGER PESEL	Project.ProjectManagerPesel	Dimension attribute
MATERIAL PRICE CATEGORY	Project.MaterialsPriceCategory	Dimension attribute

WAREHOUSE	Warehouse	Dimension
WAREHOUSE LOCATION	Warehouse.Location	Dimension attribute
TRANSFER DATE	Date	Dimension
TRANSFER YEAR	Date.Year	Dimension attribute
TRANSFER MONTH	Date.Month	Dimension attribute
TRANSFER DAY OF WEEK	Date.DayOfWeek	Dimension attribute
TRANSFER DATE HIERARCHY	• Date.Year •• Date.Month ••• Date.Date	Hierarchical dimension

Checking the feasibility of queries based on the multidimensional model

1. Compare the quantity of materials transferred for each project relative to previous months.
   Measure: Total quantity transferred
   Dimension: Project (dimension attribute: Project name)
   Dimension: Transfer Date (dimension attribute: Transfer date)

2. What were the most expensive materials ordered this month compared to previous months?
   Measure: Total material cost
   Dimension: Order date (dimension attribute: Order month)
   Dimension: Material (dimension attribute: Material name)

3. Compare the material purchase price trends over the last 12 months.
   Measure.
   Measure: Total material cost
   Dimension: Material (dimension attributes: MaterialName, PriceCategory)
   Dimension: Order Date (dimension attributes: Order month, Order year)

4. Compare the quantity of materials ordered from each supplier.
   Measure: Total quantity ordered
   Dimension: Supplier (dimension attribute: SupplierName)
   Dimension: Material (dimension attributes: MaterialName, Unit)

5. How much material is wasted due to damage during transportation from each warehouse to project sites?
   Measure: Total Lost Quantity
   Dimension: Warehouse (dimension attribute: WarehouseName)
   Dimension: Project (dimension attribute: ProjectName)
   Dimension: Material (dimension attribute: MaterialName, Unit)

6. Determine which materials are ordered most frequently.
   Measure: Total Quantity Ordered
   Dimension: Material (dimension attribute: MaterialName)
   Dimension: Order Date (dimension attributes: Order year, Order month)

7. What is the relationship between material wastage and total material costs?
   Measure: Total Lost Quantity
   Measure: Total Material Cost
   Dimension: Material (dimension attributes: MaterialName, Unit)
   Dimension: Warehouse (dimension attributes: WarehouseName)

8. Which materials are most prone to damage or loss during transportation, and how does this impact project costs?
   Measure: Total Lost Quantity
   Measure: Total Transfer Cost
   Dimension: Material (dimension attributes: MaterialName, Unit)
   Dimension: Project (dimension attributes: ProjectNamel)
   Dimension: Warehouse (dimension attributes: WarehouseName)

9. Which projects experience the highest material transportation costs, and how can we optimize them?
   Measure: Total Transfer Cost
   Dimension: Project (dimension attributes: ProjectName)
   Dimension: Warehouse (dimension attributes: WarehouseName)
   Dimension: Material (dimension attributes: MaterialName, Unit)

10. How frequently do materials arrive on time, and what percentage is delayed.
    Measure: Number of order facts
    Dimension: Junk(dimension attributes: IsDelayed)

Checking if there are Data in the Data sources needed to fill the Data warehouse

TABLE NAME	COLUMN	SOURCE

MaterialOrders	**One tuple describes one fact of material ordering.**	
	ID_Material	Foreign key from Material dimension table. Based on Material ordering system - Material table (MaterialID column).
	ID_Supplier	Foreign key from Supplier dimension table. Based on the Material ordering system – Supplier table (SupplierID column).
	ID_OrderDate	Foreign key from Date dimension table. Based on Material ordering system – Order table (OrderDate column)
	ID_DeliveryDate	Foreign key from Date dimension table. Based on the Material ordering system – Order table (DeliveryDate column).
	ID_ExpectedDeliveryDate	Foreign key from Date dimension table. Based on Material ordering system – Order table (ExpectedDeliveryDate column)
	ID_Warehouse	Foreign key from Warehouse dimension table. Based on Material ordering system – Order table (WarehouseID column)
	ID_Junk	Foreign key from Junk dimension table. Based on Material ordering system - Order table (DeliveryDate and ExpectedDeliveryDate).
	Quantity	Based on Material ordering system – Project_Material table (Quantity column)

	TransferCost	Based on Material ordering system – Order table (OrderCost column)
	Material Cost	Calculated as Quantity* UnitPrice, derived from Material ordering system – Material table (UnitPrice column)
MaterialTransfers	**One tuple describes one fact of material transfer from warehouse to project.**	
	ID_Material	Foreign key from Material dimension table. Based on Inventory management system – Material table (MaterialID column)
	ID_Warehouse	Foreign key from Warehouse dimension table. Based on Inventory management system – Warehouse table (WarehouseID column)
	ID_Project	Foreign key from Project dimension table. Based on Material ordering system – Project table (ProjectID column)
	ID_TransferDate	Foreign key from Date dimension table. Based on Inventory management system – Transfer table (TransferDate column)
	QuantityTransferred	Based on Inventory management system – Part_of_transfer table (QuantityMoved column)
	LostQuantity	Based on Inventory management system – Part_of_transfer table (LostQuantity column)
	TransferCost	Based on Inventory management system – Transfer table

		(TransferCost column)
Material	**One tuple describes one material in the specified price category with specified name.**	
	ID_Material	Material ID. Surrogate key – generated by the data warehouse.
	MaterialName	Based on Material ordering system – Material table (MaterialName column)
	Unit	Based on Material ordering system – Material table (Unit column).
	PriceCategory	Category based on UnitPrice from Material ordering system – Material table (UnitPrice column). <10 - very cheap 10 to 50 - cheap 50 to 100 - moderate 100 to 500 - expensive >500 - exclusive
Supplier	**One tuple describes one supplier with the specified name.**	
	ID_Supplier	Supplier ID. Surrogate key – generated by the data warehouse.
	SupplierName	Based on Material ordering system – Supplier table (SupplierName column).
Warehouse	**One tuple describes one warehouse with the specified name and location.**	
	ID_Warehouse	Warehouse ID. Surrogate key – generated by the data warehouse.
	WarehouseName	Based on Inventory management system – Warehouse table (WarehouseName column).
	Location	Based on Inventory

		management system – Warehouse table (Location column).
Project	**One tuple describes one construction project with the specified name, price category, location and project manager. (Implementation of SCD 2)**	
	ID_Project	Project ID. Surrogate key – generated by the data warehouse.
	ProjectName	Based on the Material ordering system – Project table (ProjectName column).
	Location	Based on Material ordering system – Project table (Location column).
	ProjectManagerPesel	Based on Material ordering system – Project table (ProjectManagerPesel column).
	MaterialPriceCategory	Calculated based on total MaterialCost in the data warehouse's MaterialOrders table, categorized as. [Economical, Standard, Premium] <10000 - economical 10000 to 100000 - standard >100000 - premium
	IsCurrent	"1" if information is current, otherwise "0" (SCD2 implementation).
Junk	**The tuples correspond to "all" possible combinations of values for isDelayed.**	
	ID_Junk	Junk ID. Surrogate key – generated by the data warehouse.
	IsDelayed	Indicates if the delivery was delayed by calculating the

		difference between DeliveryDate and ExpectedDeliveryDate. If the result is greater than 0 (i.e., DeliveryDate is later than ExpectedDeliveryDate), the IsDelayed flag in the Junk