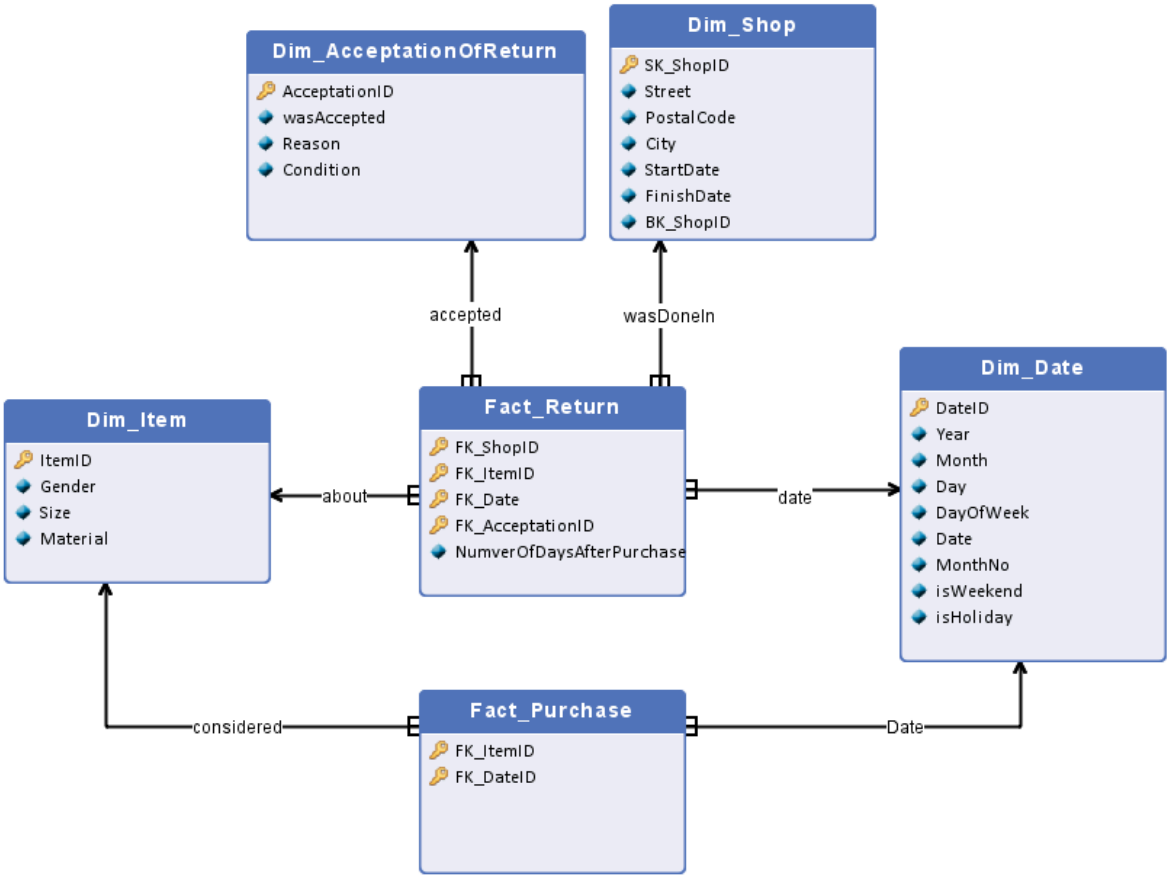
“Outspire” clothing store – Data Warehouse Design

Business process

The data warehouse is designed for business process “management of returns”. This process is described in the document Specification of business processes in Outspire clothing stores.



Fact_Purchase (fact table) One tuple describes one fact of purchasing an item.		
Attribute	*Attribute Type*	*Description*
FK_ItemID	Numerical	FK Item Indicate the item that was purchased
FK_DateID	string characters up to 20 digits	FK Date Indicate date of the purchase

Dim_Item (dimension table) One tuple describes one item in a shop.		
Attribute	*Attribute Type*	*Description*
ItemID	numerical	primary key
Gender	string characters up to 20 digits	Gender for which the product is aimed for. Allowed values: male, female
Size	char - only capital letters XS, S, M, L, XL	size of the item: XS - extra small, S - small, M - medium, L - large, XL - extra large
Material	string characters up to 20 digits	type of material the product is made of

Fact_Returns (fact table) One tuple describes one fact of return of previous purchase.		
Attribute	*Attribute Type*	*Description*
FK_ShopID	Numerical	FK Shop Indicate the shop that the return was made
FK_ItemID	Numerical	FK Item Indicate the item that was returned
FK_DateID	Numerical	FK Date Indicate the date of return
FK_AcceptationOfReturnID	Numerical	FK AcceptationOfReturn Indicate if the returned item was accepted
NumberOfDaysAfterPurchase	Numerical	Number of days after purchase, 1-30

Dim_Shop (dimension table) One tuple describes one shop.		
Attribute	*Attribute Type*	*Description*
ShopID	Numerical	PK – identification number
Street	string characters up to 20 digits	name of the street
PostalCode	string characters up to 20 digits	postal code
City	string characters up to 20 digits	City where shop is located
StartDate	DateTime	Start Date of functioning of the shop in the form YYYY-MM-DD e.g. 2022-04-12
FinishDate	DateTime	End Date of functioning of the shop in the form YYYY-MM-DD e.g. 2022-04-12

Dim_Date (dimension table) One tuple describes one date of return.		
Attribute	*Attribute Type*	*Description*
DateID	Numerical	PK – identification number
Date	DateTime	Date of return in form YYYY-MM-DD e.g. 2022-04-12
Year	Numerical, 4 digits	Year
Month	Varchar(10)	Month. Allowed values: January, February,March, April, May, June, July, August, September, October, November and December.
MonthNo	Numerical, 2 digits	Month’s numeric value

Day	Numerical, 2 digits	Day's numeric value
DayOfTheWeek	string characters up to 20 digits	Day of week. Allowed values: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday
isWeekend	Varchar(12)	Allowed values: weekend, workingDay
isHoliday	Varchar(12)	Allowed values: isChistmas, inNotChristmas

Dim_AcceptationOfReturn (dimension table) One tuple describes one acceptance of return.		
Attribute	*Attribute Type*	*Description*
AcceptationID	Numerical	PK – identification number
wasAccepted	string characters up to 20 digits	Allowed values: isAccepted, isNotAccepted
Reason	string characters up to 20 digits	Reason of the return
Condition	string characters up to 20 digits	Condition of returned item

Dimensional model

Fact definitions

1. Fact 1: Return fact: Fact describing return of a specific item, done on a specific date, to a particular shop. Return is also characterized by criteria such as acceptance of return, reason of return and condition of returned item(s).

   Fact table: Return

   Granularity:
   - a specified date of return
   - a specified shop where return is made
   - a specified characteristic of the return with its acceptance, reason and condition
   - a specified item for particular gender

Measures and aggregation functions:
Number of returns - COUNT (1)
Number of returns done at the last possible day (30th day) in comparison to all returns - Sum()
Average number of returns - SUM(FK_ShopID)/ COUNT(1)

2. Fact 2: Purchase fact: Fact describing purchase of a particular item on a specific date

   Fact table: Purchase

   Granularity:
   - a specified item for particular gender
   - a specified date of purchase.

   Measures and aggregation functions:
   Number of purchases - COUNT (1)
   Average number of purchases - SUM(FK_ItemID)/ COUNT(1)

Dimension definitions

1. Dimensions for fact 1: Return Fact

DIMENSION/ DIMENSION ATTRIBUTE	*TABLE/ COLUMN*	*TYPE*
RETURN DATE HIERARCHY	• Dim_Date.Year •• Dim_Date.Month ••• Dim_Date.Date	Hierarchical dimension
RETURN DATE	Dim_Date	Dimension
RETURN YEAR	Dim_Date.Year	Dimension attribute
RETURN MONTH	Dim_Date.Month	Dimension attribute
RETURN DAY	Dim_Date.Day	Dimension attribute
RETURN DAY OF THE WEEK	Dim_Date.DayOfTheWeek	Dimension attribute
RETURN MONTH NUMBER	Dim_Date.MonthNo	Dimension attribute

WEEKEND DAY	Dim_Date.isWeekend	Dimension attribute
HOLIDAY DAY	Dim_Date.isHoliday	Dimension attribute
SHOP	Dim_Shop	Dimension
SHOP ADDRESS HIERARCHY	• Dim_Shop.City •• Dim_Shop.PostalCode ••• Dim_Shop.Street	Hierarchical dimension
STREET ADDRESS	Dim_Shop.Street	Dimension attribute
POSTAL CODE	Dim_Shop.PostalCode	Dimension attribute
CITY	Dim_Shop.City	Dimension attribute
START DATE OF FUNCTIONING THE SHOP	Dim_Shop.StartDate	Dimension attribute
FINISH DATE OF FUNCTIONING THE SHOP	Dim_Shop.FinishDate	Dimension attribute
RETURNED ITEM	Dim_Item	Dimension
RETURNED ITEM GENDER	Dim_Item.Gender	Dimension attribute
ITEM SIZE	Dim_Item.Size	Dimension attribute
ITEM MATERIAL	Dim_Item.Material	Dimension attribute
RETURN ACCEPTATION	Dim_AcceptationOfReturn	Dimension
WAS RETURN ACCEPTED	Dim_AcceptationOfReturn.wasAccepted	Dimension attribute
REASON OF RETURN	Dim_AcceptationOfReturn.Reason	Dimension attribute
CONDITION OF RETURNED ITEM	Dim_AcceptationOfReturn.Condition	Dimension attribute

2. Dimensions for fact 2: Purchase Fact

DIMENSION/ DIMENSION ATTRIBUTE	*TABLE/ COLUMN*	*TYPE*
PURCHASE DATE HIERARCHY	• Dim_Date.Year •• Dim_Date.Month ••• Dim_Date.Date	Hierarchical dimension
PURCHASE DATE	Dim_Date	Dimension
PURCHASE YEAR	Dim_Date.Year	Dimension attribute
PURCHASE MONTH	Dim_Date.Month	Dimension attribute
PURCHASE DAY	Dim_Date.Day	Dimension attribute
PURCHASE DAY OF THE WEEK	Dim_Date.DayOfTheWeek	Dimension attribute
PURCHASE MONTH NUMBER	Dim_Date.MonthNo	Dimension attribute
ITEM	Dim_Item	Dimension
ITEM GENDER	Dim_Item.Gender	Dimension attribute
ITEM SIZE	Dim_Item.Size	Dimension attribute
ITEM MATERIAL	Dim_Item.Material	Dimension attribute

Checking the feasibility of queries based on the multidimensional model

I. What is the number of returns during the analyzed month compared to the previous one?
Measure: Number of returns,
Dimension: Return date (dimension attributes: Return month)
Dimension: Return date (dimension attributes: Return month)

II. What percentage of returned items went back again to the sale in the analyzed month?
Measure: Percentage of returned items - BackOnSale/ Number of returned items,
Dimension: Number of returned items (dimension attributes: ReturnID)
Dimension: BackOnSale (dimension attributes: Condition)

III. Are there more returns on female or male clothes in a particular year?
Measure: Number of returns,
Dimension: Item (dimension attributes: Gender)
Dimension: Return date (dimension attributes: Return Year)

IV. Is there a relation between the shop and the number of returns to this shop in a given month?
Measure: Number of returns in shops,
Dimension: Shop (dimension attributes: ShopID)
Dimension: Number of returns (dimension attributes: ReturnID)
Dimension: Return Date (dimension attributes: Return Month)

V. In which cities the stores accept the most returns in a given month?
Measure: Number of returns
Dimension: Shop (dimension attributes: Shop ID?)
Dimension: Number of returns (dimension attributes: Return ID)
Dimension: Return (dimension attributes: WasAccepted)
Dimension: Return Date (dimension attributes: Return Month)

VI. In which cities there are more than average number of returns on weekends during the last 6 months?
Measure: Number of returns
Dimension: Shop (dimension attributes: Shop ID?)
Dimension: Number of returns (dimension attributes: Return ID)
Dimension: Return Date (dimension attributes: Return Month)
Dimension: Return Date (dimension attributes: DayOfReturn)

VII. After how many days the items are returned after purchase to shops in the last 2 months?
Measure: Number of days after purchase
Dimension: Shop (dimension attributes: Shop ID)
Dimension: Number of returns (dimension attributes: Return ID)
Dimension: Return Date (dimension attributes: Return Month)

VIII. In which cities there are more than average number of returns per month in comparison to the same month in the previous year?
Measure: number of returns
Dimension: Shop (dimension attributes: Shop ID)
Dimension: Number of returns (dimension attributes: Return ID)
Dimension: Return Date (dimension attributes: Return Month)

IX. Compare the average number of returns on particular working days in the last week.
Measure: Average number of days after purchase
Dimension: Shop (dimension attributes: Shop ID?)
Dimension: Number of returns (dimension attributes: Return ID)
Dimension: Return Date (dimension attributes: Return Month)
Dimension: Return Date (dimension attributes: Day of the week)

X. Is there an average number of returns bigger/smaller the week after Christmas in comparison to the week before last year?
Measure: number of returns
Dimension: Shop (dimension attributes: Shop ID)
Dimension: Number of returns (dimension attributes: Return ID)
Dimension: Return Date (dimension attributes: Return Day)

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

Fact_Purchase One tuple describes one fact of purchase.	
Column	*Source*
FK_ItemID	Item ID. Foreign key from dimension. Based on Item ID in relational data source.
FK_DateID	DateID. Foreign key from dimension. Based on Date ID in relational data source.

Dim_Item One tuple describes one item in a shop.	
Column	*Source*
ItemID	Item Id. Surrogate key - generated by database.
Size	Size of the item taken from the Size column from the Item table in relational data source.
Material	Material of the item taken from the Material column from the Item table in relational data source.
Gender	Gender as a category of product taken from the Gender column from the Item table in relational data source.

Fact_Returns One tuple describes one fact of return of previous purchase.	
Column	*Source*
FK_ShopID	Shop ID. Foreign key from dimension. Based on Shop ID in relational data source.
FK_ItemID	Item ID. Foreign key from dimension. Based on Item ID in relational data source.
FK_DateID	Date ID. Foreign key from dimension. Based on DateID in relational data source.
FK_AcceptationOfReturnID	Acceptation ID. Foreign key from dimension. Based on Acceptation ID in relational data source.
NumberOfDaysAfterPurchase	Number Of Days After Purchase taken from NumberOfDaysAfterPurchase column from Return table in relational data source.

Dim_Shop One tuple describes one shop.	
Column	*Source*
SK_ShopID	Shop Id. Surrogate key - generated by database.
BK_ShopID	Business key. Based on value stored in column A of Sheet 1
Street	Street where the shop is located. Based on value stored in column B of Sheet 1
PostalCode	Postal code of the place where the shop is located. Based on value stored in column C of Sheet 1
City	City where the shop is located. Based on value stored in column D of Sheet 1
StartDate	Start Date of functioning of the shop. Based on value stored in column E of Sheet 1

FinishDate	Start Date of functioning of the shop. Based on value stored in column F of Sheet 1

Dim_Date One tuple describes one date of return.
One tuple describes one day. All the data in this table are generated tuple by tuple based on any calendar, before ETL process.

Dim_AcceptationOfReturn One tuple describes one acceptance of return.	
Column	*Source*
AcceptationID	Shop Id. Surrogate key - generated by database.
wasAccepted	Allowed values: isAccepted, isNotAccepted Taken from wasAccepted column from Return table in relational data source.
Reason	Reason for the return. Taken from Reasoncolumn from Return table in relational data source.
Condition	Condition of returned item. Taken from Condition column from Return table in relational data source.