DESIGN DOCUMENT

NopCommerce Group 2

IST 722

Abhijeet Mali, Dennis Oh, Gauri Komawar, Ishani Jariwala, Swati Nibban

Table of Contents

OVERVIEW

PROJECT SCOPE / MISSION

The scope of this project is to create a data warehouse, provide data analysis, visualizations and insights, and to arrange a system for employees of nopCommerce to make further analysis, insights, and visualizations via user-friendly dashboards.

TEAM MEMBERS & ROLES

Role	Description	Assigned to		Phase
Business Lead	In Charge of Initiative	Swati Nibban		1
Project Manager	Manages Project	Gauri Komawar	Ishani Jariwala	1
Business Analyst	Collects Requirements	Dennis Oh	Abhijeet Mali	1
Data Architect	Dimensional Modeling / Implementation	Dennis Oh	Gauri Komawar	2
ETL Architect	ETL Design / Implementation	Abhijeet Mali	Swati Nibban	2
BI Architect	BI Design / Implementation	Ishani Jariwala		2

OVERVIEW

KEY STAKEHOLDERS

```
┌─────────────────┐        ┌─────────────────┐
│    Internal     │        │    External     │
│  Stakeholders   │        │  Stakeholders   │
└─────────────────┘        └─────────────────┘

    Business                  Manufacturers
    Owner[s] of               / Suppliers
    nopCommerce
    Store

    Business                  Customers
    Manager[s] of
    nopCommerce
    Store

    Business                  Shareholders
    Lead[s]

    Project                   Competitors
    Manager[s]

    Business
    Analyst[s]

    Data
    Architect[s]

    ETL
    Architect[s]

    BI Architect[s]

    Other
    Employees
```
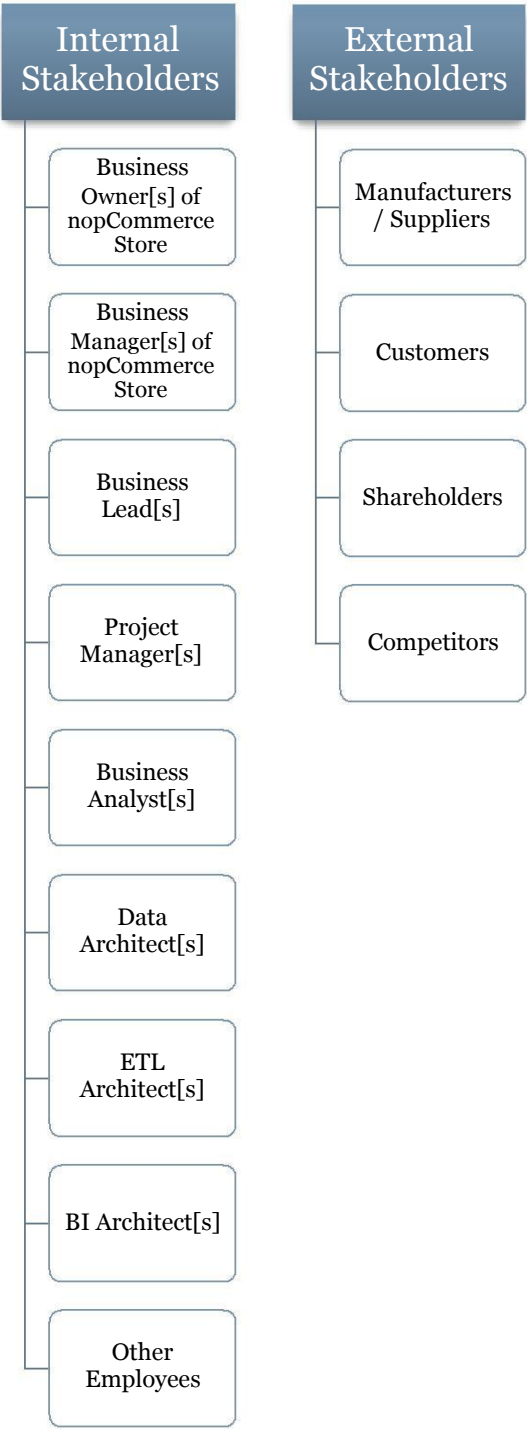
ANALYSIS

BUSINESS PROCESSES

The nopCommerce (Group 2) offers a selection of various product categories; they include desktops, notebooks, software, cameras, cell phones, other electronics, shoes, clothing, other apparel, digital downloads for music, books, jewelry and gift cards. For each category, the store provides around 15 products. The store experiences global sales; customers shop from countries such as Austria, India, United States, France and Germany. Customers also have the ability to leave ratings and reviews for products they purchase. Finally, the store offers various discounts/promotions. For example, there was a "Spring" discount code during spring break. The following are the business processes:

1. Sales Reporting:

   Sales are the top priority for an e-commerce business like nopCommerce. Therefore, it is important to analyze sales data in order to study, understand and make insights to optimize business and increase profit. For this particular nopCommerce store, we focus on exploring purchasing habits (which product categories are they buying?) of customers based on their varying demographics (age, gender, location). This will ultimately reveal where the most sales are coming from. Note: This business process has two fact tables: one that allows analysis at an order item level and another summarized one which allows analysis at order level.

   Data Questions to Explore:

   - *Which are the best selling products, based on quantity sold?*
   - *Which are the best selling products, based on amount (sales) sold?*
   - *Which are the popular product categories/products based on customer age, gender and location?*
   - *Which products have the highest average rating?*
   - *Which products do customer leave the most ratings for?*
   - *Which customers (age, gender, location) are the frequent buyers?*
   - *How do discounts affect sales?*

2. Shipping Orders Reporting:

   Shipping lag is the time elapsed between the time customer places an order and the time the order is shipped and delivered to the customer. Ideally, a customer's order should be shipped promptly after the minimal processing time. However, there are many cases when a lag may occur: shipping to certain countries, errors by customers during

checkout, unavailable products, etc. An analysis of shipping lag times will reveal longer than average times and will initiate a fix.

Data Questions to Explore:

> *What is the average ship lag?*
> *Which locations experience the most ship lag?*
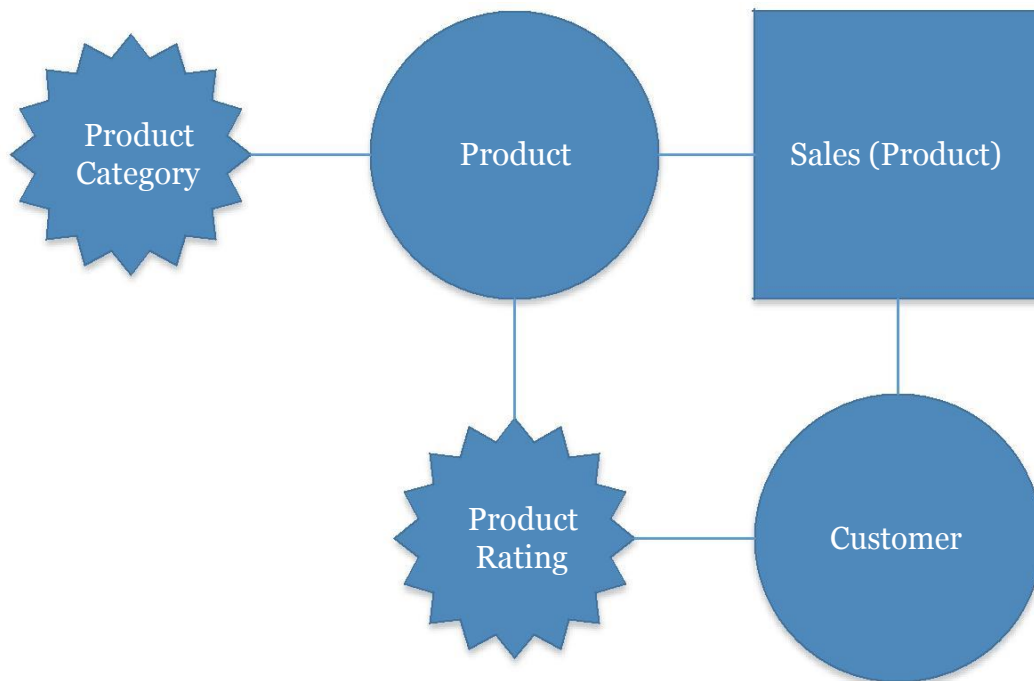> *Do certain shipping methods experience a longer ship lag?*

BUS MATRIX

The 'Sales Reporting' business process has two fact tables: *FactSalesProduct* and *FactSalesOrder*. The former allows analysis to be performed on products, while the latter allows analysis for orders. The *Product Category* and *Product Rating* dimensions are not listed on the bus matrix because of the snowflake schema: *Product Category* dimension is linked to the *Product* dimension, and the *Product Rating* dimension acts as a bridge table for the *Product* and *Customer* dimensions. The 'Ship Lag' business process uses the *FactSalesOrder* fact table, which was also used in the 'Sales Reporting' business process. This is possible because the two processes have the same granularity.
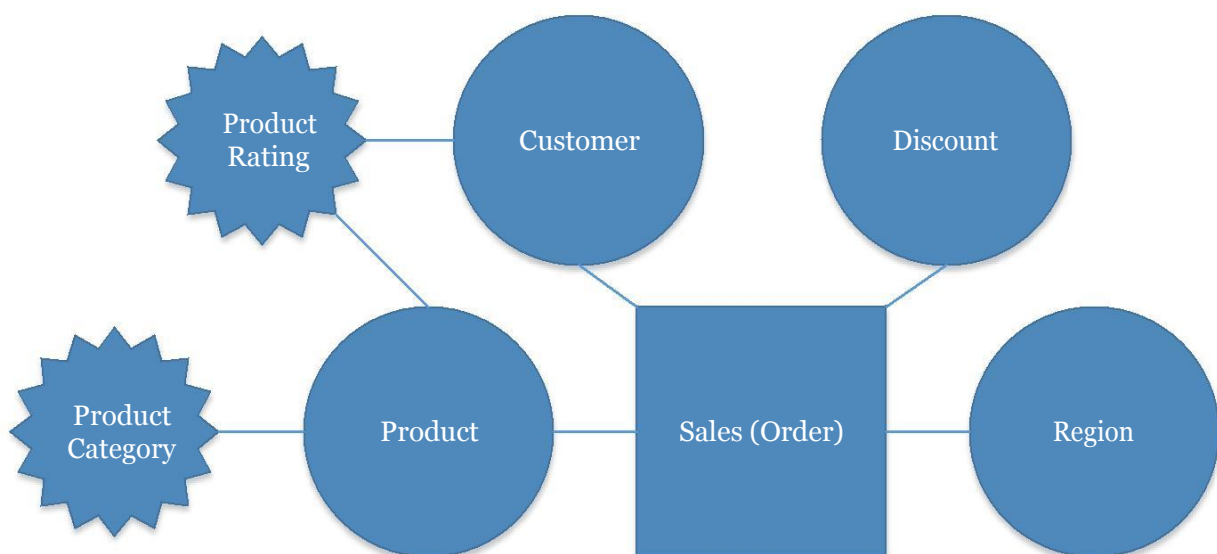
Please refer to the "IST722_GROUP2_HIGH_LEVEL.xlsx" file.

ANALYSIS

BUBBLE CHART FOR SALES (PRODUCT)

```
Product Category — Product — Sales (Product)
                     |            |
                Product Rating — Customer
```

BUBBLE CHART FOR SALES (ORDER)

```
Product Rating — Customer        Discount
           |          |             |
Product Category — Product — Sales (Order) — Region
```

ANALYSIS

ATTRIBUTE LIST

Dimension / Fact Table	Attribute / Fact Name	Description
DimProduct	ProductId	Business key from source system (aka natural key)
DimProduct	ProductName	Name of Product
DimProduct	Price	Price of Product
DimProduct	ApprovedRatingSum	Sum of all Ratings
DimProduct	ApprovedTotalReviews	Total Number of Ratings
DimProduct	NotApprovedRatingSum	Sum of all Ratings not approved
DimProduct	NotApprovedTotalReviews	Total Number of Ratings not approved
DimProduct	IsGiftCard	Used a gift card or not
DimProduct	AdditionalShippingCharge	Additional shipping charge
DimProduct	StockQuantity	Quantity in stock
DimProduct	OldPrice	The old price of a product
DimProduct	ProductCost	Purchased cost
DimProduct	HasDiscountsApplied	Any discounts applied?
DimProduct	CategoryKey	Key to Category
DimProduct	ShortDescription	Description of a product
DimCustomer	CustomerId	ID of Customer
DimCustomer	City	City of Customer
DimCustomer	Address1	Address line 1 of Customer
DimCustomer	Address2	Address line 2 of Customer
DimCustomer	ZipPostalCode	Zip Code of Customer
DimCustomer	Username	Username of Customer
DimCustomer	FirstName	First Name of Customer
DimCustomer	LastName	Last Name of Customer
DimCustomer	Gender	Gender of Customer
DimCustomer	Email	Email of Customer
DimCustomer	Birthdate	Birth Date of Customer
DimCustomer	Company	Company of Customer
DimDiscount	DiscountID	Business key from source system (aka natural key)
DimDiscount	DiscountName	Name of Discount
DimDiscount	DiscountPercentage	Percentage of Discount
DimDiscount	DiscountAmount	Amount of Discount
DimDiscount	DiscountMaximumAmount	Maximum Amount of Discount
DimDiscount	CouponCode	The Code for the Discount
DimProductReview	ReviewId	Business key from source system (aka natural key)
DimProductReview	IsApproved	Was the review approved?
DimProductReview	Rating	Rating

ANALYSIS

Dimension / Fact Table	Attribute / Fact Name	Description
DimProductCategory	CategoryId	Business key from source system (aka natural key)
DimProductCategory	CategoryName	Name of Category
DimProductCategory	ParentCategoryId	Id of Parent Category
DimProductCategory	Deleted	Has it been deleted?
DimRegion	StateId	Business key from source system (aka natural key)
DimRegion	CountryId	Business key from source system (aka natural key)
DimRegion	StateName	Name of State
DimRegion	Abbreviaion	Abbreviation of State
DimRegion	CountryName	Name of Country
FactSalesOrder	OrderID	The natural key for the fact table (degenerate dimension)
FactSalesOrder	CustomerKey	Key to Customer
FactSalesOrder	RegionKey	Key to Region
FactSalesOrder	OrderSubtotalInclTax	Subtotal of Order Including Tax
FactSalesOrder	OrderSubtotalExclTax	Subtotal of Order Excluding Tax
FactSalesOrder	OrderShippingInclTax	Shipping Cost Including Tax
FactSalesOrder	OrderShippingExclTax	Shipping Cost Excluding Tax
FactSalesOrder	DiscountKey	Key to Discount
FactSalesOrder	OrderDiscount	Discount Amount of Order
FactSalesOrder	OrderTotal	Total Amount after Discounts
FactSalesOrder	RewardPointsWereAdded	Were Reward Points Added
FactSalesOrder	ShippingMethod	Method of Shipping
FactSalesOrder	OrderDateKey	Key to Date
FactSalesOrder	ShippedDateKey	Key to Date
FactSalesOrder	DeliveryDateKey	Key to Date
FactSalesOrder	ShipLag	ShippedDate - OrderDate
FactSalesProduct	ProductKey	Key to Product

ANALYSIS

Dimension / Fact Table	Attribute / Fact Name	Description
FactSalesProduct	OrderId	Order ID, Primary Key, Natural Key
FactSalesProduct	CustomerKey	Key to Customer
FactSalesProduct	Quantity	Quantity Sold of Item on Order
FactSalesProduct	UnitPriceInclTax	Price of Unit of unique product on Order - Including Tax
FactSalesProduct	UnitPriceExclTax	Price of Unit on Order - Excluding Tax
FactSalesProduct	PriceInclTax	Total Price of Units on Order (for a unique product) - Including Tax
FactSalesProduct	PriceExclTax	Total Price of Units on Order (for a unique product) - Excluding Tax
FactSalesProduct	DiscountAmountInclTax	Discount Amount Including Tax
FactSalesProduct	DiscountAmountExclTax	Discount Amount Excluding Tax
FactSalesProduct	OriginalProductCost	Original cost of a Product
FactSalesProduct	ItemWeight	Weight of Item
FactSalesProduct	AverageRatings	ApprovedRatingSum / ApprovedTotalReviews

ANALYSIS

ISSUES LIST

Please refer to the "IST722_GROUP2_HIGH_LEVEL.xlsx" file.

DESIGN

DESIGN

DETAILED BUS MATRIX

Please refer to the "IST722_GROUP2_HIGH_LEVEL.xlsx" file.

FACT / DERIVED FACT WORKSHEET FOR FACTSALESPRODUCT

Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.

FACT / DERIVED FACT WORKSHEET FOR FACTSALESORDER

Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.

DETAILED DIMENSIONAL DESIGN WORKSHEET (PRODUCT)

Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.

DETAILED DIMENSIONAL DESIGN WORKSHEET (DISCOUNT)

Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.

DETAILED DIMENSIONAL DESIGN WORKSHEET (CATEGORY)

Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.

DETAILED DIMENSIONAL DESIGN WORKSHEET (CUSTOMER)

Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.

DETAILED DIMENSIONAL DESIGN WORKSHEET (REGION)

Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.

DETAILED DIMENSIONAL DESIGN WORKSHEET (PRODUCTREVIEW)

*Please refer to the "IST722_GROUP2_DETAILED_MODELING.xlsx" file.*

DESIGN

DIMENSIONAL HIERARCHIES

Example of Year-Quarter-Month hierarchy: 2016 → 1st Quarter → February → 2/12/2016.

Example of Weekday-Day (right): Y → Monday → 2/12/2016.

⚠ Year-Quarter-Month ⊠
• Year
•• Quarter
⁜ Month
⁑ Date ⤓
<new level>

⚠ Weekday-Day ⊠
• Is A Weekday
•• Day Of Week
⁜ Date ⤓
<new level>
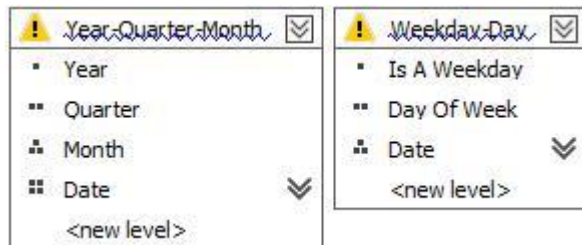
Example of Country-State: USA → New York

Example of Gender-Customer hierarchy : Male → India

⚠ Country-State ⊠
• Country Name
•• State Name
<new level>

⚠ Gender-Customer ⊠
• Gender
•• Country Name
<new level>
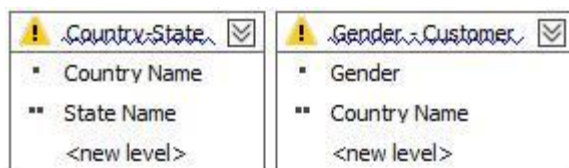
Example of Approved-Rating hierarchy: Y → 5

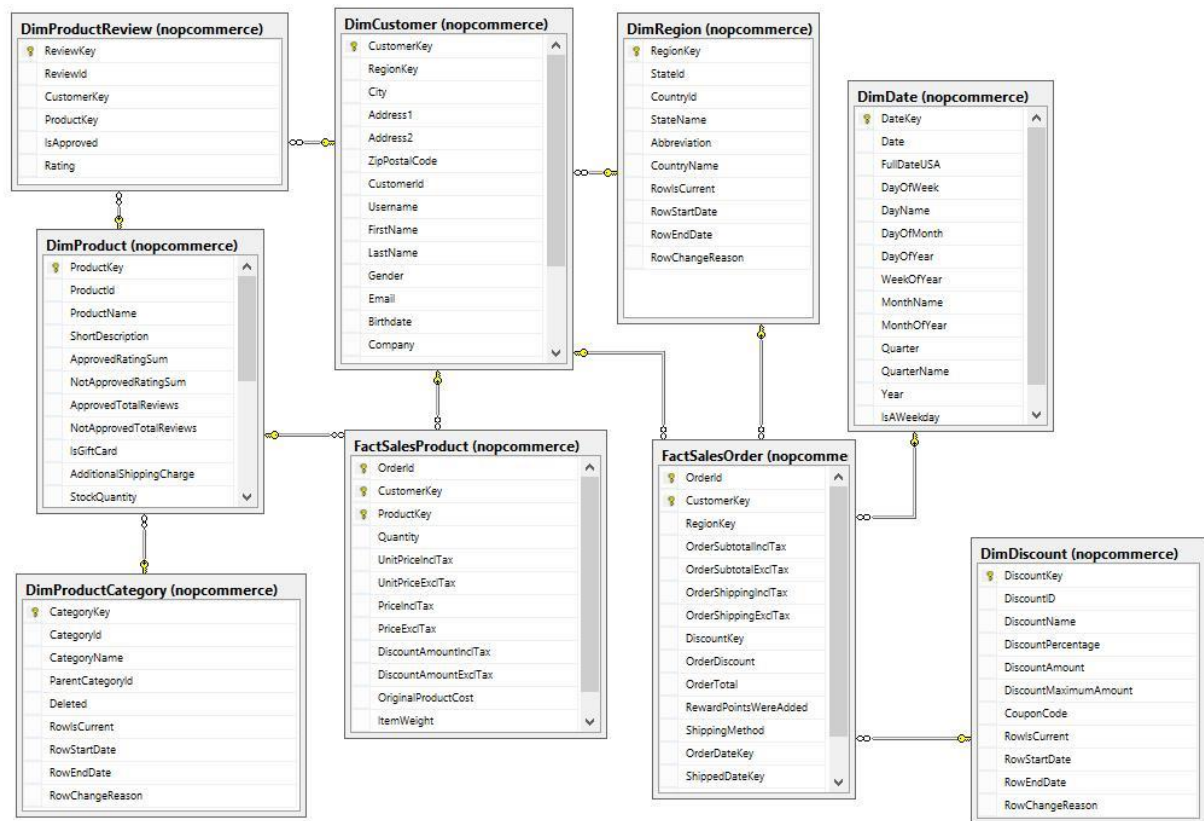| Approved-Rating ⊠ |
|---|
| • Is Approved |
| •• Review ⤓ |
| <new level> |

Example of Discount-Product hierarchy: Y → Apple iPhone

Example of Category hierarchy: Notebooks → Apple MacBook

Discount-Product ⊠
• Has Discounts Applied
•• Product ⤓
<new level>

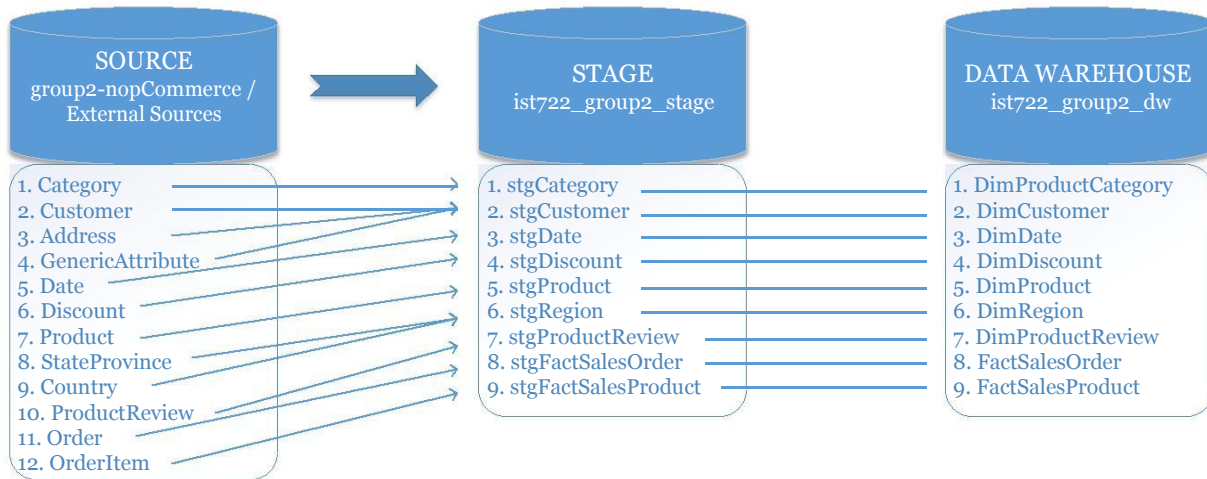| Category-Hierarchy ⊠ |
|---|
| • Category |
| •• Product ⤓ |
| <new level> |
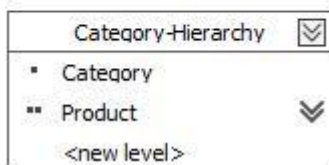
DESIGN

SNOWFLAKE MODEL DIAGRAMS

DESIGN

ETL SPECIFICATIONS

DETAILED ETL FLOW

Please refer to the "IST722_GROUP2_DETAILED_SRC_TGT.xlsx" file.
Further details can be observed in the SSIS ETL File.

TEAM CONTRIBUTION REPORT

TASK	Abhijeet Mali	Dennis Oh	Gauri Komawar	Ishani Jariwala	Swati Nibban
Determine Business Processes	5	5	5	5	5
Bus Matrix	5	5	4	4	4
Detailed Design Matrix	5	5	4	4	4
SSIS - ETL	4	4	5	5	5
SSAS - MOLAP	4	4	5	5	5
BI	4	4	5	5	5
Project Design Document	5	5	4	4	4
Poster	5	5	4	4	4