

Assignment 6 - Part 1

DATAWAREHOUSING BY DR IMRAN KHAN

AMIN NIZAR ALI

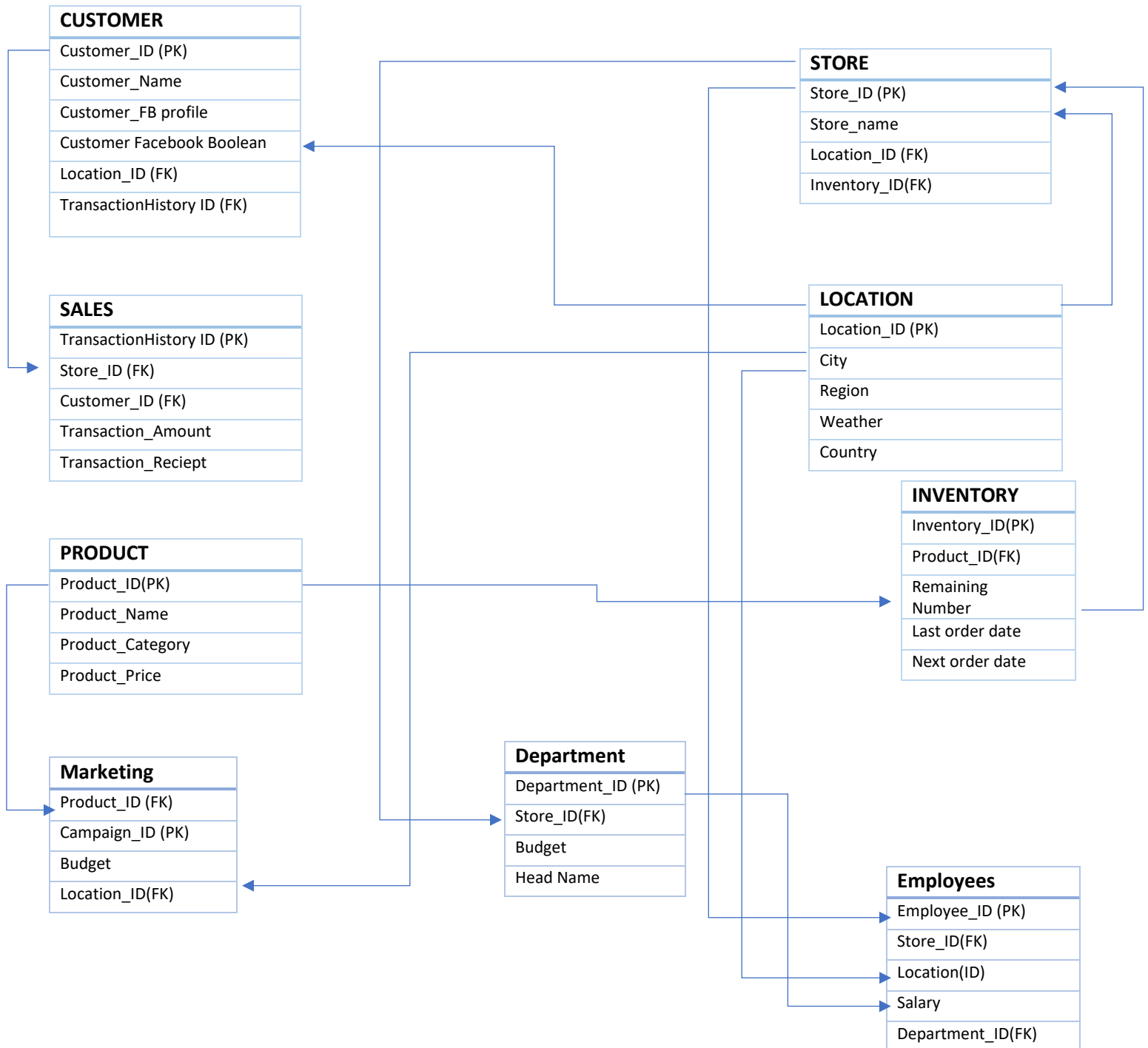
13183

DELIEVERABLE 1: SOURCES

The data sources I have come out with will be coming through 3 different places

1. From the Enterprise/shops
 - a. UPC Scanners
 - b. Inventory/Stock
 - c. Transactional Data
 - d. Shelf Space Data
 - e. CCTV data (for customer behavior)
 - f. Departmental Data
 - g. Footfalls
 - h. Segmentations
 - i. Sales Data
2. Online
 - a. App Data
 - b. Website Data
 - c. Geographical data
 - d. Weather data
3. Customer Data
 - a. Unique customer details(name, address etc)
 - b. Loyalty card data
 - c. Social media presences
 - d. Emotional behavior
 - e. Purchase history

DELIEVERABLE 2: ERD DIAGRAM



DELIVERABLE 2: DATAWAREHOUSE CONCEPTUAL DESIGN

A: DATAWAREHOUSE BUS MATRIX

The enterprise data warehouse bus matrix is the essential tool for designing and communicating the enterprise data warehouse bus architecture. The rows of the matrix are business processes and the columns are dimensions.

The following Enterprise Data Warehouse Bus Matrix for our retail store, will be a key design tool representing the organization's core business processes and associated dimensionality.

BUSINESS PROCESSES	Product	Warehouse	Store	Customer	Promotions
Stock Receiving	X	X			
Store Delivery	X	X	X		X
Store Inventory	X		X		
Retail Sales	X		X	X	X
Customer Returns	X		X	X	X
Loyalty Card Usage	X		X	X	X
Receipt Making	X		X	X	X
Request New Stock	X	X	X		

B: INFORMATION PACKAGE FOR EACH PROCESS

1. Stock Receiving:
Product, sender, receiver and date.
2. Store Delivery:
Product, Sender, Store , Cost and the date
3. Store Inventory:
Product, Number of Units, Store, and date
4. Retail Sales:
Product, Customer, Transaction and date
5. Customer Returns:
Product, Customer, Product condition, purchase date and return date.
6. Loyalty Card Usage
Customer, Customer History, and promotions
7. Receipt Making
Transaction, customer, product, amount and date.
8. Request New Stock
Inventory, Units, product and date.

C: SCHEMA

Tables with same ERD but with 3 added fact tables

FACT TABLE MARKETING
Campaign_ID (PK)
Cost
Social Media presence
New Customers

FACT TABLE CEO
Campaign_ID (FK)
Location_ID(FK)
Department_ID(FK)
Store_ID(FK)
Product_ID (FK)
TransactionHistory ID (FK)
Inventory_ID(FK)
Customer_ID (PK)

FACT TABLE PRODUCTION
Inventory_ID(FK)
Store_ID(FK)
Product_ID (FK)
Remaing Units at shop
Stocked Units
Last delivery Date
Next Delivery date

D: FACT TABLES

No	FACT TABLE NAME	FACT GRANULARITY	FACT TABLE TYPE	BRIEF JUSTIFICATION
1	FACT TABLE CEO	Campaign_ID (FK)	FACTLESS TABLE	Marketing campaign unique ID
		Location_ID(FK)		City and Country
		Department_ID(FK)		Unique ID of the departments
		Store_ID(FK)		Unique ID of all the stores
		Product_ID (FK)		Unique ID of all the products sold by the enterprise
		TransactionHistory ID (FK)		Unique ID of all the transactions taking place
		Inventory_ID(FK)		ID of all the stocks and deliveries at the inventory
		Customer_ID (PK)		Unique ID of all our previous and current customers
2	FACT TABLE MARKETING	Campaign_ID (FK)	SEMI ADDITIVE	Marketing campaign unique ID
		Cost		Cost of the campaign done
		Social Media presence		Boolean for FB, Insta etc.
		New customers		The return of the campaign

<u>3</u>	<u>FACT TABLE PRODUCTION</u>	Inventory_ID(FK)	<u>SEMI ADDITIVE</u>	Unique ID of all the transactions taking place
		Store_ID(FK)		Unique ID of all the stores
		Product_ID (FK)		Unique ID of all the products sold by the enterprise
		Remaining Units at shop		Numbers of units left at the shop
		Stocked Units		Number of units stored in the inventory
		Last delivery Date		-
		Next Delivery date		-

E: DIMENSION TABLES

<u>No</u>	<u>DIMENSION TABLE NAME</u>	<u>BRIEF JUSTIFICATION</u>	<u>HIERARCHIES</u>
<u>1</u>	<u>Customers</u>	Data for all the unique Customers	Customer_FB profile Customer Facebook Boolean
<u>2</u>	<u>Product</u>	Data for all the unique Products with their categories	Product Category
<u>3</u>	<u>Location</u>	Geographical data for all the shops and customers	City Region
<u>4</u>	<u>Sales</u>	Data for all the unique transactions with customers	None
<u>5</u>	<u>Store</u>	Data for all the stores present in the enterprise	None
<u>6</u>	<u>Inventory</u>	Unit based data for all the products at store and at stock	None
<u>7</u>	<u>Employees</u>	Salary, dept and location of all the employees working for the enterprise	None
<u>8</u>	<u>Department</u>	Store wise Data and details of all the departments.	None
<u>9</u>	<u>Marketing</u>	Current and previous marketing data.	None

F: DESIGN FEATURES

<u>DESIGN FEATURES</u>	<u>BRIEF DESCRIPTION</u>	<u>BRIEF JUSTIFICATION</u>
Store and Inventory	The store dimension can contain a reference to Inventory dimension table.	The store will need to trigger the inventory if units are about finish.
Product and Marketing	The Marketing dimension can contain a reference to product dimension table.	The Marketing dimension must know the best and worst performing products
Location and Marketing	The Marketing dimension can contain a reference to location dimension table.	The Marketing team must know the best areas and the worst areas performing by the sales and profit.
Store and Department	The department dimension can contain a reference to store dimension table.	We will need to know the progress and performance of every department functioning on each store.

G: SOLUTION TO THE PROBLEMS

1:What are top 3 stores have the highest sales across the country?

Answer - To answer this question we need to select store table and fact sales for retrieving all records of stores who's having highest sales.

2:Is there any geographical or regional impact on the stores?

Answer - To answer this question we need to select store table and their location for retrieving all records of stores and comparing them with their location to gather insights.

3: How many customers are regular using our specific products? Which product is the most selling product?

Answer - The customer dimension and fact sales will select and the we need retrieve the product detail which is purchased by customers frequently.

4: Are there certain times of the year when more products are sold? Which day of the week more products sold?

Answer - For this question we need to select the dimensional table Sales .So that we can know the month or time in which that particular product is in high demand .

Deliverable 4: Procedure

Step 1: Sources were identified through the diagram of a modern Datawarehouse provided in the PDF file.

Step 2: An ODS was designed on Word, first the tables were designed and connected. Primary and foreign keys were designed then the other attributes were added

Step 3: The Bus Matrix was designed after thinking about all the possible processes that can take place at a retail shop and its inventory.

Step 4: The information packages were made for all the process by noting all the people and sub processes or what data would be collected or used by each process.

Step 5: Fact tables were made for the CEO, Marketing Head and the Production head. Relevant attributes were added based on their usage.

Step 6: The solution to the problems were provided

Deliverable 5: Presentation Layer

