Project Proposal

Background:

Warmate is a global supermarket company based in Southeast Asia that owns a network of hypermarkets, discount department stores, and convenience stores. In a typical case, it is generally very capable of balancing demand and supply. However, the situation has changed, and the company is now having difficulty releasing goods from both the supermarket and the factory due to unforeseen circumstances. Excess goods spoil and become obsolete, causing financial hardship.

Warmate has recently disclosed their inventory and discovered that there is a significant amount of excess inventory, which could result in bad cash flow and unnecessary costs. As a result, the corporation wishes to have a more comprehensive understanding of the problem of supply and demand alignment. One particular issue that contributed to this misalignment of demand and supply is that sales/marketing-driven pricing incentives do not necessarily correspond to procurement and manufacturing capability. As a consequence, there is a stockout due to an increase in demand. After the stock was heavily replenished, the firm had already lost buyers, resulting in surplus inventory in supermarkets and warehouses. Demand shaping is one approach the team wanted to use in order to change consumer demand for their surplus goods and re-attract them.

Following a sudden shift in the situation, the business requires integrated reports in order to produce real-time advertisements and purchases. As a result, the next phase of Warmate's data warehouse and business intelligence platform will have to incorporate sales, manufacturing, and distribution, as well as financials, so that solutions can be evaluated financially. This will give the corporation a bird's-eye view of the effect of a transition as a consequence of a product promotion on the organization's capacity to satisfy the new demand and how to optimize supply and demand accordingly.

Problems:

As mentioned above, users want information to be shared in a timely, reliable, and safe manner in order to make the right decision. However, due to the large number of databases from various departments, obtaining the data in a timely and efficient manner necessitates a concerted effort among the current structures. Instead of attempting to navigate the various databases that exist, a data warehouse can be used to provide a centralized repository for the collection and exchange of data that employee can conveniently use to make better business decisions.

One of the most common problem for this Warmate of not providing a Data Warehouse is the way employee access data in the form of reports. Many that need information or knowledge from the data they use are often forced to wait for a response depending on someone else's timeline. Furthermore, they may have to manipulate data inside an application such as Microsoft Excel to suit their needs after they receive the report, which can lead to errors or miscommunication. Apart from challenging of business perspectives, following technical issues could also be addressed in this Warmate organization.

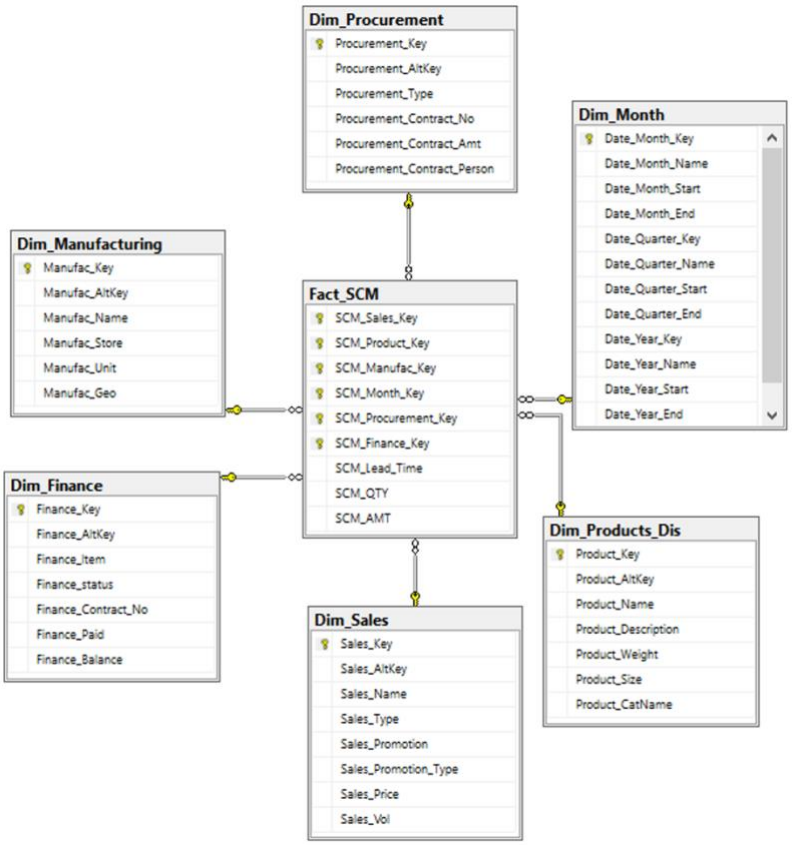
Aspects	Technical challenges
Data Collection	• Data is stored in various databases from different departments which are sales, manufacturing, procurement and financial services. • Extraction of usable data necessitates manual operations, such as downloading CSV files from procurement Ariba program for data manipulation in Excel sheet.
Reporting	• Administrators are scrambling to gather data from various locations since many sources deploy different devices and they are usually occupied with their daily tasks. • Job is time-consuming, repetitive, and deadline-driven. This demotivates people to get tasks done since tools and excel sheets perplex them.
Analysis	• There is a lack of standard process or the identification of data that is needed. This unstandardized data causes misinterpretation due to some naming issue like ExxonMobil vs Exxon Mobil Corp. or Exxon. • Data analysis is limited and inconclusive. Generally, administrators will pull all data from various sources for users without specific user's needs.
Information distribution	• Information is manually compiled and duplicated on occasion. So, when employee tries to integrate data within excel file, the vlookup technique can be used only by 1 column at a time. This not only results in time-consuming but data duplication across columns. • Many of the data and outcomes was inconclusive and untraceable. Without proper tools, it is possible that the outputs could be misjudged.

By tackling the issues above, I wish to generate the complete view of information from different departments: sales, manufacturing, procurement and finance in order to provide insight for decision makers to execute the right and fast business decision. In addition, this project also hopes to reduce workloads from employee who act as data gathering coordinators.

Deliverables:

1) Implementation of a dimensional model with queries (SQL) to show its validity

The current dimensional model as below is derived from the fact related data. This aggregate fact table is used when the number of specific records to be retrieved is high and query processing starts to affect the BI application's output. Users of business intelligence tools want applications to respond quickly to their queries. Aggregates are the response to this requirement. This table would be able to solve the above problems in terms of data collection and reporting. It should be able to display product manufacturing month vs month sales for the past 12 months including month-to-date. The user may sometimes choose to dig down and see the top items sold during the month, as well as the possibility of surplus from manufacturing. Then, decision maker can respond to this excess products with active sales promotion immediately. This will significantly reduce the number of records in the fact table by aggregating this data by month and removing some other dimension, thus enhancing the accuracy of the most frequent request by business users.



ACADSQLDBPRD01....NPRO - Diagram_0* SQLQuery1.sql - ac...PRO (nsornpro (57))*

```sql
CREATE TABLE Dim_Sales(
Sales_Key int NOT NULL PRIMARY KEY,
Sales_AltKey varchar(10) NOT NULL,
Sales_Name varchar(50) NULL,
Sales_Type  varchar(50) NULL,
Sales_Promotion varchar(50) NULL,
Sales_Promotion_Type varchar(50) NULL,
Sales_Price int NULL,
Sales_Vol int NULL);

CREATE TABLE Dim_Products_Dis(
Product_Key int NOT NULL PRIMARY KEY,
Product_AltKey varchar(10) NOT NULL,
Product_Name varchar(50) NULL,
Product_Description varchar(50) NULL,
Product_Weight int NULL,
Product_Size int NULL,
Product_CatName varchar(50) NULL);

CREATE TABLE Dim_Manufacturing(
Manufac_Key int NOT NULL PRIMARY KEY,
Manufac_AltKey varchar(10) NOT NULL,
Manufac_Name varchar(50) NULL,
Manufac_Store varchar(50) NULL,
Manufac_Unit int NULL,
Manufac_Geo int NULL);
```

```sql
CREATE TABLE Dim_Month(
Date_Month_Key int NOT NULL PRIMARY KEY,
Date_Month_Name varchar(50) NULL,
Date_Month_Start int NULL,
Date_Month_End int NULL,
Date_Quarter_Key int NULL,
Date_Quarter_Name varchar(50) NULL,
Date_Quarter_Start int NULL,
Date_Quarter_End int NULL,
Date_Year_Key int NULL,
Date_Year_Name varchar(50) NULL,
Date_Year_Start int NULL,
Date_Year_End int NULL);

CREATE TABLE Dim_Procurement(
Procurement_Key int NOT NULL PRIMARY KEY,
Procurement_AltKey varchar(10) NOT NULL,
Procurement_Type int NULL,
Procurement_Contract_No int NULL,
Procurement_Contract_Amt int NULL,
Procurement_Contract_Person varchar(50));

CREATE TABLE Dim_Finance(
Finance_Key int NOT NULL PRIMARY KEY,
Finance_AltKey varchar(10) NOT NULL,
Finance_Item varchar(50) NULL,
Finance_status varchar(50) NULL,
Finance_Contract_No int NULL,
Finance_Paid int NULL,
Finance_Balance int NULL);
```

```sql
CREATE TABLE Fact_SCM(
SCM_Sales_Key int NOT NULL REFERENCES Dim_Sales(Sales_Key),
SCM_Product_Key int NOT NULL REFERENCES Dim_Products_Dis(Product_Key),
SCM_Manufac_Key int NOT NULL REFERENCES Dim_Manufacturing(Manufac_Key),
SCM_Month_Key int NOT NULL REFERENCES Dim_Month(Date_Month_Key),
SCM_Procurement_Key int NOT NULL REFERENCES Dim_Procurement(Procurement_Key),
SCM_Finance_Key int NOT NULL REFERENCES Dim_Finance(Finance_Key),
SCM_Lead_Time int NOT NULL,
SCM_QTY int NOT NULL,
SCM_AMT int NOT NULL

CONSTRAINT[PK_Fact_SCM]PRIMARY KEY
([SCM_Sales_Key],[SCM_Product_Key],[SCM_Manufac_Key],[SCM_Procurement_Key],[SCM_Finance_Key],[SCM_Month_Key]))
```

2) ETL that shows input of source data into the dimensional model

1. Sources data from different departments: sales, manufacturing, distribution, procurement and finance, these files will be loaded to staging area.
2. Data from staging area is loaded to data mart using Dimension and Fact load jobs in a cleaned and integrated manner. I will be using Talend Open DI Studio for ETL purposes.

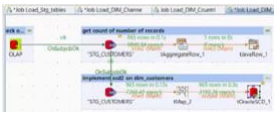
Stage Load Jobs

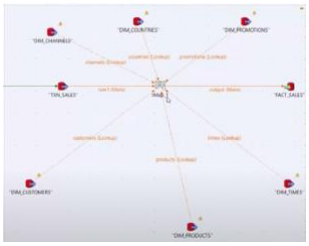
Job Name	Source	Target	Action
Load_Stg_Tables	Files	Staging Tables	Insert

Base Load Jobs

Job Name	Source	Target	Action
Load_dim_sales	STG_SALES	DIM_SALES	Update/Insert
Load_dim_products_dis	STG_PRODUCTS_DIS	DIM_PRODUCTS_DIS	Insert
Load_dim_manufacturing	STG_MANUFACTURING	DIM_MANUFACTURING	Update/Insert
Load_dim_month	STG_MONTH	DIM_ MONTH	Insert
Load_dim_procurement	STG_PROCUREMENT	DIM_PROCUREMENT	Insert
Load_dim_finance	STG_FINANCE	DIM_ FINANCE	Insert
Load_fact_SCM	TXN_SCM	FACT_SCM	Update/Insert

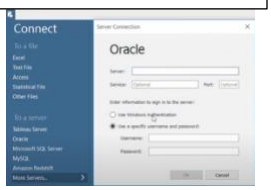
Loading each staging table → Loading dim table → Loading fact table using Talend → Perform Join to target





3) **Use of BI (BI tools) on existing model – *Tableau***

Connect Tableau to Oracle → Consider Join → Data Manipulation/ Add Alias → Build BI Dashboard



This tableau will retrieve all data from Oracle database where the data will be loaded using Talend tool. This BI tool should be able to perform basic computation and display dashboard according to KPI requirement such as percentage of sale cycle time by calculating the amount of time from first touch with a prospect to closing the deal. In addition, percentage of demand variability could also be shown by finding the mean of all demand levels for an item during over a period of time.