Data Warehouses – Project01

Wroclaw University of Science and Technology, Date: April 6, 2022

Student:	----------	Grade
Identifier	245784	
First name	Rahul	
Last name	Vijayvargiya	

This mini project assignment consists of 1 task. You should focus on providing a complete solution; however, if you cannot solve a particular step of the test, try to give at least a partial solution or provide a justification for the reason for the lack of a solution.

The process of creating a data warehouse should be preceded by an understanding of the "business needs" and reality (domain problems) represented by the available data resources. The implementation of the following task is to make you aware of the problems occurring in a specific (selected) part of reality and then enable the identification (determination) of the needs, purposes, and capabilities of business analysis to support decision-making processes (making the right business decisions).

Prepare the scope of the project according to the specifications below and discuss the project proposal with the lecturer. Record the arrangements made in the form of applications. During laboratory classes, you should present your project topic proposal to the group (justifying and arguing the purposefulness of such an undertaking and its main elements 1.1 – 1.6) using PowerPoint. In the end, after all presentations you will vote and choose two the most interesting project propositions that will be implemented in the next classes.

Task 1

1.1 Title of the project

Wide World Importers Database

1.2 General description of the domain

Introduce the selected domain.

Wide World Importers (WWI) is a wholesale novelty goods importer and distributor operating from the San Francisco Bay area.

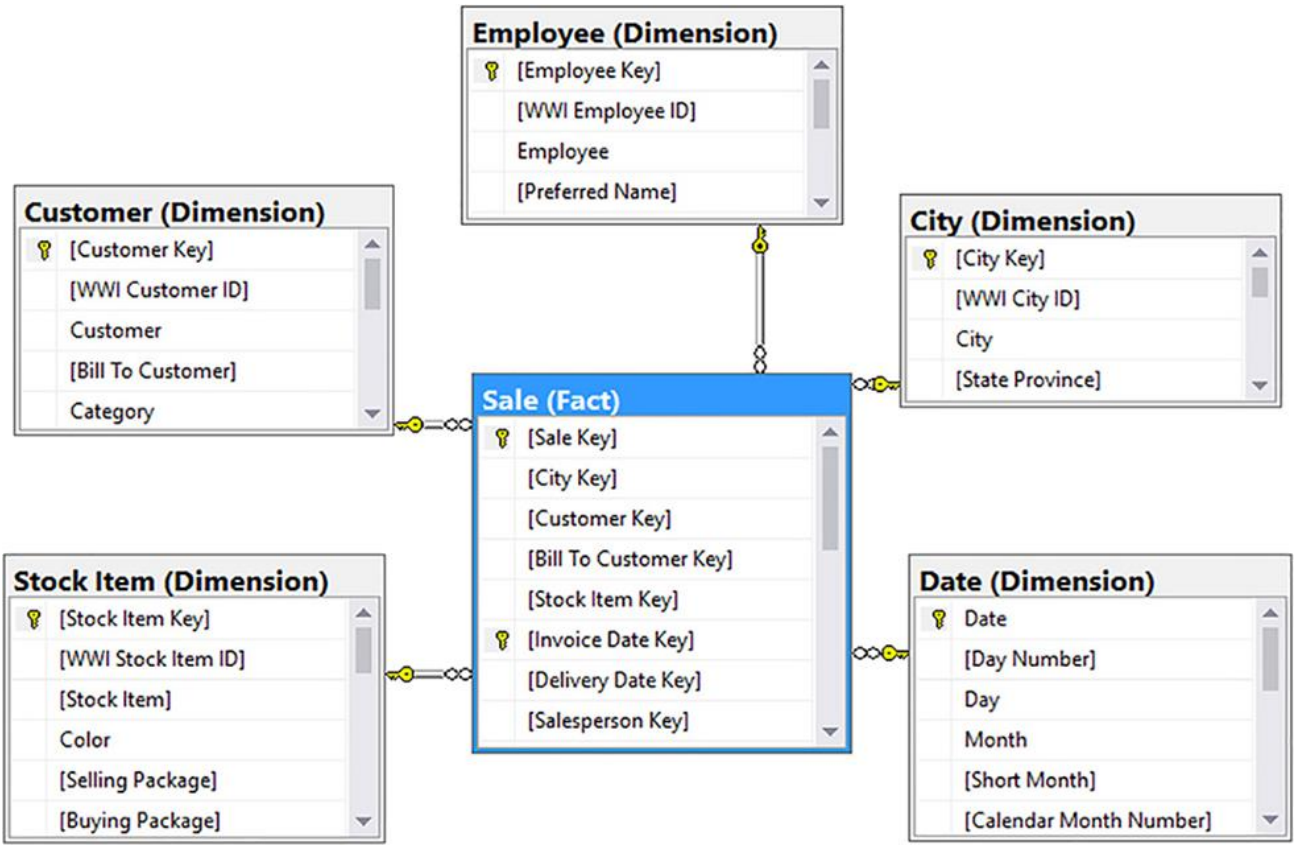
As a wholesaler, WWI's customers are mostly companies who resell to individuals. WWI sells to retail customers across the United States including specialty stores, supermarkets, computing stores, tourist attraction shops, and some individuals. WWI also sells to other wholesalers via a network of agents who promote the products on WWI's behalf. While all of WWI's customers are currently based in the United States, the company is intending to push for expansion into other countries.

WWI buys goods from suppliers including novelty and toy manufacturers, and other novelty wholesalers. They stock the goods in their WWI warehouse and reorder from suppliers as needed to fulfill customer orders. They also purchase large volumes of packaging materials, and sell these in smaller quantities as a convenience for the customers.

1.3 Description of the analysis area with justification

Please state and comment on the selected fragment of the domain intended for the data warehouse. Focus on the identification and introduction of the underlying business processes (events).

- *AverageNumberOfCustomerOrdersPerDay: The average number of orders per day. This is the most important parameter to influence the size of the generated data.*
- *SaturdayPercentageOfNormalWorkDay and SundayPercentageOfNormalWorkDay: The respective percentages of the average number of customer orders, set by the previous parameter. This also influences the size of the data, but only for the weekends.*
- *IsSilentMode and AreDatesPrinted: The parameters are for debugging purposes and to keep track of the status of the data generation*



1.4 identified problems

Identify and comment on the identified decision problem(s) in the assumed domain.

…

1.5 Project goal

Identify basic users (types) – provide a brief description of each type – and state expectations of each type. Do not select too many types, as it might be beneficial to focus on more details for 3-4 user types. Furthermore, we present detailed needs by specifying 10-15 OLAP user query types. Note that these should not be very specific queries (using specific attribute values), but rather general queries/need types. For each query/need, please indicate how a user can use the resultant information. Match each query to the identified user types. Provide the results in Section 1.5.1.

Perform a basic analysis of these query types (1.5.1), trying to infer some general user requirements. Focus on identifying event(s) and perspectives that the user is interested in. Further, perform a basic analysis of user query types (1.5.1), and try to infer some detailed user requirements. In particular, list all possible measures and all possible individual dimensions. Provide results in Section 1.5.2

This Project aims to examine Sales insights of Importers and Distributors which are operating from San Francisco according to its Invoice date and Delivery date and no. orders placed in one day and so on.

1. Higher Selling Product
2. Current Stock Left in warehouse
3. Order management
4. Product delivery date
5. Billing and Stock price
6. Maintaining day to day sell orders
7. Maintaining Day to Day purchase orders
8. Demand and supply
9. Aiming to get more profit to our distributors and importers who is currently operating from San Francisco Bay

Table	***Source tables***
City	*Application.Cities, Application.StateProvinces, Application.Countries.*
Customer	*Sales.Customers, Sales.BuyingGroups, Sales.CustomerCategories.*
Date	*New table with information about dates, including financial year (based on November 1st start for financial year).*
Employee	*Application.People.*

StockItem	*Warehouse.StockItems, Warehouse.Colors, Warehouse.PackageType.*
Supplier	*Purchasing.Suppliers, Purchasing.SupplierCategories.*
PaymentMethod	*Application.PaymentMethods.*
TransactionType	*Application.TransactionTypes.*

1.5.1 Expectations and detailed needs for decision support

Need 1 : City and Sales insights in fiscal year

- Ø User Types : Importers and Distributers
- Ø Detailed User Requirements: Query will be examining specific all locations and sales vol according to geographic locations
- Ø General User Requirements : Importers and distributors require this data to understand all circumstances and planning everything according to its sales data

Need 2 : Customers and product group in fiscal year

- Ø User types : Distributors and Importers

- Ø Detailed User Requirements : to get specific type of product according to the need and massive buying of particular according to group of customers, which help to target on specific part and will reduce burden of our users

- Ø General User Requirements : Forecast Sales of particular product group and based on targeted customers

Need 3: Date and Total sales b/w whole financial year

- Ø User Type : Importers and Distributors

- Ø Detailed user Requirements : To get total sales insight of particular product and group product and target customers, Repeated buying and selling good in whole Fiscal year for accounting

- Ø General User Requirement : to use all data to generate more in sales and profit

Need 4: Employee and there ID, total sales in fiscal year

- Ø User Type : Importers and distributors

- Ø Detailed user Requirements : To get to examine which employee is doing best in sales

- Ø General User Requirements : to use data to examine is this employee is generating sale or not

Need 5: Stock Items in warehouse in fiscal year

- Ø User type : Importers and Distributors

- Ø Detailed User Requirements : To get to know about purchasing goods, how many items are currently left in the warehouse and how many of it we need according to the sales insights.

- Ø General User Requirements : Our end users use this data to forecast sales increased or decreased over a particular time period

Need 6: Suppliers and Purchasing in fiscal year

- Ø User Type : Importers and distributors

- Ø Detailed User Requirements : to keep the records of purchasing and suppliers maintained and for the accounting b/w them

- Ø General User Requirements : end users will this data to maintain accounting

Need 7 : Payment and mode of payment throughout fiscal year

- Ø User type : Importers and Distributors
- Ø Detailed user requirements : to keep up with all transactions and figure out how much profit has made b/w this time
- Ø General user type : end user will use data insight for area of growth

Need 8 : after all this we will get know to area of improvisation and out best selling product and total profit has generated in fiscal year → End users(importers and distributors)

Need 9 : To see how the product and group affects the sales → End Users

Need 10 : To see how massively customers are buying products → End Users

Perspective that End Users are interested In :

1. Demand and supply
2. Total supply
3. Total demand
4. Total market share of product
5. Customers Need

1.5.2 Scope of analysis – aspects examined

While the team at WWI use SQL Server Reporting Services to generate operational reports from the WideWorldImporters database, they also need to perform analytics on their data and need to generate strategic reports. The team have created a dimensional data model in a database WideWorldImportersDW. This database is populated by an Integration Services package.

SQL Server Analysis Services is used to create analytic data models from the data in the dimensional data model. SQL Server Reporting Services is used to generate strategic reports directly from the dimensional data model, and also from the analytic model. Tableau is used to create dashboards from the same data

1.6 Data sources

Prepare a brief description of the selected data source, focusing on general structure and its usage, access options, update characteristics, volume, time span for facts, etc. Further, perform initial assessment of the quality of available data – focus on number of records, number of valid records, number of missing records.

https://github.com/Microsoft/sql-server-samples/tree/master/samples/databases/wide-world-importers

1.6.1 Location, format, availability

WideWorldImporters is a sample for SQL Server and Azure SQL Database. It showcases database design, as well as how to best leverage SQL Server features in a database.

WideWorldImporters is a wholesale company. Transactions and real-time analytics are performed in the database WideWorldImporters. The database WideWorldImportersDW is an OLAP database, focused on analytics.

The sample includes the databases that can be explored, as well as sample applications and sample scripts that can be used to explore the use of individual SQL Server features in the sample database.

1.6.2 Initial assessment (number of records, time span for facts, etc.)

…

1.6.3 Facts and measures:

Database Source	Key	Name	Data type	References	Description
WWI DATABASE	1	Sale Key	bigint		DW key for a row in the Sale fact
WWI DATABASE	2	City Key	int	Dimension.City	City for this invoice
WWI DATABASE	3	Customer Key	int	Dimension.Customer	Customer for this invoice
WWI DATABASE	4	Bill To Customer Key	int	Dimension.Customer	Bill To Customer for this invoice
WWI DATABASE	5	Stock Item Key	int	Dimension.Stock Item	Stock item for this invoice
WWI DATABASE	6	Invoice Date Key	date	Dimension.Date	Invoice date for this invoice
WWI DATABASE	7	Delivery Date Key	date	Dimension.Date	Date that these items were delivered
WWI DATABASE	8	Salesperson Key	int	Dimension.Employee	Salesperson for this invoice
WWI DATABASE	9	WWI Invoice ID	int		InvoiceID in source system
WWI DATABASE	10	Description	nvarchar(100)		Description of the item supplied (Usually the stock item name but can be overridden)
WWI DATABASE	11	Package	nvarchar(50)		Type of package supplied
WWI DATABASE	12	Quantity	int		Quantity supplied

Database Source	Key	Name	Data type	References	Description
WWI DATABASE	13	Unit Price	decimal(18, 2)		Unit price charged
WWI DATABASE	14	Tax Rate	decimal(18, 3)		Tax rate applied
WWI DATABASE	15	Total Excluding Tax	decimal(18, 2)		Total amount excluding tax
WWI DATABASE	16	Tax Amount	decimal(18, 2)		Total amount of tax
WWI DATABASE	17	Profit	decimal(18, 2)		Total amount of profit
WWI DATABASE	18	Total Including Tax	decimal(18, 2)		Total amount including tax
WWI DATABASE	19	Total Dry Items	int		Total number of dry items
WWI DATABASE	20	Total Chiller Items	int		Total number of chiller items
WWI DATABASE	21	Lineage Key	int		Lineage Key for the data load for this row

General Conclusions:

Use this section to provide your general conclusions:

1. This is an overview of the fictitious company Wide World Importers and the workflows that are addressed in the WideWorldImporters
2. The company operates with a financial year that starts on November 1st.
3. Some of the information from given task is missing because of lack in information and don't know what to put it there

REMARKS:

- *A report without final conclusions will not be checked and results in a negative score!*
- *The report file should be named **Proj01DW-StudentID-Last name-2022**, please use the PDF format*
- *You should use MS SQL SERVER 2019 (or 2017), Visual Studio and Tableau Prep (available at https://www.tableau.com/academic/students)*