Warehousing Project

Supply Chain Models: Manufacturing & Warehousing (ISyE 3104) - Fall 2001
Due December 6, 2001

You are strongly encouraged to complete this project in groups of 4 students.

Purpose
Warehouse profiling is a special case of *data-mining*, which is simply jargon for sifting through historical data for opportunities and insights that might confer advantage. The data of most enterprises resides in large relational database management systems, such as Oracle, Informix, Sybase, and others. The data in a database is stored as a collection of tables, which are similar in some ways to huge spreadsheets: Every row describes some object, such as a sku; and every column describes some aspect of the object, such as its name.

To mine data requires first that you manage large datasets. The main tool you will need is some program that will allow you to query multiple tables and to perform joins, which connect the data in one table with that in another through some common key. For example, if both the file of sku's and the file of order-lines both contain a field Unique_sku_id, then these two tables may be joined to form a new table that combines the data of the two. Now each order-line will contain, in addition to the Unique_sku_id, all the information about that sku from the sku's file.

This project is based on data from a wholesale distributor of office products. In this warehouse every item has a single location. Small stuff is in shelving upstairs on a mezzanine (zones A, B); larger cartons are downstairs in shelving (zones C, D); and the largest items are in pallet rack (zone G). Small, expensive items are picked from a security area (zone E).

Data files
You must download two files:
http://www.isye.gatech.edu/people/faculty/John_Bartholdi/wh/book/profile/projects/projects/spr-lines.txt.tgz
and
http://www.isye.gatech.edu/people/faculty/John_Bartholdi/wh/book/profile/projects/projects/ spr-skus.txt.tgz
They can be uncompressed by programs such as tar and gunzip (UNIX) or winzip (Windows). Each compressed file fits on a 1.4MB floppy disk. Each file starts with two lines of header: The first names the columns and the second is a line of dashes that separates the column names from the data. You may discard these two lines if importing into a relational database, such as MS Access. The file of sku descriptions is in tab-delimited ASCII format and contains the following information:

Sku_id: A tag that uniquely identifies each sku
Desc: A brief description of the product
Vendor: Abbreviation identifying the supplier of the sku

Zone: In which zone of the warehouse is the product stored
Aisle: In which aisle is the product stored
Bay: In which bay of rack is the product stored
Sell unit: The smallest level of packaging shipped to the customer
Pack_1: The finest level of packaging that the warehouse might handle
Pack_2: A coarser level of packaging
Pack_3: A still coarser level of packaging
DOFT: Date of first transaction, to help identify new products

The file of pick-lines is in tab-delimited ASCII format and contains the following information:
Order_id: A tag that uniquely identifies this order and customer
Date: The date the order was picked and shipped (YYMMDD)
Sku_id: A tag that uniquely identifies the sku. NB: This corresponds to a similar entry in the file of sku's.
Order_qty: How many selling-units were requested by the customer in this order
Ship_qty: How many selling units were actually shipped to the customer

Questions
Describe the warehouse and its activity based on the data set. Look for interesting patterns that reveal something potentially useful or interesting about the warehouse. Think especially about how to display the results in succinct, elegant graphical format. Amidst your explorations discuss the following.

1. Analyze the sku's. For example: Where are they? Which are most popular? On what percentage of orders do they appear? How many sku's are never requested? How does the population of sku's change over time?
2. Analyze the work. How much is there and where is it? For example, how is it distributed among sku's, among zones, among aisles, among bays, among orders, among vendors, among days, etc.?
3. Analyze the orders. For example: What percentage of orders are for a single line? What percentage are for more than 20 lines? What is the distribution of lines per order? What is the distribution of lines per order within each zone? What percentage of the labor do single-line orders represent? Orders with more than 20 lines? What percentage of orders touch the mezzanine (zones A, B)?

You will be evaluated based on the content of your report and the clarity of writing. Pretend that you are team of consultants and writing this report for a client. So, do not make your report a list of questions and answers.