Design Document

Kennedy Holifield

A I. Business Problem:

EcoMart's mission is to offer ethically sourced and sustainably sourced products to its community. A variety of categories are available, including groceries, apparel, home foods, and care items. Because of this, EcoMart needs a relational database that will support the variety of data that they have. Currently, product, customer, and transaction data are inside one spreadsheet. This consolidation increases file size and make it more difficult to track metrics such as product availability, price changes, customer orders and status, and generate accurate sales and review reports. A relational database can centralize these different types of data, ensuring organization, faster queries, and reliable analytics.

II. Proposed Data Structure:

To solve this business problem, a structure that is consistent, organized, and distinguished is necessary. The structure will be broken down into five main tables. The first is products. This will store product details such as product name, unit price, and unit cost. The second table will be the categories for the product names. These will include groceries, apparel, home foods, and care items. The third table will be for location, where all the customer details for their location can be seen. A fourth table will be dedicated to orders that show when customers have purchased, including customer ID, order date, and ship date. Lastly, the fifth table will be ordering details that store information about the quantity, unit price, total revenue, total cost and total profit of the orders.

III. Database Justification

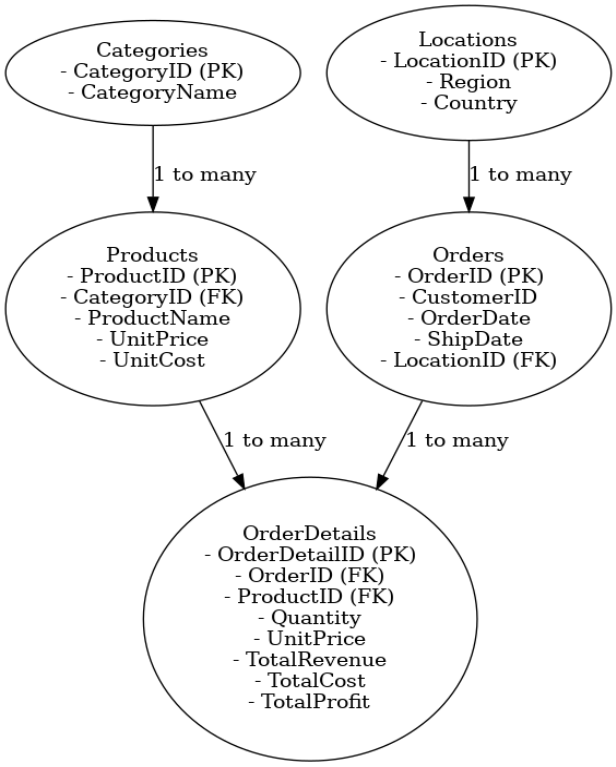
The solution to the business problem lies in the use of a relational database for several reasons. The first advantage is that it eliminates duplication across spreadsheets. This is important because it reduces inconsistent data and provides a lower risk of errors. Furthermore, a relational database allows for flexible reporting and clear insights. For example, if EcoMart

wants to see which is the best-selling product, a well-made query can provide insight into that. This type of database will also support scalability with indexing and normalization. In addition, it enforces data integrity with primary and foreign keys.

III. How Data Will Be Used

Specifically, EcoMart will use product and category data for product management. It will store and retrieve detailed product information. The data will also be used for customer analytics. This will track customer preferences, purchase patterns and sustainability engagement. Personalized marketing will be able to be done from the data by targeting high demand products in specific regions. From a financial perspective, data such as total revenue and profit can be used to improve pricing strategies. Lastly, orders and order details data will be great for sales tracking and inventory management.

B.



C. In this database design, the tables are going to store the actual data. This categories, locations, products, orders, and order details. Additionally, this database design will in include an index containing the file attributes that will help queries run faster. File attributes will be included in all the tables. In the Categories table, the file attributes are CategoryID and CategoryName. In the Locations table, the file attributes are the LocationID, region, and country. The Order table's file attributes are OrderId, CustomerId, LocationID, order date, and ship date. In the Products table, the file attributes are ProductID, CategoryID, product name, unit price, and unit cost. Lastly, the file attributes in the Order Details table are OrderDetailID, OrderId, ProductId, quantity, unit price, total revenue, total cost, and total profit. Additionally, this database will be including primary keys and foreign keys. Primary keys are the unique IDs for each record. In my database design, I have five primary keys: CateogoryID, LocationId, ProductId, OrderId, and OrderDetailID. Foreign keys are keys that link data between tables and in my design CategoryID, LocationID, OrderID, and ProudctID are my foreign keys. This database will be stored in WGU's VM relational database management system on PostgreSQL.

D. This database design will address scalability concerns for several reasons. To start, it has a normalized structure. This will assist in avoiding duplicate data so that data storage and queries are more efficient. Since the database design includes indexes, it will speed up the searching for queries, even with large and growing datasets. Additionally, this database design has a flexible schema. Adding new products, categories, or attributes won't fault the existing tables because the design separates product information, categories, and orders. This design also allows for partitioning so if a table gets large, it can split so that queries only scan part of the data.

E. The privacy and security measures that should be implemented in the database design are encryption, access control, audit logging, and compliance standards. Encryption will help store sensitive customer information such as email or phone number. Access control will give role-based permissions meaning that only certain roles are able to see certain datasets within the database. Audit logging will track all data changes to ensure that all alterations to the database is being made within reason. Lastly, the database should be compliant with the standards of the data rights and privacy rules like those stated in the CCPA.