**Report**

**Methodology:**

After discussing with the vice president of Guac4U on the business requirements, and carefully considering all the information provided. We have opted for the Kimball Approach of data warehouse, The Kimball approach utilizes dimensional models such as star schema and snowflake schema to arrange the data into different business classified data, in order to quickly enable business processes. Data from all the different sources like employee information, retail transaction, store, promotion tracker, order fulfillment, inventory, and customer relations management systems are to be loaded into the staging area using ETL (Extract, Transform, and Load) Software. From there, this data is moved to the data warehouse. Using the data present in the dimensional model data marts like sales, finance, etc. are created based on the business/client requirements. This model partitions data into the fact table and dimension table. The combination of a fact table connected with dimensional tables is known as a star schema. Here we construct many star schemas which help to provide various reporting needs. This dimensional model provides data to analyze and extract business intelligence insights.

Diagram

Description automatically generated

**Activities:**

* Identified the key business processes and the key business questions that are needed for the installation of the data warehouse.
* Dimension tables are denormalized so that the users can drill up and down without joining another table.
* Created ‘Enterprise Bus Matrix’ where the different business processes are listed vertically, and the shared/common dimensions are listed horizontally. This serves as a document to build the star schemas and the relationship between dimensions and business processes.
* Data is ingested into the data dimensional model using ETL software/tools.
* Created Entity Relationship Diagram to identify all the dimensions and facts tables.

**Limitations:**

* This data model is business process-oriented, instead of focusing on the enterprise, it cannot manage all the Business Intelligence layer reporting.
* As Promotions are to be store specific and cannot be different for every store, but they are to be the same in every store. It will help to save storage and reduce the complexity of the Promotions dimension.
* Fact tables are quite detailed, new columns are not to be added to expand the fact table dimensions, which will affect the performance.
* The process of ingesting huge amounts of historic data is complex.

**Outcomes:**

* This data warehouse system focuses on individual business areas. So, it takes less space in the database, making it easy for system management.
* This data warehouse lifecycle is referred to as the business dimensional lifestyle methodology because it allows BI tools to have deeper access across all the star schemas and generates reliable insights.
* By using star schema because of its denormalized structure, it simplifies querying and analysis of data for all users.

**APPENDIX:**

A picture containing text, sign, reading

Description automatically generated

**Entity Relationship Diagram**

**APPENDIX:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **BUSINESS PROCESS** | **SHARED DIMENSIONS** | | | | | | |
|  |
|  |
| **Customer** | **Employee** | **Store** | **Product** | **Promotion** | **Supplier** | **Shipping** |  |
| **Retail Sales** | **X** | **X** | **X** | **X** | **X** |  | **X** |  |
| **Retail Sales Forecast** | **X** | **X** | **X** | **X** |  |  |  |  |
| **Supplier Inventory** |  |  | **X** |  | **X** | **X** |  |  |
| **Retail Promotion** |  |  | **X** | **X** | **X** |  |  |  |
| **Online Sales** | **X** |  | **X** | **X** | **X** | **X** | **X** |  |
| **Customer Billing** | **X** | **X** | **X** | **X** | **X** |  | **X** |  |
| **Employee Labor** |  | **X** | **X** |  |  |  |  |  |
| **Human Resource** |  | **X** | **X** |  |  |  |  |  |
| **Store Operation** | **X** | **X** | **X** | **X** | **X** | **X** | **X** |  |
| **Purchase Order** | **X** | **X** | **X** | **X** | **X** |  | **X** |  |
| **Product Support** |  |  |  | **X** |  | **X** |  |  |
| **Cost acquisition** |  |  |  | **X** | **X** | **X** |  |  |
| **Store Deliveries** | **X** | **X** | **X** | **X** | **X** |  |  |  |
| **Supplier Delivery** | **X** |  | **X** | **X** |  | **X** | **X** |  |

**Enterprise Bus Matrix**

**APPENDIX:**

**Interview Questions:**

1. In which format data is stored?
2. Before the installation of the data warehouse, how are you managing to store data?
3. What are the different items/products sold in Guac4U other than avocados?