End-to-End Data Warehouse

The IST 722 project was a comprehensive group effort to develop a data warehouse, integrating Fudgeflix and Fudgemart for enhanced business intelligence. This involved creating an intricate ETL process, crucial for staging, transforming, and loading data into the warehouse.

Overview

Objective

Create a data warehouse, merging the fictional companies Fudgeflix and Fudgemart into a single source for business intelligence in order to derive insight for the Fulfillment Team.

Team Collaboration

Worked alongside Daniel Caley, Michael Johnson, Jennifer Lammers Zimmer, and demonstrated teamwork and collaborative problem-solving skills.

Business Processes Addressed:

Scope of the project encompassed various business aspects, including sales, inventory, customer service, and order fulfillment. We chose to make our key goal to analyze order fulfillment for Fudgemart and Fudgeflix from a unified data source, focusing on product lead times, departmental performance, and shipping destinations.

ETL

Dimensional Modeling

Created a single fact table representing the order fulfillment process, complemented by date, product, and customer dimensions. This structure was crucial for measuring order fulfillment performance, particularly focusing on the order-to-ship lag.

Date Dimension and Staging

Developed a control flow for date dimension, involving ETL as a single control flow. This was crucial for pulling date fields and placing them into the staging schema.

Staging Customer and Product Dimensions

Managed data flows for customer and product dimensions, merging data from different sources and resolving data type mismatches.

Staging Order

Built a single table order in Stage, incorporating data from various sources and ensuring consistency across the data flow.

Load Dimension & Fact Table

Transformed data from Stage to the Data Warehouse for multiple dimensions, focusing on creating slowly changing dimensions for customer and product data.

Challenges and Problem-Solving

Faced challenges in merging data from different sources, requiring careful data conversion and cleansing.
Addressed issues related to data type mismatches, especially in customer and product data.

Impact and Results

Successful Merger

Successfully established a robust and efficient ETL process, facilitating effective data management in a warehouse environment.
Enhanced data accessibility and reliability, supporting informed decision-making.

Business Intelligence

With our BI, which we created using PowerBI, we aimed to empower the business with analytics capabilities, providing tools for data visualization and decision-making. This included dashboards for monitoring fulfillment health and order fulfillment processes. Our suggestions, derived from analyzing our model, included reducing lead times for movie orders, capturing received dates for end-to-end BI, offering promotional discounts for high lead-time orders, and transforming Fudgeflix into a full-fledged streaming service.

Reflection on Data Warehousing

This project highlighted the intricacies of data warehousing, particularly the importance of an efficient ETL process in managing large volumes of data.
It also underscored the significance of teamwork and collaborative efforts in tackling complex data challenges.
Building a data warehouse required managing data from multiple sources, resolving data mismatches, and ensuring data consistency, which supported the project's broader goal of providing a unified platform for order fulfillment analysis and business intelligence. The project not only consolidated data for operational efficiency but also set the stage for strategic business recommendations and improvements in customer service and fulfillment operations.