

IST - 722 Data Warehouse

Final Project Report

Topic : Fudge Enterprise

Dataset : Fudgemart / Fudgeflix

Team: Data Divas and a Dude

Team Members:-

1. Adarsh Shetty Andar
2. Abirami Rajalingam
3. Hiral Paghadal
4. Lahari Chowtoori

Project Overview

We are building an end to end integrated data warehouse for fudge enterprise using tools such as dbt, Snowflake and Power BI. We have used the data centric approach and followed the Kimbal enterprise bus matrix. We identified 4 business processes to answer key questions in different areas such as finance, ratings, sales and reviews. Project focuses on extracting, transforming and loading data and building a data warehouse and ultimately business intelligence. We integrated the customers from both fudgemart and fudgeflix into a common customer base for the finance business process, which will help us to determine the total revenue generated across and the common customers in both fudgemart and fudgeflix.

Project Objective

Our objective is to integrate customers from both fudgemart and fudgeflix datasets to create a common customer base. The purpose of creating an integrated data warehouse was to have a unified warehouse, so that the data is accessible by any department. We aim to focus on the customer sales, customer ratings, product reviews and finance analysis to analyze various facts and attributes.

Business Process

1) Sales Order

For the Sales Order Business process we analyze and capture data belonging to orders. The fact_sales records transactional data for each item ordered and granularity is transaction level. The facts included are quantity, sold amount, product name and product department. These fields provide information on the quantity of items ordered, their total amount sold and the descriptive fields such as product name and department which are degenerate fields.

2) Product Review Analysis

For the product review analysis business process we focused on analyzing customer reviews stars on various products and their department and product name. The fact table fact_product_review is an accumulating snapshot which captures one row per customer product review. The fact review stars is the rating provided by customers to the products. This business process helps to determine the highly rated products and the departments to increase sales.

3) Finance Analysis

Finance Analysis business process is focused on providing the revenue and sales analysis of both fudgemart and fudgeflix. It helps the finance team of fudge mart to analyze the total revenue, across, frequent buyer across both sales area, and total customer spendings. This business process used integrated customer dimension to provide a comprehensive view of customer spending behavior and overall revenue across geographically as well. It allows tracking customer buying patterns in both fudgemart and fudgeflix.

4) Customer Rating/Title/Genre Analysis

The customer rating/title/genre analysis is focused on analyzing customer ratings for titles and genres enabling to track customer preferences and trends on titles and

genres of fudgeflix. The fact_customer_ratings is an accumulating snapshot providing facts about title average rating, title rating and customer ratings.

Bus Matrix

Instructions!									
Business Process Name	Fact Table	Facts	dim_products	dim_customers	dim_date	dim_account	dim_plans	dim_combined_customers	dim_titles
Sales Order	fact_sales	quantity, soldamount, productname productdepartment	x	x	x				
Product Review Analysis	fact_product_review	review stars, review date	x	x	x				
Finance Analysis	fact_finance	total customer spending , plan price, plan name, total billed amount till date, frequent buyers across all business , total revenue across	x		x		x	x	
Customer Rating/title/genre Analysis	fact_customer_ratings	titleavgrating, titlerrating, customerrating				x			x

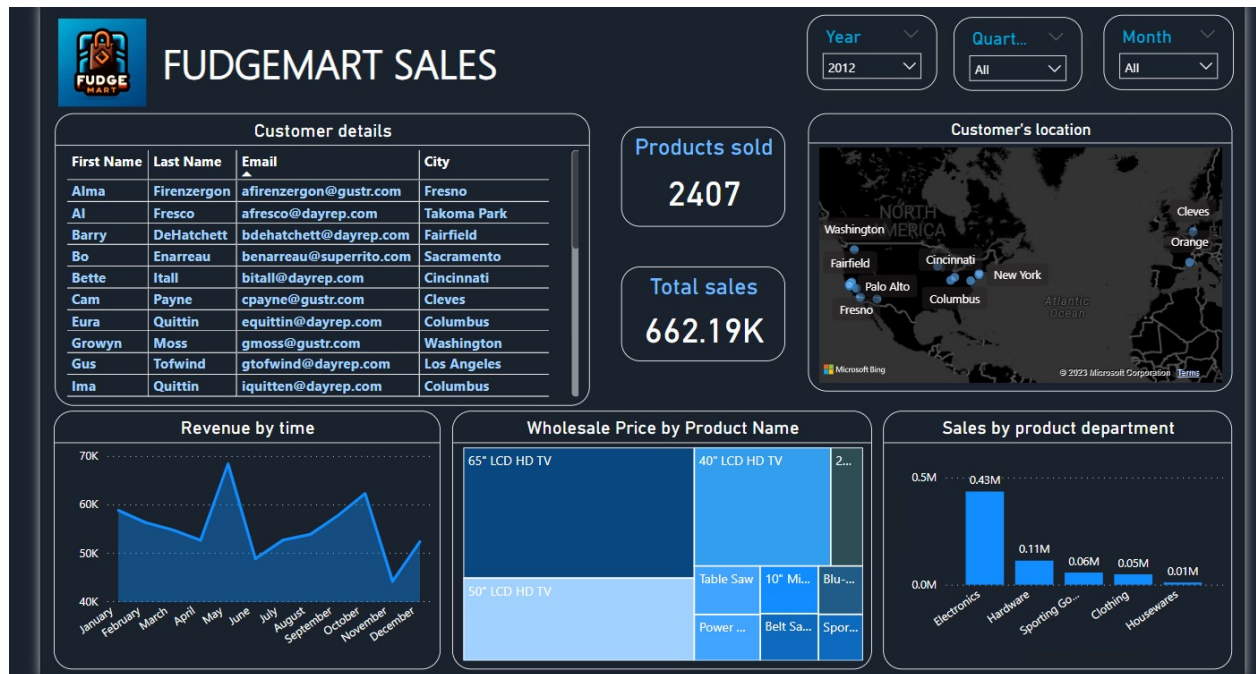
Business Intelligence Dashboard

1) FudgeMart Sales



This dashboard provides insights on fudgemart sales. Our map shows the distribution of customers globally. Two cards show the products sold and the total sales till date. The text table is the customers data and the line charts show the trend of revenue over time. The Line chart is also affected by the slicers on top year, month and quarter. The tree map wholesale price by product name displays the wholesale price for the highly sold products. The bar chart shows the total sales by product department, indicating Electronics has the highest sales and houseware has the lowest sales.

Slicer :- When we select the year 2012, the graphs and cards changes indicating that for year 2012, products sold were 2407, the sales 662.1 The revenue by time graph shows that for year 2012, May had the highest sales



2) Product Review



Product Ratings dashboard provides insights on review stars given by customers on fudgemart products.

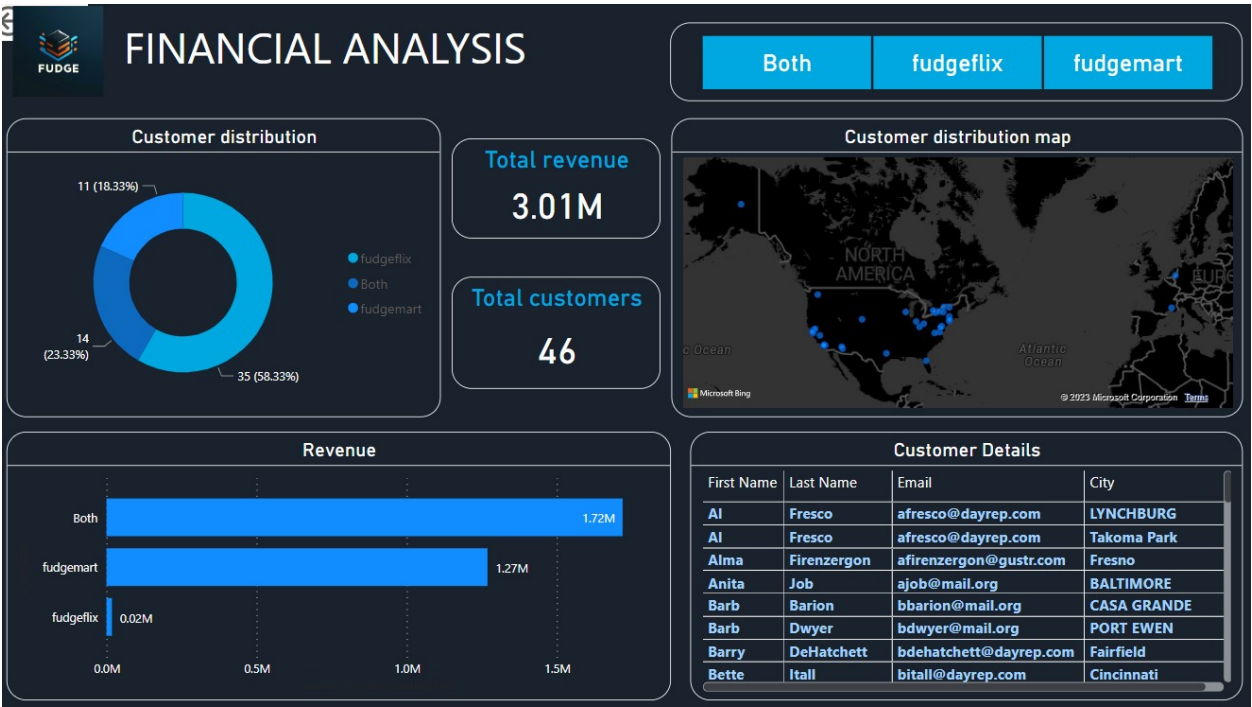
The customer text table is the list of customers who reviewed the products. The cards display information on total products and total reviews. If a particular customer is selected the cards and all the graphs change. We have two slicers for the product department and product name.

Slicer:- when slicer housewares is selected the product and review Cards changes indicating only 5 products exists in Houseware department and the total number of reviews are 10 with Crack Pot the highest rated



3) Finance Analysis (Integrated Data Warehouse (Customers))

One stop destination for finance teams to analyze the revenue for fudge enterprise comprising of both fudgemart and fudgeflix. The Customer distribution chart tells us that we have 14 customers who are both in fudgemart and fudgeflix. Total revenue for the fudge enterprise is 3.01 M and the total customers are 46 excluding the common customers in both fudgemart and fudgeflix.

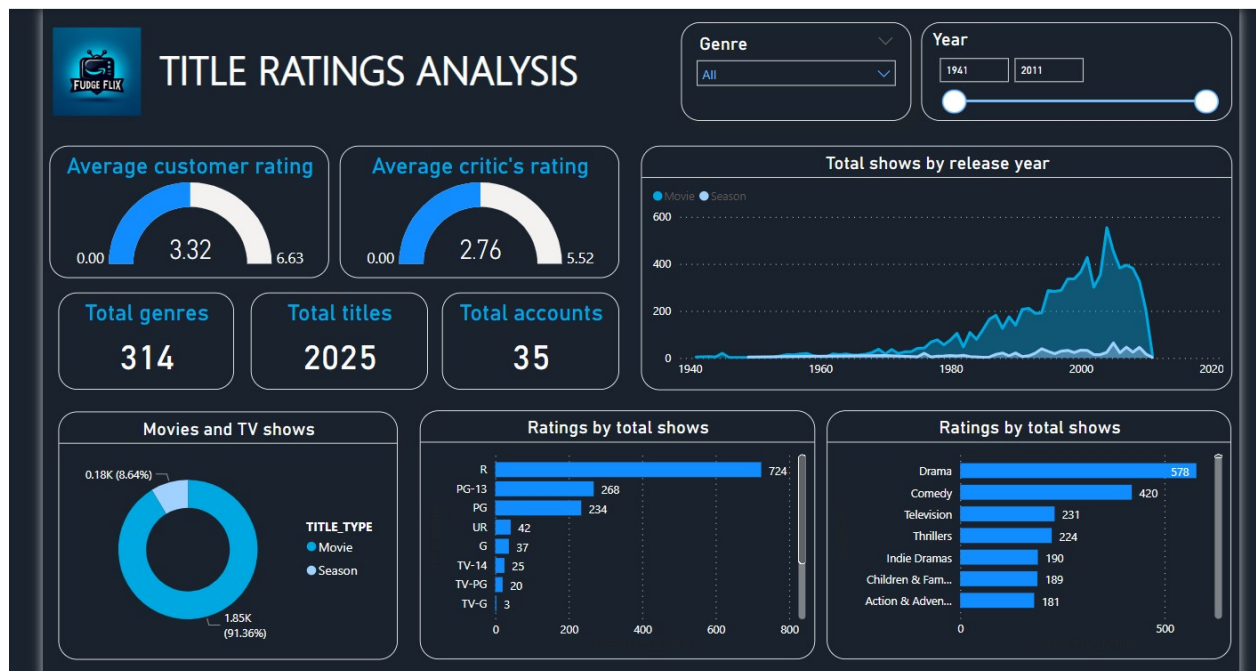


Both, Fudgemart and Fudgeflix are the slicers, when each selected the revenue and customers changes

4) Title Rating Analysis

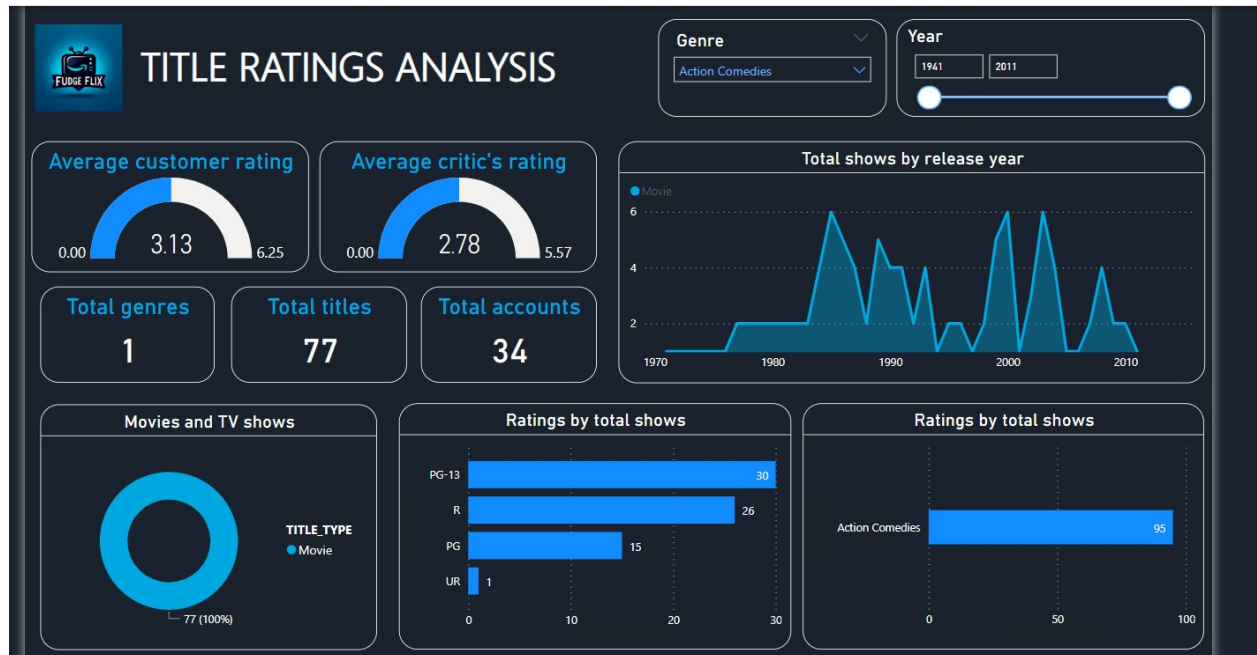
This dashboard analyzes the fudgeflix ratings on various titles. The dashboards help to analyze the most rated shows, critics and customer

ratings and shows over time. The gauge helps us to compare the critics rating with the customer ratings. If we change the genre and year from slicer the customer and critics rating changes. The cards shows the total genres, total titles and total accounts. The total release year shows the area chart indicating the peak started in 1980 and lasted till 2011. The below bar charts show the rating by total shows.



Slicer:-

Selecting particular genre and year, will change all graphs in the dashboard. After selecting Genre as Action Comedies all the graphs changed.



Reflections / Lessons Learned

1. IST 722 helped us to gain very deep understanding of the nuances of data warehouse
2. We learned to develop, design and understand core concepts of data warehouse and business intelligence
3. The various architecture of data warehouse will serve as a foundation for our career in data engineering
4. Along with the technical concepts we also learned the managerial concepts to data warehouse
5. The latest technologies such as snowflake and dbt provided a deeper understanding of a cloud based data warehouse integrated with dbt
6. The milestones helped us to reflect on our project and also allowed us to incorporate feedback after every milestone.
7. The project gave an end to end understanding of implementing and building hands on data warehouse and business intelligence
8. We faced several challenges while building our dbt. Our aim was to build an integrated data warehouse combining both fudgemart and fugdeflix into a single data warehouse.

9. We combined the customers from both fudgemart and fudgeflix into a single customer dimension giving us the opportunity to analyze the customer behavior and revenue from both.
10. The 3 exams were very challenging but quite analytical and made us think out of the box and apply our knowledge and concepts learned.