

## **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 Fudgelix 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.

# **IST 722 Project Fudgemart & Fudgeflix**

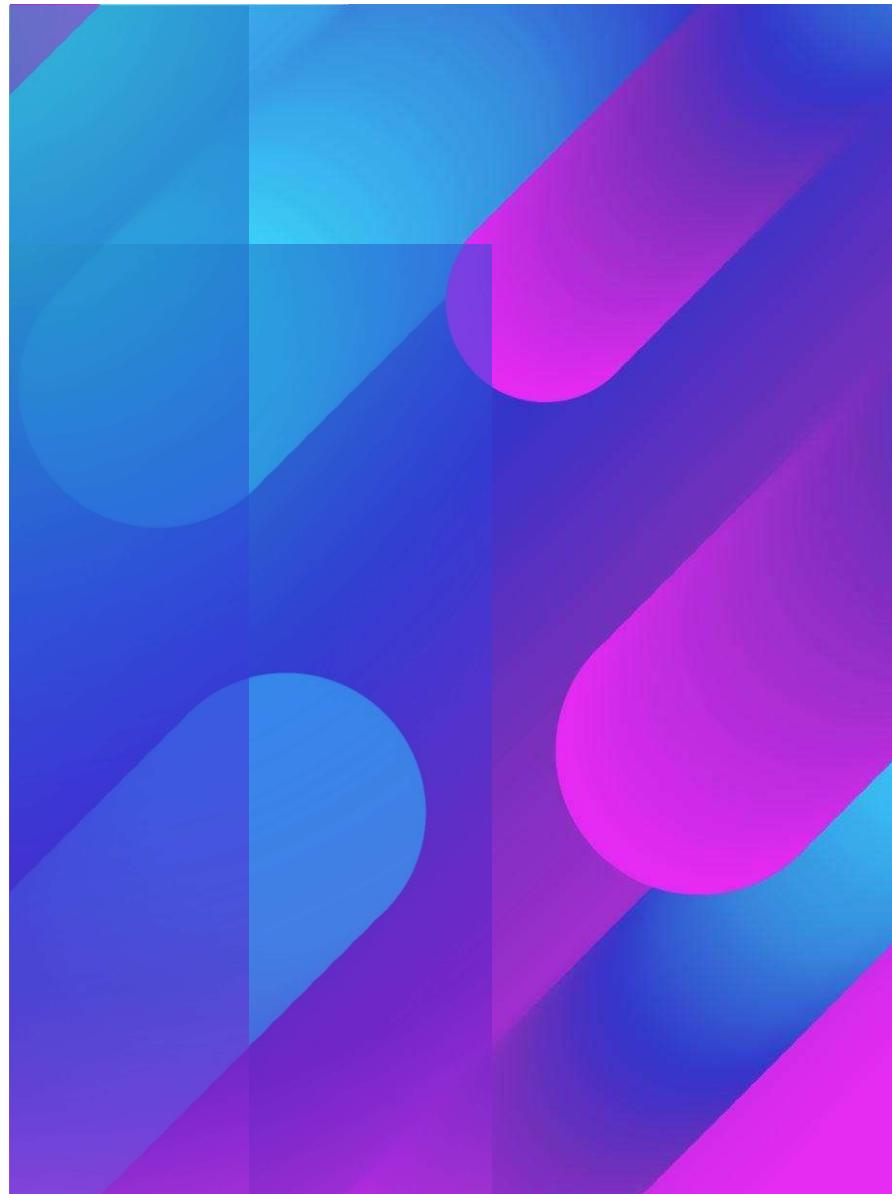
Order Fulfillment – Group 2

Jennifer Lammers Zimmer

Samuel Deery-Schmitt

Michael Johnson

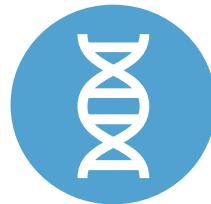
Dan Caley



# Project Objective



Create a Data Warehouse



Create a Business Intelligence Platform



Bring Fudgeflix and Fudgemart to a single source for the business



Derive insights for the Fulfillment Team

# **Business Objectives**

Successfully guide merger of Fudgflix and Fudgemart data sources to ensure minimal loss in business processes throughout the transition.

Ensure all aspects of the business remain operational from sales to inventory to order fulfillment and everything in-between.

# Business Processes



Order  
Fulfillment



Sales



Inventory



Customer  
Service

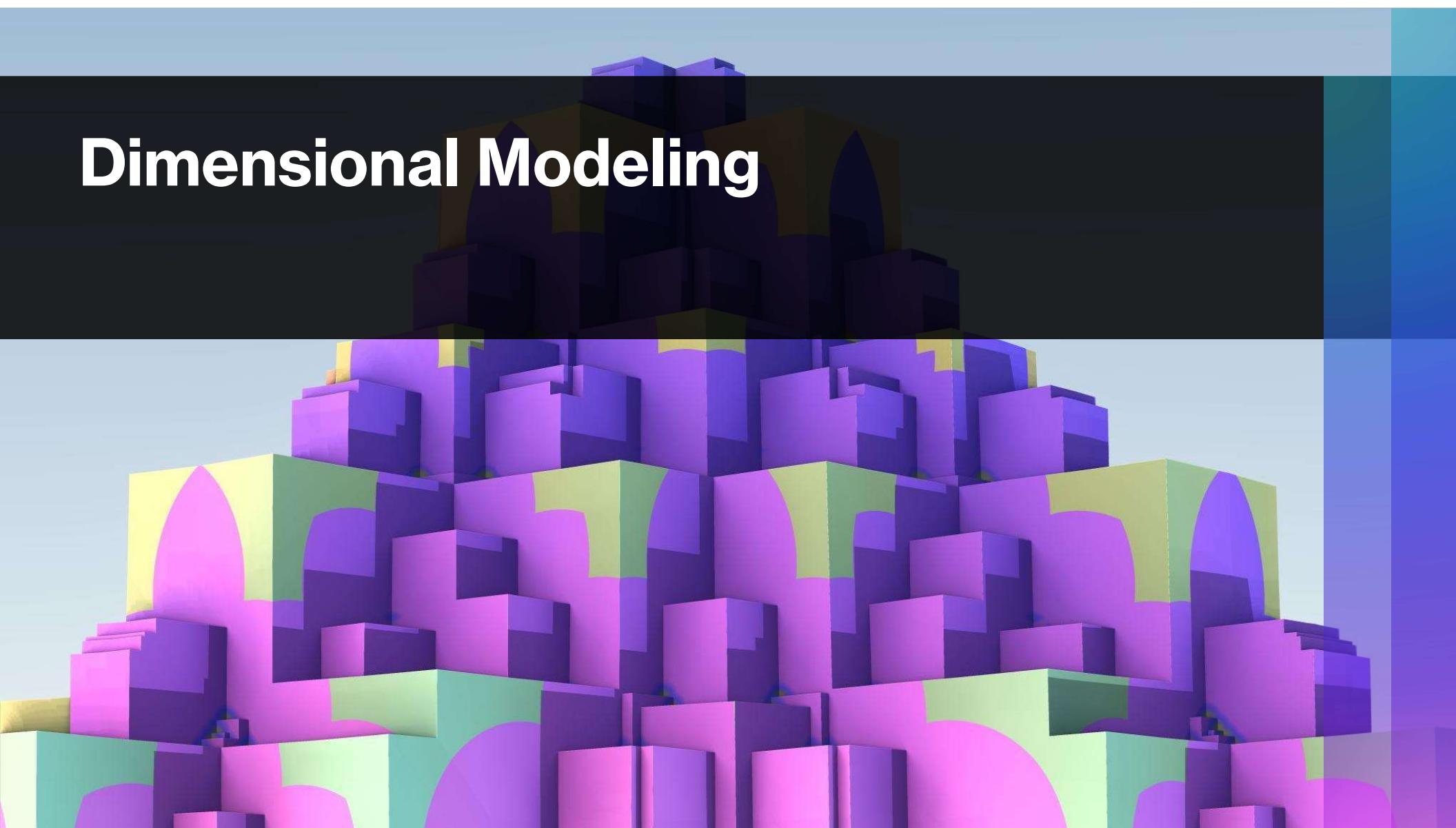


Sales  
Coverage

## Project Goal:

The business should be able to perform all their analytical needs regarding order fulfillment for Fudgmart and Fudgeflix from a single source. This will include the ability to analyze lead times of products by day of week, month of the year, department, and where the product was shipped to.

# Dimensional Modeling



# Bus Matrix

| Business Process  | Fact Table             | Fact Grain Type       | Granularity          | Facts  |
|-------------------|------------------------|-----------------------|----------------------|--|
| Order Fulfillment | fact_order_fulfillment | Accumulating Snapshot | One row per order    | order_date, shipped_date, carrier, order_to_ship_lag     |
| Sales             | fact_sales             | Transaction           | One row per sale     | product_retail_price, order_qty, ab_billed_amount        |
| Inventory         | fact_inventory         | Periodic Snapshot     | One row per product  | qty, return_date   |
| Customer Service  | fact_customer_service  | Periodic Snapshot     | One row per review   | rating, review_date                                      |
| Sales Region      | fact_sales_region      | Periodic Snapshot     | One row per location | city, state, zip, product_retail_price, ab_billed_amount |

# Detailed Dimensional Model (part 2)

Example Fact Table

| Column Name          | Display Name         | Description   | Example Values | SCD Type | ETL Rules                               |
|----------------------|----------------------|---|----------------|----------|---|
| ProductKey           | ProductKey           | Key to DimProduct   | 1, 2, 3        |          | Key lookup from DimProduct.ProductKey   |
| CustomerKey          | CustomerKey          | Key to DimCustomer  | 1, 2, 3        |          | Key lookup from DimCustomer.CustomerKey |
| CarrierID            | CarrierID            | Business key from source system (aka natural key)                                     | 1, 2, 3...     | key      |   |
| OrderDateKey         | OrderDateKey         | Key to DimDate  | 20120108       |          | Key lookup from DimDate.DateKey         |
| ShippedDateKey       | ShippedDateKey       | Key to DimDate  | 20120108       |          | Key lookup from DimDate.DateKey         |
| OrderID              | OrderID              | The natural key for the fact table, which represents an order that is being fulfilled | 1, 2, 3        |          |   |
| OrderToShipLagInDays | OrderToShipLagInDays | shipped_date - order_date   | 1, 22, 45      |          |   |

# Detailed Dimensional Model (part 2)

## Example Dimension

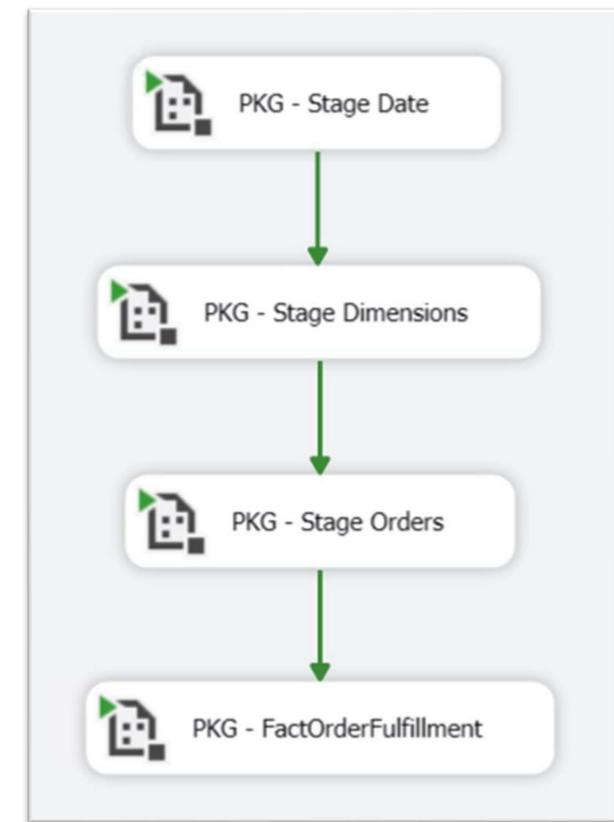
| Column Name        | Display Name       | Description   | Unknown Member | Example Values        | SCD Type | Display Folder    | ETL Rules      |
|--------------------|--------------------|---|----------------|-----------------------|----------|-------------------|----------------|
| ProductKey         | ProductKey         | Surrogate primary key   | -1             | 1, 2, 3...            | key      |                   |                |
| ProductID          | ProductID          | Business key from source system (aka natural key)             | -1             | 1, 2, 3...            | key      |                   |                |
| product_department | product_department | Department for product  | Unk Department | Electronics           | 2        |                   |                |
| product_name       | product_name       | Name of product   | Unk Product    | DVD Player            | 2        |                   |                |
| RowIsCurrent       | Row Is Current     | Is this the current row for this member (Y/N)?                | 1              | TRUE, FALSE           | n/a      | Exclude from cube | Standard SCD-2 |
| RowStartDate       | Row Start Date     | When did this row become valid for this member?               | 1/1/00         | 1/24/11               | n/a      | Exclude from cube | Standard SCD-2 |
| RowEndDate         | Row End Date       | When did this row become invalid? (12/31/9999 if current row) | 12/31/99       | 1/14/1998, 12/31/9999 | n/a      | Exclude from cube | Standard SCD-2 |
| RowChangeReason    | Row Change Reason  | Why did the row change last?                                  | N/A            |                       | n/a      | Exclude from cube | Standard SCD-2 |

# ETL Process



# ETL Process Overview

- The ETL process consisted of loading Date, Product, Customer, and Orders into the staging database.
- During staging, data conversion and derived columns fixed any differences in data type between companies.
  - Example: truncating the Zip Code from Fudgemart to the standard 10-character length and increasing Fudgeflix to 10 characters.
- Linking primary keys to the correct Fudgemart and Fudgeflix databases was integral to this process for mapping the correct orders to the appropriate customers and products.

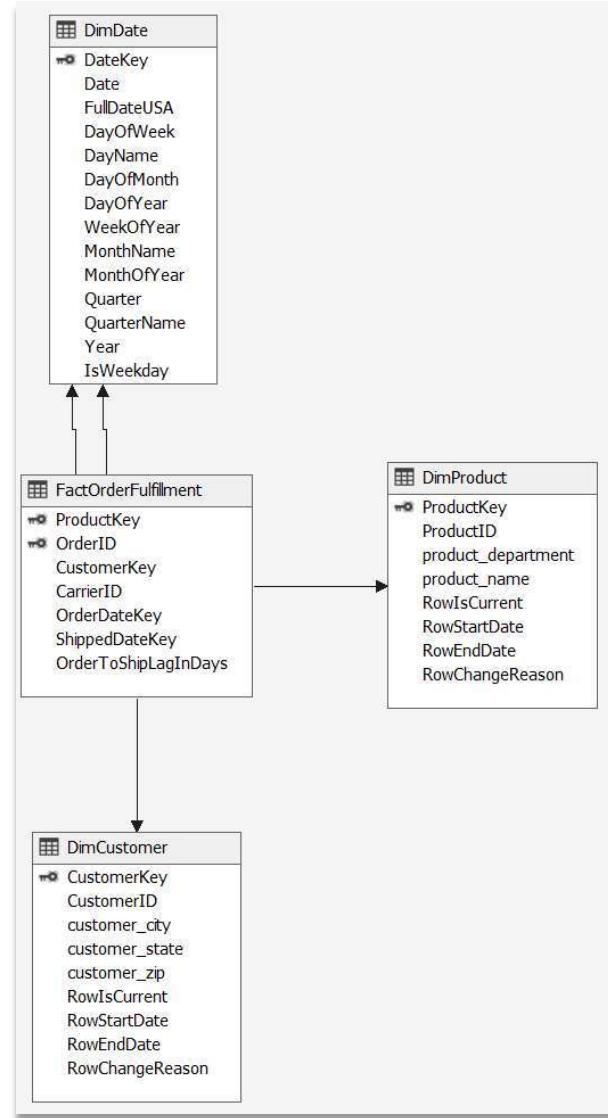


The background of the slide features a stunning image of a nebula in deep space. The nebula is composed of wispy, translucent clouds of gas and dust, primarily in shades of pink, purple, and blue. A dense cluster of stars is visible in the upper left, while a bright, central star cluster is located in the lower left. The overall effect is one of vastness and celestial beauty.

# Star Schema & Cube

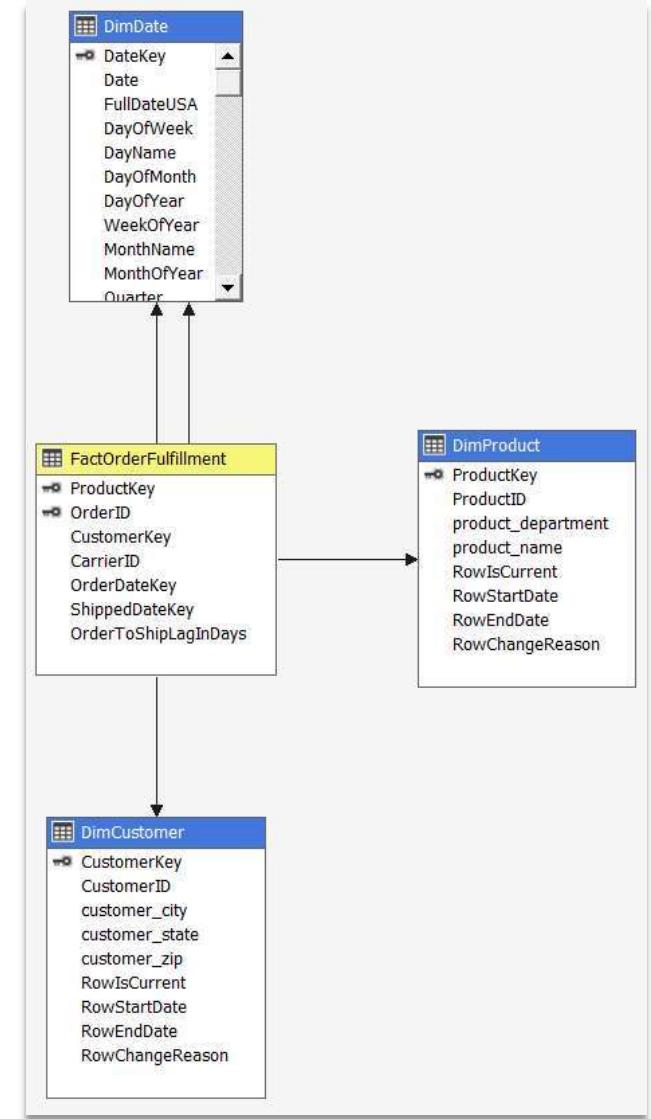
# Star Schema

- The single fact table represents the order fulfillment business process for the merged organization
- Date, Product and Customer dimensions provide meaningful context
- Our sole measure of order fulfillment performance is order to ship lag



# MOLAP Order Fulfillment Cube

- Fact table measures order fulfillment process
- Date, Product and Customer dimensions allow us to focus on year, department, and order
- Primary difference to ROLAP is a calculated measure, average order to ship lag



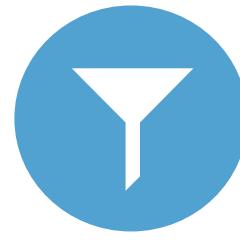
A close-up photograph of a person's hand, wearing a white shirt cuff, pointing with their index finger towards a series of 3D white bars on a dark, semi-transparent grid. The background is a blurred blue and purple, suggesting a digital or analytical environment. The overall composition conveys a sense of data analysis or business strategy.

# Business Intelligence

# Business Intelligence Goal



Empower the business  
to perform Analytics



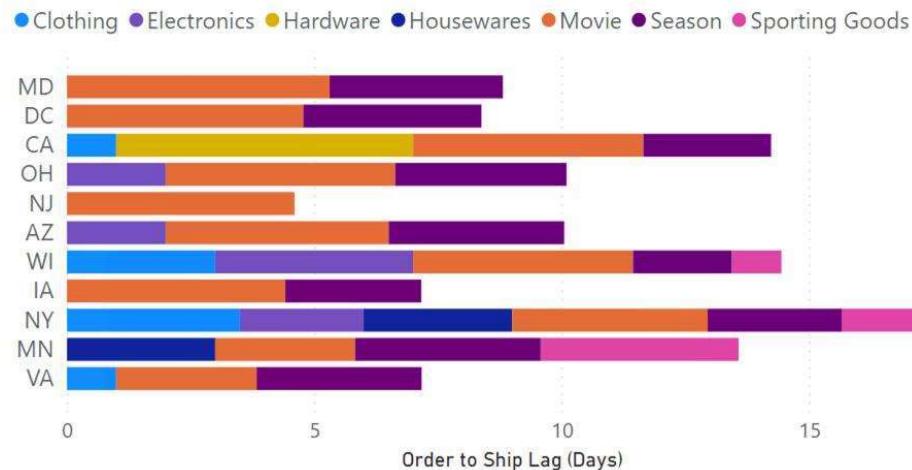
Allow access to  
multiple dimensions to  
slice and filter data



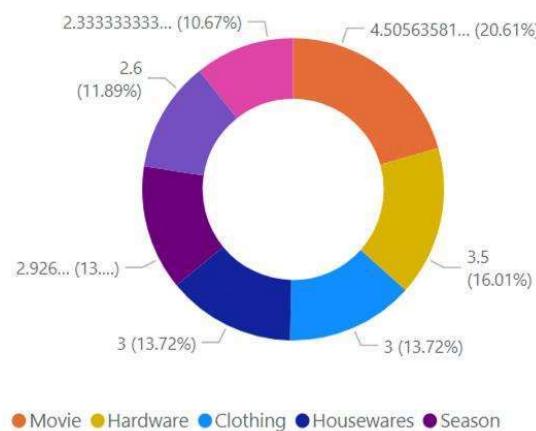
Provide tools  
necessary for  
visualization

# Order Fulfillment Dashboard

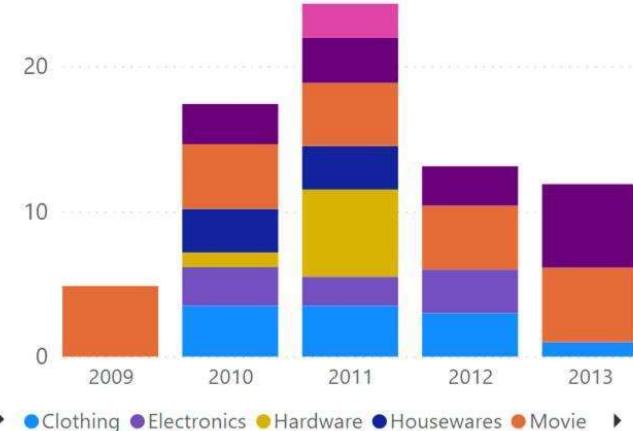
Order Lag Times by State & Department - Average



Order Lag Times by Department - Average



Order Lag Times by Year - Average



Order Lag Times by State - Average



1/1/2009 12/31/2014

Quarter

All

Order Lag Times Overall - Average

**4.48**

Average

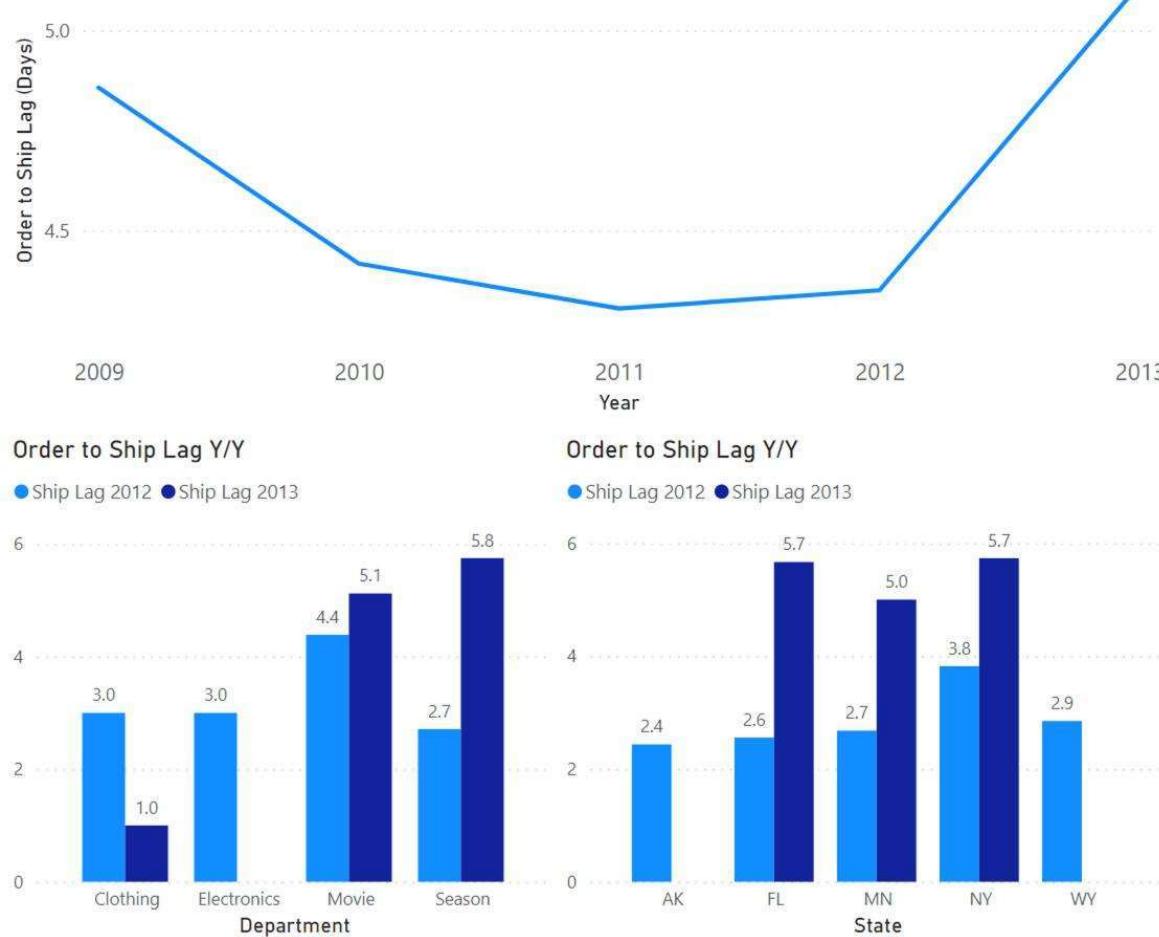
Median

Count

Stdev

# Fulfillment Health

Order to Ship Lag (Days) by Year



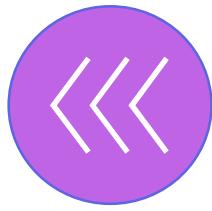
| Quarter   | 2009 | 2010 | 2011 | 2012 |
|-----------|------|------|------|------|
| Qtr 1     | 4.52 | 4.19 | 4.34 | 4.02 |
| January   | 4.97 | 4.41 | 4.45 | 4.58 |
| February  | 5.65 | 4.58 | 3.86 | 4.73 |
| March     | -    | -    | -    | -    |
| Qtr 2     | 4.21 | 4.43 | 3.95 | 4.44 |
| April     | 4.72 | 4.65 | 4.22 | 4.04 |
| May       | 5.17 | 4.36 | 4.29 | 4.38 |
| June      | -    | -    | -    | -    |
| Qtr 3     | 5.06 | 4.40 | 4.94 | 4.52 |
| July      | 4.52 | 4.60 | 4.45 | 4.46 |
| August    | 5.69 | 4.26 | 4.12 | 4.06 |
| September | -    | -    | -    | -    |
| Qtr 4     | 4.63 | 4.21 | 4.42 | 4.20 |
| October   | 4.41 | 4.23 | 4.42 | 4.56 |
| November  | 5.03 | 4.70 | 4.09 | 4.24 |
| December  | -    | -    | -    | -    |

|         |        |
|---------|--------|
| Average | Median |
| Count   | Stdev  |

# Business Recommendations



Capture received date to perform end-to end BI



Reduce lead time for processing movie orders



Offer promotional discounts to customers who experience high lead time



Transform Fudgelix into a full-fledged streaming service