

IST 722 Project Charter

Fudgeflix and Fudgemart Merger

June 2019

Haley Brown

Rebecca Karunakaran

Kyle Wojtaszek

Table of Contents

Overview	2
Objective	2
Scope	2
Team Roles	2
Schedule	3
Functional Business Requirements	3
Identified Process Outline	4
Consolidate Business Processes	4
Simple Business Processes	4
Models	5
Source-to-Target Map	5
SSIS Packaging - Staging	6
Dimensions	6
Facts	8
SSIS Packaging - Loading to DW	10
Dimensions	10
Facts	12
ETL Notes	13
Star Schema	13
Business Intelligence	13

Overview

Fudgemart, a retail corporation and Fudeflix, a movie rental store need to merge to form a new company, Fudge Inc. The merger has triggered a need to consolidate the data from the 2 companies into a single system for better management and organization.

A new project was kicked off to review, design and implement a Consolidated Database for better reporting and analysis.

Objective

The objective of this project is to implement an integrated Data Warehouse, bringing together data from both enterprises. Additionally, we will use that data warehouse to create BI reports to meet the needs of our business processes. The project will be deployed for use by late June 2019.

Scope

Activities that *are* in the scope of this project include:

- Identifying essential functional requirements and business processes
- Creating both a high-level and detailed dimensional modeling worksheet
- Implementing the models into a SQL schema/DW
- Creating an initial ETL, including a source-to-target map and SSIS package to stage and load data to DW
- Building SSAS cubes and BI dashboard
- Presenting project outcomes to executive management

Activities that *are not* in the scope of this project include:

- Full and actual deployment of any of the above activities
- Follow-up and maintenance of the DW/BI tool

Team Roles

Haley Brown: Business Lead, Data Architect

Rebecca Karunakaran: Program Manager, ETL Architect

Kyle Wojtaszek: Business Analyst, BI Architect

Schedule

Milestone	Dates
Identify functional requirements	27-Apr
High-level dimensional modeling worksheet	4-May
Detail-level dimensional modeling worksheet	11-May
Data warehouse on SQL Server	18-May
Initial ETL done in SSIS	25-May
Business intelligence reporting	8-Jun
Deployment	16-Jun

Functional Business Requirements

1. Fudgeflix Business Users need to:
 - a. Sorts Fudgeflix plans by highest sales, state
 - b. Pull Customer Review Reports by star rating, product, product type
 - c. Pull monthly and yearly Customer Reports that detail total purchases
 - d. View the most popular movies, cast members by genre
 - e. Query the number of movies queued per day of week
 - f. Query the number of rentals by movie rating (PG, PG-13, R)
 - g. View sales by plan type, region and movie genre
 - h. View movie demand by review/genre.
2. Fudgemart Business Users need to:
 - a. View sales by department, vendor
 - b. Pull monthly status reports by shipment, product, vendor, department, employee
 - c. View sales by popularity of the product (based on reviews)
 - d. View monthly and yearly Customer Reports by that detail total purchases
 - e. Query the average time it takes to ship an order from its order date
3. Both kinds of Users need to:
 - a. View Sales Reporting for all products over both subsidiaries
 - b. View Customer Information over both subsidiaries by various filters (region, product interaction, etc) for cross-subsidiary marketing purposes

Identified Process Outline

**processes in blue are fully modelled through DW and BI reporting*

Consolidate Business Processes

- **Consolidated Sales Reporting:** Senior management would like to be able to track sales (dollar amount and order count) across both FudgeMart and FudgeFlix by customer, product, supplier, time period, and/or region.
 - Business justification
 - This process will allow management to identify top selling items, suppliers, and regions, with special interest to what times of year Fudge Inc. sales are lagging.
 - With sales from both subsidiaries consolidated into a single warehouse, the marketing team will more easily be able to identify opportunities for cross-marketing promotion in certain regions.
 - Possible BI reporting
 - Top selling products by quantity and by dollars per month, quarter, year
 - Sales (dollar amount and order count) by state
 - Sales (dollar amount and order count) by time period (year, quarter)
 - Percent of total sales
- **Consolidated Customer Satisfaction Integration:** Senior management would like to map where customer base is located and what products are popular with specific customers.
 - Business justification
 - This information will be used for cross-marketing purposes to get more customers to engage with products from both Fudgemart and Fudgeflix, based on their past interaction with products.

Simple Business Processes

- **FudgeFlix Queue Reporting:** FudgeFlix management wants to track movie queues/rental volume based on factors such as movie rating, day of the week, region, and genre
 - Business justification
 - This process will help IT management make recommendations about server capacity planning and future movie rights purchases.
 - Possible BI reporting
 - Queue volume by day of the week, rating, genre
 - Time to ship per day of week

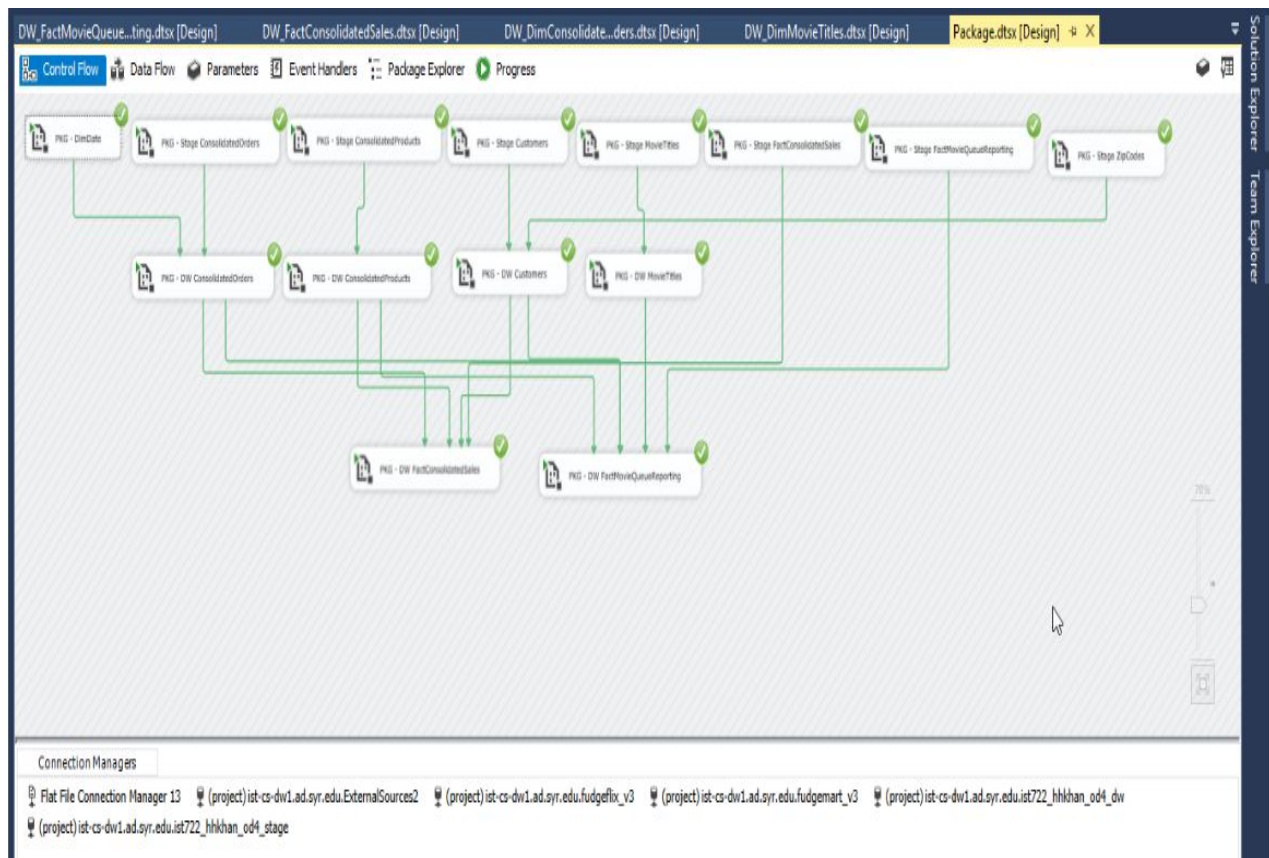
- **FudgeMart Shipment Processing:** FudgeMart management would like to see metrics on the time between when an order is placed and when an order is shipped.
 - Business justification
 - This ability would provide the shipping department helpful information about where ship times can improve.

Models

Please consult the attached documentation for the following deliverables:

- High-level dimensional modeling worksheet
- Detail-level dimensional modeling worksheet
- SQL schema generation script

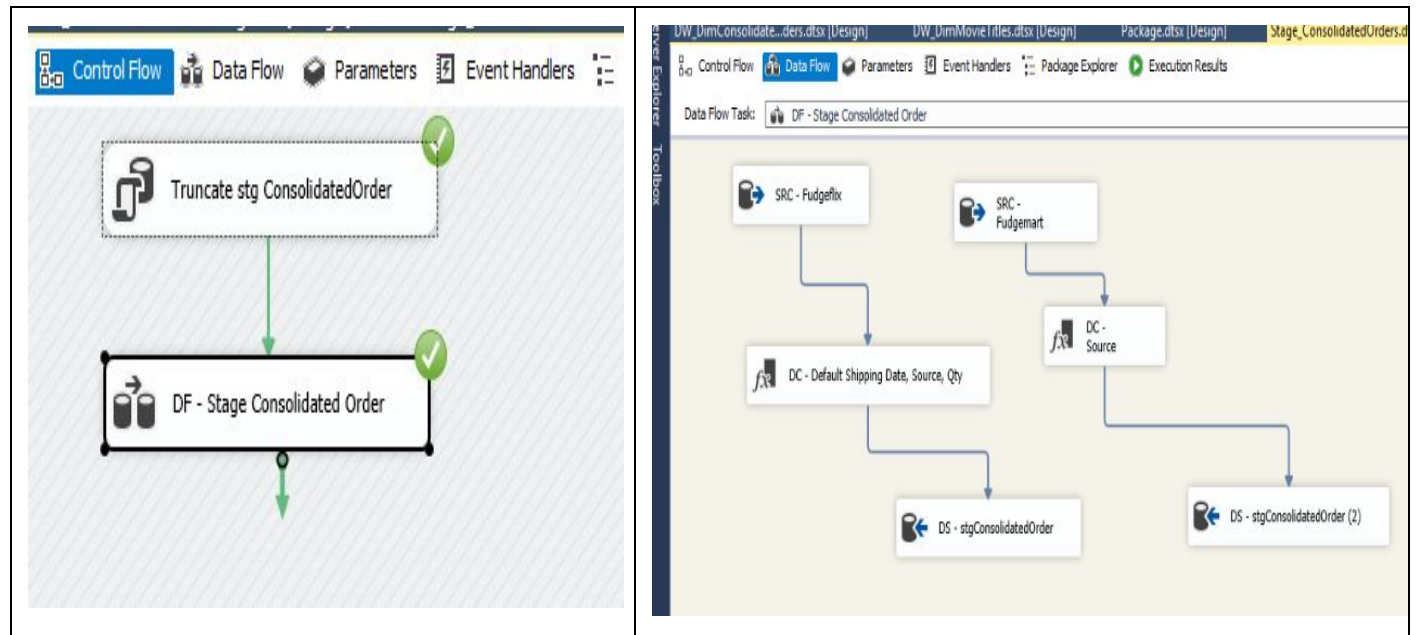
Source-to-Target Map



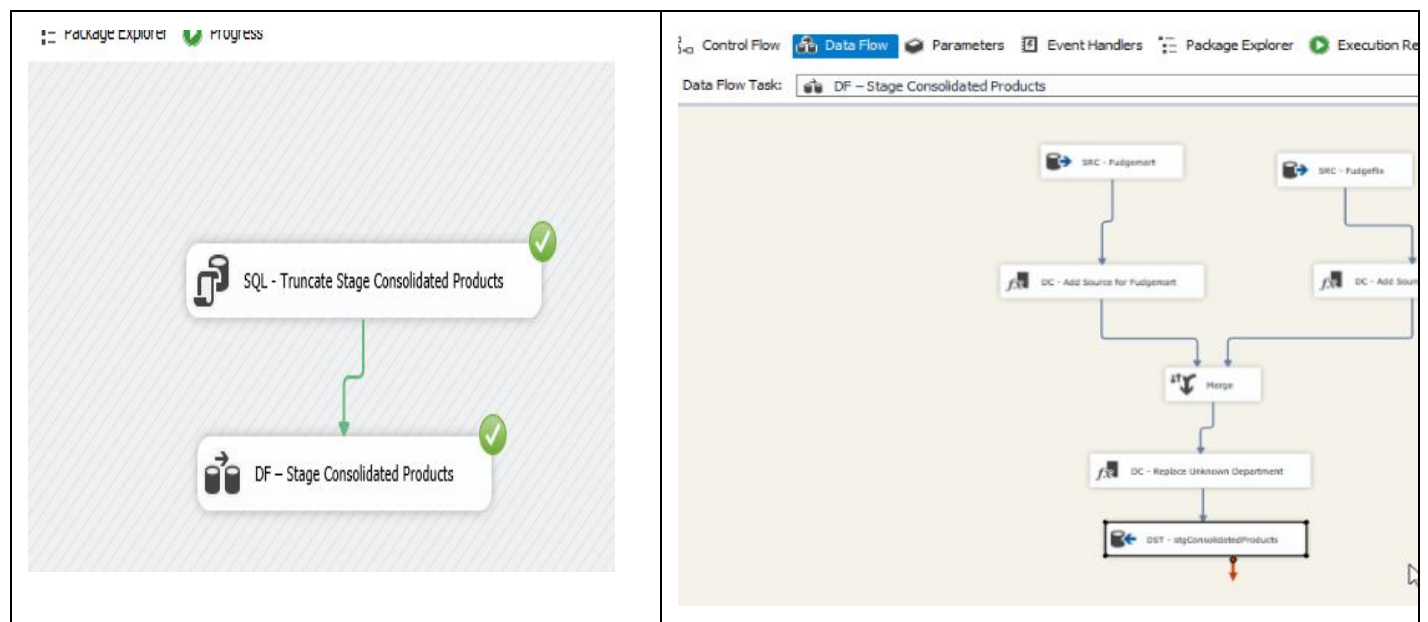
SSIS Packaging - Staging

Dimensions

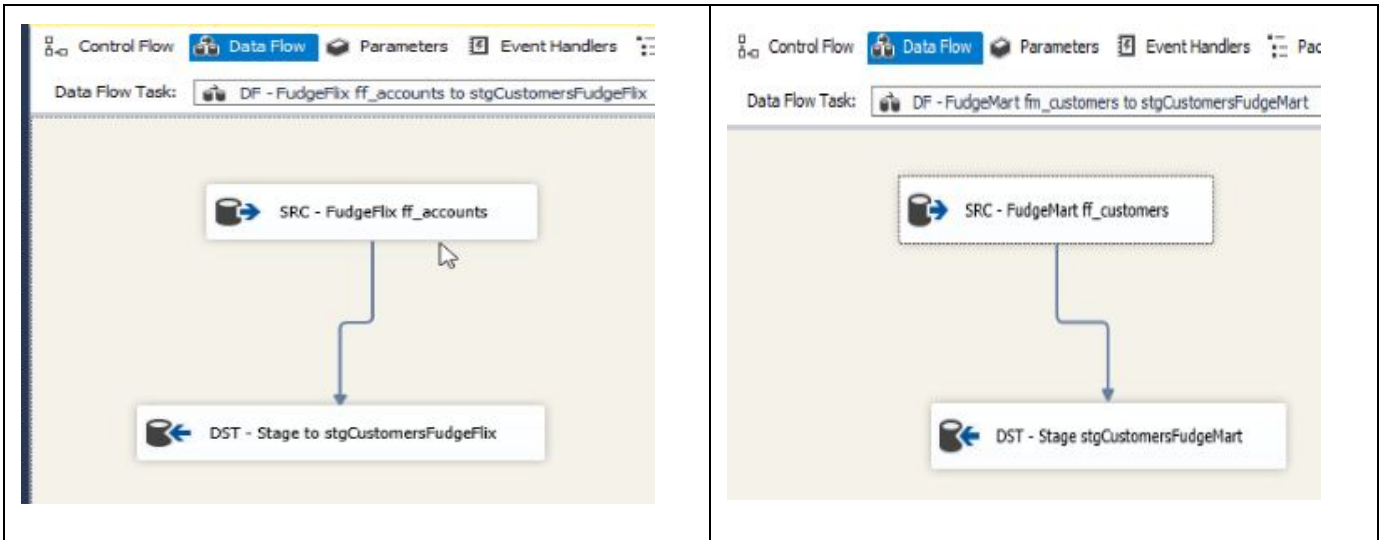
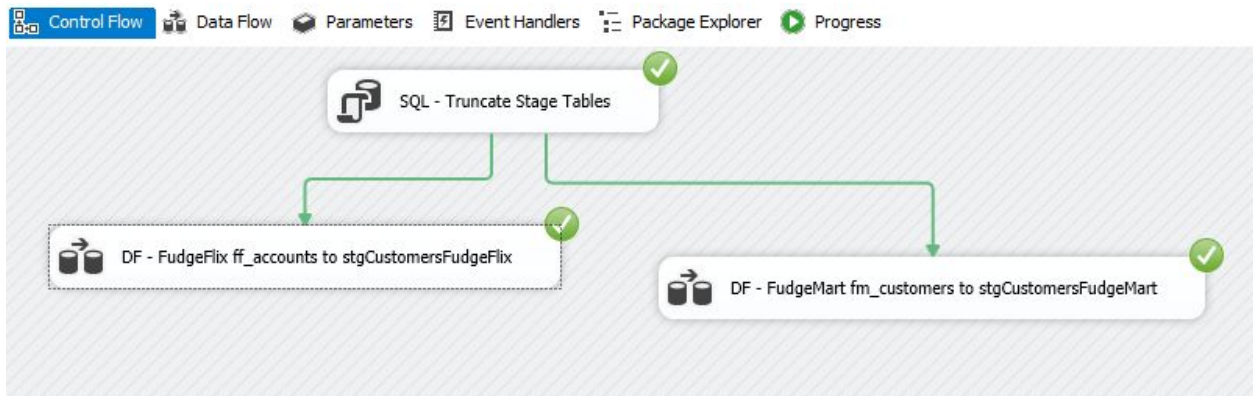
Consolidated Orders



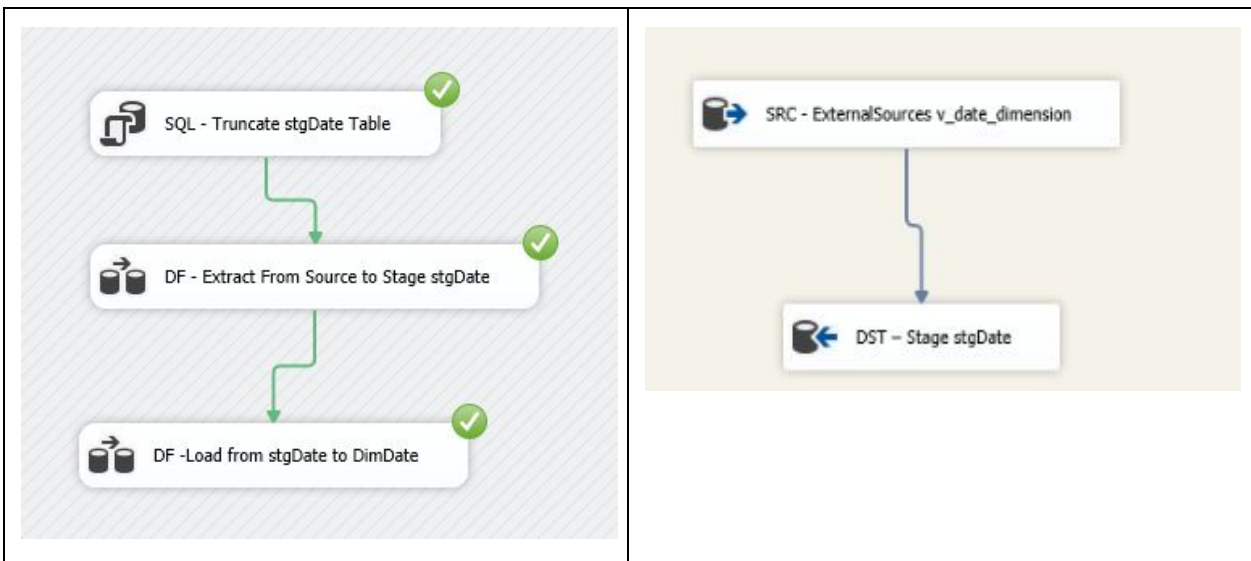
Consolidated Products



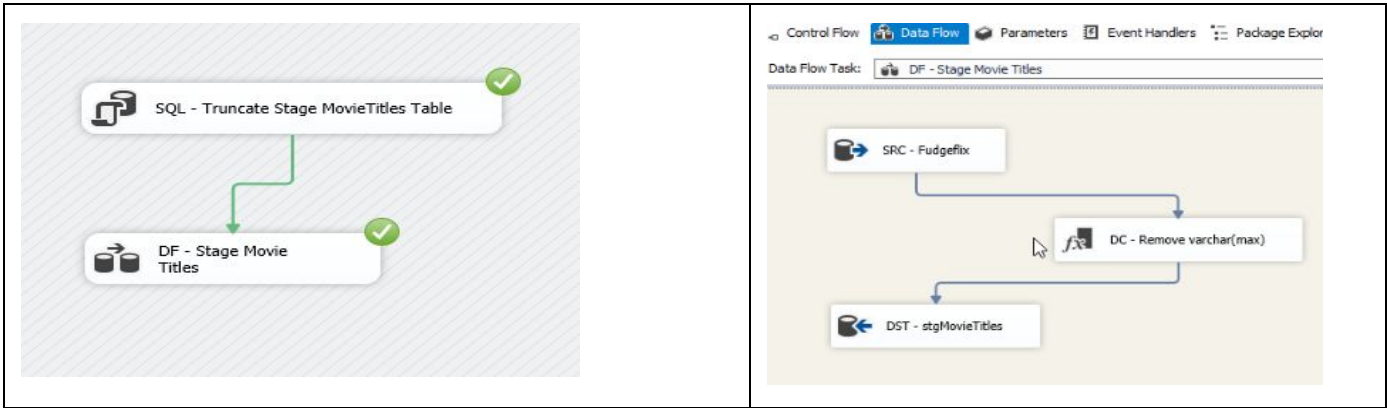
Consolidated Customers



Dates



Movie Title

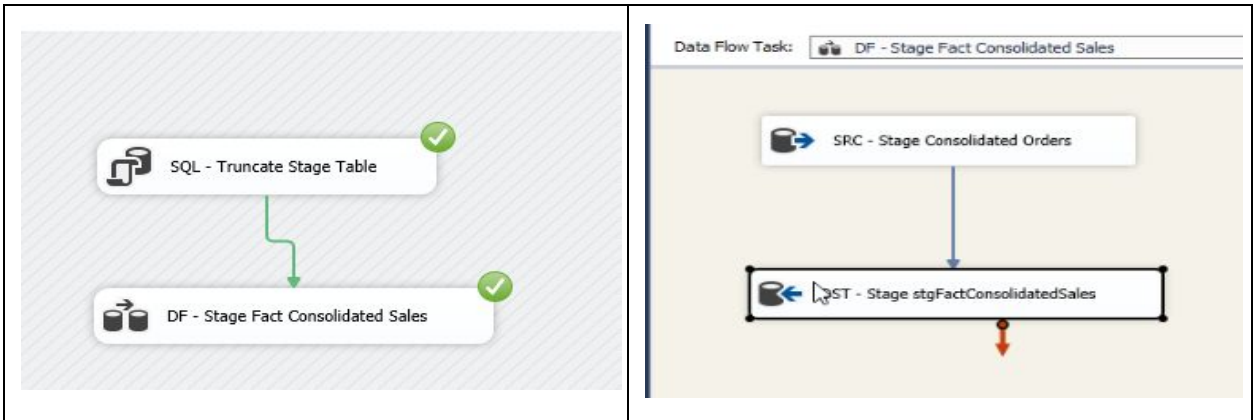


ZipCodes



Facts

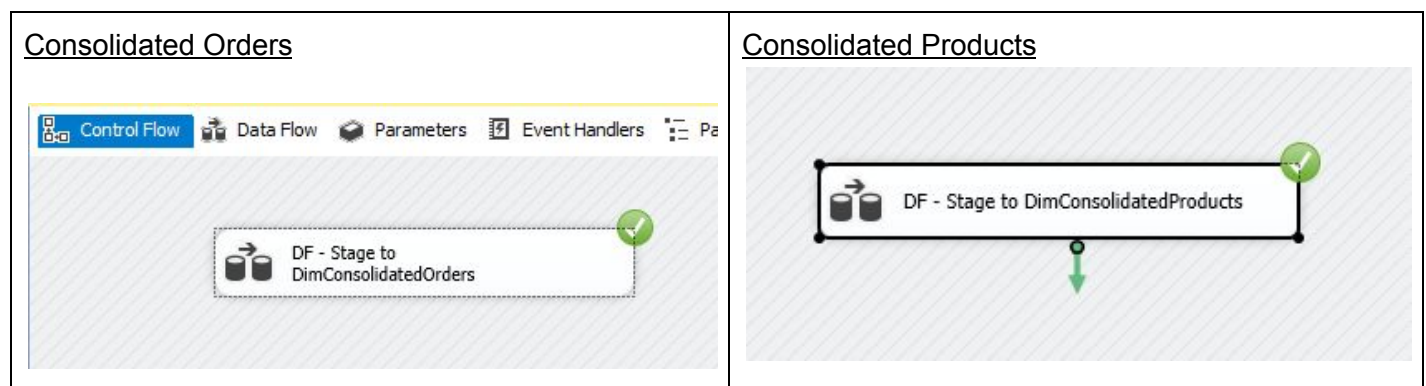
FactConsolidatedSales

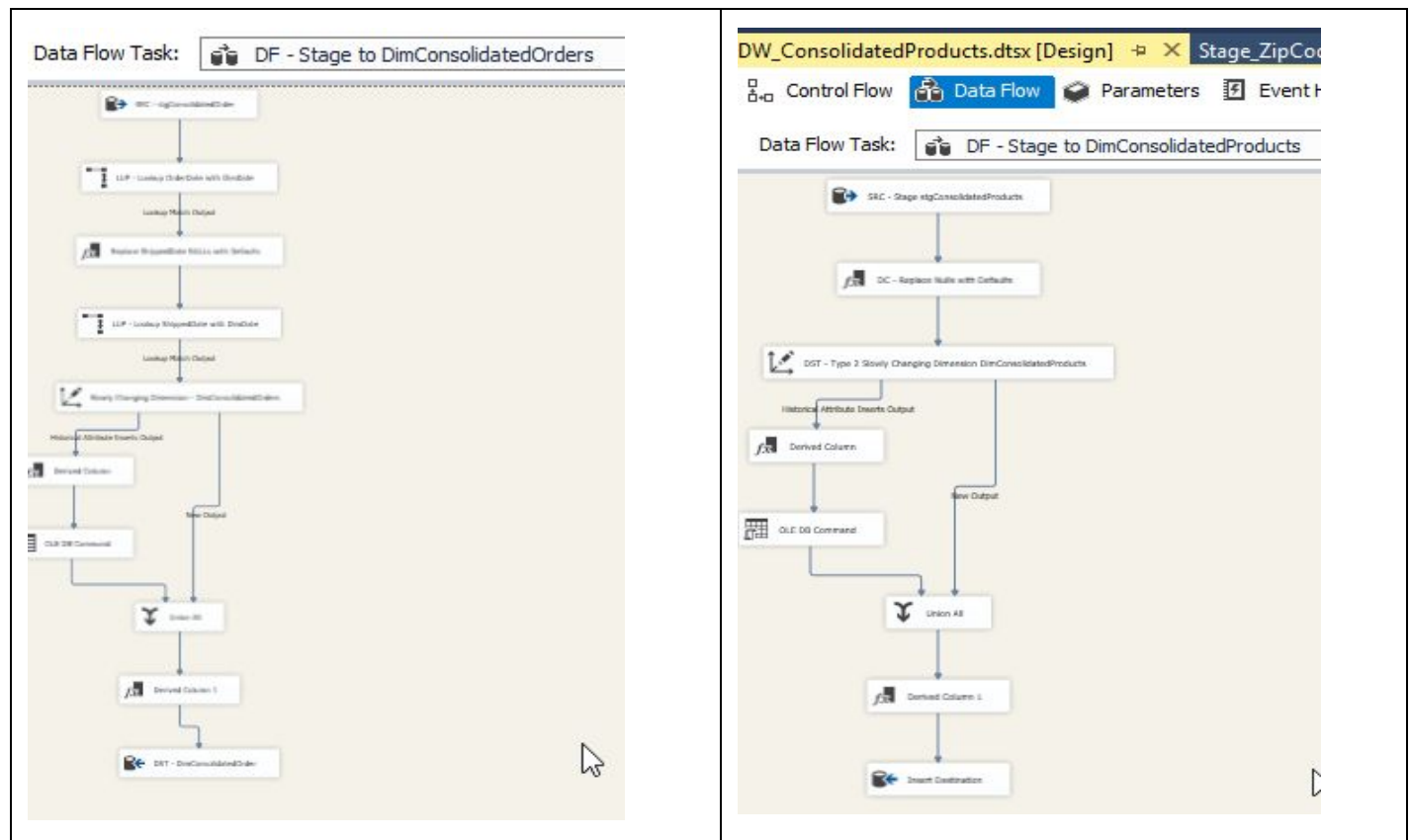


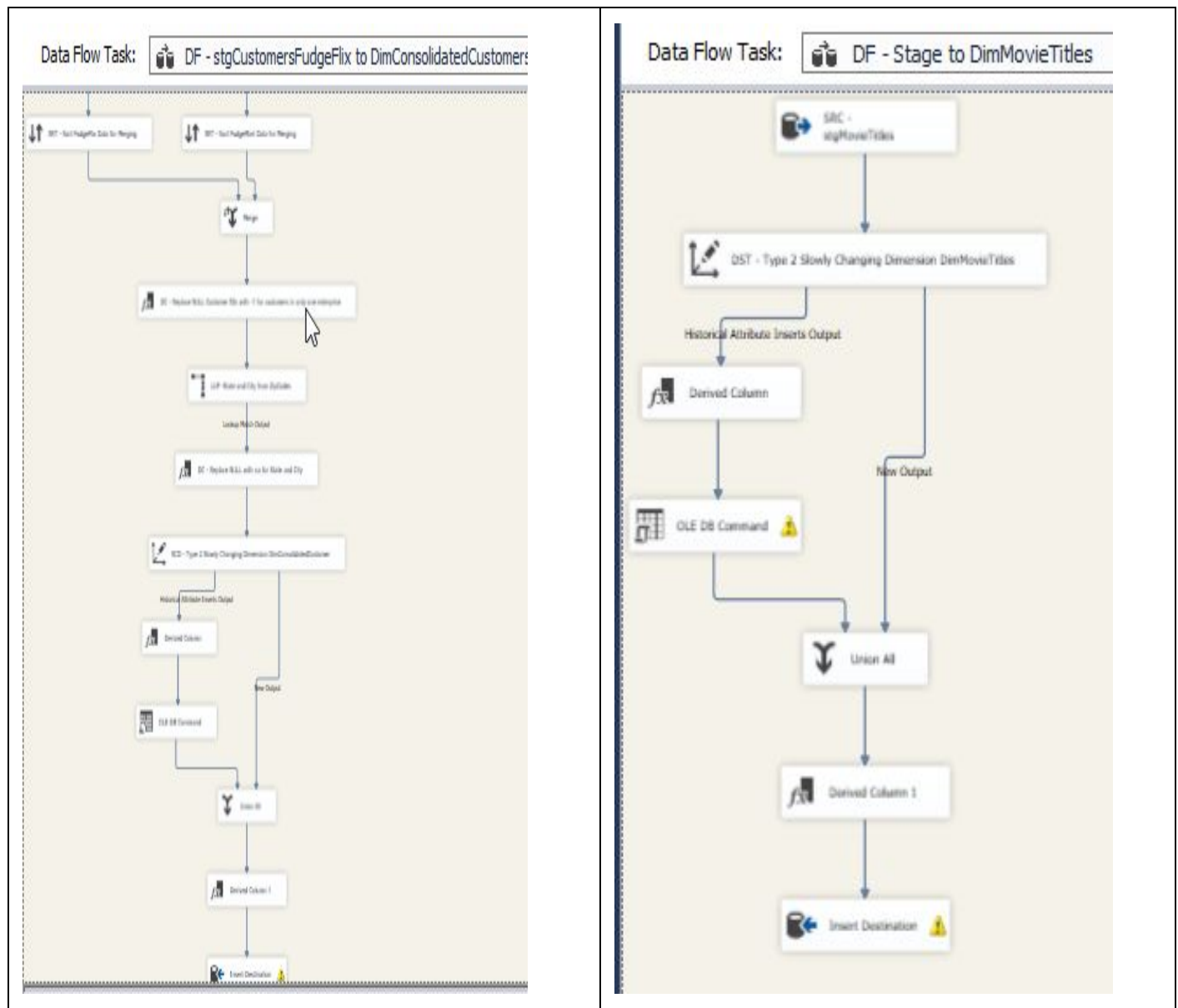
FactMovieQueueReporting

SSIS Packaging - Loading to DW

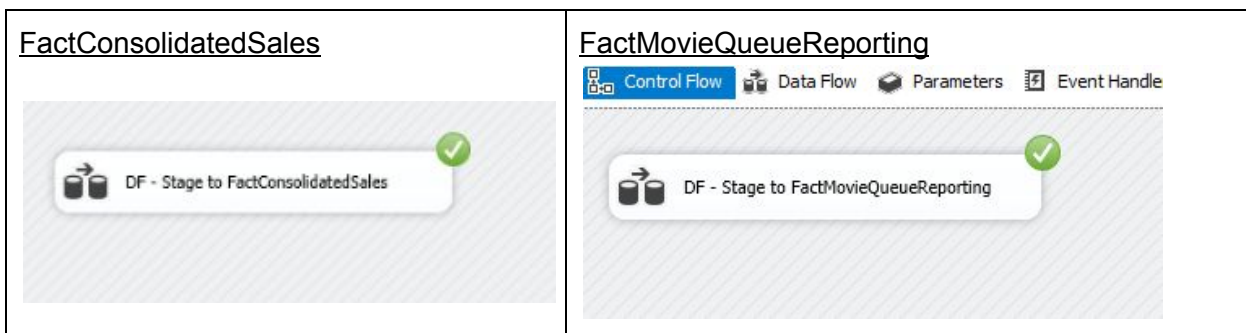
Dimensions

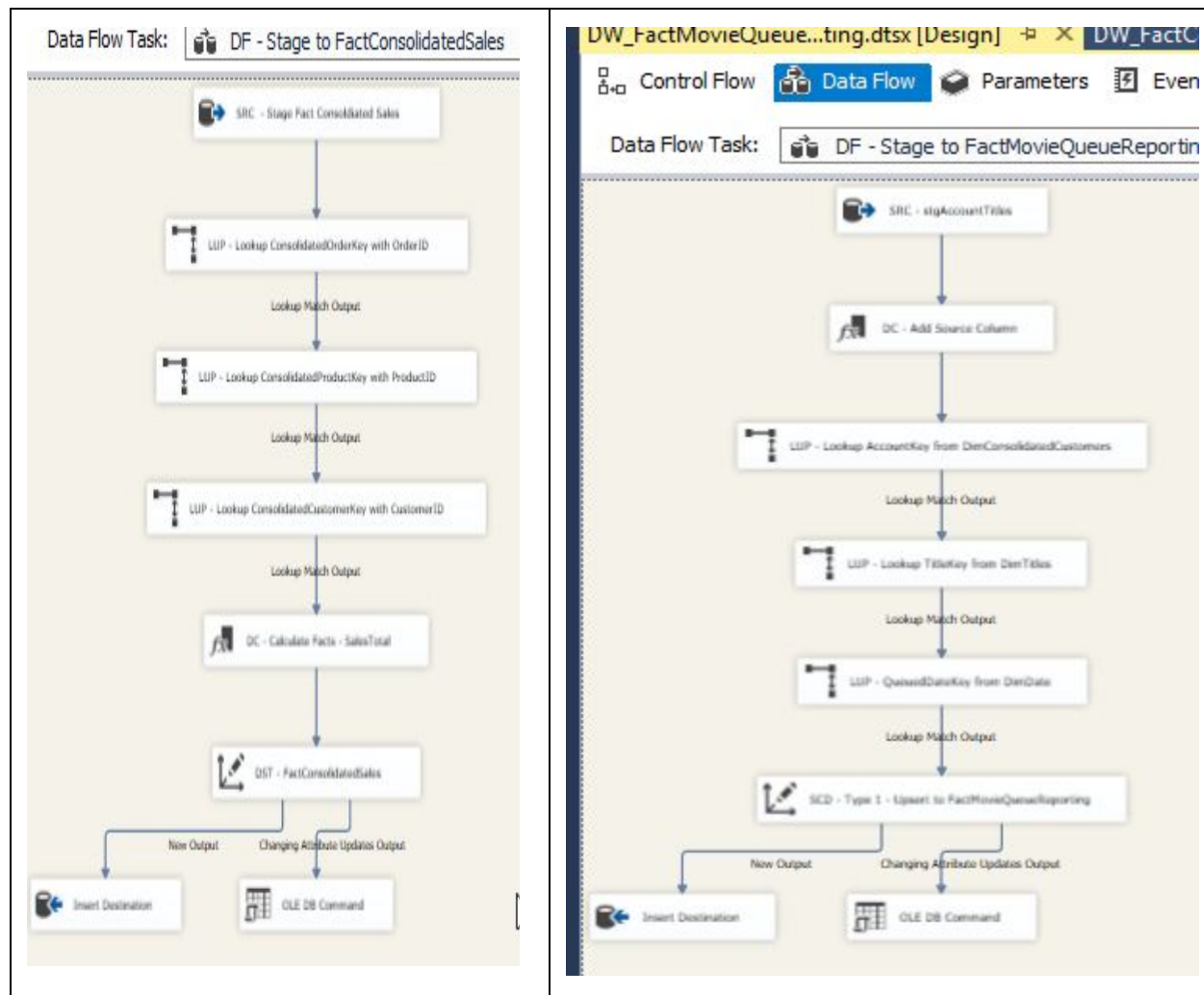






Facts





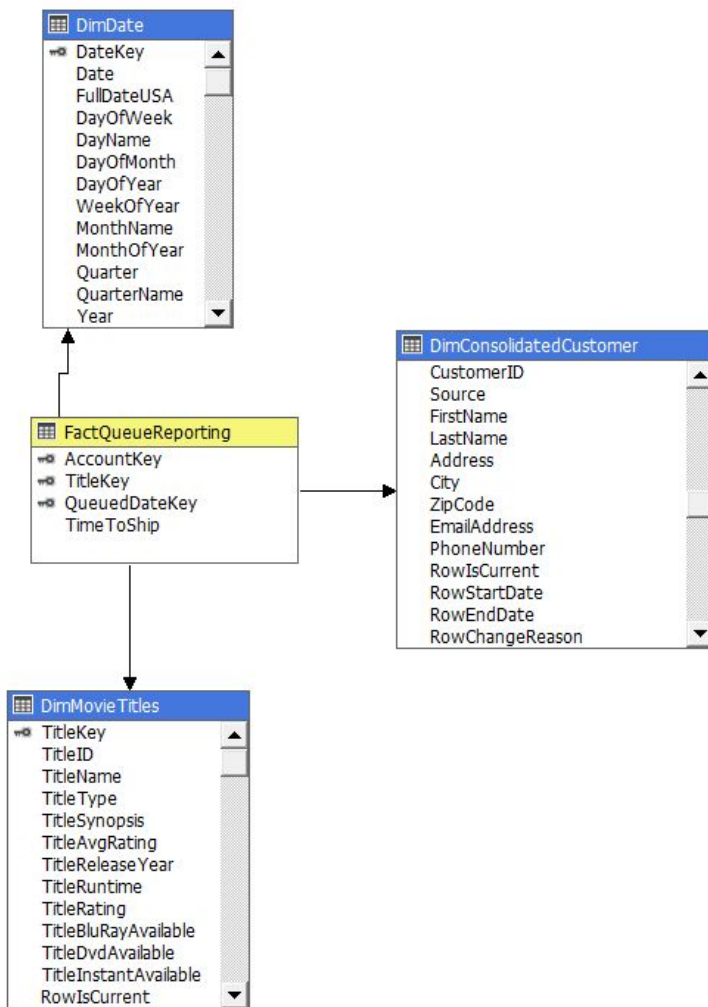
ETL Notes

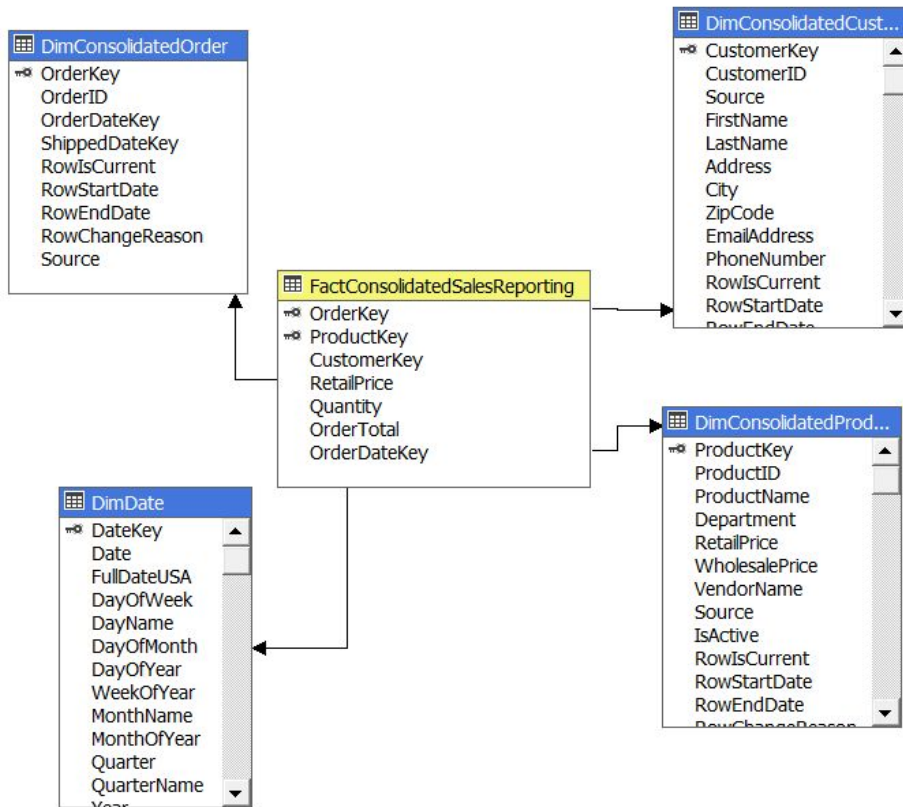
- For our consolidated tables, we added a Source column which indicates which OLTP the data came from. In order to make this look nice for reporting purposes, we added a DimSource table which translates the Source ID to a human readable name. This table is hardcoded and has 3 rows: An 'unknown' row, a FudgeFlix row, and a FudgeMart row.
- When combining the customer tables from FudgeMart and FudgeFlix, we had to decide what to do with customers who use both services. At first, we were going to combine them and have two columns in each row, one for the FudgeFlix ID and one for the FudgeMart ID, which would allow a user to quickly determine which companies a customer used. However, after further considering this option we decided instead to copy all columns and just include the aforementioned Source column to indicate which company the data came from. If knowing which customers that make orders from both

subsidiaries was important for a certain project, a layer to indicate this would need to be added to the BI layer.

- The reasoning for this choice was that it was unclear how to determine if two customers were the 'same'. Obviously name is not enough, name and zip code is reasonably good, but still not perfect. And even if we had a perfect solution for matching, it is unclear if all business units would want us to match them. For example, if a person has a FudgeMart account for business purposes and a FudgeFlix account for personal purposes, do we want to count that as one customer or two? Different business units may answer that question differently, so we felt it best to leave the data in a form where each business unit can answer the question how they would like by adjusting their queries.

Star Schema





Business Intelligence

Please consult the attached documentation for live Power BI dashboards.

Screenshots and detailed recommendations are outlined below.

- **Global Sales Dashboard.**
 - **Goal :** Senior Management would like a dashboard that would show high level Sales Volume by Year/Quarter, Product/Product Category and by Region/State.
 - **Tool :** Power BI
 - **Recommendations:**
 - Sales numbers have dropped since 2009 which indicates a need to expand Customer Base. Based on the Sales Volume by Region, the middle of the country would be a good target area!
 - New Marketing Campaign and Sales Promotions to boost Sales Numbers for all Product Departments other than Electronics like Housewares/Clothing/Sporting goods which together currently comprise less than 35% of the total sales volume.

High Level Sales Dashboard

Global Sales Dashboard

This dashboard shows an overview of worldwide sales for Fudge Inc

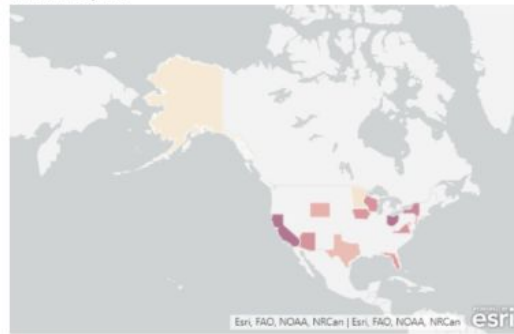
On the right is a chart which shows sales by product department/product.

Below from left to right -
Raw Sales report by State/City
Map of the Sales by State
Sales Volume by Year

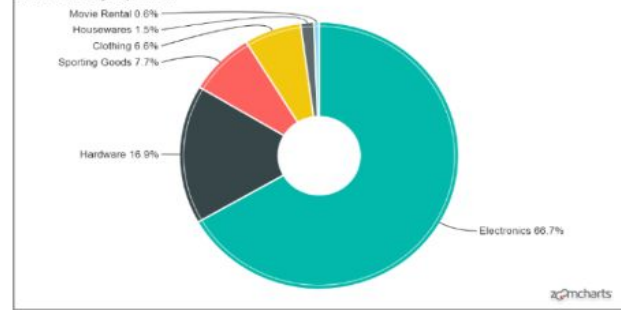
Sales Volume Raw Data

State	City	Order/Total
CA	FRESNO	\$356,576.20
OH	COLUMBUS	\$342,811.23
CA	LOS ANGELES	\$226,017.68
DC	WASHINGTON	\$220,905.66
NY	NEW YORK	\$151,188.19
CA	ANAHEIM	\$151,017.83
CA	VALENCIA	\$137,796.60
OH	CLEVELS	\$136,998.03
CA	MANHATTAN BEACH	\$133,123.78
WI	MADISON	\$132,782.43
AZ	PHOENIX	\$132,575.58
VA	LYNCHBURG	\$130,581.23
NJ	FAIRFIELD	\$118,097.80
NY	TARRYTOWN	\$113,857.05
CA	SANTA ANA	\$113,363.10
CA	SAN FRANCISCO	\$111,065.90
CA	SACRAMENTO	\$101,716.43
IA	FAIRFIELD	\$97,271.13
OH	CINCINNATI	\$88,432.68
Total		\$3,006,200.51

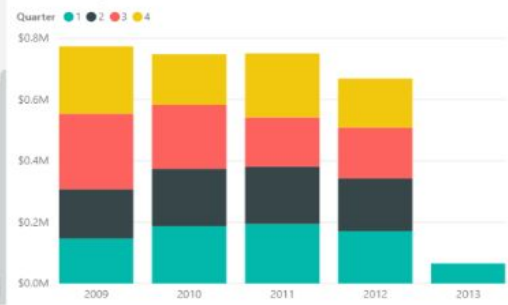
Sales Volume By State



Sales Volume By Department



Sales Volume by Year



Drill-down by Product Department for Detail Sales Numbers by Product

Global Sales Dashboard

This dashboard shows an overview of worldwide sales for Fudge Inc

On the right is a chart which shows sales by product department/product.

Below from left to right -
Raw Sales report by State/City
Map of the Sales by State
Sales Volume by Year

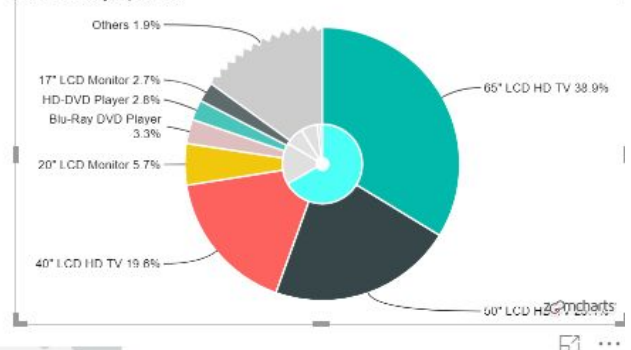
Sales Volume Raw Data

State	City	Order/Total
CA	FRESNO	\$243,014.00
OH	COLUMBUS	\$220,041.00
DC	WASHINGTON	\$145,710.00
CA	LOS ANGELES	\$143,480.00
CA	ANAHEIM	\$112,105.00
NY	NEW YORK	\$103,060.00
OH	CLEVELS	\$97,533.00
CA	VALENCIA	\$92,530.00
WI	MADISON	\$91,691.00
AZ	PHOENIX	\$90,057.00
NJ	FAIRFIELD	\$85,534.00
CA	MANHATTAN BEACH	\$84,625.00
CA	SANTA ANA	\$83,633.00
VA	LYNCHBURG	\$81,398.00
NY	TARRYTOWN	\$78,081.00
CA	SAN FRANCISCO	\$71,273.00
CA	SACRAMENTO	\$70,269.00
IA	FAIRFIELD	\$60,929.00
OH	CINCINNATI	\$50,558.00
Total		\$2,005,521.00

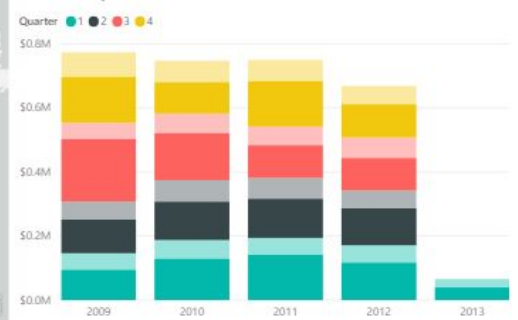
Sales Volume By State



Sales Volume By Department



Sales Volume by Year



Quarterly and States Dashboards

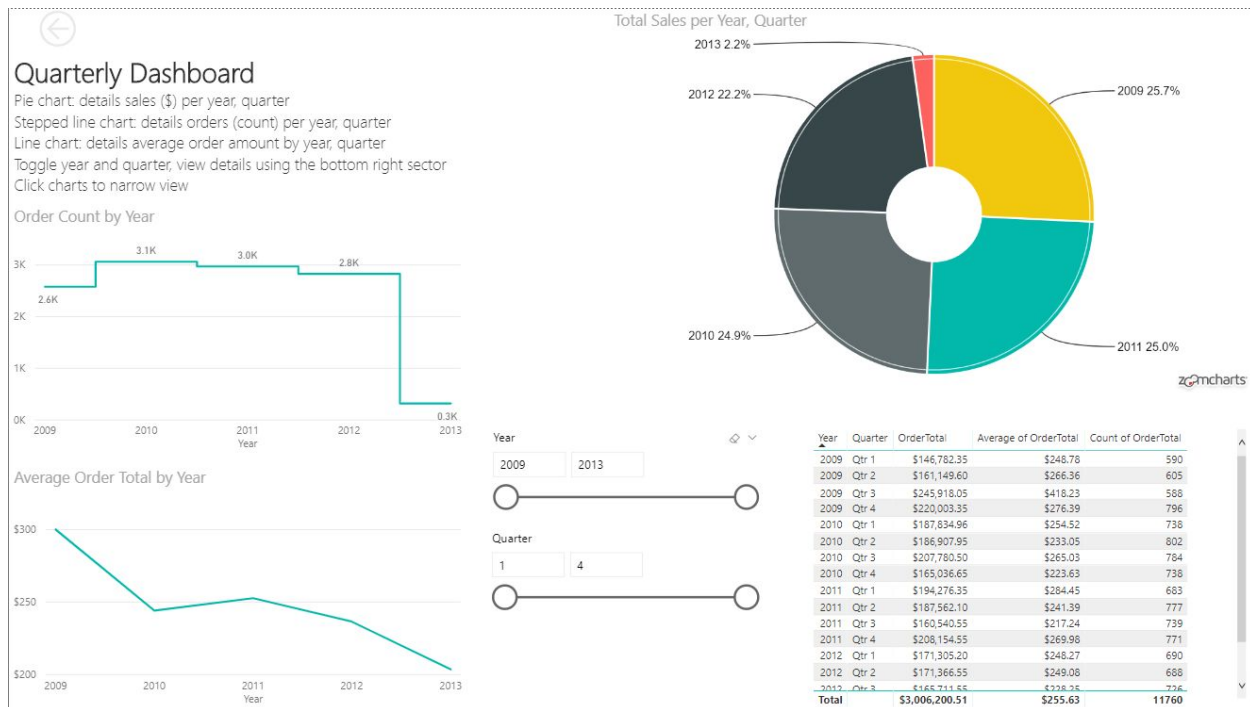
- Goal: Senior management would like to drill down further using time (year/quarter) and state. The following dashboards visualize average order amount (in dollars) and order count as well, providing further detail to support the Global Sales Dashboards. They will provide information about when sales are

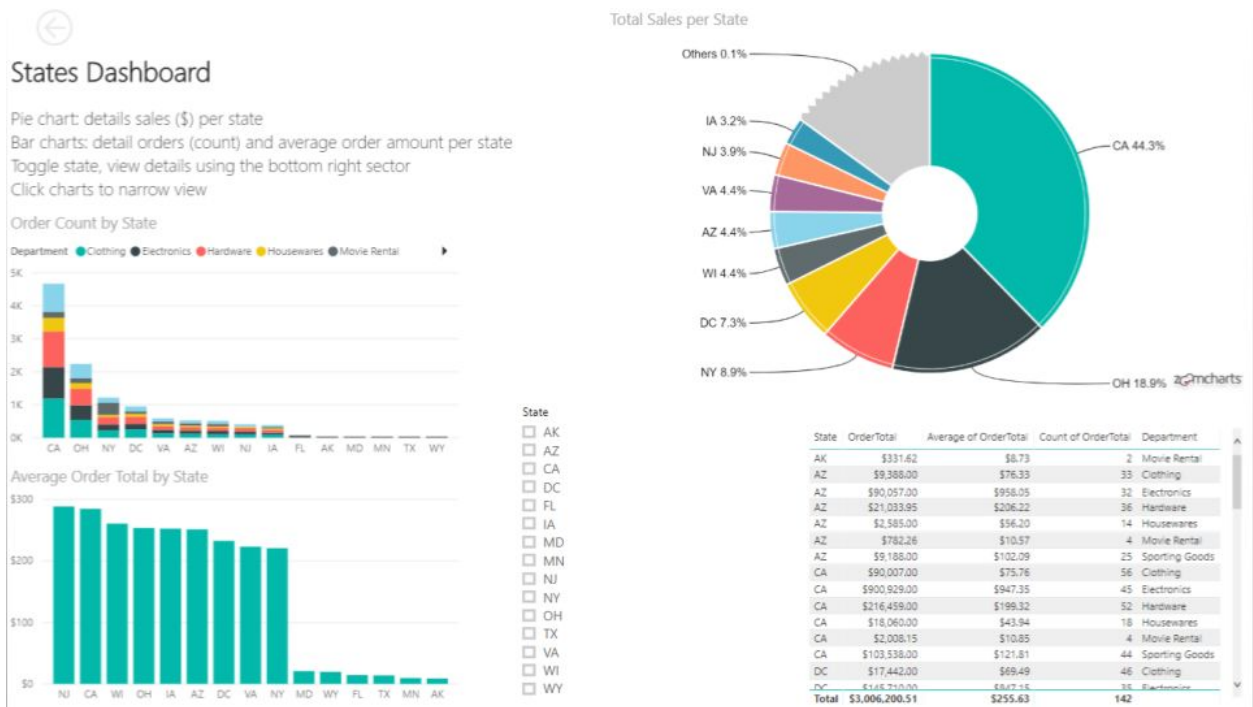
lagging or booming and what specific activity is happening in our different customer base states.

○ **Tool:** Power BI

○ **Recommendations**

- The average order total in 2012 dropped to about \$236. If this trend continues into 2013, marketing should consider promotions and incentives that will boost the average order amount. Possible incentives include free shipping for orders over \$250 or BOGO deals
- The number orders per year is also down slightly from 2011 to 2012. If the trend continues, marketing can also create incentives for customers to make future orders. Possible incentives include discounts on the next order or implementation of a points program where customers earn points for each order.
- California, Ohio, and New York are states with the top sales in both dollar amount and order count. Fudge Inc has several up and coming states however. Management needs to boost sales in these states and states not currently represented by our customer base.





- **FudgeFlix Queue and Media Data**

- **Goal:** To allow management and our supply chain teams to get insight into queuing demand and time to ship performance in order to make proper adjustments to ensure that our movie inventory and supply chain are serving customers needs
- **Tool:** Power BI
- **Recommendations:**
 - Need to improve warehouse staffing on Monday and Tuesday, as our time to ship is lagging behind on those days
 - Need to consider investing in new marketing strategies to connect with families, as our PG and PG-13 rentals seem to be lagging behind what one might expect.
 - Current ratio of movies available instantly seems to be working, but we need to continue to monitor as streaming becomes more popular.

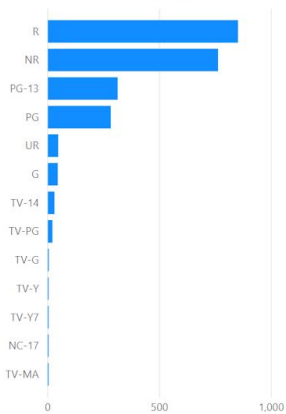
FudgeFlix Queue Data

Below is aggregate data around FudgeFlix's queue and shipping process.

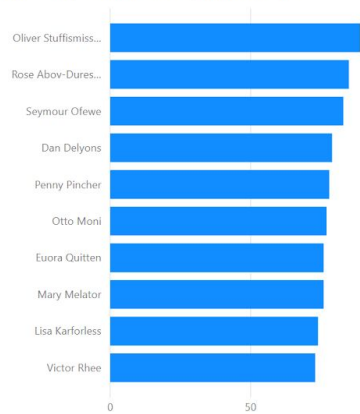
We can see that our Time to Ship takes a hit on Mondays and Tuesdays. Recommend increasing our warehouse capability on those days.

R rated movies remain our most popular movies and we should continue to invest in stocking more R movies than other titles.

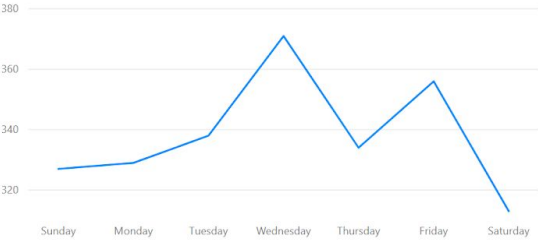
Queued Titles By MPAA Rating



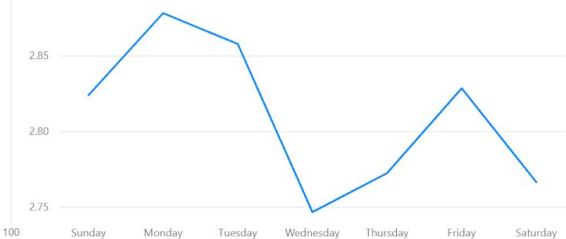
Top 10 FudgeFlix Customers by Queued Movies



Queued Titles Per Day of Week



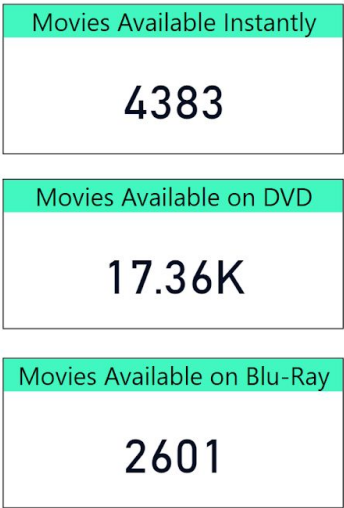
Average Time to Ship Per Day of Week



FudgeFlix Media Report

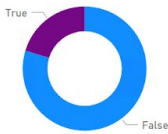
To the right we can see the distribution of various media types in our video inventory vs. the same media types in our customer's orders.

Currently our inventory is well inline with customer demands, however as our customers continue to become more digitally driven, we should keep an eye on these distributions and act accordingly.

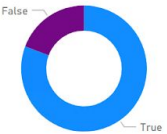


Movie Stock vs. Customer Orders

Instant Available



DVD Available



Blu-Ray Available

