# IST 722- DATA WAREHOUSE PROJECT REPORT

# Fudgemart Data Warehouse

Team Members:
Aditya Chauhan
Anisha Sudhesh Kumar
Khushabh Thakar
Niti Saluja

**PROJECT CHARTER**

## Objectives:

We have been presented with a conglomerate, Fudgemart Inc., consisting of two subsidiary companies, Fudgemart and Fudgeflix.
Fudgemart- It is an online retailer with 6 departments, 53 primary products being sold across them with 33 employees working these 6 departments.
Fudgeflix- It is an online DVD by mail and video on demand service providing users with 7185 movies across 485 different genres.
Currently Fudgemart and FudgeFlix have two independent databases along with a third supporting database. Fudgemart has commissioned the creation of a centrally integrated DW/BI solution from scratch, which will involve:

- staging the required data
- performing ETL
- transferring the data onto the warehouse
- using BI tools to generate actionable insights

## Business case:

The Data Warehouse is being created in order to enable the seamless integration of the two given databases into a single data warehouse, compatible with Business Intelligence tools. Essentially, this is done in order to achieve Subject- orientation, Non-volatility, Integration, and Time- Variance, hence fulfilling the very need for a Data Warehouse.

## Project plan:

Functional requirements:
Our main intention of this project is to various business questions revolving around Fudgemart and Fudgeflix, some of them being:

- Understanding how marital status and income levels affect customer behavior under different product departments
- Identifying regions that have a higher difference between requested and shipped dates, to improve functioning in those areas
- Evaluating customer trends to see which customers fall into the category of high spenders per year for both Fudgemart and Fudgeflix
- Assessing quality of various products to suggest best vendor to purchase the product from
- Understanding details like movies watched, number of movies shipped, which can help us make recommendations
- Understanding positive and negative views for both business fronts
- Revenue generated by each of the individual subscription plans

- Evaluating the top 10 movies by year, ranked by average customer reviews

**Business Processes:**

The business processes we have chosen are as follows:

- Shopping behavior: Shopping behavior of different customers based on the product categories. The customers can be grouped by the marital status and income. This helps us understand how the departments cater differently to different sections of people, and if certain products need to be ordered more than others.
- Delivery efficiency: Determining the shipments which have a lot of difference between the requested and shipped date, based on their zip codes. This will help us understand areas where the delivery process needs to be improved.
- Revenue reporting: Evaluating the customer trends to see which customers fall into the category of high spenders per year for both Fudgemart and Fudgeflix. This will eventually help us provide customer vouchers, extra discounts and loyalty points to loyal customers in order to make sure they remain faithful to the brand in the long term.
- Account reporting: This helps us understand the details of a particular customer, like the movies watched, the number of movies to be shipped. This will help us make suggestions based on the movie genre and optimize the shipping process in case many movies are to be sent to the customer.
- Product quality: Assesses the quality of a particular product provided by different vendors in order to suggest the best vendor to purchase the product from. This gives the customer a better idea of the service provided by the different vendors, and facilitates the process of choosing the most suitable one.

**Tasks Performed:**

The various tasks performed in order to accomplish the project are as follows:

- Creation of High level dimensional model
- Creation of Detailed level dimensional model
- Execution of ETL in SSIS
- Creation MOLAP cubes using SSAS and using PowerBI to generate dashboards

**Roles:**

- Aditya Chauhan: Dimensional Modeling, ETL in SSIS, Data Warehouse on SQL Server, Business Intelligence
- Anisha Sudhesh Kumar: Dimensional Modeling, Data Warehouse on SQL Server, Business Intelligence
- Khushabh Thakar: Dimensional Modeling, Data Warehouse on SQL Server, Business Intelligence
- Niti Saluja: Dimensional Modeling, ETL using SSIS, Data Warehouse on SQL Server, Business Intelligence

- **Data Profiling:**

| Database | Table Name | Type of Data? | Row Count | One row in this table means.... | Business Keys | Notes / Issues ? |
|---|---|---|---|---|---|---|
| ExternalSources | fm_customer_survey | Master Data | 25 | Customer demographic | Email | |
| ExternalSources | fm_customer_tweets | Business Process | 625 | Tweet details and tweet | TweetID | |
| fudgeflix_v3 | ff_account_billing | Business Process | 1294 | Billing details as per account id | ab_id | |
| fudgeflix_v3 | ff_account_titles | Business Process | 2368 | Shipping details of every ordered title | at_id | |
| fudgeflix_v3 | ff_accounts | Master Data | 35 | Customer details | account_email | |
| fudgeflix_v3 | ff_cast | Business Process | 70543 | Actor of a title | cast_people_id + cast_title_id | |
| fudgeflix_v3 | ff_directors | Business Process | 7136 | Director of a title | director_people_id + director_title_id | |
| fudgeflix_v3 | ff_genres | Master Data | 485 | A genre available in fudgeflix | genre_name | |
| fudgeflix_v3 | ff_people | Master Data | 25347 | Person in a title | people_id | |
| fudgeflix_v3 | ff_plans | Master Data | 7 | Plan details | plan_id | |
| fudgeflix_v3 | ff_title_genres | Business Process | 27841 | Genre of a title | tg_genre_name + tg_title_id | |
| fudgeflix_v3 | ff_titles | Master Data | 7185 | Title details | title_name | |
| fudgeflix_v3 | ff_zipcodes | Master Data | 41968 | City and state of every zipcode | zip_code | |
| fudgemart_v3 | fm_creditcards | Master Data | 49 | Details of credit card used | creditcard_number | |
| fudgemart_v3 | fm_customer_creditcards | Business Process | 54 | Credit card details used by customer | customer_id, creditcard_id | |
| fudgemart_v3 | fm_customer_product_reviews | Business Process | 1039 | Product rating given by a customer | customer_id, product_id | |
| fudgemart_v3 | fm_customers | Master Data | 25 | Complete customer details | customer_email | |
| fudgemart_v3 | fm_departments_lookup | Master Data | 6 | Department type in fudgemart | department_id | |
| fudgemart_v3 | fm_employee_timesheets | Business Process | 6275 | Details filled by an employee in the timesheet | timesheet_id | |
| fudgemart_v3 | fm_employees | Master Data | 33 | Complete employee details | employee_ssn | |
| fudgemart_v3 | fm_jobtitles_lookup | Master Data | 4 | Position in fudgemart | jobtitle_id | |
| fudgemart_v3 | fm_order_details | Business Process | 10466 | Quantity of every product in an order | order_id, product_id | |
| fudgemart_v3 | fm_orders | Business Process | 3516 | Important order details | order_id | |
| fudgemart_v3 | fm_products | Master Data | 53 | All details of a product | product_name | |
| fudgemart_v3 | fm_shipvia_lookup | Master Data | 4 | Shipping vendor | ship_via | |
| fudgemart_v3 | fm_vendors | Master Data | 9 | Details of product supplier | vendor_name, vendor_phone, vendor_website | |

## High-level Dimensional Modeling:

We created a High-level dimensional model for two of our business processes: FactShoppingBehavior and FactDeliveryEfficiency, and defined the grain type, granularity, facts and dimensions composing each of the business processes.

Bus Matrix:

| Instructions! Business Process Name | Fact Table | Fact Grain Type | Granularity | Facts | Dim_Accesst | Dim_Order_D | Dim_Order_D | Dim_ab_date | Dim_Product | Dim_Sappho | Dim_Marital | Dim_Housek | Dim_Zipcode | Dim_Queue_ | Dim_Shipped | Dim_Titles | Dim_Custome | Dim_Returne | Dim_Accesst |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Revenue Generation | Fact_RevenueGeneration | Periodic Snapshot | One row per customer | Total Billed Amount for FudgeFlix, Total billed amount for FudgeMart, Number of customers in a specific year | x | x | x | x | x | x | | | | | | | | | |
| Shopping Behavior | Fact_ShoppingBehavior | Transaction | One row per order | OrderQuantity,ProductPrice,EstimatedTitleCost , FlixTotalBill, TotalViewCount,ProductCost | | x | | | x | | x | x | | | | x | x | | x |
| DeliveryEfficiency | Fact_Delivery Efficiency | Accumulating Snapshot | One row per order | ShippedCount, ReturnedCount, ShipLag,ReturnLag | | | | | | | | | x | x | x | x | | x | x |
| Account Reporting | Fact_ Account Reporting | Accumulating Snapshot | One row per account and title Showing the movies rented | Days taken to receive the movie, Days taken to return the movies , Count of movies in Queue, Rating given by the customer | | | | | | | | | | x | x | x | x | x | |

## Attributes and metrics

| *Instructions!* Dimension / Fact Table | Attribute / Fact Name | Description | Alternate Names | Sample Values |
|---|---|---|---|---|
| Fact_RevenueGeneration | Total Billed Amount for FludgeFlix | Sum of the billed amount of each account | | |
| | Total Billed Amount for Fudge Mart | Sum of the ordered amount(Qty*Retail Price) for each customer | | |
| | Number of Customers of Fudge Mart in a specific year | | | |
| | Number of Customers of Fudge Flix in a specific year | | | |
| Fact_ShoppingBehavior | Order Quantity | | | |
| | ProductPrice | | | |
| | EstimatedTitleCost | | | |
| | FlixTotalBill | | | |
| | TotalViewCount | | | |
| | ProductCost | | | |
| Fact_DeliveryEfficiency | ShippedCount | | | |
| | ReturnedCount | | | |
| | ShipLag | | | |
| | ReturnLag | | | |
| Fact_ Account Reporting | Days taken to receive the movie | | | |
| | Days taken to return the movies | | | |
| | Count of movies in Queue | | | |
| | Rating given by the customer | | | |
| Dim_Account_Biling | ab_date | | | |
| | ab_billed_amount | | | |
| Dim_Order_Details | order_qty | | | |
| Dim_Products | Product_Name | | | |
| | Product_Department | Category of the Product | | Clothing, Household |
| | Vendor_ID | | | |
| | Vendor_Name | | | |
| | Product_Retail_Price | | | |
| Dim_Snapshot Year | | Used to track the date when the revenue generated was recorded for fudgeflix and fludge mart | | |
| Dim_Zipcode | Zipcode city | | | |
| | Zipcode State | | | |
| Dim_Queue Date | | Date when the movie was requested/ordred by the customer | | |
| Dim_Returned_Date | | Date when the movie was returned back by the customer | | |
| Dim_OrderDate | | | | |
| Dim_Shipped Date | | | | |
| Dim_Titles | Title_Name | | | |
| | Title_Type | | | |
| | Title_avg_rating | | | |
| | Title_release_date | | | |
| | Title_rating | | | |
| Dim_Customers | Customer_First_Name | | | |
| | Customer_Last_Name | | | |
| | Customer_Email | | | |
| | Customer_Address | | | |
| | Customer_Zipcode | | | |
| | Customer_Phone | | | |
| | Account_opened_on | | | |
| | Maritial Status | | | Single/Married |
| | Household Income | | | |
| | Gender | | | |
| Dim Accounts | Account_First_Name | | | |
| | Account_Last_Name | | | |
| | Account_Email | | | |
| | Account_Address | | | |
| | Account_Zipcode | | | |
| | Account_Plan_Id | | | |

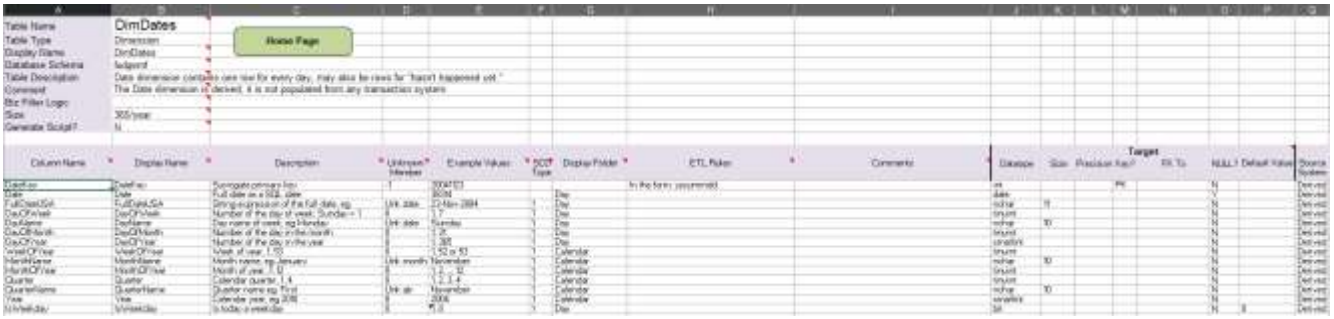
## Issues list:

| Issue # | Task / Topic | Issue | Identified Date | Reported By | Respon-sible | Status | Priority | Date Closed |
|---|---|---|---|---|---|---|---|---|
| 1 | Limited Data | There are very few customers common in both Fudge Mart and Fudge Flix. Due to the limited data, analysis would be difficult | 2/13/2018 | Team 7 | | WIP | Medium | |
| 2 | Incorrect Dates | The Queue Date and Shipped Date have redundant values with no exact timestamp. Due to this we are unable to track the amount of time required to return the movie | 2/13/2018 | Team 7 | | WIP | High | |

**Detail-level Dimensional Modeling:**

The Detail-level Dimensional Modeling sheet was used in order to define each of the dimensions and facts. This macros file defined by Kimball Group, helped us generate the SQL codes for the same, which in-turn helped us create the star schema for our business processes.

DimDates:



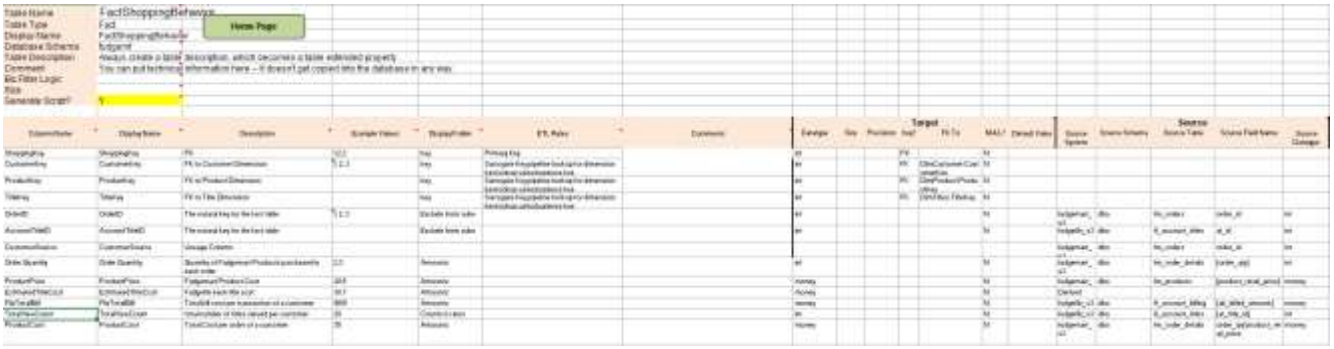
DimFudgeProduct:

DimFudgeTitles:

## DimFudgeCustomer

## DimFudgeAccounts

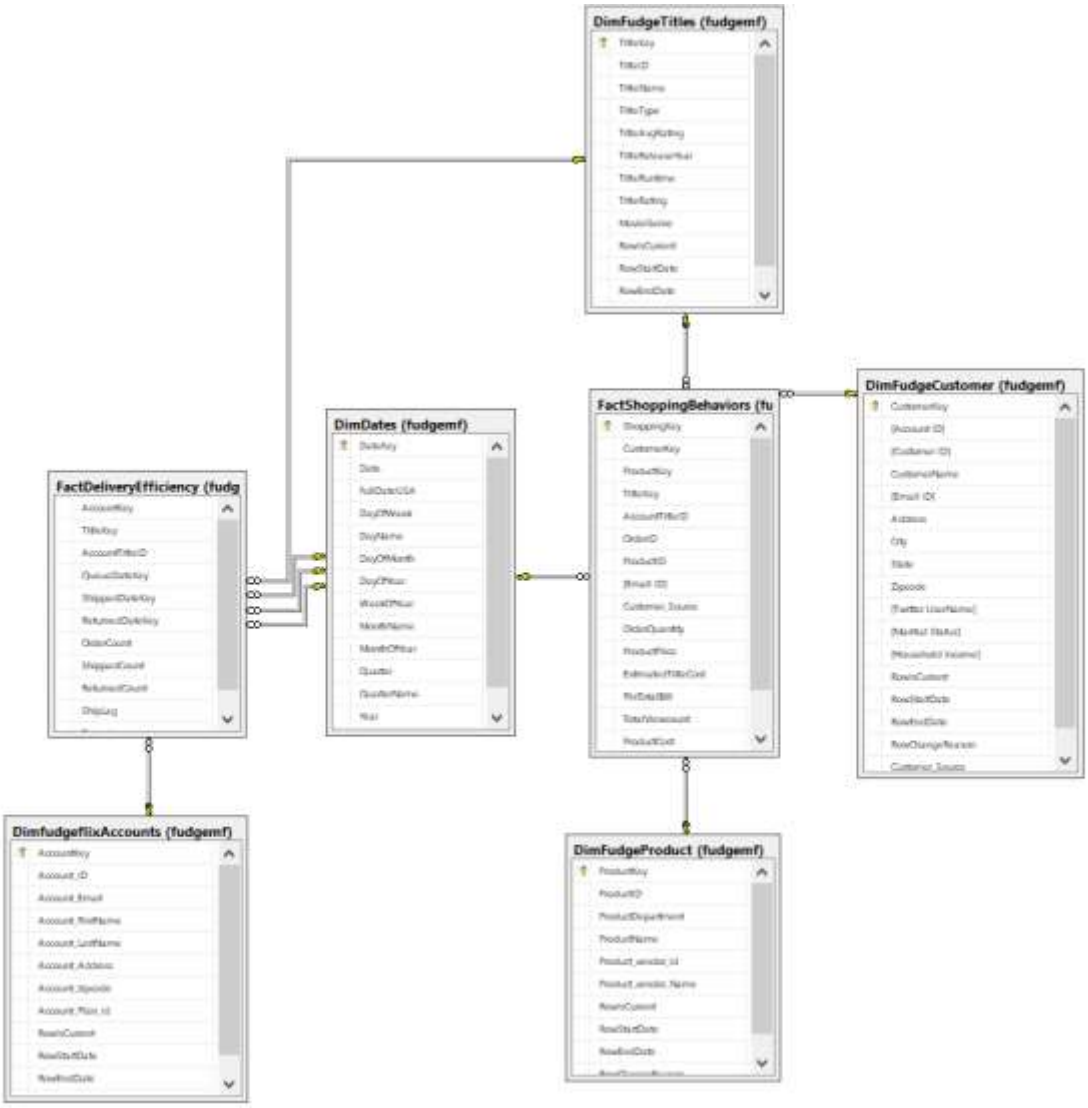
## FactShoppingBehavior



## FactDeliveryEfficiency

## DATA WAREHOUSE ON SQL SERVER

The Detail- Level Dimensional Modeling sheet is used to create the dimensions and facts on the SQL Server. Our Dimensional Schema:



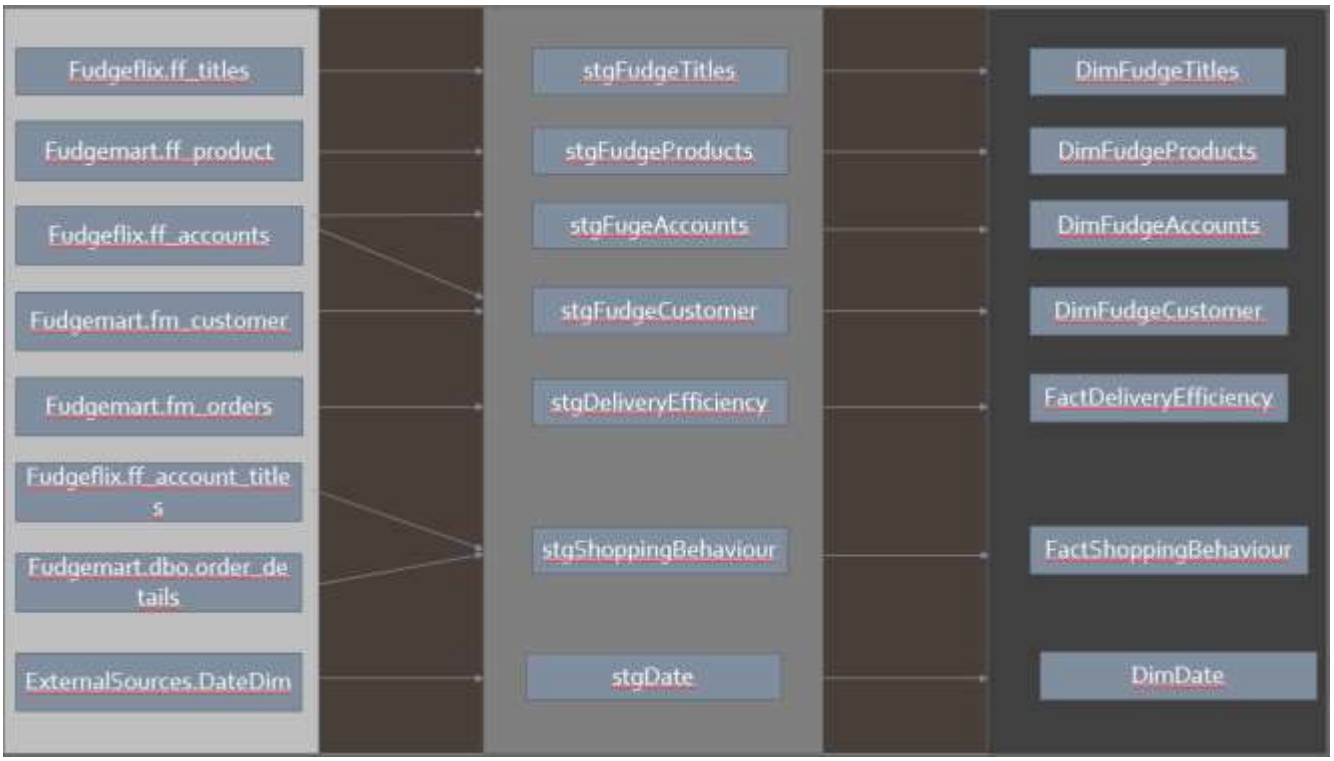
**ETL USING SQL SERVER INTEGRATION SERVICES**

**Source to Target Map:**
The main objective of the S-T-M is to document ETL processing done using tools. The processes documented :

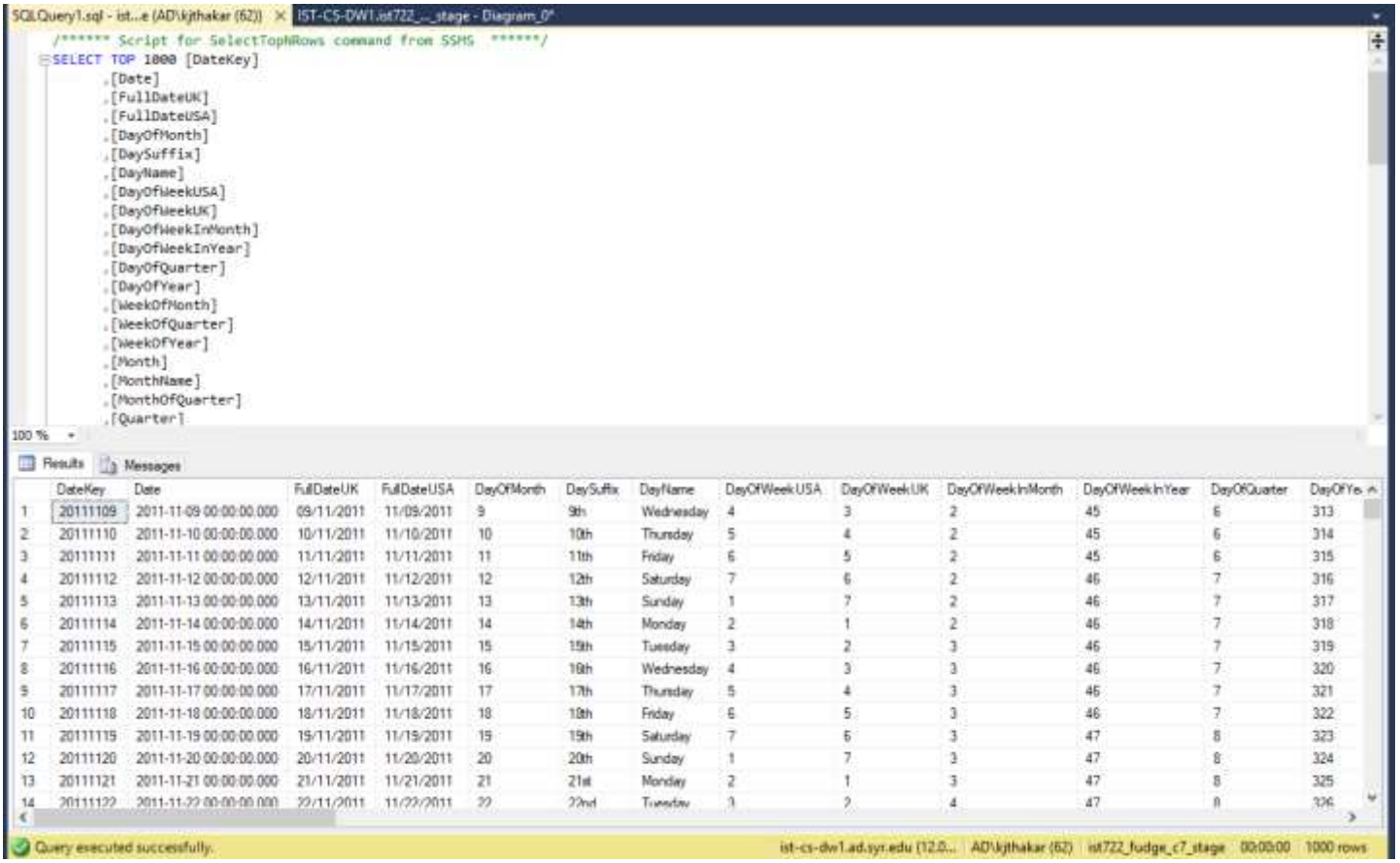
OLTP Source -> ETL Stage -> DW

Source to target map

**Stage:**
We staged OLTP source data from Fudgemart, Fudgeflix, and external sources. Staging data allows us to address data sizing issues and freeze data to a point in time during the initial development.



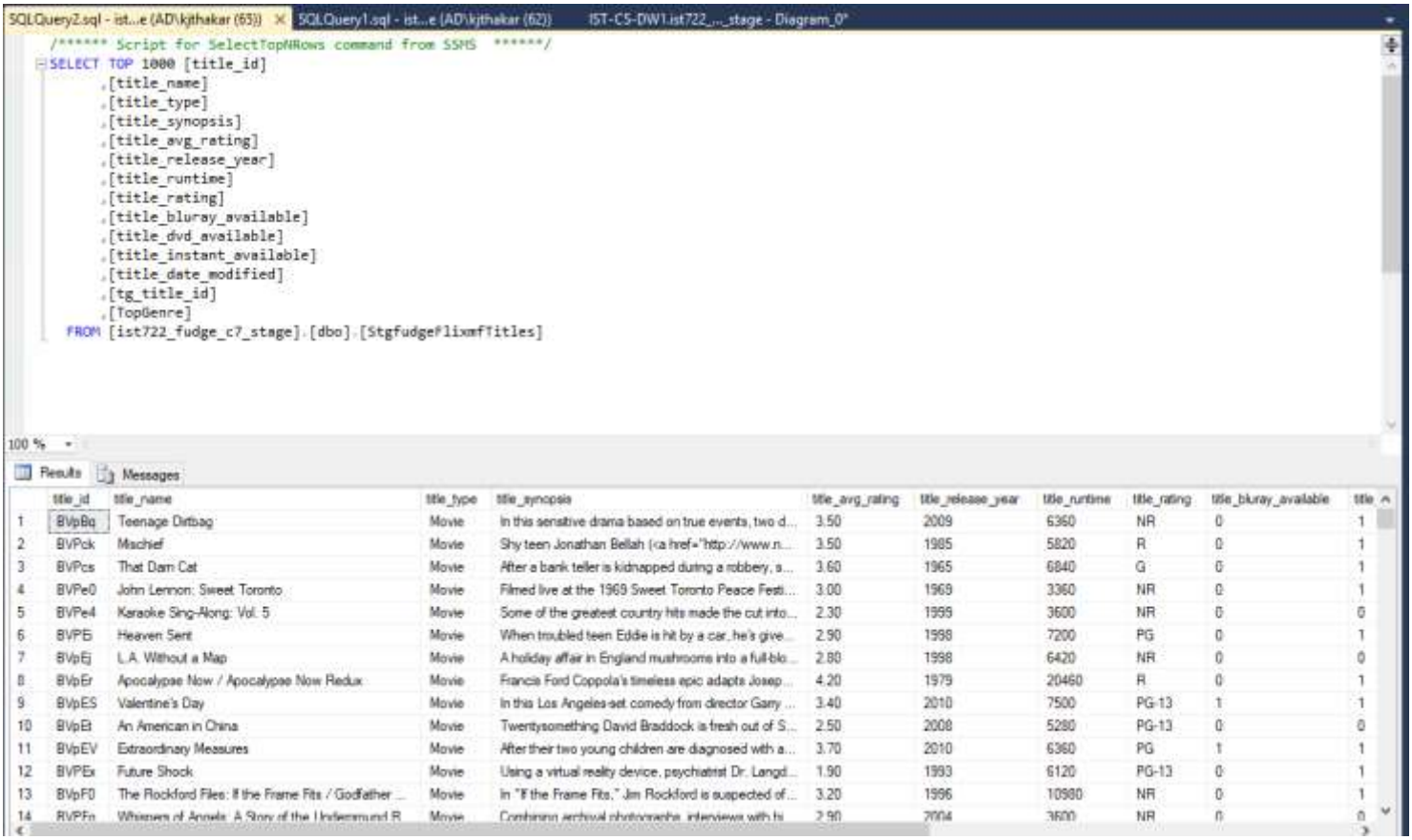
Stage Date Dimension:



```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [DateKey]
      ,[Date]
      ,[FullDateUK]
      ,[FullDateUSA]
      ,[DayOfMonth]
      ,[DaySuffix]
      ,[DayName]
      ,[DayOfWeekUSA]
      ,[DayOfWeekUK]
      ,[DayOfWeekInMonth]
      ,[DayOfWeekInYear]
      ,[DayOfQuarter]
      ,[DayOfYear]
      ,[WeekOfMonth]
      ,[WeekOfQuarter]
      ,[WeekOfYear]
      ,[Month]
      ,[MonthName]
      ,[MonthOfQuarter]
      ,[Quarter]
```

| | DateKey | Date | FullDateUK | FullDateUSA | DayOfMonth | DaySuffix | DayName | DayOfWeekUSA | DayOfWeekUK | DayOfWeekInMonth | DayOfWeekInYear | DayOfQuarter | DayOfYe |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 20111109 | 2011-11-09 00:00:00.000 | 09/11/2011 | 11/09/2011 | 9 | 9th | Wednesday | 4 | 3 | 2 | 45 | 6 | 313 |
| 2 | 20111110 | 2011-11-10 00:00:00.000 | 10/11/2011 | 11/10/2011 | 10 | 10th | Thursday | 5 | 4 | 2 | 45 | 6 | 314 |
| 3 | 20111111 | 2011-11-11 00:00:00.000 | 11/11/2011 | 11/11/2011 | 11 | 11th | Friday | 6 | 5 | 2 | 45 | 6 | 315 |
| 4 | 20111112 | 2011-11-12 00:00:00.000 | 12/11/2011 | 11/12/2011 | 12 | 12th | Saturday | 7 | 6 | 2 | 46 | 7 | 316 |
| 5 | 20111113 | 2011-11-13 00:00:00.000 | 13/11/2011 | 11/13/2011 | 13 | 13th | Sunday | 1 | 7 | 2 | 46 | 7 | 317 |
| 6 | 20111114 | 2011-11-14 00:00:00.000 | 14/11/2011 | 11/14/2011 | 14 | 14th | Monday | 2 | 1 | 2 | 46 | 7 | 318 |
| 7 | 20111115 | 2011-11-15 00:00:00.000 | 15/11/2011 | 11/15/2011 | 15 | 15th | Tuesday | 3 | 2 | 3 | 46 | 7 | 319 |
| 8 | 20111116 | 2011-11-16 00:00:00.000 | 16/11/2011 | 11/16/2011 | 16 | 16th | Wednesday | 4 | 3 | 3 | 46 | 7 | 320 |
| 9 | 20111117 | 2011-11-17 00:00:00.000 | 17/11/2011 | 11/17/2011 | 17 | 17th | Thursday | 5 | 4 | 3 | 46 | 7 | 321 |
| 10 | 20111118 | 2011-11-18 00:00:00.000 | 18/11/2011 | 11/18/2011 | 18 | 18th | Friday | 6 | 5 | 3 | 46 | 7 | 322 |
| 11 | 20111119 | 2011-11-19 00:00:00.000 | 19/11/2011 | 11/19/2011 | 19 | 19th | Saturday | 7 | 6 | 3 | 47 | 8 | 323 |
| 12 | 20111120 | 2011-11-20 00:00:00.000 | 20/11/2011 | 11/20/2011 | 20 | 20th | Sunday | 1 | 7 | 3 | 47 | 8 | 324 |
| 13 | 20111121 | 2011-11-21 00:00:00.000 | 21/11/2011 | 11/21/2011 | 21 | 21st | Monday | 2 | 1 | 3 | 47 | 8 | 325 |
| 14 | 20111122 | 2011-11-22 00:00:00.000 | 22/11/2011 | 11/22/2011 | 22 | 22nd | Tuesday | 3 | 2 | 4 | 47 | 8 | 326 |

Stage Titles table:



SQLQuery2.sql - ist...e (AD\kjthakar (63)) | SQLQuery1.sql - ist...e (AD\kjthakar (62)) | IST-CS-DW1.ist722_..._stage - Diagram_0*

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [title_id]
      ,[title_name]
      ,[title_type]
      ,[title_synopsis]
      ,[title_avg_rating]
      ,[title_release_year]
      ,[title_runtime]
      ,[title_rating]
      ,[title_bluray_available]
      ,[title_dvd_available]
      ,[title_instant_available]
      ,[title_date_modified]
      ,[tg_title_id]
      ,[TopGenre]
  FROM [ist722_fudge_c7_stage].[dbo].[StgfudgeFlixmfTitles]
```

100 %

Results | Messages

| | title_id | title_name | title_type | title_synopsis | title_avg_rating | title_release_year | title_runtime | title_rating | title_bluray_available | title |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | BVpBq | Teenage Dirtbag | Movie | In this sensitive drama based on true events, two d... | 3.50 | 2009 | 6360 | NR | 0 | 1 |
| 2 | BVPck | Mischief | Movie | Shy teen Jonathan Bellah (<a href="http://www.n... | 3.50 | 1985 | 5820 | R | 0 | 1 |
| 3 | BVPcs | That Darn Cat | Movie | After a bank teller is kidnapped during a robbery, a... | 3.60 | 1965 | 6840 | G | 0 | 1 |
| 4 | BVPe0 | John Lennon: Sweet Toronto | Movie | Filmed live at the 1969 Sweet Toronto Peace Festi... | 3.00 | 1969 | 3360 | NR | 0 | 1 |
| 5 | BVPe4 | Karaoke Sing-Along: Vol. 5 | Movie | Some of the greatest country hits made the cut into... | 2.30 | 1999 | 3600 | NR | 0 | 0 |
| 6 | BVPEi | Heaven Sent | Movie | When troubled teen Eddie is hit by a car, he's give... | 2.90 | 1998 | 7200 | PG | 0 | 1 |
| 7 | BVpEj | L.A. Without a Map | Movie | A holiday affair in England mushrooms into a full-blo... | 2.80 | 1998 | 6420 | NR | 0 | 0 |
| 8 | BVpEr | Apocalypse Now / Apocalypse Now Redux | Movie | Francis Ford Coppola's timeless epic adapts Joseph... | 4.20 | 1979 | 20460 | R | 0 | 1 |
| 9 | BVpES | Valentine's Day | Movie | In this Los Angeles-set comedy from director Garry ... | 3.40 | 2010 | 7500 | PG-13 | 1 | 1 |
| 10 | BVpEt | An American in China | Movie | Twentysomething David Braddock is fresh out of S... | 2.50 | 2008 | 5280 | PG-13 | 0 | 0 |
| 11 | BVpEV | Extraordinary Measures | Movie | After their two young children are diagnosed with a... | 3.70 | 2010 | 6360 | PG | 1 | 1 |
| 12 | BVPEx | Future Shock | Movie | Using a virtual reality device, psychiatrist Dr. Langd... | 1.90 | 1993 | 6120 | PG-13 | 0 | 1 |
| 13 | BVpF0 | The Rockford Files: If the Frame Fits / Godfather ... | Movie | In "If the Frame Fits," Jim Rockford is suspected of... | 3.20 | 1996 | 10980 | NR | 0 | 1 |
| 14 | BVPFn | Whispers of Angels: A Story of the Underground R... | Movie | Combining archival photographs, interviews with hi... | 2.90 | 2004 | 3600 | NR | 0 | 0 |

Stage Account Billing table:

SQLQuery3.sql - ist...e (AD\kjthakar (71)) | SQLQuery2.sql - ist...e (AD\kjthakar (65)) | SQLQue

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [ab_id]
      ,[ab_date]
      ,[ab_account_id]
      ,[ab_plan_id]
      ,[ab_billed_amount]
  FROM [ist722_fudge_c7_stage].[dbo].[stgFudgemfAccountBilling]
```

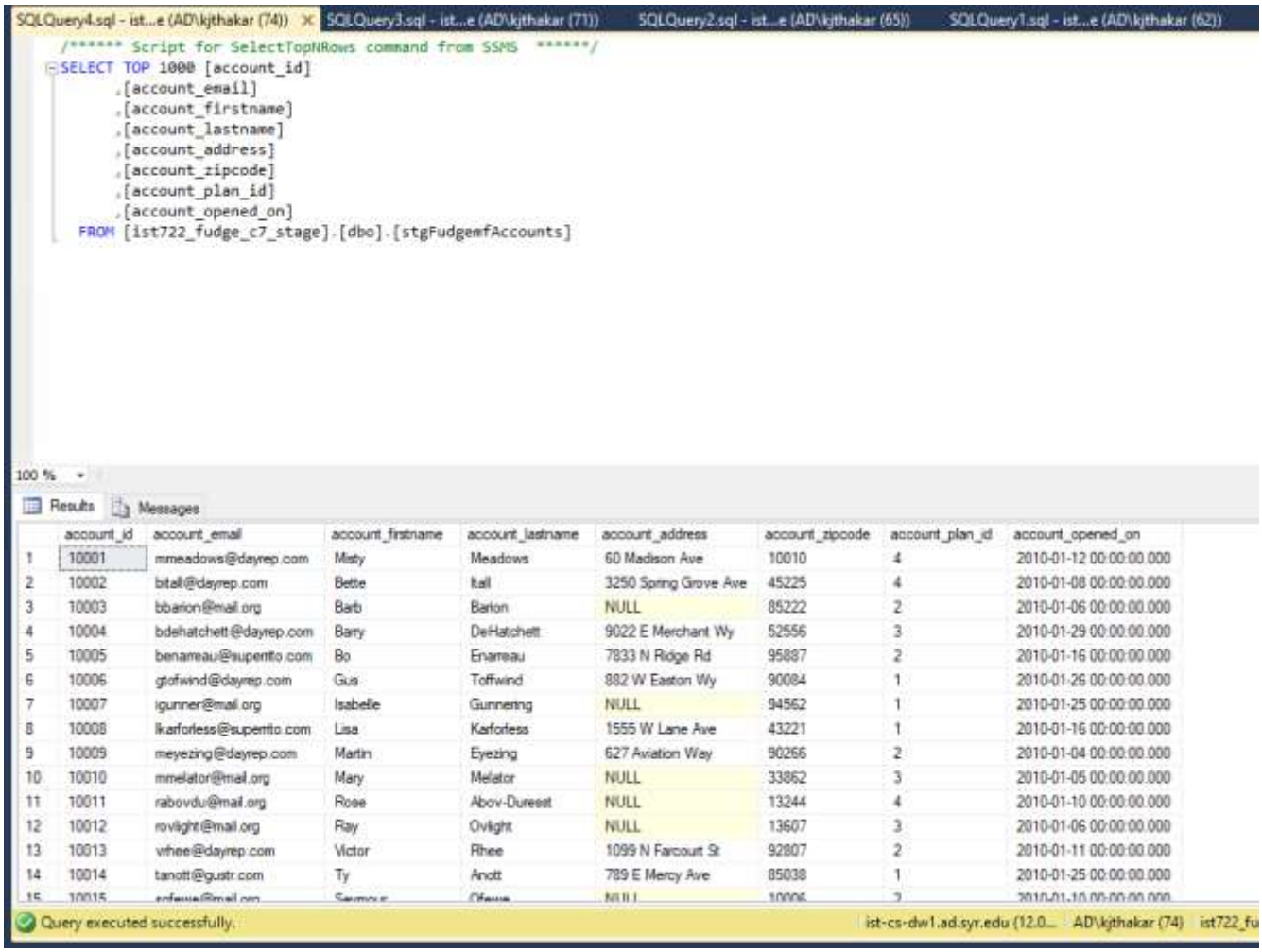
100 %

Results | Messages

| | ab_id | ab_date | ab_account_id | ab_plan_id | ab_billed_amount |
|---|---|---|---|---|---|
| 1 | 1 | 2010-01-01 00:00:00.000 | 10016 | 5 | 7.99 |
| 2 | 2 | 2010-01-01 00:00:00.000 | 10018 | 5 | 7.99 |
| 3 | 3 | 2010-01-01 00:00:00.000 | 10031 | 5 | 7.99 |
| 4 | 5 | 2010-02-01 00:00:00.000 | 10001 | 5 | 7.99 |
| 5 | 6 | 2010-02-01 00:00:00.000 | 10002 | 5 | 7.99 |
| 6 | 7 | 2010-02-01 00:00:00.000 | 10003 | 5 | 7.99 |
| 7 | 8 | 2010-02-01 00:00:00.000 | 10004 | 6 | 19.99 |
| 8 | 9 | 2010-02-01 00:00:00.000 | 10005 | 5 | 7.99 |
| 9 | 10 | 2010-02-01 00:00:00.000 | 10006 | 5 | 7.99 |
| 10 | 11 | 2010-02-01 00:00:00.000 | 10007 | 5 | 7.99 |
| 11 | 12 | 2010-02-01 00:00:00.000 | 10008 | 5 | 7.99 |
| 12 | 13 | 2010-02-01 00:00:00.000 | 10009 | 7 | 12.99 |
| 13 | 14 | 2010-02-01 00:00:00.000 | 10010 | 6 | 19.99 |
| 14 | 15 | 2010-02-01 00:00:00.000 | 10011 | 6 | 19.99 |
| 15 | 16 | 2010-02-01 00:00:00.000 | 10012 | 6 | 19.99 |

Query executed successfully.

Stage Accounts table:



Stage Account Titles:

SQLQuery5.sql - ist...e (AD\kjthakar (75)) | SQLQuery4.sql - ist...e (AD\kjthakar (74)) | SQLQuery3.sql - ist...e (AD\kjthakar (71)) | SC

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [at_id]
      ,[at_account_id]
      ,[at_title_id]
      ,[at_queue_date]
      ,[at_shipped_date]
      ,[at_returned_date]
      ,[at_rating]
  FROM [ist722_fudge_c7_stage].[dbo].[stgfudgemfAccountTitles]
```

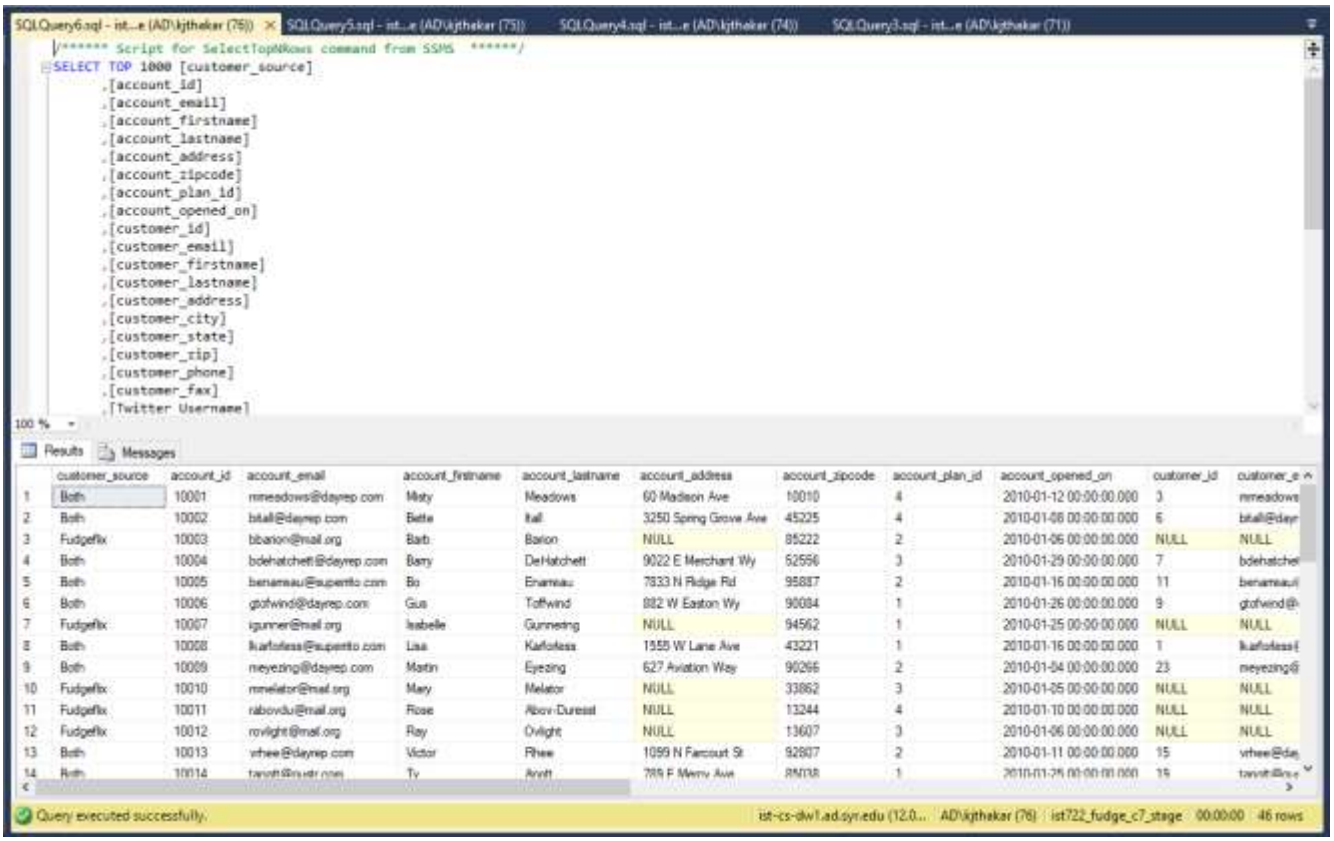
100 %

Results | Messages

| | at_id | at_account_id | at_title_id | at_queue_date | at_shipped_date | at_returned_date | at_rating |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 10001 | 7NG6 | 2013-02-01 00:00:00.000 | 2013-02-02 00:00:00.000 | 2013-02-02 00:00:00.000 | 3 |
| 2 | 2 | 10004 | BVNdO | 2013-02-01 00:00:00.000 | 2013-02-02 00:00:00.000 | 2013-02-02 00:00:00.000 | 2 |
| 3 | 3 | 10004 | knkj | 2013-02-01 00:00:00.000 | 2013-02-02 00:00:00.000 | 2013-02-02 00:00:00.000 | NULL |
| 4 | 4 | 10004 | ApNwh | 2013-02-02 00:00:00.000 | NULL | NULL | 3 |
| 5 | 5 | 10004 | ApQhl | 2013-02-03 00:00:00.000 | NULL | NULL | 4 |
| 6 | 6 | 10005 | 7tEu | 2013-02-01 00:00:00.000 | 2013-02-02 00:00:00.000 | 2013-02-02 00:00:00.000 | 4 |
| 7 | 7 | 10005 | knkj | 2013-02-02 00:00:00.000 | NULL | NULL | NULL |
| 8 | 8 | 10012 | 5prm | 2013-02-01 00:00:00.000 | 2013-02-02 00:00:00.000 | 2013-02-02 00:00:00.000 | 4 |
| 9 | 9 | 10012 | ApO1H | 2013-02-02 00:00:00.000 | NULL | NULL | 4 |
| 10 | 10 | 10012 | ApQrJ | 2013-02-02 00:00:00.000 | NULL | NULL | 4 |
| 11 | 11 | 10012 | BVV1C | 2013-02-03 00:00:00.000 | NULL | NULL | 3 |
| 12 | 12 | 10016 | BVODD | 2013-02-01 00:00:00.000 | 2013-02-02 00:00:00.000 | 2013-02-02 00:00:00.000 | NULL |
| 13 | 13 | 10016 | BVUNd | 2013-02-02 00:00:00.000 | NULL | NULL | NULL |
| 14 | 14 | 10020 | 8XnPX | 2013-02-01 00:00:00.000 | 2013-02-02 00:00:00.000 | 2013-02-02 00:00:00.000 | NULL |
| 15 | 15 | 10020 | BVUOH | 2013-02-02 00:00:00.000 | NULL | NULL | 3 |

Query executed successfully. ist-cs-dw1.ad.syr.

Stage Customers:

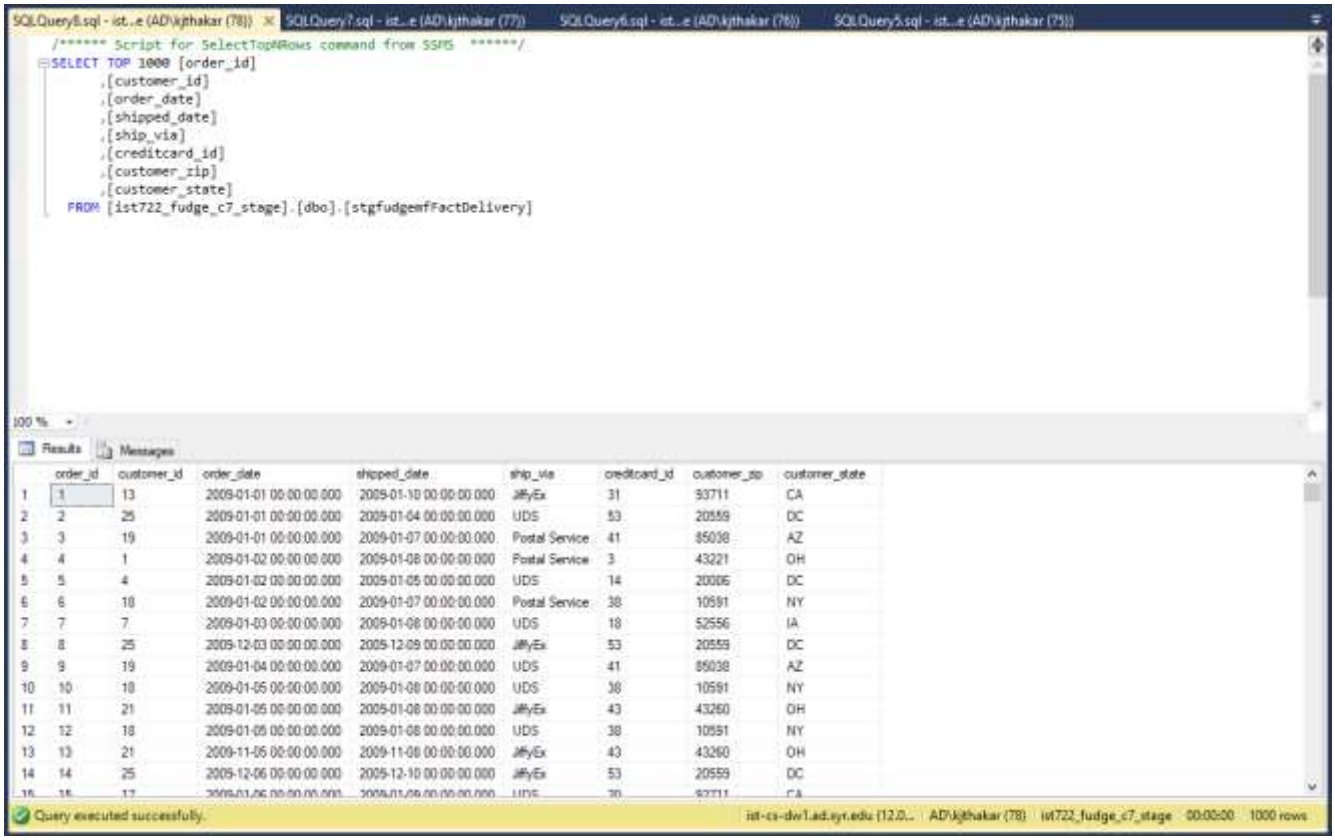


```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [customer_source]
      ,[account_id]
      ,[account_email]
      ,[account_firstname]
      ,[account_lastname]
      ,[account_address]
      ,[account_zipcode]
      ,[account_plan_id]
      ,[account_opened_on]
      ,[customer_id]
      ,[customer_email]
      ,[customer_firstname]
      ,[customer_lastname]
      ,[customer_address]
      ,[customer_city]
      ,[customer_state]
      ,[customer_zip]
      ,[customer_phone]
      ,[customer_fax]
      ,[Twitter Username]
```

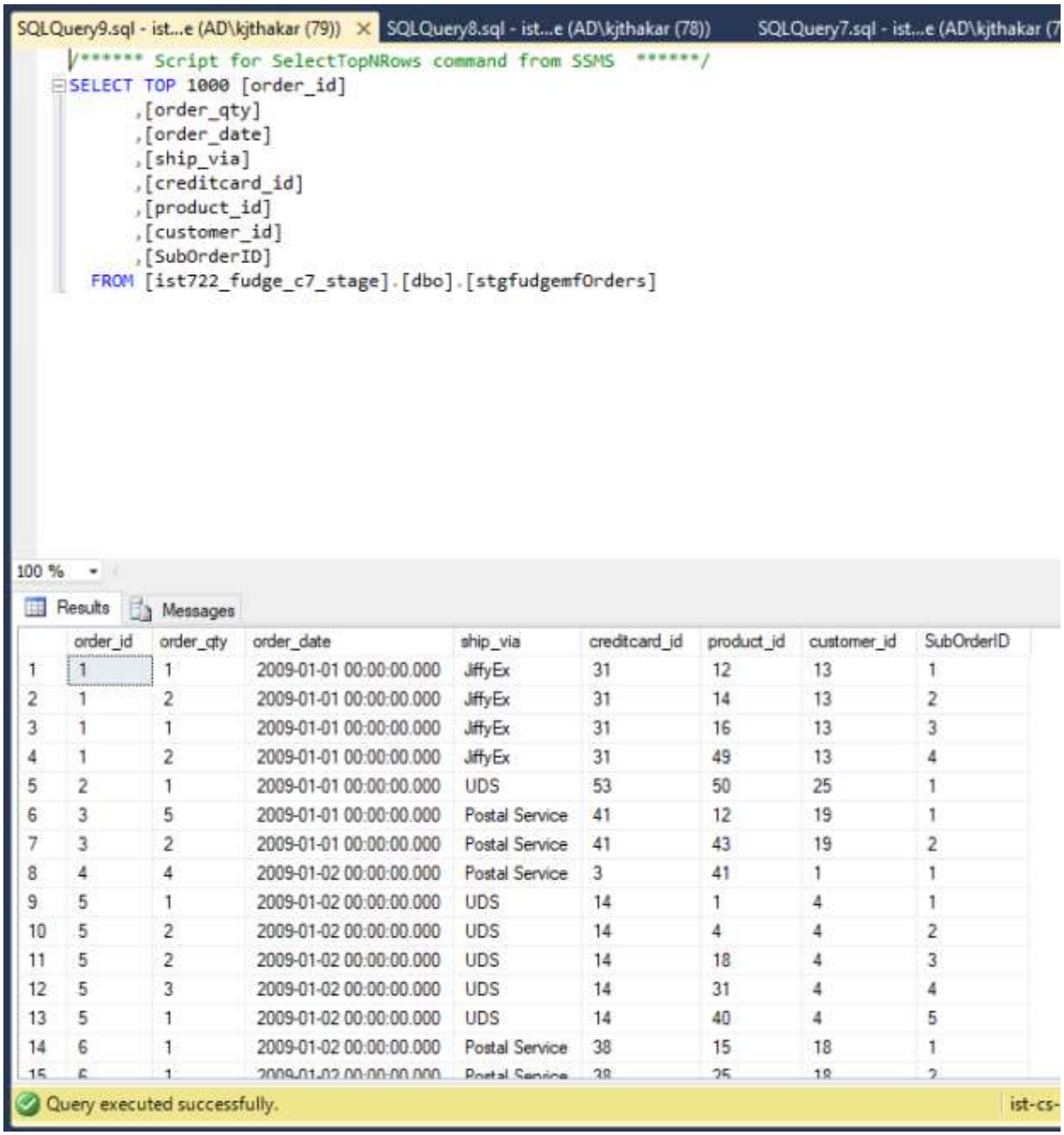
| | customer_source | account_id | account_email | account_firstname | account_lastname | account_address | account_zipcode | account_plan_id | account_opened_on | customer_id | customer_e |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Both | 10001 | mmeadows@dayrep.com | Misty | Meadows | 60 Madison Ave | 10010 | 4 | 2010-01-12 00:00:00.000 | 3 | mmeadows |
| 2 | Both | 10002 | btall@dayrep.com | Bette | Itall | 3250 Spring Grove Ave | 45225 | 4 | 2010-01-08 00:00:00.000 | 6 | btall@dayr |
| 3 | Fudgeflix | 10003 | bbarion@mail.org | Barb | Barion | NULL | 85222 | 2 | 2010-01-06 00:00:00.000 | NULL | NULL |
| 4 | Both | 10004 | bdehatchett@dayrep.com | Barry | DeHatchett | 9022 E Merchant Wy | 52556 | 3 | 2010-01-29 00:00:00.000 | 7 | bdehatchet |
| 5 | Both | 10005 | benarneau@supertito.com | Bo | Enarneau | 7833 N Ridge Rd | 95887 | 2 | 2010-01-16 00:00:00.000 | 11 | benarneau@ |
| 6 | Both | 10006 | gtofwind@dayrep.com | Gus | Toffwind | 882 W Easton Wy | 90084 | 1 | 2010-01-26 00:00:00.000 | 9 | gtofwind@r |
| 7 | Fudgeflix | 10007 | igunner@mail.org | Isabelle | Gunnering | NULL | 94562 | 1 | 2010-01-25 00:00:00.000 | NULL | NULL |
| 8 | Both | 10008 | lkarfofess@supertito.com | Lisa | Karfofess | 1555 W Lane Ave | 43221 | 1 | 2010-01-16 00:00:00.000 | 1 | lkarfofess@ |
| 9 | Both | 10009 | meyezing@dayrep.com | Martin | Eyezing | 627 Aviation Way | 90266 | 2 | 2010-01-04 00:00:00.000 | 23 | meyezing@ |
| 10 | Fudgeflix | 10010 | mmelator@mail.org | Mary | Melator | NULL | 33862 | 3 | 2010-01-05 00:00:00.000 | NULL | NULL |
| 11 | Fudgeflix | 10011 | rabovdu@mail.org | Rose | Abov-Duressl | NULL | 13244 | 4 | 2010-01-10 00:00:00.000 | NULL | NULL |
| 12 | Fudgeflix | 10012 | rovlight@mail.org | Ray | Ovlight | NULL | 13607 | 3 | 2010-01-06 00:00:00.000 | NULL | NULL |
| 13 | Both | 10013 | vrhee@dayrep.com | Victor | Rhee | 1099 N Farcout St | 92807 | 2 | 2010-01-11 00:00:00.000 | 15 | vrhee@day |
| 14 | Both | 10014 | tarvatt@dayrep.com | Ty | Arvatt | 789 E Merry Ave | 89108 | 1 | 2010-01-25 00:00:00.000 | 19 | tarvatt@day |

Query executed successfully. ist-cs-dw1.ad.syr.edu (12.0... AD\kjthakar (76) ist722_fudge_c7_stage 00:00:00 46 rows

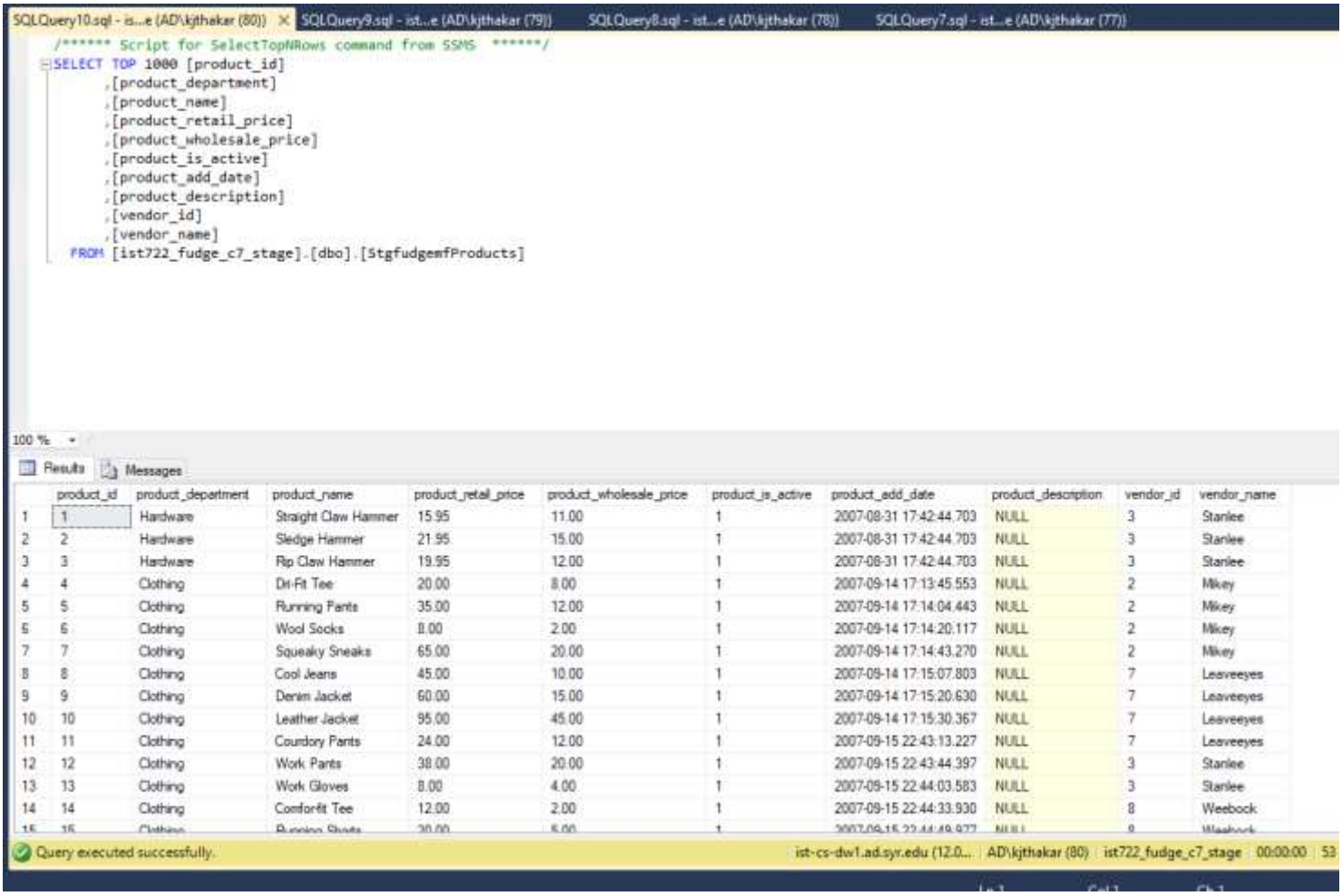
Stage Fact Delivery:



Stage Fudge Orders:



Stage Products Table:



Stage Order Table:

SQLQuery11.sql - is...e (AD\kjthakar (81)) | SQLQuery10.sql - is...e (AD\kjthakar (80)) | SQLQuery9.sql - ist...e (A

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [order_id]
      ,[customer_id]
      ,[order_date]
      ,[shipped_date]
      ,[ship_via]
      ,[creditcard_id]
  FROM [ist722_fudge_c7_stage].[dbo].[stgsrcOrder]
```

100 %

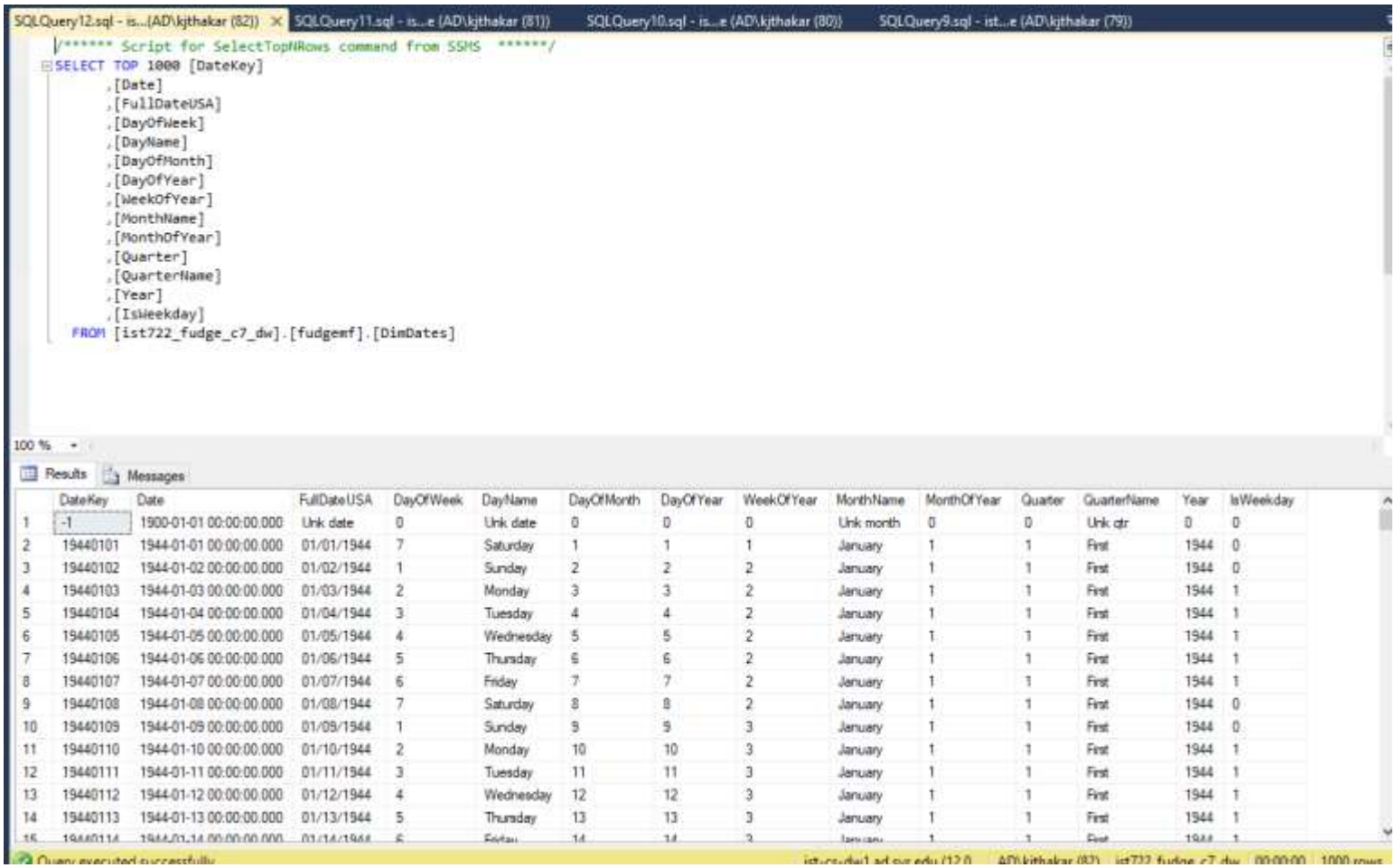
Results | Messages

| | order_id | customer_id | order_date | shipped_date | ship_via | creditcard_id |
|---|---|---|---|---|---|---|
| 1 | 1 | 13 | 2009-01-01 00:00:00.000 | 2009-01-10 00:00:00.000 | JiffyEx | 31 |
| 2 | 2 | 25 | 2009-01-01 00:00:00.000 | 2009-01-04 00:00:00.000 | UDS | 53 |
| 3 | 3 | 19 | 2009-01-01 00:00:00.000 | 2009-01-07 00:00:00.000 | Postal Service | 41 |
| 4 | 4 | 1 | 2009-01-02 00:00:00.000 | 2009-01-08 00:00:00.000 | Postal Service | 3 |
| 5 | 5 | 4 | 2009-01-02 00:00:00.000 | 2009-01-05 00:00:00.000 | UDS | 14 |
| 6 | 6 | 18 | 2009-01-02 00:00:00.000 | 2009-01-07 00:00:00.000 | Postal Service | 38 |
| 7 | 7 | 7 | 2009-01-03 00:00:00.000 | 2009-01-08 00:00:00.000 | UDS | 18 |
| 8 | 8 | 25 | 2009-12-03 00:00:00.000 | 2009-12-09 00:00:00.000 | JiffyEx | 53 |
| 9 | 9 | 19 | 2009-01-04 00:00:00.000 | 2009-01-07 00:00:00.000 | UDS | 41 |
| 10 | 10 | 18 | 2009-01-05 00:00:00.000 | 2009-01-08 00:00:00.000 | UDS | 38 |
| 11 | 11 | 21 | 2009-01-05 00:00:00.000 | 2009-01-08 00:00:00.000 | JiffyEx | 43 |
| 12 | 12 | 18 | 2009-01-05 00:00:00.000 | 2009-01-08 00:00:00.000 | UDS | 38 |
| 13 | 13 | 21 | 2009-11-05 00:00:00.000 | 2009-11-08 00:00:00.000 | JiffyEx | 43 |
| 14 | 14 | 25 | 2009-12-06 00:00:00.000 | 2009-12-10 00:00:00.000 | JiffyEx | 53 |
| 15 | 15 | 17 | 2009-01-06 00:00:00.000 | 2009-01-09 00:00:00.000 | UDS | 20 |

Query executed successfully.

**LOADING STAGE DATA TO DATA WAREHOUSE**

We use a similar ETL for loading the data to the Data Warehouse, like we used for Staging.



```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [DateKey]
      ,[Date]
      ,[FullDateUSA]
      ,[DayOfWeek]
      ,[DayName]
      ,[DayOfMonth]
      ,[DayOfYear]
      ,[WeekOfYear]
      ,[MonthName]
      ,[MonthOfYear]
      ,[Quarter]
      ,[QuarterName]
      ,[Year]
      ,[IsWeekday]
  FROM [ist722_fudge_c7_dw].[fudgemf].[DimDates]
```

100 %

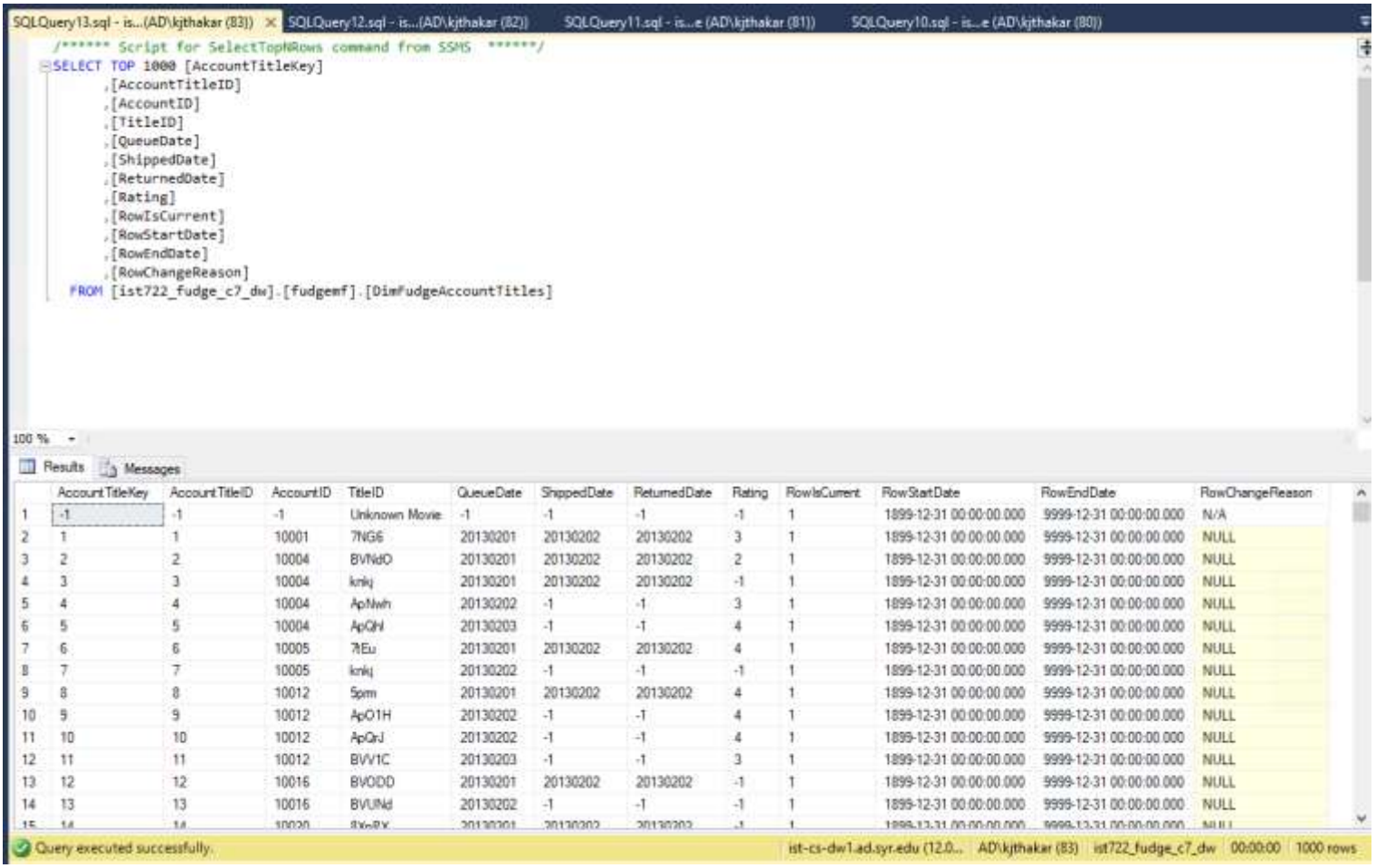
Results | Messages

| | DateKey | Date | FullDateUSA | DayOfWeek | DayName | DayOfMonth | DayOfYear | WeekOfYear | MonthName | MonthOfYear | Quater | QuarterName | Year | IsWeekday |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -1 | 1900-01-01 00:00:00.000 | Unk date | 0 | Unk date | 0 | 0 | 0 | Unk month | 0 | 0 | Unk qtr | 0 | 0 |
| 2 | 19440101 | 1944-01-01 00:00:00.000 | 01/01/1944 | 7 | Saturday | 1 | 1 | 1 | January | 1 | 1 | First | 1944 | 0 |
| 3 | 19440102 | 1944-01-02 00:00:00.000 | 01/02/1944 | 1 | Sunday | 2 | 2 | 2 | January | 1 | 1 | First | 1944 | 0 |
| 4 | 19440103 | 1944-01-03 00:00:00.000 | 01/03/1944 | 2 | Monday | 3 | 3 | 2 | January | 1 | 1 | First | 1944 | 1 |
| 5 | 19440104 | 1944-01-04 00:00:00.000 | 01/04/1944 | 3 | Tuesday | 4 | 4 | 2 | January | 1 | 1 | First | 1944 | 1 |
| 6 | 19440105 | 1944-01-05 00:00:00.000 | 01/05/1944 | 4 | Wednesday | 5 | 5 | 2 | January | 1 | 1 | First | 1944 | 1 |
| 7 | 19440106 | 1944-01-06 00:00:00.000 | 01/06/1944 | 5 | Thursday | 6 | 6 | 2 | January | 1 | 1 | First | 1944 | 1 |
| 8 | 19440107 | 1944-01-07 00:00:00.000 | 01/07/1944 | 6 | Friday | 7 | 7 | 2 | January | 1 | 1 | First | 1944 | 1 |
| 9 | 19440108 | 1944-01-08 00:00:00.000 | 01/08/1944 | 7 | Saturday | 8 | 8 | 2 | January | 1 | 1 | First | 1944 | 0 |
| 10 | 19440109 | 1944-01-09 00:00:00.000 | 01/09/1944 | 1 | Sunday | 9 | 9 | 3 | January | 1 | 1 | First | 1944 | 0 |
| 11 | 19440110 | 1944-01-10 00:00:00.000 | 01/10/1944 | 2 | Monday | 10 | 10 | 3 | January | 1 | 1 | First | 1944 | 1 |
| 12 | 19440111 | 1944-01-11 00:00:00.000 | 01/11/1944 | 3 | Tuesday | 11 | 11 | 3 | January | 1 | 1 | First | 1944 | 1 |
| 13 | 19440112 | 1944-01-12 00:00:00.000 | 01/12/1944 | 4 | Wednesday | 12 | 12 | 3 | January | 1 | 1 | First | 1944 | 1 |
| 14 | 19440113 | 1944-01-13 00:00:00.000 | 01/13/1944 | 5 | Thursday | 13 | 13 | 3 | January | 1 | 1 | First | 1944 | 1 |

**Dim Dates:**

**Dim Account Titles:**

SQLQuery13.sql - is...(AD\kjthakar (83)) | SQLQuery12.sql - is...(AD\kjthakar (82)) | SQLQuery11.sql - is...e (AD\kjthakar (81)) | SQLQuery10.sql - is...e (AD\kjthakar (80))



```sql
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [AccountTitleKey]
      ,[AccountTitleID]
      ,[AccountID]
      ,[TitleID]
      ,[QueueDate]
      ,[ShippedDate]
      ,[ReturnedDate]
      ,[Rating]
      ,[RowIsCurrent]
      ,[RowStartDate]
      ,[RowEndDate]
      ,[RowChangeReason]
  FROM [ist722_fudge_c7_dw].[fudgemf].[DimFudgeAccountTitles]
```

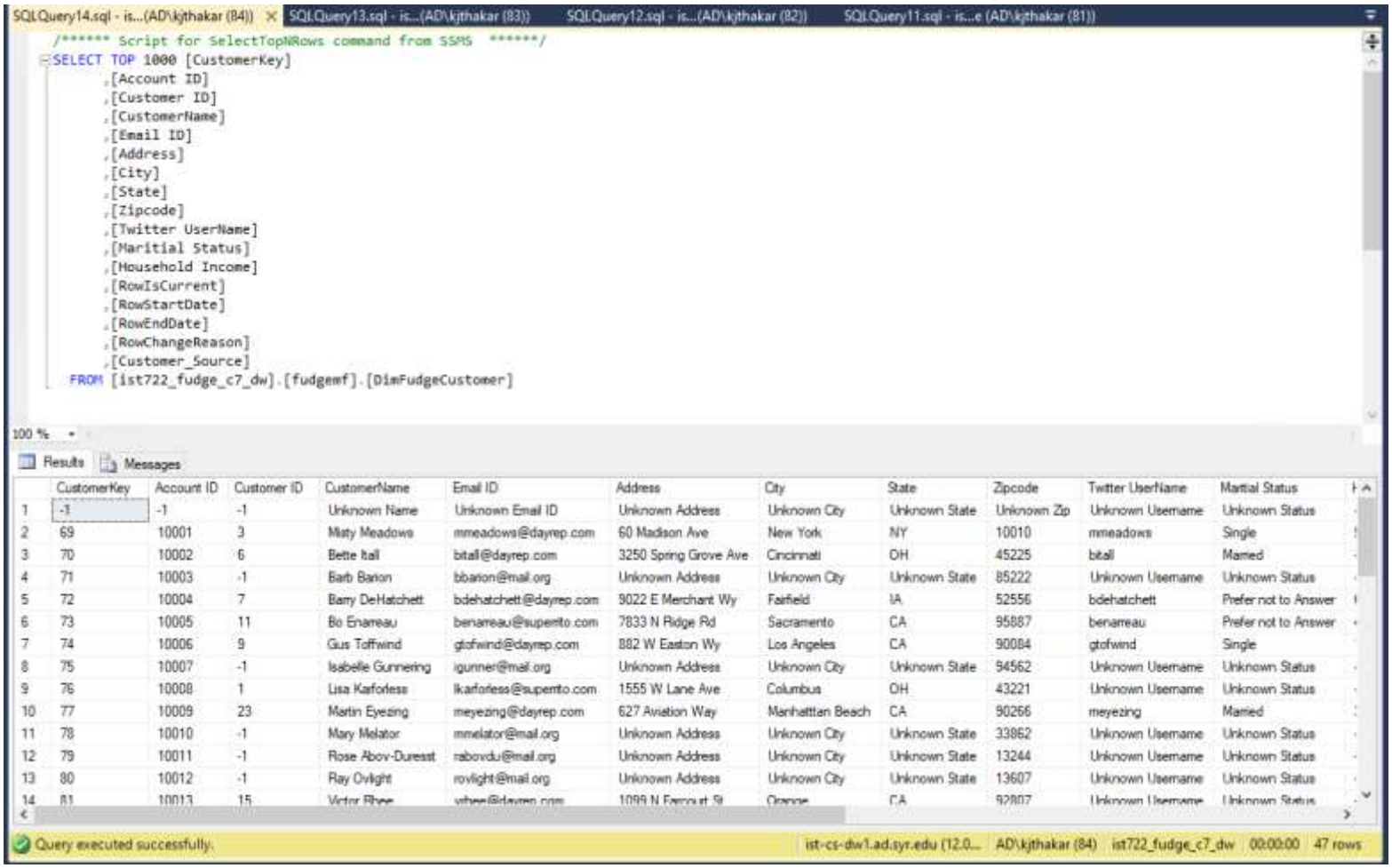
100 %

Results | Messages

| | AccountTitleKey | AccountTitleID | AccountID | TitleID | QueueDate | ShippedDate | ReturnedDate | Rating | RowIsCurrent | RowStartDate | RowEndDate | RowChangeReason |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -1 | -1 | -1 | Unknown Movie | -1 | -1 | -1 | -1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | N/A |
| 2 | 1 | 1 | 10001 | 7NG6 | 20130201 | 20130202 | 20130202 | 3 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 3 | 2 | 2 | 10004 | BVNdO | 20130201 | 20130202 | 20130202 | 2 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 4 | 3 | 3 | 10004 | knkj | 20130201 | 20130202 | 20130202 | -1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 5 | 4 | 4 | 10004 | ApNwh | 20130202 | -1 | -1 | 3 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 6 | 5 | 5 | 10004 | ApQhl | 20130203 | -1 | -1 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 7 | 6 | 6 | 10005 | 7tEu | 20130201 | 20130202 | 20130202 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 8 | 7 | 7 | 10005 | knkj | 20130202 | -1 | -1 | -1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 9 | 8 | 8 | 10012 | 5prm | 20130201 | 20130202 | 20130202 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 10 | 9 | 9 | 10012 | ApO1H | 20130202 | -1 | -1 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 11 | 10 | 10 | 10012 | ApQrJ | 20130202 | -1 | -1 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 12 | 11 | 11 | 10012 | BVV1C | 20130203 | -1 | -1 | 3 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 13 | 12 | 12 | 10016 | BVODD | 20130201 | 20130202 | 20130202 | -1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 14 | 13 | 13 | 10016 | BVUNd | 20130202 | -1 | -1 | -1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 15 | 14 | 14 | 10020 | 8XnRX | 20130201 | 20130202 | 20130202 | -1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |

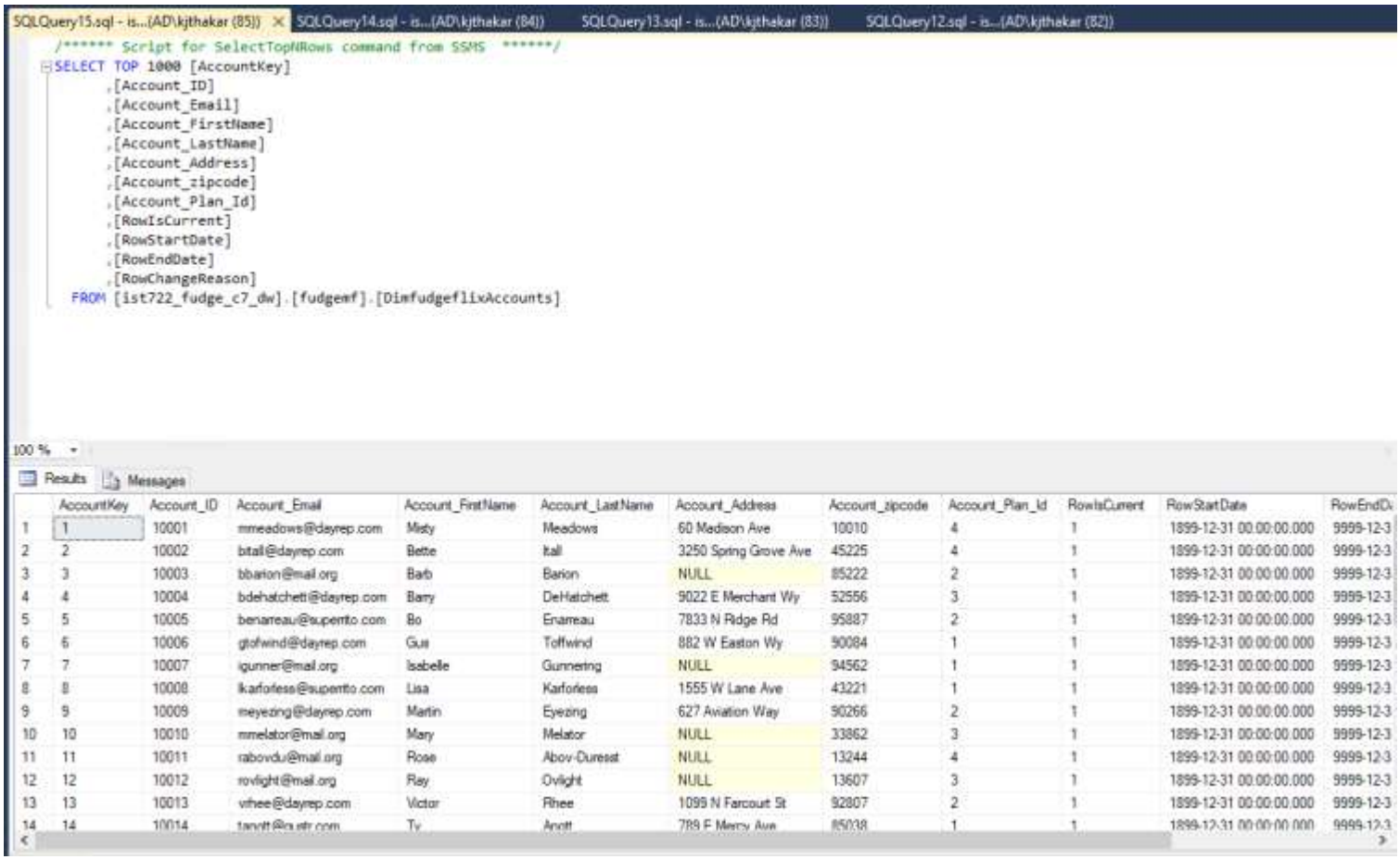
Query executed successfully. | ist-cs-dw1.ad.syr.edu (12.0... | AD\kjthakar (83) | ist722_fudge_c7_dw | 00:00:00 | 1000 rows

**Dim Customer:**



**Dim Accounts:**

SQLQuery15.sql - is...(AD\kjthakar (85)) SQLQuery14.sql - is...(AD\kjthakar (84)) SQLQuery13.sql - is...(AD\kjthakar (83)) SQLQuery12.sql - is...(AD\kjthakar (82))



```sql
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [AccountKey]
      ,[Account_ID]
      ,[Account_Email]
      ,[Account_FirstName]
      ,[Account_LastName]
      ,[Account_Address]
      ,[Account_zipcode]
      ,[Account_Plan_Id]
      ,[RowIsCurrent]
      ,[RowStartDate]
      ,[RowEndDate]
      ,[RowChangeReason]
  FROM [ist722_fudge_c7_dw].[fudgemf].[DimfudgeflixAccounts]
```

100 %

Results | Messages

| | AccountKey | Account_ID | Account_Email | Account_FirstName | Account_LastName | Account_Address | Account_zipcode | Account_Plan_Id | RowIsCurrent | RowStartDate | RowEndD |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 10001 | mmeadows@dayrep.com | Misty | Meadows | 60 Madison Ave | 10010 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 2 | 2 | 10002 | bitall@dayrep.com | Bette | Itall | 3250 Spring Grove Ave | 45225 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 3 | 3 | 10003 | bbarion@mail.org | Barb | Barion | NULL | 85222 | 2 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 4 | 4 | 10004 | bdehatchett@dayrep.com | Barry | DeHatchett | 9022 E Merchant Wy | 52556 | 3 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 5 | 5 | 10005 | benarreau@superrito.com | Bo | Enarreau | 7833 N Ridge Rd | 95887 | 2 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 6 | 6 | 10006 | gtofwind@dayrep.com | Gus | Toffwind | 882 W Easton Wy | 90084 | 1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 7 | 7 | 10007 | igunner@mail.org | Isabelle | Gunnering | NULL | 94562 | 1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 8 | 8 | 10008 | lkarforless@superrito.com | Lisa | Karforless | 1555 W Lane Ave | 43221 | 1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 9 | 9 | 10009 | meyezing@dayrep.com | Martin | Eyezing | 627 Aviation Way | 90266 | 2 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 10 | 10 | 10010 | mmelator@mail.org | Mary | Melator | NULL | 33862 | 3 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 11 | 11 | 10011 | rabovdu@mail.org | Rose | Abov-Duresst | NULL | 13244 | 4 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 12 | 12 | 10012 | rovlight@mail.org | Ray | Ovlight | NULL | 13607 | 3 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 13 | 13 | 10013 | vrhee@dayrep.com | Victor | Rhee | 1099 N Farcourt St | 92807 | 2 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |
| 14 | 14 | 10014 | tanott@gustr.com | Ty | Anott | 789 E Mercy Ave | 85038 | 1 | 1 | 1899-12-31 00:00:00.000 | 9999-12-3 |

**Dim Products:**

SQLQuery16.sql - is...(AD\kjthakar (86)) | SQLQuery15.sql - is...(AD\kjthakar (85)) | SQLQuery14.sql - is...(AD\kjthakar (84)) | SQLQuery13.sql - is...(AD\kjthakar (83))

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [ProductKey]
      ,[ProductID]
      ,[ProductDepartment]
      ,[ProductName]
      ,[Product_vendor_id]
      ,[Product_vendor_Name]
      ,[RowIsCurrent]
      ,[RowStartDate]
      ,[RowEndDate]
      ,[RowChangeReason]
  FROM [ist722_fudge_c7_dw].[fudgemf].[DimFudgeProduct]
```

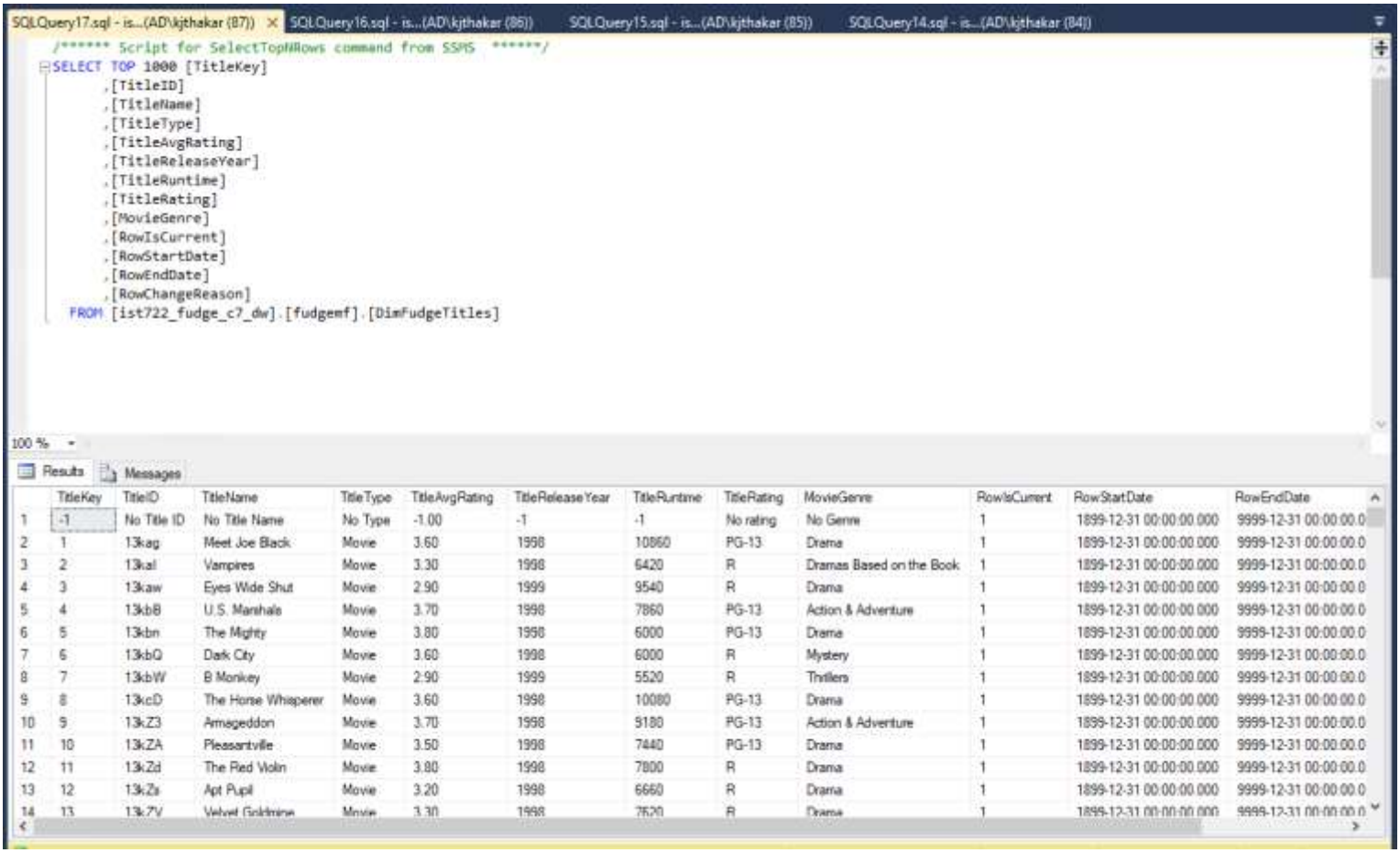
100 %

Results | Messages

| | ProductKey | ProductID | ProductDepartment | ProductName | Product_vendor_id | Product_vendor_Name | RowIsCurrent | RowStartDate | RowEndDate | RowChangeReason |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -1 | -1 | Unknown Department | Unknown Name | -1 | No Vendor | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | N/A |
| 2 | 1 | 1 | Hardware | Straight Claw Hammer | 3 | Stanlee | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 3 | 2 | 2 | Hardware | Sledge Hammer | 3 | Stanlee | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 4 | 3 | 3 | Hardware | Rip Claw Hammer | 3 | Stanlee | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 5 | 4 | 4 | Clothing | Dri-Fit Tee | 2 | Mikey | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 6 | 5 | 5 | Clothing | Running Pants | 2 | Mikey | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 7 | 6 | 6 | Clothing | Wool Socks | 2 | Mikey | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 8 | 7 | 7 | Clothing | Squeaky Sneaks | 2 | Mikey | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 9 | 8 | 8 | Clothing | Cool Jeans | 7 | Leaveeyes | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 10 | 9 | 9 | Clothing | Denim Jacket | 7 | Leaveeyes | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 11 | 10 | 10 | Clothing | Leather Jacket | 7 | Leaveeyes | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 12 | 11 | 11 | Clothing | Courdory Pants | 7 | Leaveeyes | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 13 | 12 | 12 | Clothing | Work Pants | 3 | Stanlee | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 14 | 13 | 13 | Clothing | Work Gloves | 3 | Stanlee | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |
| 15 | 14 | 14 | Clothing | Comfort Tee | 8 | Weebob | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.000 | NULL |

Query executed successfully. | ist-cs-dw1.ad.syr.edu (12.0... | AD\kjthakar (86) | ist722_fudge_c7_dw | 00:00:00 | 54 rows

**Dim Titles:**




| | TitleKey | TitleID | TitleName | TitleType | TitleAvgRating | TitleReleaseYear | TitleRuntime | TitleRating | MovieGenre | RowIsCurrent | RowStartDate | RowEndDate |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -1 | No Title ID | No Title Name | No Type | -1.00 | -1 | -1 | No rating | No Genre | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 2 | 1 | 13kag | Meet Joe Black | Movie | 3.60 | 1998 | 10860 | PG-13 | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 3 | 2 | 13kal | Vampires | Movie | 3.30 | 1998 | 6420 | R | Dramas Based on the Book | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 4 | 3 | 13kaw | Eyes Wide Shut | Movie | 2.90 | 1999 | 9540 | R | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 5 | 4 | 13kbB | U.S. Marshals | Movie | 3.70 | 1998 | 7860 | PG-13 | Action & Adventure | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 6 | 5 | 13kbn | The Mighty | Movie | 3.80 | 1998 | 6000 | PG-13 | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 7 | 6 | 13kbQ | Dark City | Movie | 3.60 | 1998 | 6000 | R | Mystery | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 8 | 7 | 13kbW | B Monkey | Movie | 2.90 | 1999 | 5520 | R | Thrillers | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 9 | 8 | 13kcD | The Horse Whisperer | Movie | 3.60 | 1998 | 10080 | PG-13 | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 10 | 9 | 13kZ3 | Armageddon | Movie | 3.70 | 1998 | 9180 | PG-13 | Action & Adventure | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 11 | 10 | 13kZA | Pleasantville | Movie | 3.50 | 1998 | 7440 | PG-13 | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 12 | 11 | 13kZd | The Red Violin | Movie | 3.80 | 1998 | 7800 | R | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 13 | 12 | 13kZa | Apt Pupil | Movie | 3.20 | 1998 | 6660 | R | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |
| 14 | 13 | 13kZV | Velvet Goldmine | Movie | 3.30 | 1998 | 7620 | R | Drama | 1 | 1899-12-31 00:00:00.000 | 9999-12-31 00:00:00.0 |

**Fact Delivery Efficiency:**

SQLQuery18.sql - is...(AD\kjthakar (88)) | SQLQuery17.sql - is...(AD\kjthakar (87)) | SQLQuery16.sql - is...(AD\kjthakar (86)) | SQLQuery15.sql - is...(AD\kjthakar (85))

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [AccountKey]
      ,[TitleKey]
      ,[AccountTitleID]
      ,[QueueDateKey]
      ,[ShippedDateKey]
      ,[ReturnedDateKey]
      ,[OrderCount]
      ,[ShippedCount]
      ,[ReturnedCount]
      ,[ShipLag]
      ,[ReturnLag]
  FROM [ist722_fudge_c7_dw].[fudgemf].[FactDeliveryEfficiency]
```

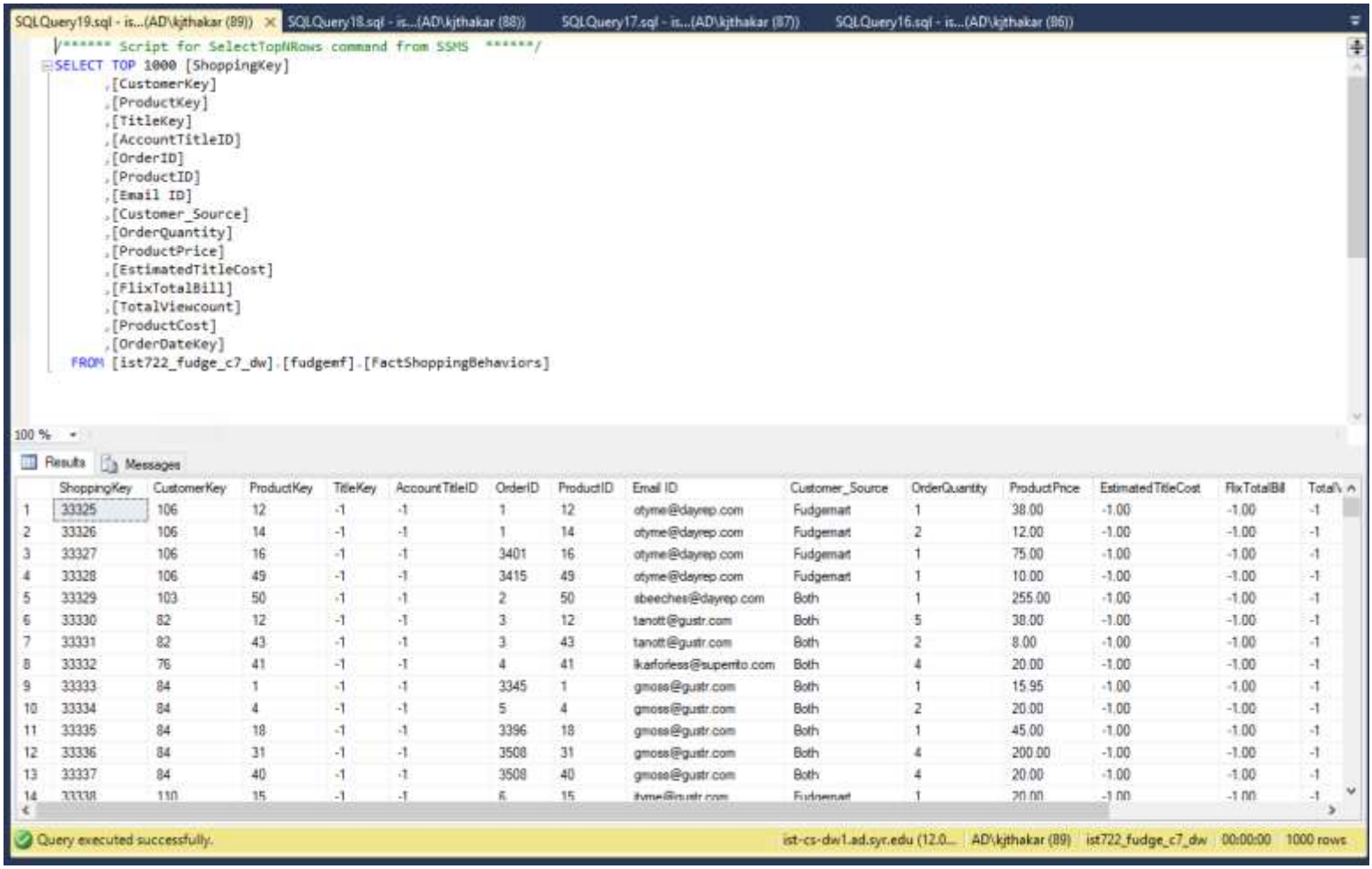
100 %

Results | Messages

| | AccountKey | TitleKey | AccountTitleID | QueueDateKey | ShippedDateKey | ReturnedDateKey | OrderCount | ShippedCount | ReturnedCount | ShipLag | ReturnLag |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 6 | 4390 | 10006 | 20100202 | 20100204 | 20100204 | 1 | 1 | 1 | 2 | 0 |
| 2 | 31 | 1173 | 10031 | 20100408 | 20100414 | 20100414 | 1 | 1 | 1 | 6 | 0 |
| 3 | 28 | 7095 | 10028 | 20100611 | 20100615 | 20100615 | 1 | 1 | 1 | 4 | 0 |
| 4 | 23 | 6316 | 10023 | 20100828 | 20100830 | 20100830 | 1 | 1 | 1 | 2 | 0 |
| 5 | 10 | 1917 | 10010 | 20101022 | 20101024 | 20101024 | 1 | 1 | 1 | 2 | 0 |
| 6 | 13 | 3407 | 10013 | 20110102 | 20110103 | 20110103 | 1 | 1 | 1 | 1 | 0 |
| 7 | 10 | 882 | 10010 | 20110313 | 20110314 | 20110314 | 1 | 1 | 1 | 1 | 0 |
| 8 | 19 | 2816 | 10019 | 20110523 | 20110526 | 20110526 | 1 | 1 | 1 | 3 | 0 |
| 9 | 30 | 1513 | 10030 | 20110808 | 20110809 | 20110809 | 1 | 1 | 1 | 1 | 0 |
| 10 | 25 | 6093 | 10025 | 20111012 | 20111014 | 20111014 | 1 | 1 | 1 | 2 | 0 |
| 11 | 4 | 5273 | 10004 | 20111229 | 20120103 | 20120103 | 1 | 1 | 1 | 5 | 0 |
| 12 | 7 | 844 | 10007 | 20120308 | 20120309 | 20120309 | 1 | 1 | 1 | 1 | 0 |
| 13 | 28 | 3752 | 10028 | 20120524 | 20120529 | 20120529 | 1 | 1 | 1 | 5 | 0 |
| 14 | 7 | 2116 | 10007 | 20120722 | 20120725 | 20120725 | 1 | 1 | 1 | 3 | 0 |
| 15 | 30 | 5622 | 10030 | 20120926 | 20120929 | 20120929 | 1 | 1 | 1 | 3 | 0 |

Query executed successfully. | ist-cs-dw1.ad.syr.edu (12.0... | AD\kjthakar (88) | ist7

**Fact Shopping Behaviours:**



SQLQuery19.sql - is...(AD\kjthakar (89)) | SQLQuery18.sql - is...(AD\kjthakar (88)) | SQLQuery17.sql - is...(AD\kjthakar (87)) | SQLQuery16.sql - is...(AD\kjthakar (86))

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 1000 [ShoppingKey]
      ,[CustomerKey]
      ,[ProductKey]
      ,[TitleKey]
      ,[AccountTitleID]
      ,[OrderID]
      ,[ProductID]
      ,[Email ID]
      ,[Customer_Source]
      ,[OrderQuantity]
      ,[ProductPrice]
      ,[EstimatedTitleCost]
      ,[FlixTotalBill]
      ,[TotalViewcount]
      ,[ProductCost]
      ,[OrderDateKey]
  FROM [ist722_fudge_c7_dw].[fudgemf].[FactShoppingBehaviors]
```

100 %

Results | Messages

| | ShoppingKey | CustomerKey | ProductKey | TitleKey | AccountTitleID | OrderID | ProductID | Email ID | Customer_Source | OrderQuantity | ProductPrice | EstimatedTitleCost | FlixTotalBill | Total\ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 33325 | 106 | 12 | -1 | -1 | 1 | 12 | otyme@dayrep.com | Fudgemart | 1 | 38.00 | -1.00 | -1.00 | -1 |
| 2 | 33326 | 106 | 14 | -1 | -1 | 1 | 14 | otyme@dayrep.com | Fudgemart | 2 | 12.00 | -1.00 | -1.00 | -1 |
| 3 | 33327 | 106 | 16 | -1 | -1 | 3401 | 16 | otyme@dayrep.com | Fudgemart | 1 | 75.00 | -1.00 | -1.00 | -1 |
| 4 | 33328 | 106 | 49 | -1 | -1 | 3415 | 49 | otyme@dayrep.com | Fudgemart | 1 | 10.00 | -1.00 | -1.00 | -1 |
| 5 | 33329 | 103 | 50 | -1 | -1 | 2 | 50 | sbeeches@dayrep.com | Both | 1 | 255.00 | -1.00 | -1.00 | -1 |
| 6 | 33330 | 82 | 12 | -1 | -1 | 3 | 12 | tanott@gustr.com | Both | 5 | 38.00 | -1.00 | -1.00 | -1 |
| 7 | 33331 | 82 | 43 | -1 | -1 | 3 | 43 | tanott@gustr.com | Both | 2 | 8.00 | -1.00 | -1.00 | -1 |
| 8 | 33332 | 76 | 41 | -1 | -1 | 4 | 41 | lkarforless@supermito.com | Both | 4 | 20.00 | -1.00 | -1.00 | -1 |
| 9 | 33333 | 84 | 1 | -1 | -1 | 3345 | 1 | gmoss@gustr.com | Both | 1 | 15.95 | -1.00 | -1.00 | -1 |
| 10 | 33334 | 84 | 4 | -1 | -1 | 5 | 4 | gmoss@gustr.com | Both | 2 | 20.00 | -1.00 | -1.00 | -1 |
| 11 | 33335 | 84 | 18 | -1 | -1 | 3396 | 18 | gmoss@gustr.com | Both | 1 | 45.00 | -1.00 | -1.00 | -1 |
| 12 | 33336 | 84 | 31 | -1 | -1 | 3508 | 31 | gmoss@gustr.com | Both | 4 | 200.00 | -1.00 | -1.00 | -1 |
| 13 | 33337 | 84 | 40 | -1 | -1 | 3508 | 40 | gmoss@gustr.com | Both | 4 | 20.00 | -1.00 | -1.00 | -1 |
| 14 | 33338 | 110 | 15 | -1 | -1 | 6 | 15 | dyme@gustr.com | Fudgemart | 1 | 20.00 | -1.00 | -1.00 | -1 |

Query executed successfully. | ist-cs-dw1.ad.syr.edu (12.0... | AD\kjthakar (89) | ist722_fudge_c7_dw | 00:00:00 | 1000 rows

**Data Governance:**

- If the zip code and the state do not match, we then use zip code to generate the state data.
- For all the cell phone numbers, we should clean all the symbols, such as +, (), and -.
- We should establish hierarchies of department – product, so end users can analyze data more easily.

**Data Quality Rules:**

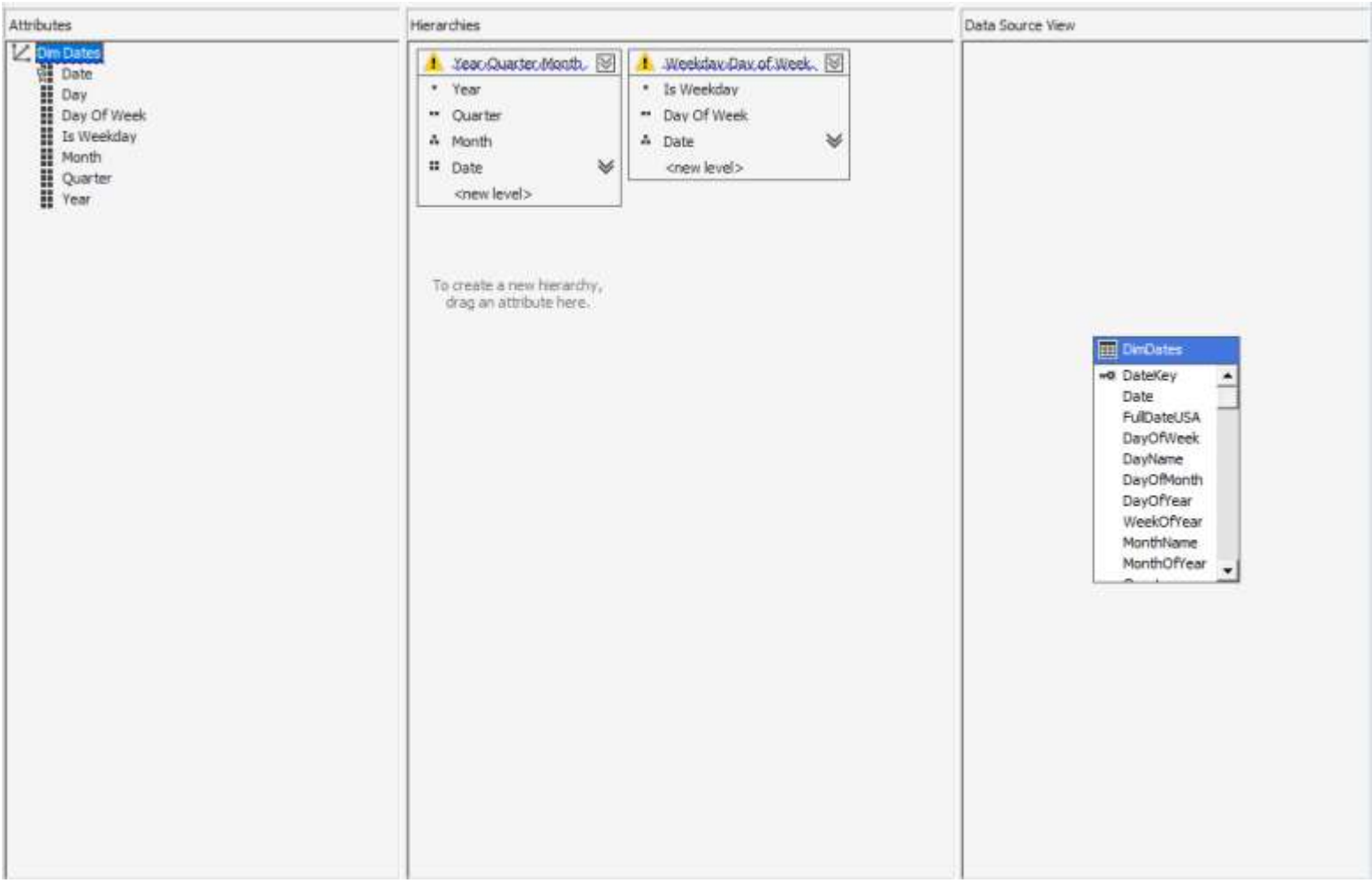
- DQ rule #1 accurate: OrderDateKey and ShippedDateKey should not later than today's date. (reject)
- DQ rule #2 completeness: SourceSystem should not contain missing values. (warn)
- DQ rule #3 accurate: SalesPrice and Quantity should be greater than or equal to 0. (reject)
- DQ rule #4 precision: ClientState can only have two letters. (fix by using zipcode)
- DQ rule #5 completeness: ClientFirstName and ClientLastName must not be null. (reject)
- DQ rule #6 consistency: OrderDateKey, ShippedDateKey, and ExpiryDate should have consistency, such as 20090101. (fix)
- DQ rule #7 timeliness: ShippedDateKey can only show up after OrderDateKey exists. (reject)

## SQL SERVER ANALYSIS SERVICES

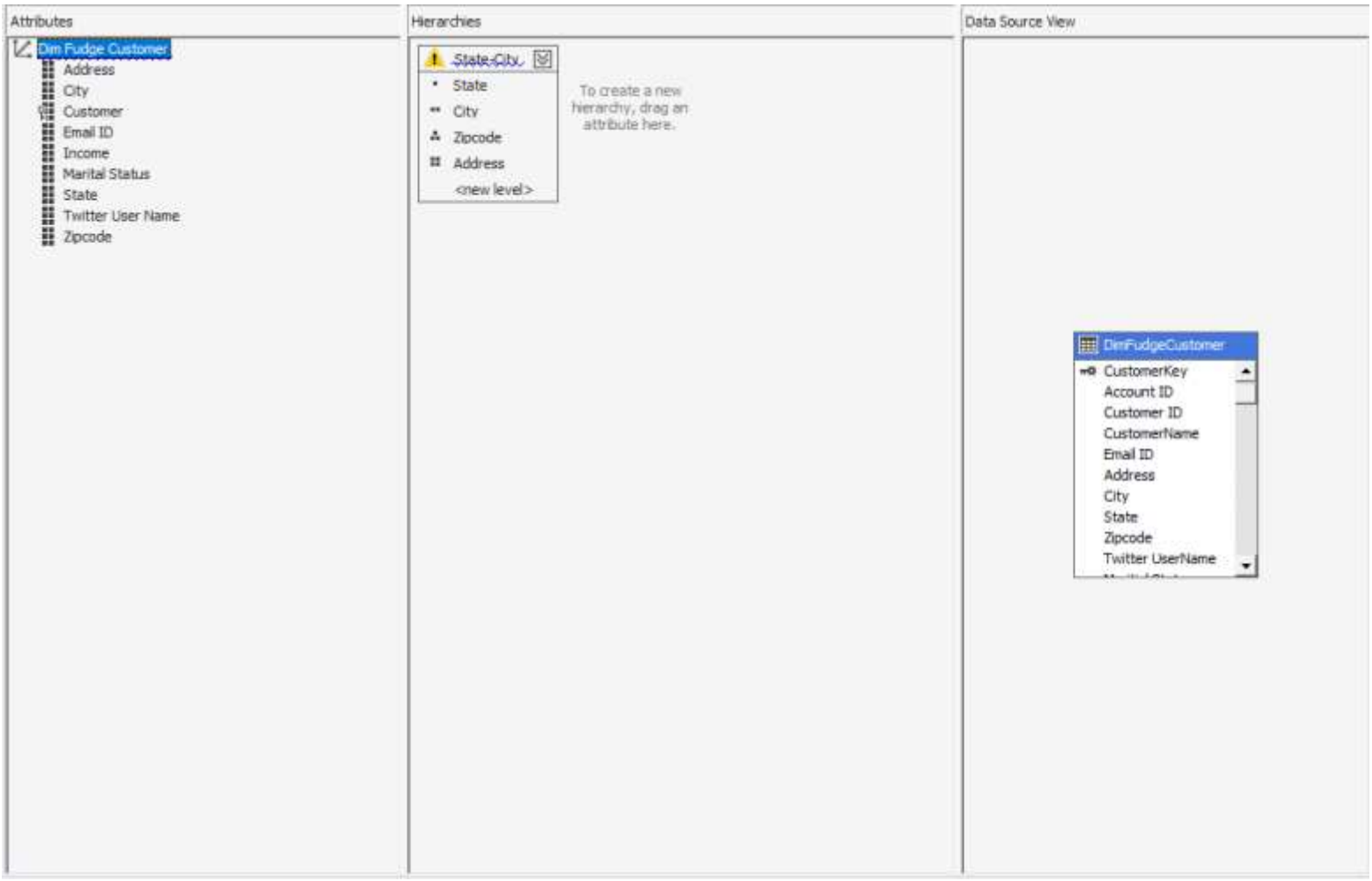
We use the SSAS for building a multidimensional database (MOLAP) from our dimensional database ( Data Warehouse). The Final end result will be a user facing data structure called as a cube. This cube has different decided dimensions along the edges, which can be used to drill across data for analytical tasks.

The properties of the dimensions can be changed to display something more simpler to the end user and also we can create the hierarchies of data so that it is organized in meaningful sequences and pre-calculate data for performance improvements. The attributes and hierarchies are as follows:

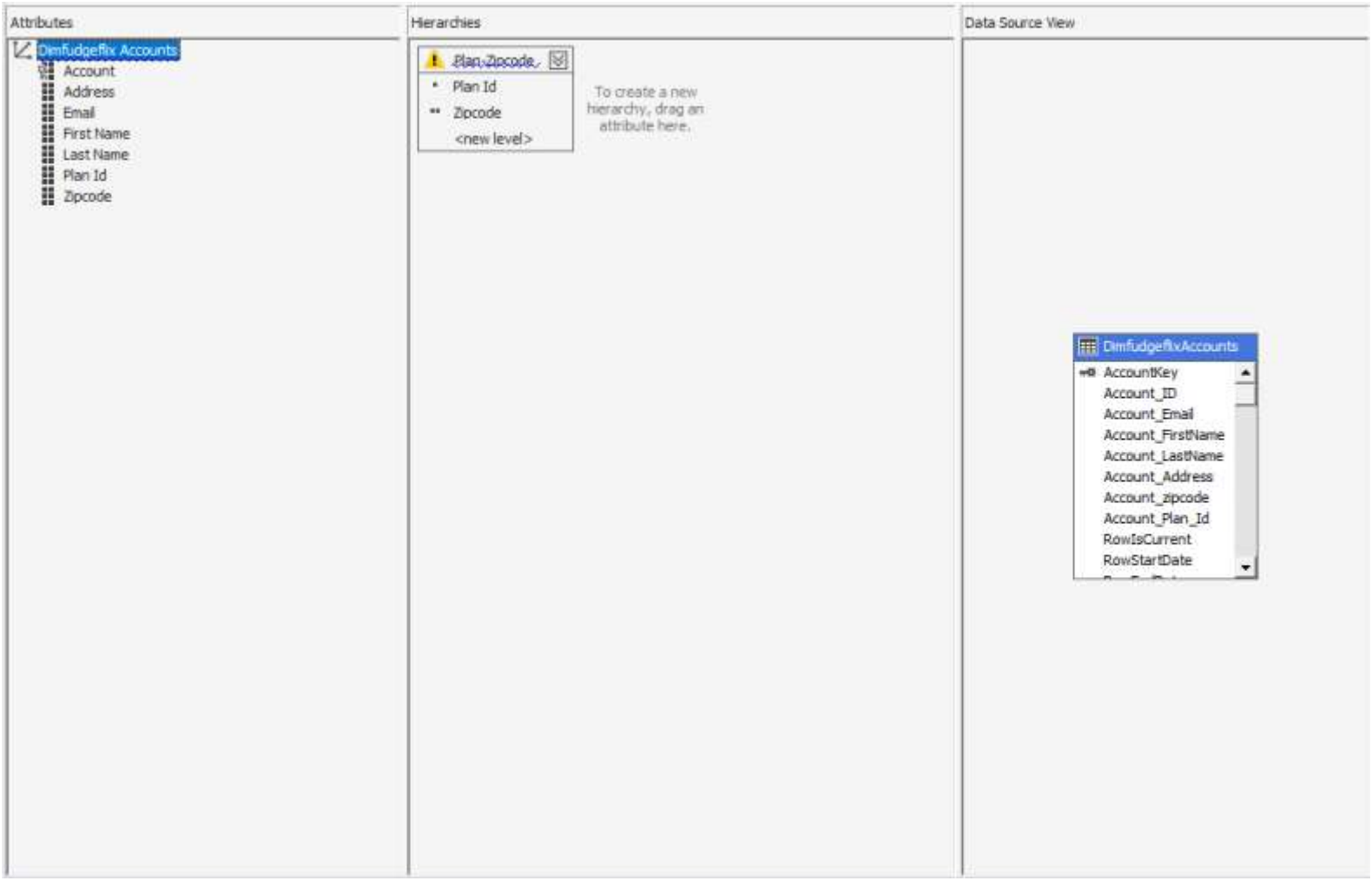
**Dimension Dates:**



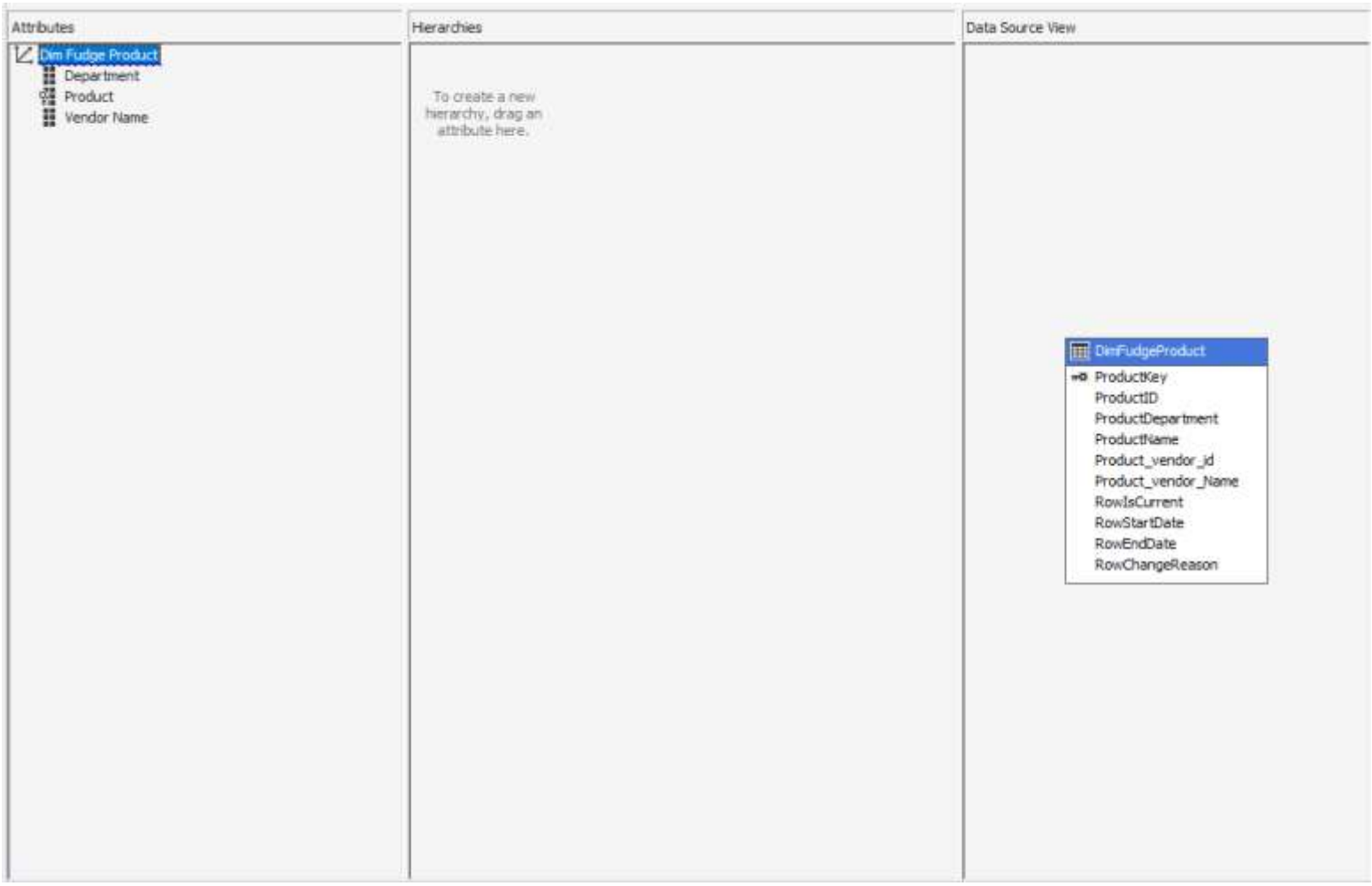
**Dimension Customer:**



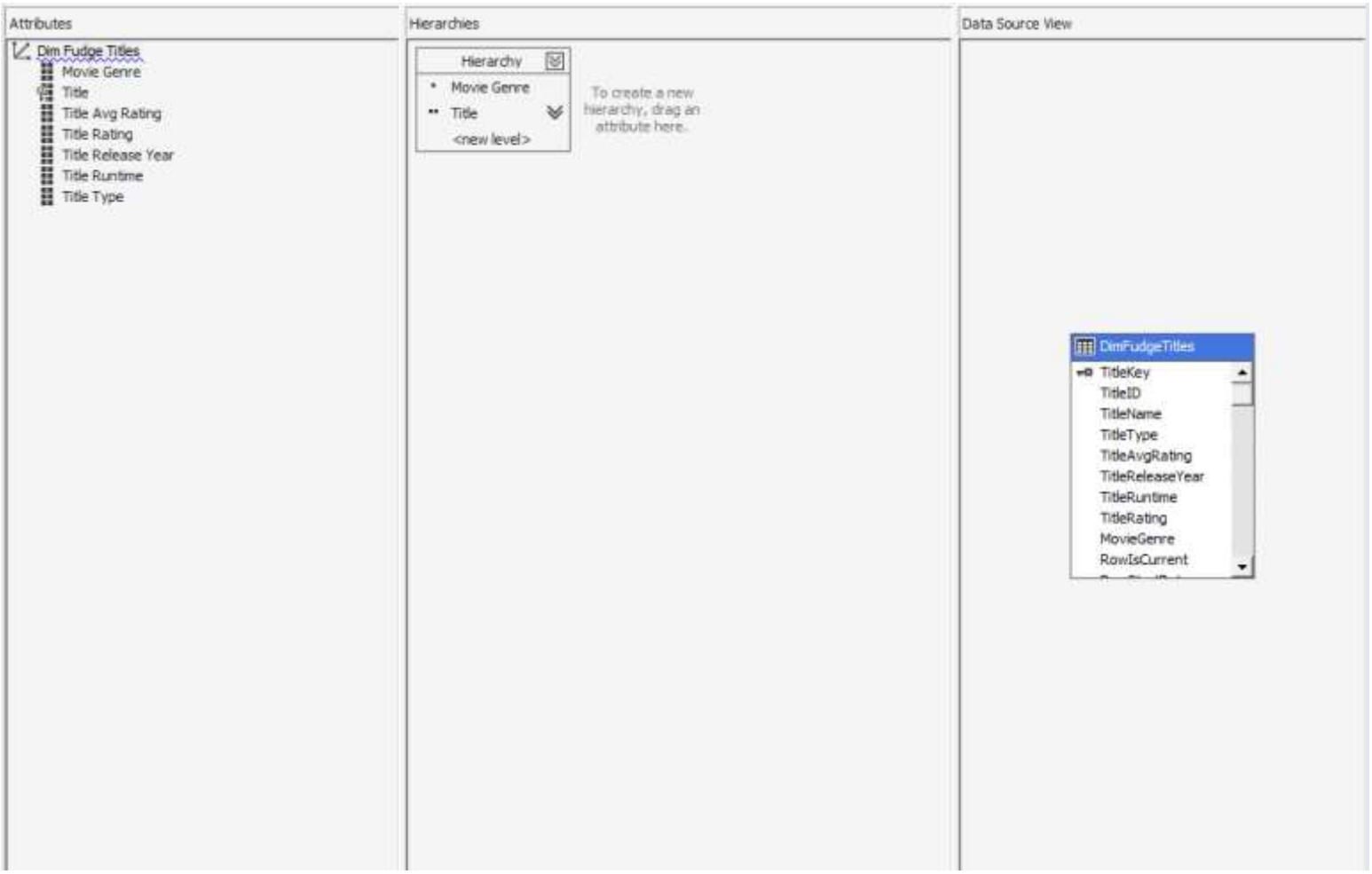
**Dimension Accounts:**



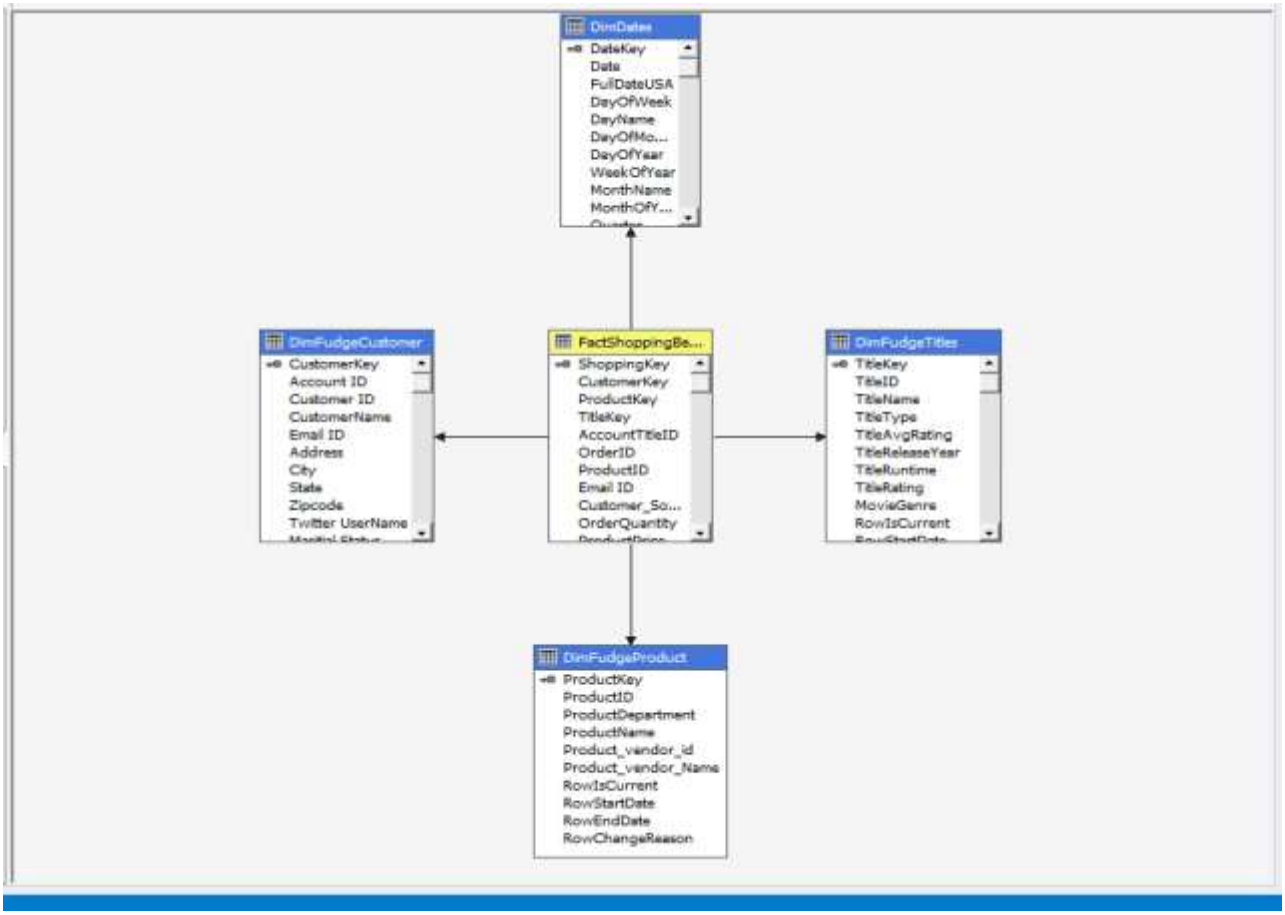
**Dimension Product:**



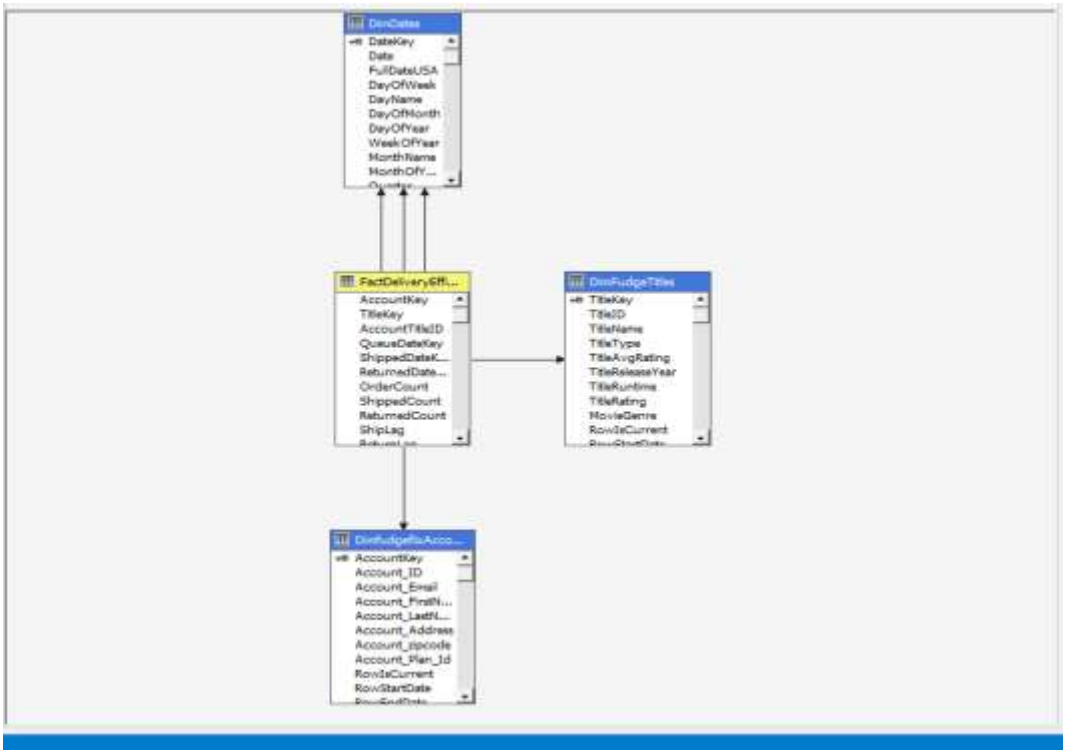
**Dimension Titles:**



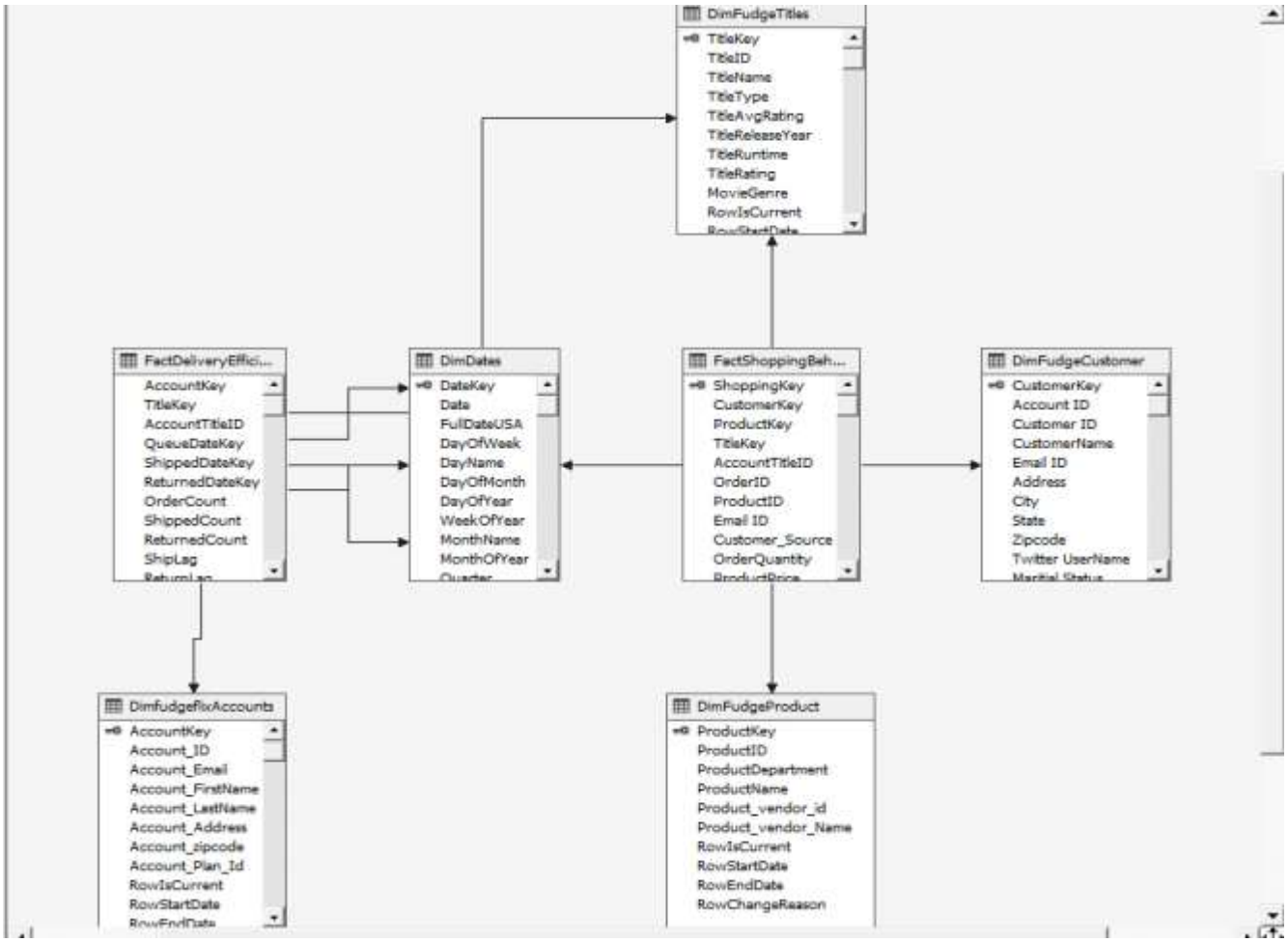
**Fact Shopping Behavior:**



**Fact Delivery Efficiency:**



SSAS Diagram



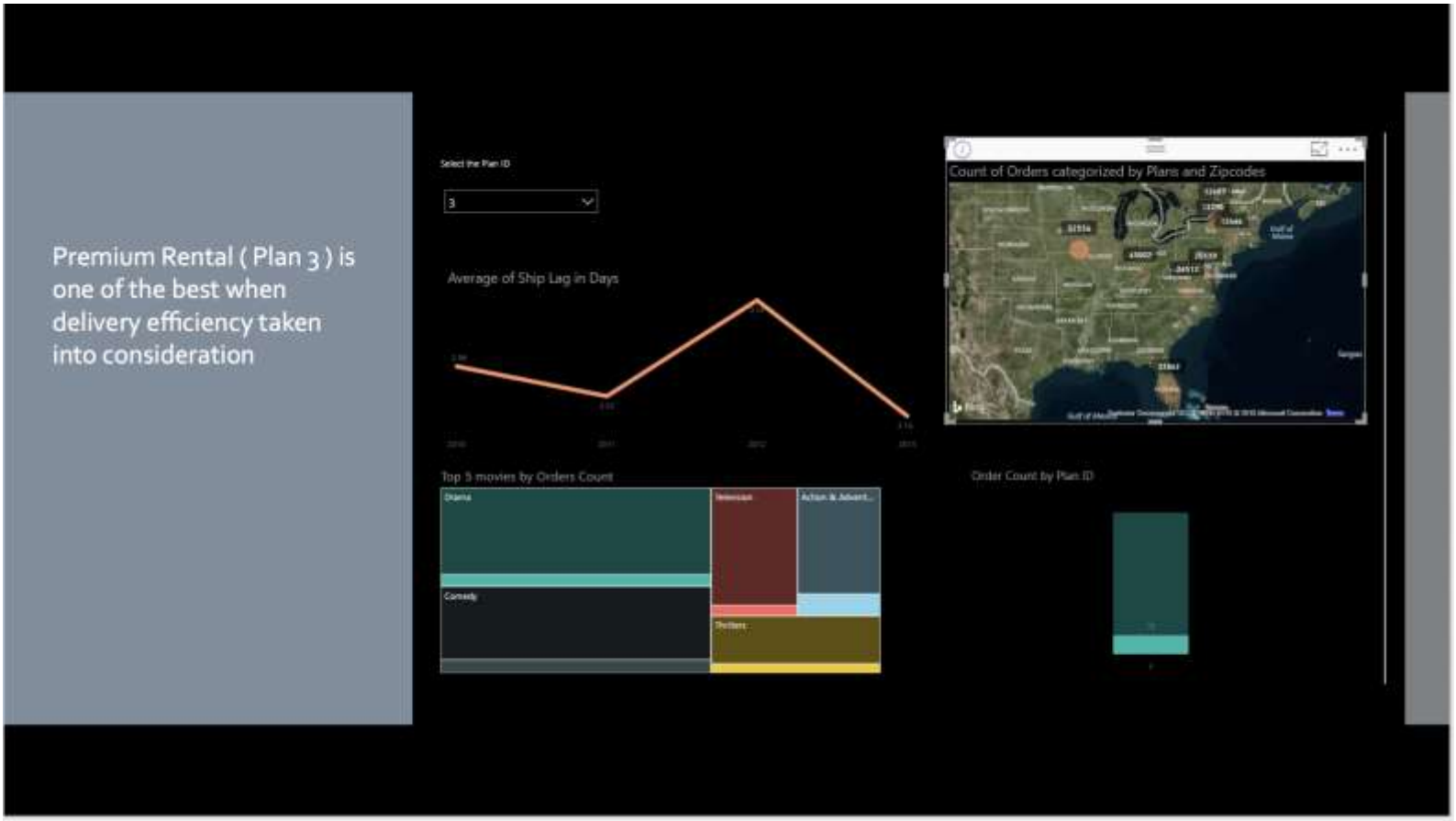
## BUSINESS INTELLIGENCE

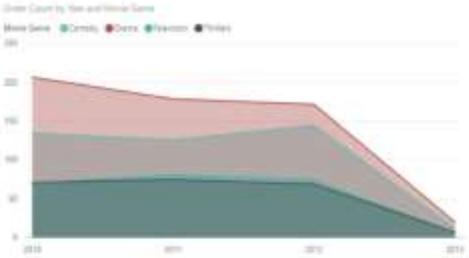
Business intelligence tools, such as Power BI and Excel, allow users to do interactive visualizations with self-service business intelligence capabilities. Our goal is to allow end users to create reports and dashboards, without having to depend on information technology staff. We loaded data from the data warehouse and used visualization charts, such as card, donut chart, stacked bar chart, gauge, line chart, and tree map, to answer the business questions. The sample questions and solutions are as follows:

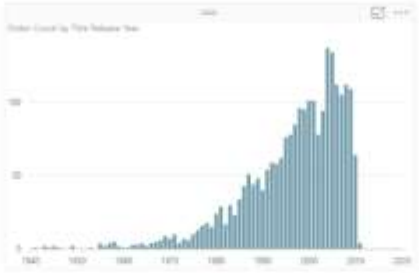
**Knowing your Customers:**

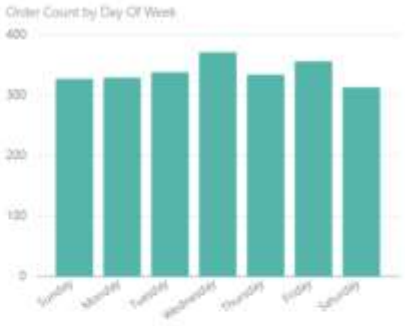
**Shopping Behavior:**

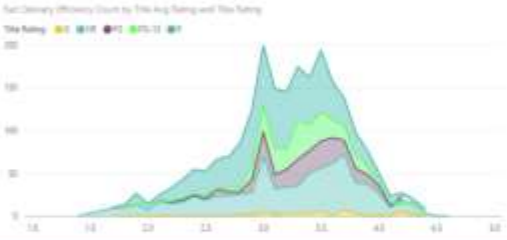
**Delivery Efficiency:**

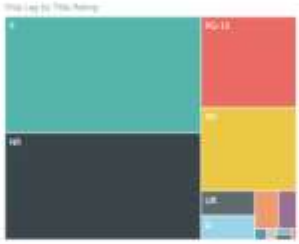










# DELIVERY EFFICIENCY

- Movies under the Drama genre have been ordered the most over the years, followed by Comedy, Television and Thrillers
- Ship Lag for Drama has been highest over the years, so we need to understand if this is influenced by Order count, or another external factor
- Order count has reduced drastically since 2011, due to other stiff competitors
- We also see that the highest orders are received on Wednesdays, and are ordered comparatively lesser on weekends
- Order count of television series is comparatively lesser than movies, hence they need to look at ways to improve television series

## SUMMARY

In the beginning, we determined the functional requirements and business processes, and documented all the dimensions and facts in the high-level and detail-level dimensional sheets. We then built the data warehouse on SQL Server and followed the source to target map integrating the data from OLTP sources, including Fudgemart, Fudgeflix, and external sources, to the data warehouse. Moreover, we created cubes in SSAS for further analysis. Finally, we extracted the data from the data warehouse and imported it to Power BI, allowing end users to answer business questions by themselves, which will improve the performance of analysis and save time.

## REFERENCE

**Fudgemart Data_Profiling.xlsx**
**High-level Dimensional Modeling.xlsx**
**Detail-level Dimensional Modeling.xlsm**
**SSIS (part1)\Fudge Inc\Fudge Inc.sln**
**SSIS (part2)\Integration Services Project1\Integration Services Project1.sln**
**SSAS\ist722_fudge_c7(IST-CS-DW1).sln**
**Power BI\IST722 Final Project.pbix**
**IST722 Final Project PowerPoint (Group 7).pptx**