**IST 722 group project report**

**Project:** IST 722 group project

**Group5:** Tessa Blankenship, Joshua Rofrano, Arthur Fu

**Date:** 2019Oct-2019 Dec

**Project portfolio and check list**

The following items are included in the Project portfolio,

**IST 722 workgroup 5 project -Submission**

1. Project document
2. High-level dimensional modeling worksheet
3. Detail-level dimensional modeling worksheet
4. Data warehouse on SQL Server
5. Initial ETL done in SSIS
6. Business intelligence
   1. Fact Account Billing.pbix

6.2 Fact Order Details.pbix

**Project background:**

* Fudgemart, Inc. is a conglomerate with two subsidiary companies: Fudgemart and Fudgeflix.
* Fudgemart is a fictitious online retailer, similar to Amazon.com or Walmart.com. The database consists of customers, products, and vendors, and has familiar business processes you would find in any online retailer. The database for Fudgemart is called Fudgemart\_v3.
* Fudgeflix is a fictitious online DVD-by-mail and video-on-demand service, similar to Amazon Instant Video or Netflix. The database for Fudgeflix is called Fudgeflix\_v3 and contains concepts such as accounts, subscriptions, and video titles, as well as other things associated with an online video-streaming service.

**Functional Requirements and Business processes**

1. FudgeMart Order Details

* OrderQty - additive
* Total Sales Amount – additive

1. FudgeFlix Account Billing

* Total Billed Amount – additive

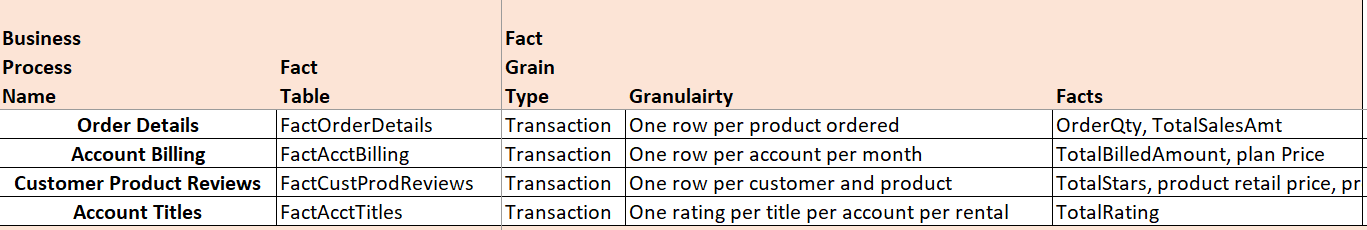
1. FudgeMart Customer Product Review

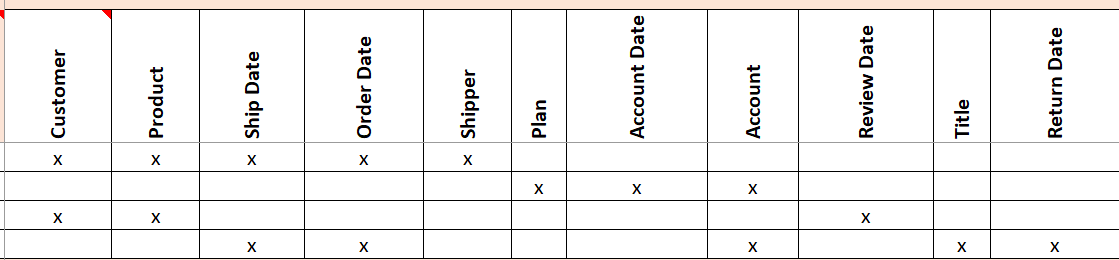
* Total Stars – semi-additive

1. FudgeFlix Account Titles

* Total Rating – semi-additive

**Bus Matrix**

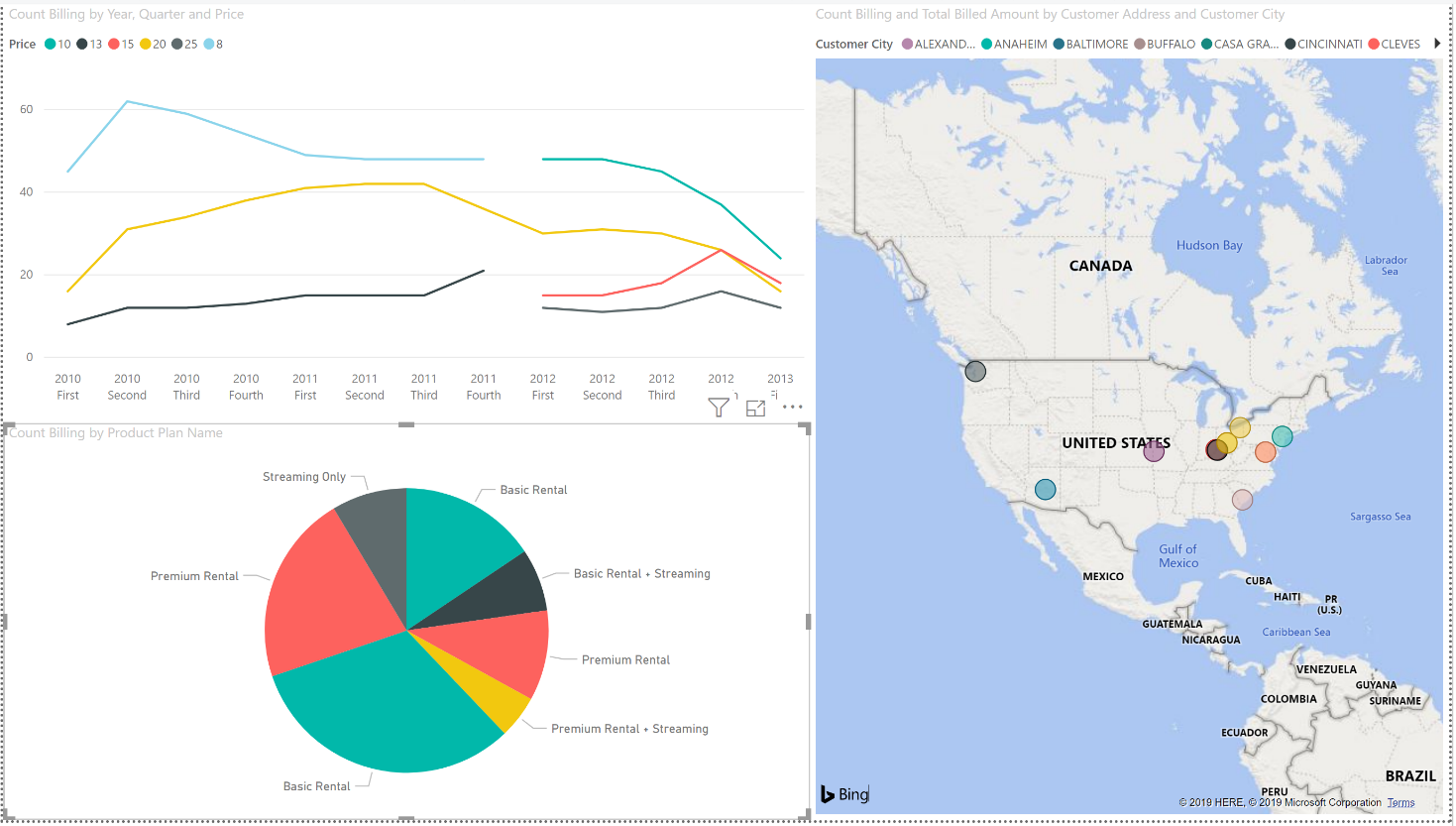




Project summary:

The most useful part of this project is to consolidate what we’ve learned throughout the course and apply that knowledge in one large project. This project is not trivial, we learned a lot in this course and the various techniques can be quite complex. Some difficulties that we encountered are summarized as follows.

1. The first challenge is identifying the functional requirements of Fudgemart. It seems simple but it’s quite tricky. For people who are new to data warehousing, correctly defining the business processes is not easy. Initially, when we designed the bus matrix, we considered using relatively complex fact table with a lot of information. We initially put orders quantity and total billed amounts together. Later we decided to separate them and move forward with the current simpler version. Start simple and develop it further seems to be the right strategy.
2. Building DW tables and making it consistent with the source data is another challenge. The main issues we encountered were the consistency of the data types. It’s ok that the data types in the DW tables are not consistent with the source data as we can use ETL methods to fix that. However, it will be much easier if we can think ahead and make them consistent when making the DW tables initially, this will save us a lot effort when doing ETL.
3. Using the right ETL tools is important. There are a few versions of visual studio available for this project, including visual studio 2015,2017 and 2019. At first, we started with the latest version, visual studio 2019, but it’s a bit buggy and caused us to have a lot of errors during ETL. After spending time on researching and communicating with other teams, we knew that the 2015 or 2017 version is more solid in terms of doing ETL.
4. Finally, it is important to make a good plan and execute it in a timely manner. Also, in view of the particularity of the data warehouse which needs to connect to the server, we should have a backup plan when the server goes down to avoid delays in delivering the data to the end users.

6.2 Account Billing 

6.1 Fact Order Details 