**Project Issues and Take-home Message**

IST722 Data Warehouse - Fall 2019 - Team#1

Xuehan Chen, Yue Wang, Yimin Xiao, Chiau Yin Yang

We would like to show our respect and thanks to Prof. Fudge. Thank you very much for your help with IST722 throughout this semester. We have learned a lot from 8 tutoring assignments and the group project.

Following are the things we have done and made great progress when we worked with this project:

1. Order table is business process, so it was not added to be an individual dimension.
2. Stage account title (Fudgefliex) and orders (Fudgemart) separately and merge into customer dimension in DW so that we can look up the orders from different sources using different keys.
3. We did not include too many dimensions - product or order details dimension, but only pull those in to calculate order amount for risk analysis fact table.
4. For shipping date, we replace null with -1 (unknown members), and calculate OrderToShipInDays using the earliest date, which cause relatively low number. This can be filtered in BI program.

However, there are several points that we could have done better for our Business intelligence process.Following are some of the lessons learned from the final project and potential improvements that we can make in future data warehouse projects.

For business users to make better BI analysis, below are a few things that we could have done better:

1. Adding the order detail dimension to analyze if certain products delay order to ship to better investigate reasons behind the delays, or estimate shipping schedule for better customer services.
2. The count of the fulfilled orders to show shipment status for each order, and more facts for business users to manipulate.
3. Integration of shipment monitoring analysis and risk analysis in order to avoid shipping unapproved orders from the risk analysis.
4. Detailed information(city and state), rather than just zip codes for better interpretation on business intelligence results.
   1. This will also resolve issue of customer address format difference from 2 data sources. So far, in the fact table, we only include zip code and address from both sources. In fudgefliex, the city and state attributes are stored in different tables, we could have included them for data consistency and semantic BI for business users.
5. Another issue that we have not resolved is that the original detailed dimension worksheet is different than the final representation in the ETL and star schema.
6. In addition, the naming for the dimensions and facts may not be consistent from the beginning to the end even though we tried to.