



MARCH 18, 2018

# A WEDDING BOUTIQUE DATABASE

KM/BI 6255

HUONG PHAM

## BACKGROUND

Kelly's Wedding Boutique is a bridal boutique business owned by my cousin. The store sells a variety of products, including bridal gowns, bridesmaid dresses, accessories and shoes. There are three stores in Seattle, Spokane and Portland. Besides selling products, Kelly's Wedding Boutique also provides rental service on selected items. The company works with six suppliers in U.S, Hong Kong and China. This project shows a small sample of store's activities from May 2014 and June 2016.

The goals of building a data mart are to answer the following questions:

1. How many units were sold at certain store in a period of time?
2. How much sales, profits, gross profit margin, average earning at certain store in a period of time?
3. What is the amount of inventory at certain store in a period of time?
4. How quick and efficient are the suppliers in a period of time?

## UNDERSTANDING THE DATA

### CUSTOMERS

This table represents information about customers, including: customer ID, first name, last name, city and state that customer are the residing at.

### ORDERS

The "Orders" table contains information about the order type, when an order is placed, discount type if applicable, discount percentage, and the store ID that the order is placed at.

### ORDER DETAIL

This table tells me information about the order quantity, the item price, product ID and order I

### RENTALS

This table helps the store manage the rental orders, including rental date, return date, amount of deposit, late fee.

## INVENTORY

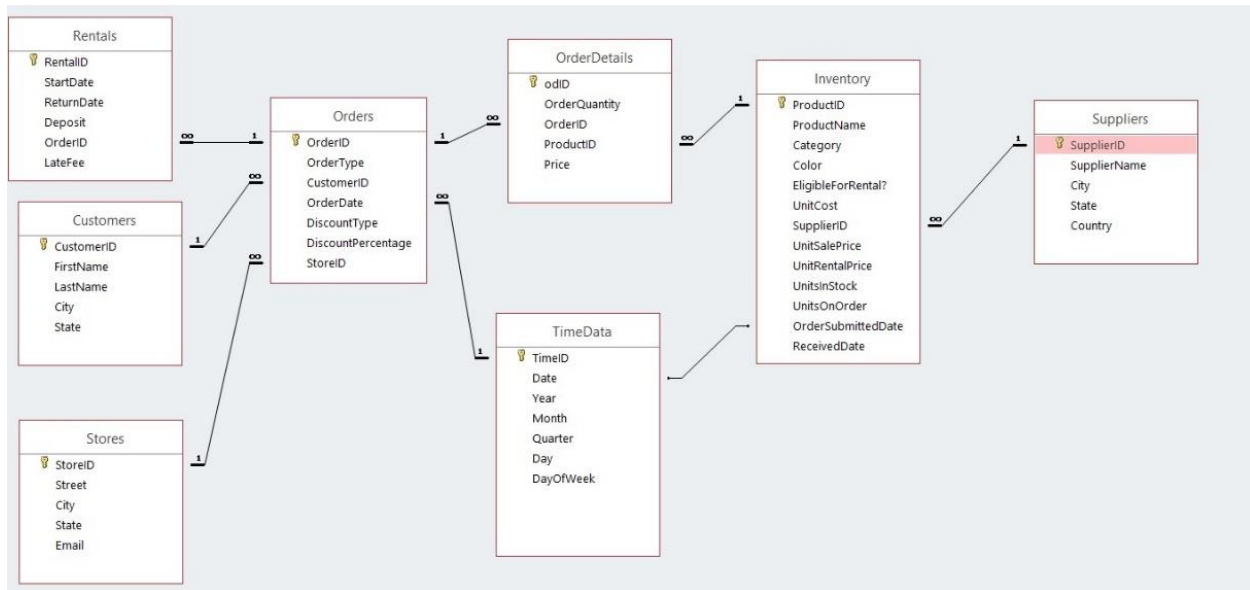
This table represents information about different products, such as: product name, eligibility for rental, category, unit cost, color and etc.

## STORES

The staffs can look up information about other store locations by using this table, including: address (Street, City, State), and email.

## DATABASE MODELING

An ER diagram was created using Access.



*Fig 1: ER Diagram for Kelly's Wedding Boutique Store*

After finalizing the ER diagram, I created one cube with five dimensions

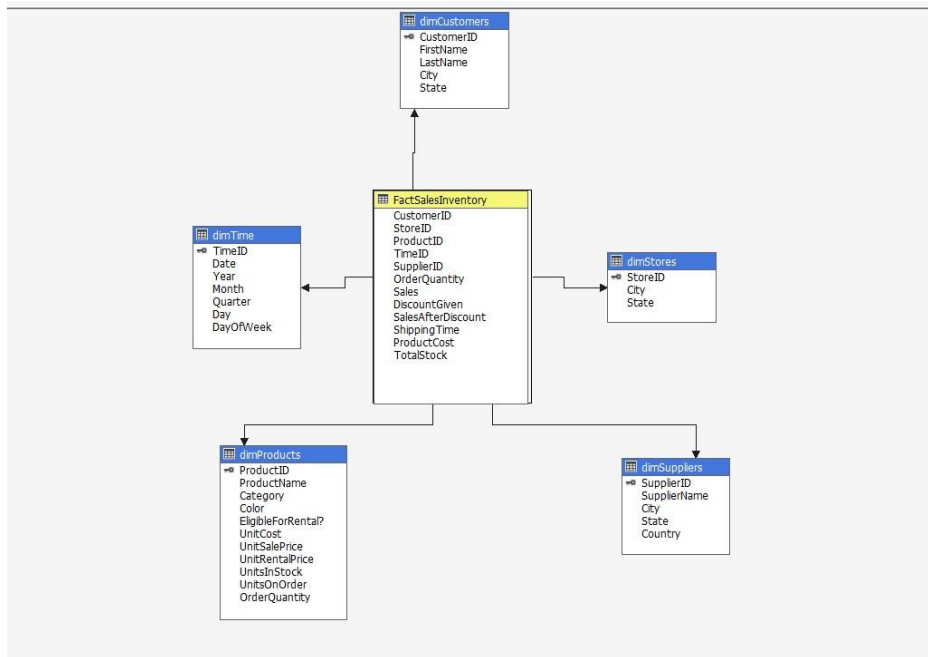


Fig 2: Data Cube with Five Dimensions in Microsoft SQL Server

There are 6 measures that I created by using the Expression Builder function in Access

## MEASURE #1: SALES

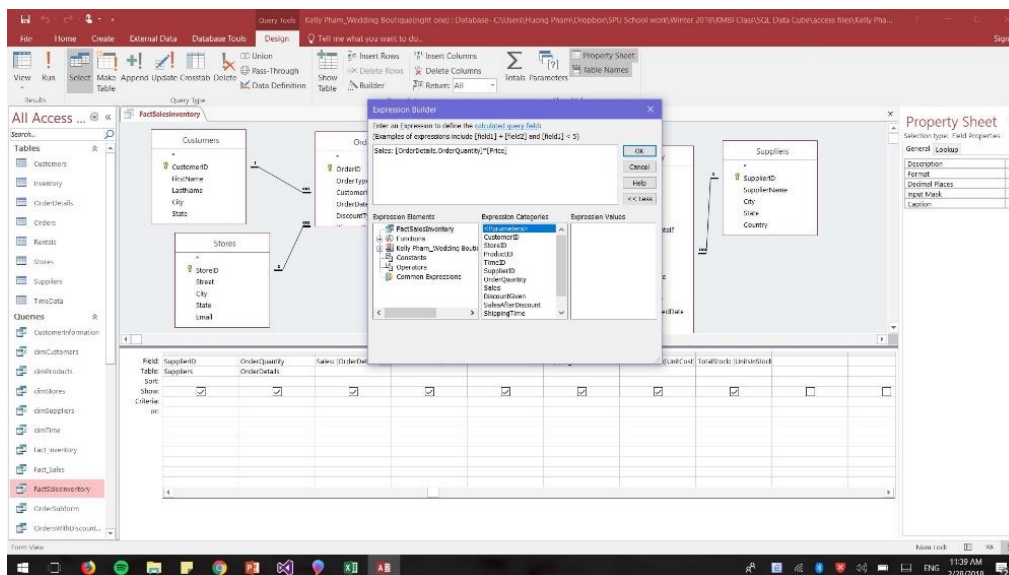


Fig 3: Measure #1 – Sales Calculation in Access

## MEASURE #2: DISCOUNTGIVEN

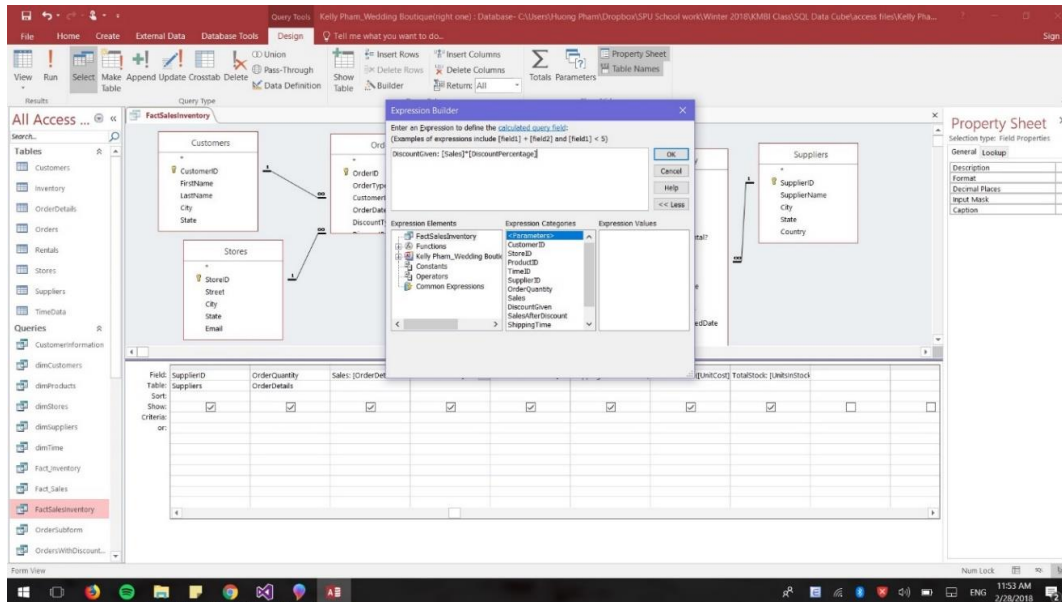


Fig 4: Measure #2 – Discount Given Calculation in Access

## MEASURE #3: SALES AFTER DISCOUNT

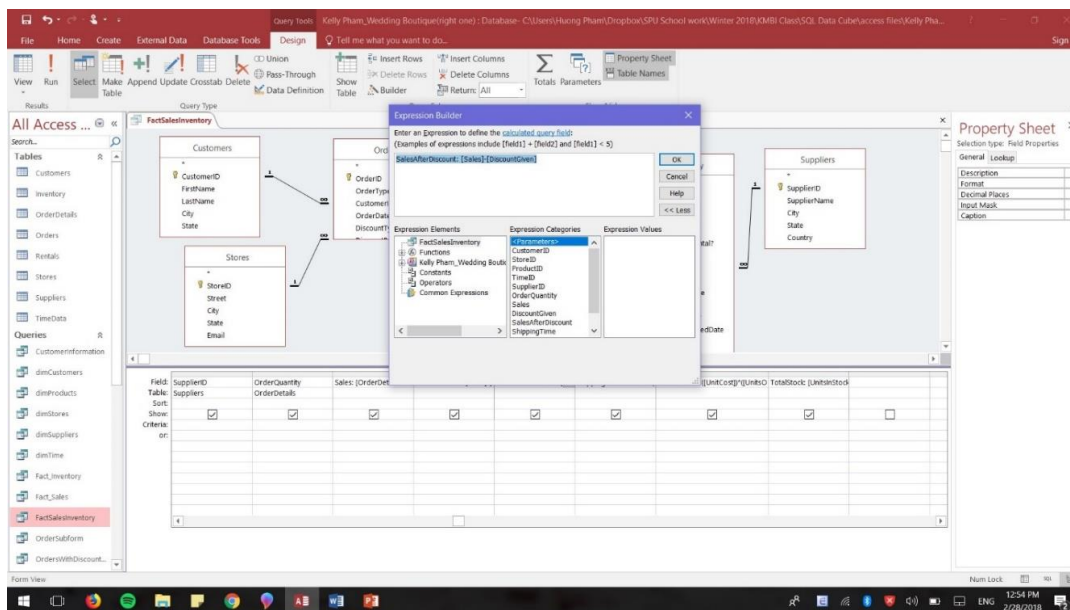


Fig 4: Measure #3 – Sales after Discount Calculation in Access

## MEASURE # 4 - SHIPPINGTIME

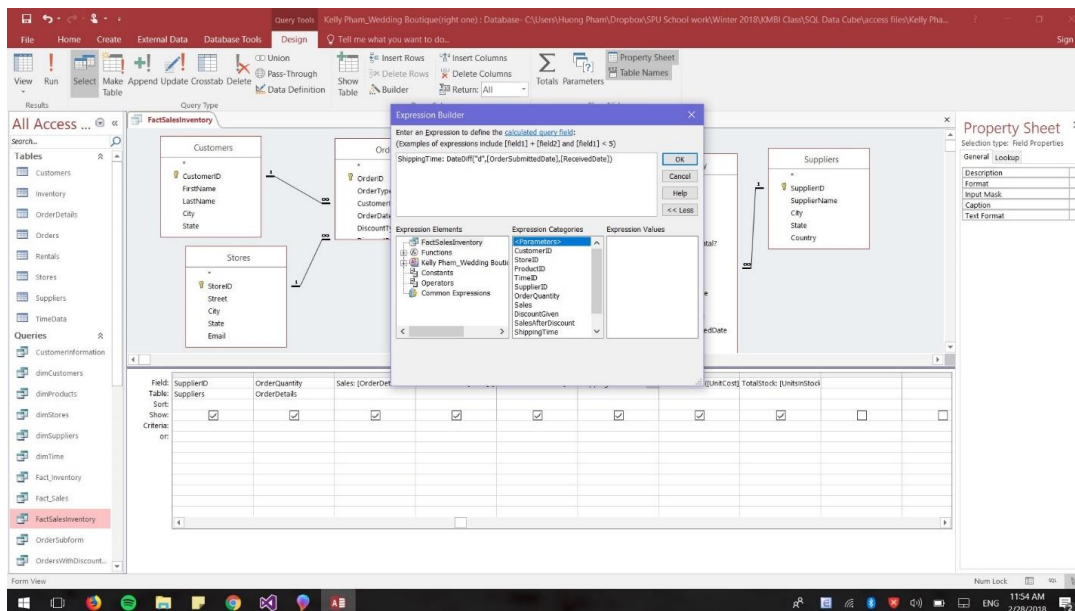


Fig 5: Measure #4 – Shipping Time Calculation in Access

I used this measure #4 to determine difference between the date when the stock order was submitted, and the date when the store received the stock. By doing this, I wanted to compare which supplier is more efficient.

## MEASURE #5: PRODUCTCOST

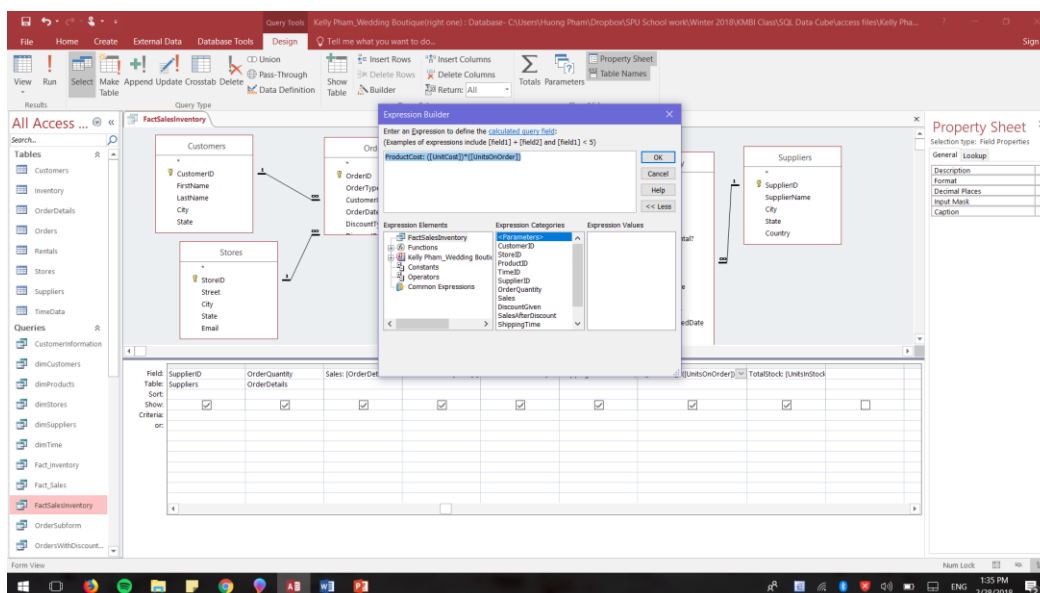
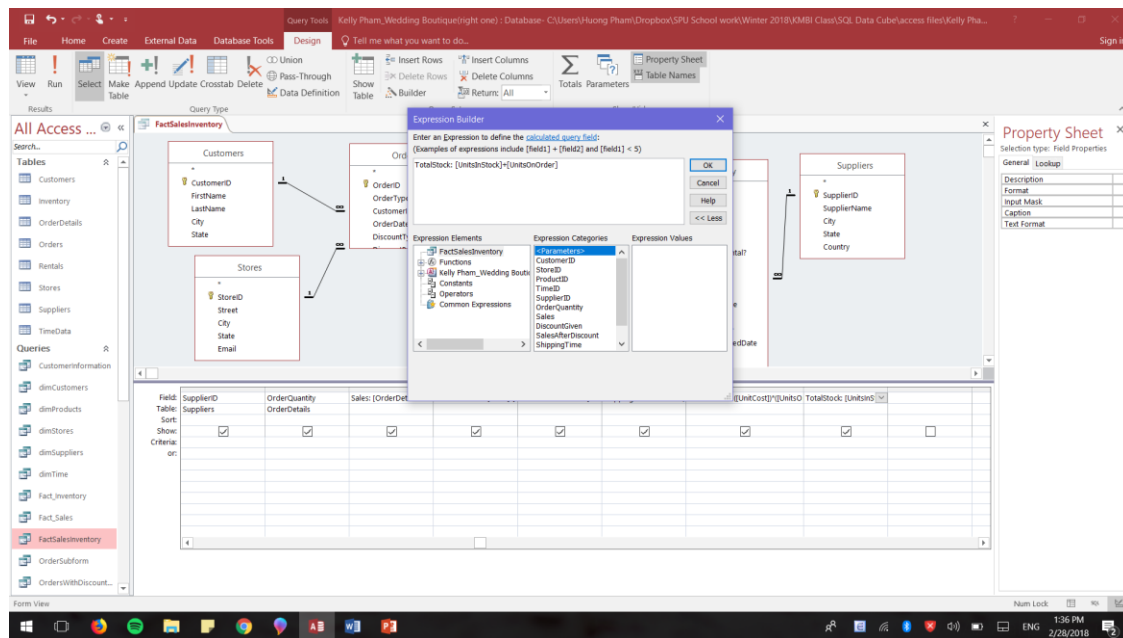


Fig 6: Measure #5 – Product Cost Calculation in Access

## MEASURE # 6: TOTALSTOCK



*Fig 7: Measure #6 – Total Stock Calculation in Access*

After I built the cube, I created more measures in SQL Server. I wanted to calculate the total profit, the average earning per period and the average shipping time per period. In order to get these, I created a new measure "Time ID Distinct Count". Originally, I wanted to do a distinct count of the date. I saw a Youtube tutorial on doing this in the Cube Structure. However, in the tutorial, this person had a date attribute in his fact table. I did not have the date in my fact table. Therefore, I decided to do a distinct count of date in the dimTime. This created a connection error in the dimTime and deleted the hierarchy in dimTime. I had to rebuild the cube and tried it again. For the second time, I decided to do a distinct count of Time ID in my fact table and it worked.

## CALCULATIONS IN SQL SERVER

First, I created "Time ID Distinct Count", and then more measures using the Calculation function in SQL Server.

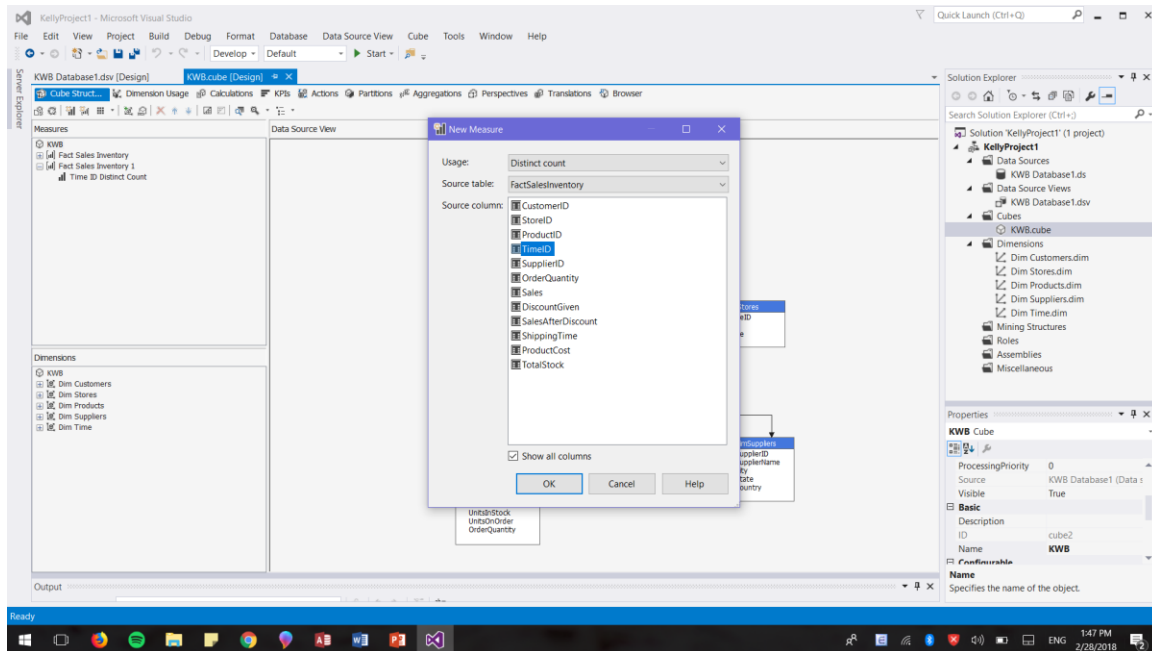


Fig 8: Create TimeID Distinct Count in Microsoft SQL Server

## CALCULATION #1: PROFIT

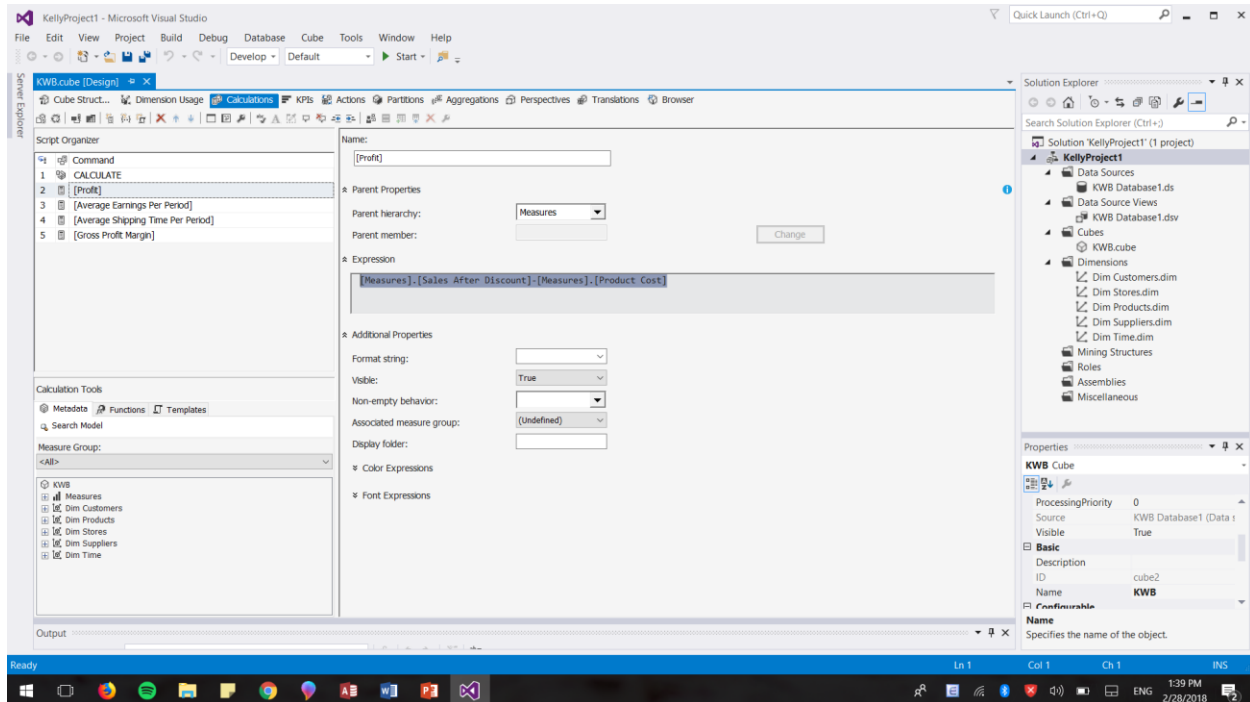


Fig 9: Profit Calculation in Microsoft SQL Server



## CALCULATION #2: AVERAGE SHIPPING TIME PER PERIOD

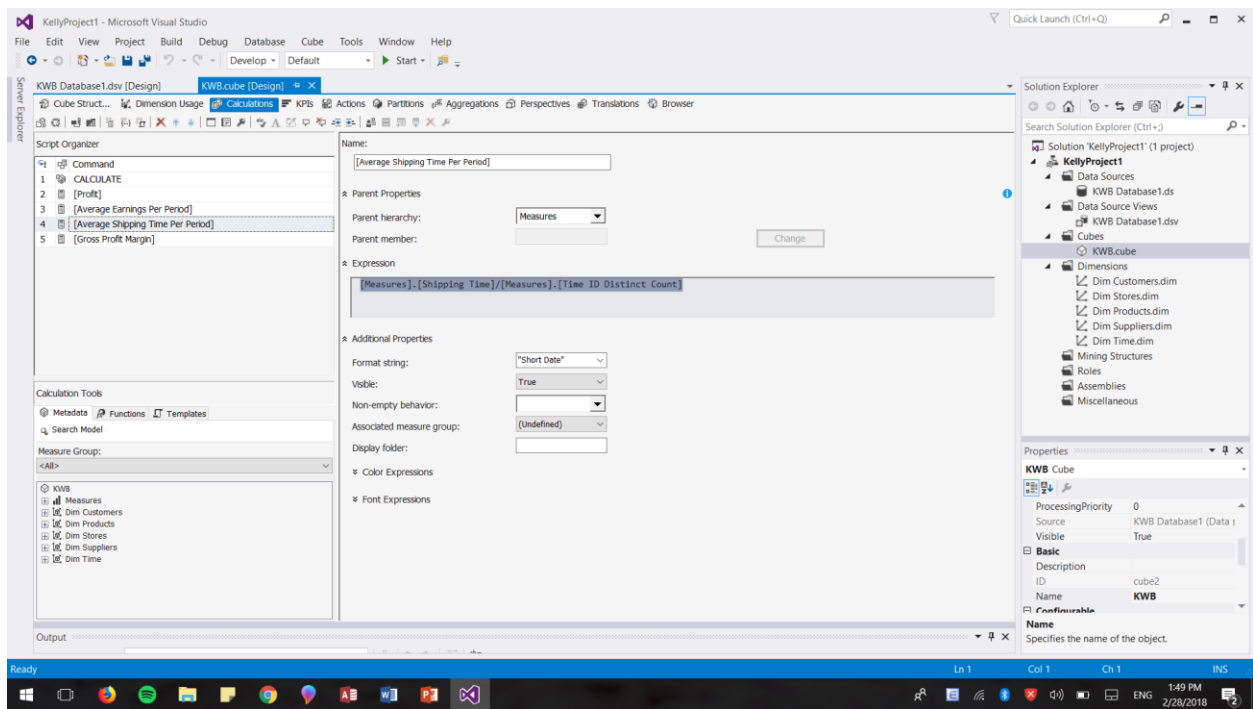


Fig 10: Profit Calculation in Microsoft SQL Server

## CALCULATION #3: GROSS PROFIT MARGIN

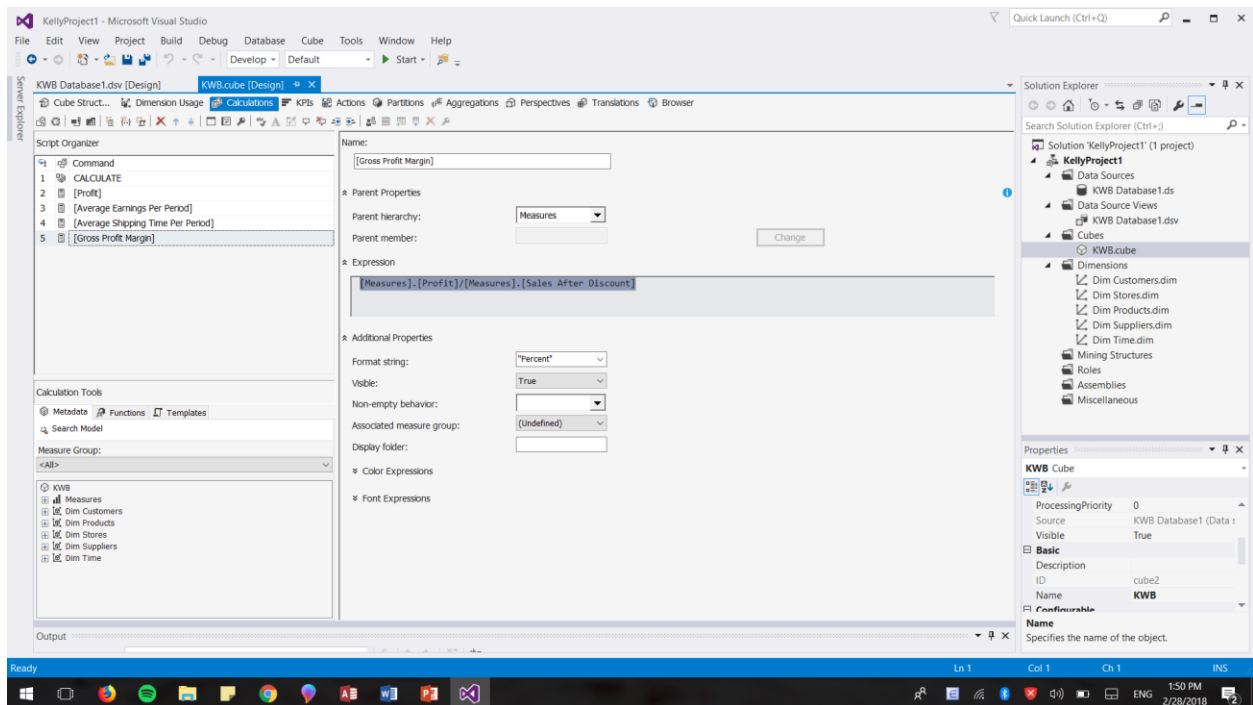


Fig 11: Gross Profit Margin Calculation in Microsoft SQL Server

## CLOSING THOUGHTS

### WHAT WENT WELL?

- It was good that I started the project early. There were so many errors and experiments during the process. I had to rebuild the cube a lot of time and having the extra time to do this was helpful.
- Expanding the database with more data was easy, but time consuming. Also, the new tool I learned in Access called Expression Builder made it easier to create calculations for my fact tables.

### WHAT WENT WRONG?

- I had a connection problem in when I created a data source view. It gave me an error that the log in user is not authorized to connect to the data source and server. This was puzzled since I followed every steps in the tutorial and never had this issue before. After trying different things, it turned out that if I chose the option "Inherit", instead of "Windows Authentication", the connection issue was solved.

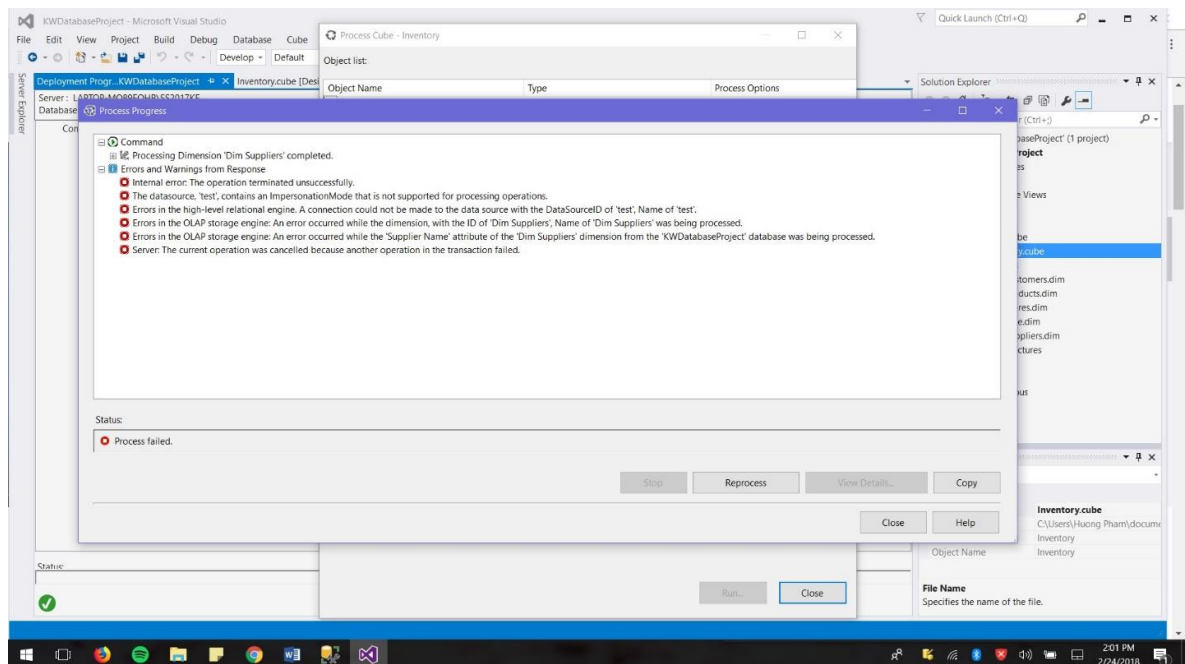


Fig 12: Error #1 - Connection Issue in Microsoft SQL Server

- I also ran into the deployment issue where I had to change the “localhost” to my server name, and choose “Use a specific Window password and username” for “Impersonation Information”.
- I encountered a problem when using the Calculation function in SQL Server. I followed a Youtube tutorial and he clicked “Start” button after he created a “Calculated Member”. I did this and saw no new calculation added to my cube. It also created an error when I tried to process the cube again. After spending a whole evening experimenting things, I still could not figure this out until Alison showed me how she did hers. We had to save the calculation first, and then process the cube again. Pressing the “Start” button or debugging the calculation was not useful.

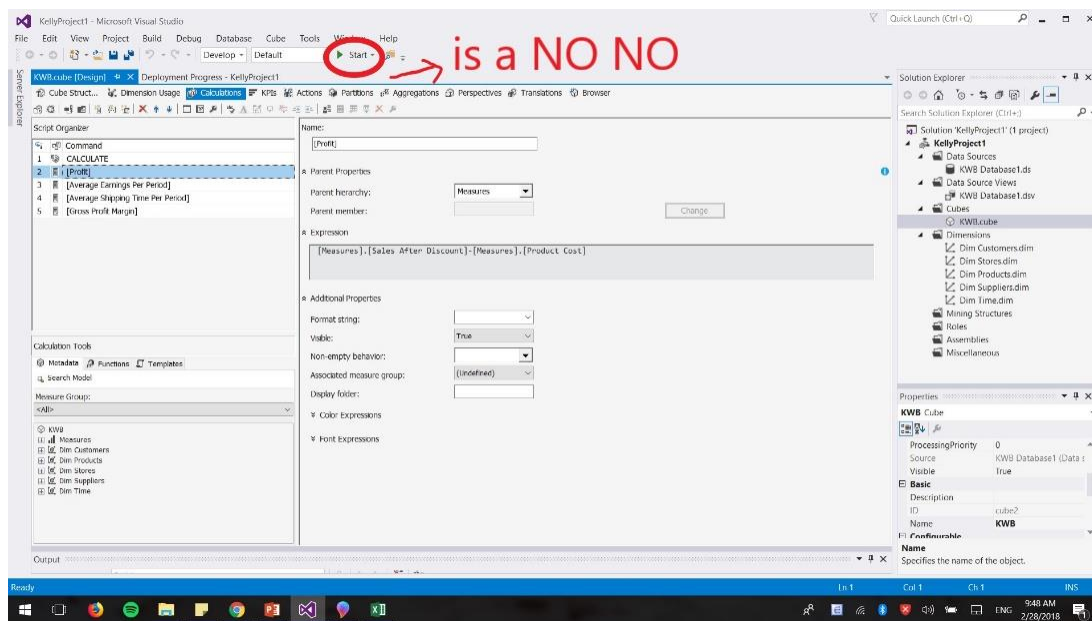


Fig 13: Error #2

- At the very last step of project, I encountered a loading error below

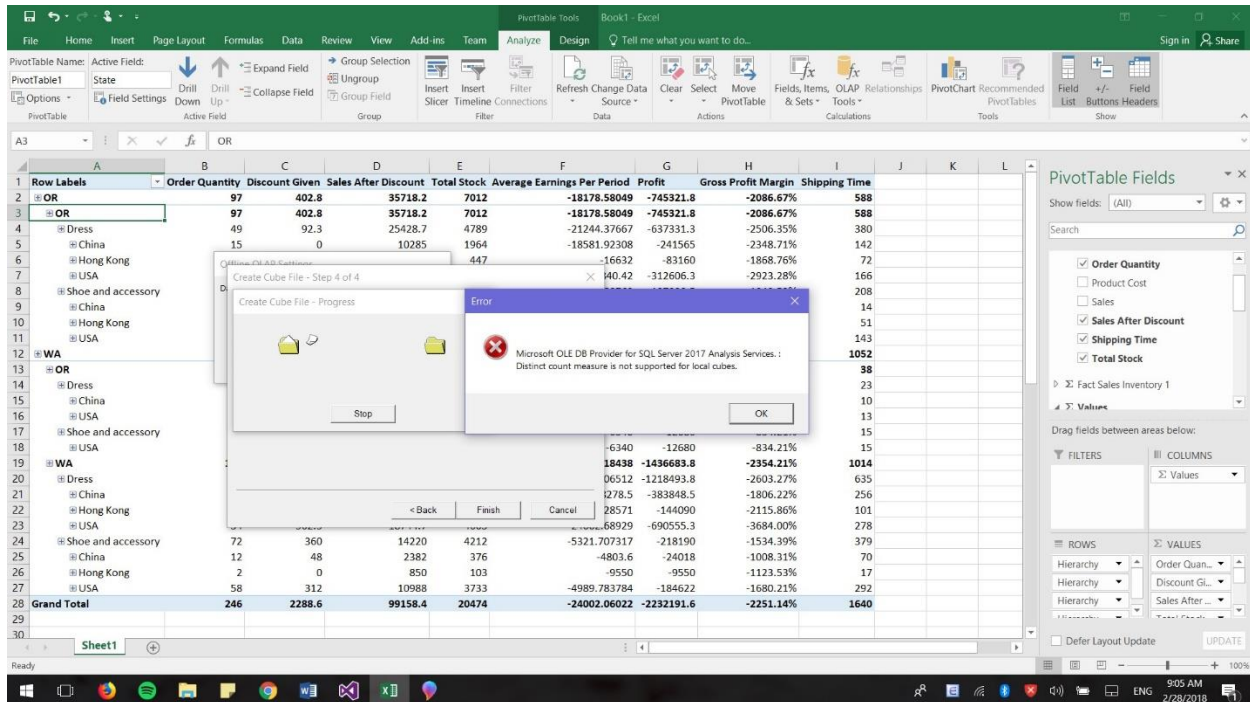


Fig 14: Error #3

Based on what I found on Google, the offline cubes do not support distinct count measures because cubes that include distinct count measures are non-additive. It seems that there is no other approach to work around this issue since it is the default setting in SSAS. The first time when I imported the cube to Excel and unchecked "Time ID Distinct Count", it also removed two measures that were based on "Time ID Distinct Count", being "Average Shipping Time Per Period" and "Average Earning Per Period". However, when I imported the cube to Excel the second time and still unchecking "Time ID Distinct Count", the two measures remained in the pivot table this time. I could not recall if I did anything different the second time, but it worked.

#### IMPROVEMENTS FOR NEXT PROJECT

- For the next project, I will need to make sure that all the relationships in Access are right before importing it to SQL Server. For example, I first created a relationship between Orders table and Time Data table (OrderDate -> Time ID). When I realized this mistake, it was time consuming go back to Access to fix it and start the whole cube again. Luckily,

while doing this, I spotted another error in the data entry as I entered the wrong state for my Stores table. Another note for future project is ensuring that the data entry is right. Quality check is necessary for every step.

- I would want to spend more time on learning how to write the SQL query in MDX for the next project. I saw a couple tutorials and followed them. However, I did not quite figure out how to make it work for my project yet. I would need more time to do this.
- I also have a Rentals table associated with the Orders table. I did not quite figure the logic to incorporate this into my star schema and have the rental orders in one cube. I know that I would have to spend a lot of time on learning how to create the measures. Therefore, I decided to not include the Rentals table, and spend more time on the measures. For next project, I would like to spend more time on building the star schema with Rentals table because it is an important aspect of this business. I think it would be helpful to drill down to the profits from the rental service.