

**Development of Business Intelligence pipeline and data-warehouse for
Sales Recommendations**
-Saurabh Sharma

INDEX:

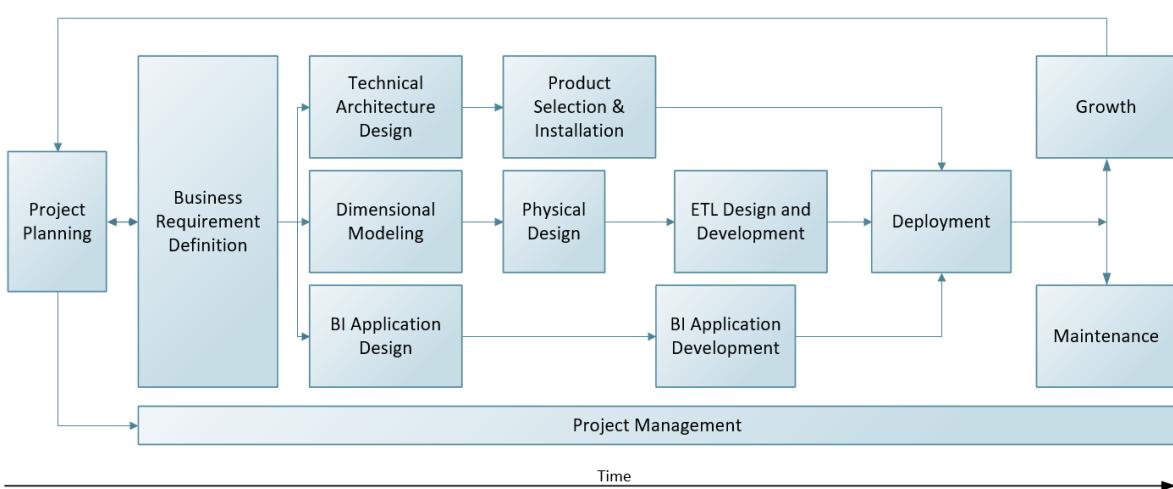
1. Outline
2. BI Project Lifecycle (Kimball)
3. Project Planning
4. Scoping
5. Dimensional Modelling
6. Process Flow for SSIS
7. Data Visualization using Tableau
8. Sales Recommendations

Outline:

In this project we used Kimball Lifecycle to design an end to end BI pipeline and data warehouse for sales recommendation to a retail chain who sale the products via various channel such as through departmental stores, online, boutiques etc. The retail chain sales goods across categories such as accessories, children's apparel, men's apparel and women's apparel. Our goal is to identify the areas where the business is doing good or bad in terms of sales and also the ways in which sales can be improved for the retail chain.

We used Kimball's BI Project lifecycle for designing our BI System.

BI Project Lifecycle (Kimball)



The steps included while designing the BI system were as follows:

1. Designing a star schema for the data set
 2. Staging of source data into staging tables
 3. Loading of all the dimension tables in star schema from the staging tables;
 4. Loading of all the fact tables in star schema from the staging tables
 5. Visualization of data get business insights through dashboard using Tableau.
-
- The entire ETL Process for done in a SSIS (SQL Server Integration Services) package using MS Visual Studio.
 - Queries for creating the schemas were executed in SSMS (SQL Server Management Studio)

Project Planning:

- (1) The first step in any BI project is to plan the project. Traditional project management involves Scope, Schedule and Budget.

SIMPLE PROJECT PLAN TEMPLATE

[Try Smartsheet for FREE](#)

PROJECT NAME	BI Project			PROJECT MANAGER	Saurabh Sharma			
PROJECT DELIVERABLE	End to end BI system							
SCOPE STATEMENT	Designing a BI system to answer business questions and give recommendations							
START DATE	03/30/2020	END DATE	06/04/2020	OVERALL PROGRESS	4%			
AT RISK	TASK NAME	ASSIGNED TO	START DATE	END DATE	Time In hours	STATUS		
	Project Planning	Saurabh	03/30/2020	04/16/2020	4	Complete		
	Dimensional Data Model ERD	Saurabh	04/16/2020	04/23/2020	5	To be started		
	<u>SSIS Package-Staging Loads</u>	Saurabh	04/23/2020	04/30/2020	5	To be started		
	SSIS Package-Dimension Loads	Saurabh	04/30/2020	05/07/2020	5	To be started		
	SSIS Package-Fact Loads	Saurabh	05/07/2020	05/14/2020	8	To be started		
	Final SSIS package and SQL Schema	Saurabh	05/14/2020	05/24/2020	12	To be started		
	Final Tableau Dashboards	Saurabh	05/24/2020	06/04/2020	12	To be started		
	Presentation: Business Question Analysis and Recommendations	Saurabh	06/04/2020	06/04/2020	12	To be started		

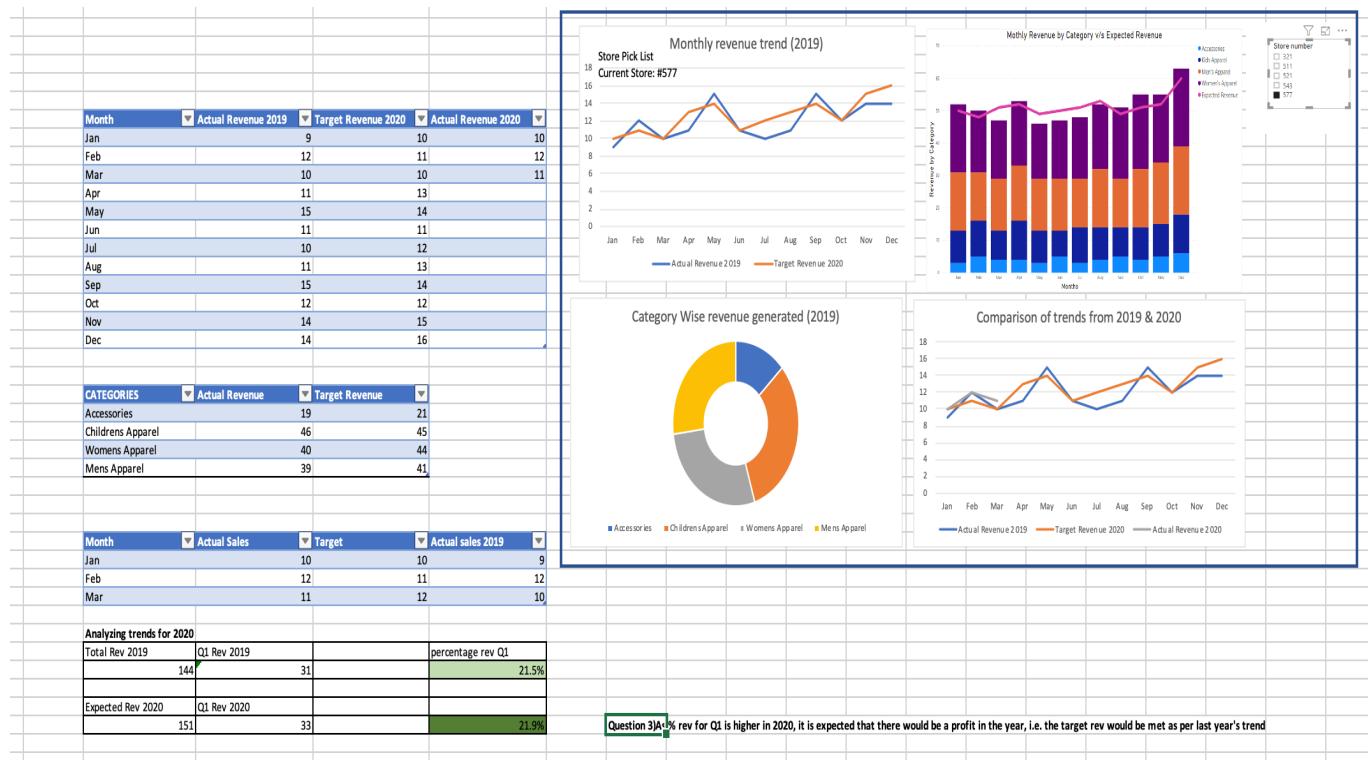
Scoping

(1) First I analysed the source data to understand what data we have to work with and what it means.

Data Profile Template					
	List each individual measure on a row			List each individual product on a row, filling in the values for each column (from your SELECT statement)	
Measure (Metric / KPI)	Product ID	Product Category	Product Type	Product	Channel Category
Price	1	Accessories	Cosmetics	Blush	Direct
WholesalePrice	2	Accessories	Cosmetics	Lipstick	Direct
Cost	3	Accessories	Jewelry	Amethyst Necklace	Direct
SalesQuantity	4	Accessories	Jewelry	Silver Ring, Gallic pattern	Direct
SalesAmount	5	Accessories	Jewelry	Silver Ring, plain	Direct
Weight	6	Children's Apparel	Baby 0-2	Blue Onesie Pajamas	Direct
	7	Children's Apparel	Baby 0-2	Thomas the Train Pajamas	Direct
	8	Children's Apparel	Baby 0-2	White Socks, 6 pack	Direct
	9	Children's Apparel	Kids 2-6	Girl's Dress	Direct
	10	Children's Apparel	Kids 2-6	Thomas the Train Pajamas	Indirect
	11	Children's Apparel	Kids 7-14	Girl's Dress	Indirect
	12	Children's Apparel	Kids 7-14	Spider-man T-shirt	Indirect
	13	Women's Apparel	Women's Casual	Blouse	Indirect
	14	Women's Apparel	Women's Casual	Skirt	
	15	Women's Apparel	Women's Eveningwear	Dressing Gown	
	16	Women's Apparel	Women's Eveningwear	Pajamas	
	17	Women's Apparel	Women's Formal	Dress	
	18	Women's Apparel	Women's Formal	Shoes, High-heel	
	19	Women's Apparel	Women's Formal	Strapless Dress	
	20	Men's Apparel	Men's Casual	Captain America T-Shirt	
	21	Men's Apparel	Men's Casual	Denim Jeans	
	22	Men's Apparel	Men's Casual	T-shirt	
	23	Men's Apparel	Men's Formal	Buttondown Shirt	
	24	Men's Apparel	Men's Formal	Formal Pants	

(2) Example of a graphic(s) we might build as part of your final BI System solution.

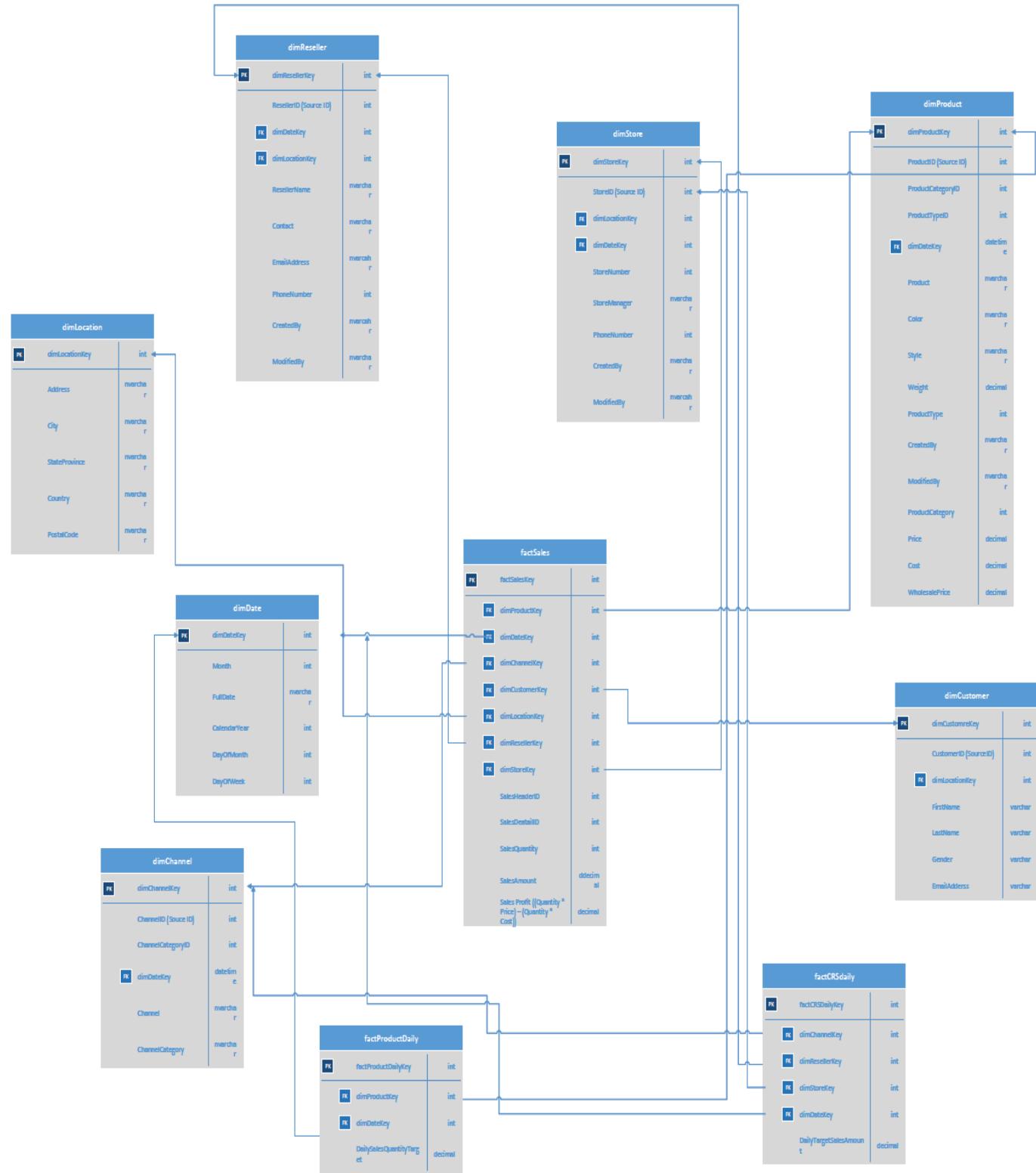
Mock-up sketches:



Dimensional Modelling:

Next I designed a data model ERD. We designed a star schema model and followed Kimball's guidelines while designing a BI system. On the next page you can see our proposed data model.

Dimensional Model ERD



Dimensional Data Model ERD

Identification of the analysis metrics:

SrNo.	Table Name	Grain	Explanation
1.	dimReseller	Reseller Name	This grain gives the data related to reseller. We have included all the data related to reseller in a single grain.
2.	dimLocation	Locations	dimLocation is the lowest grain to store all data related to location for customers, resellers and stores. It has been used as a snowflake table as it is a good practice to consider snowflaking when a dimension is used by multiple other dimensions and/or with the fact table as well.
3.	dimDate	Dates	dimDate is the lowest grain to store all data related to dates, month, year and it will allow us to map the data for particular dates. It has been used as a snowflake table as it is a good practice to consider snowflaking when a dimension is used by multiple other dimensions and/or with the fact table as well.
4.	dimStore	Store information	dimStore is the lowest grain to store all the data related to stores.
5.	dimProduct	Product information	dimProduct is the lowest grain to store all the data related to products and multiple hierarchical tables have been included in the single product table to give us product grain
6.	dimChannel	Channel information	This grain stores information related to channel which can be used to gain insights about channel and their categories
7.	dimCustomer	Customer information	This grain is used to store information about customers. A single grain gives us information like name, address, and contact details of a customer.
8.	factSales	Sales Details	It is a fact table to store sales related data. Each grain in the factSales table can be used to acquire information about products types, their sales, customers, resellers, channels of sale, quantity, sales amount and profit
9.	factProductDaily	Sales Target Product Wise	Each grain in this table stores information related to the targeted sales related to a particular product
10.	factCRSdaily	Sales Target (For a particular store)	Each grain in this table stores information related to the targeted sales related to particular store and the channel information.

List of Dimension Tables:

1. dimResller
2. dimLocation
3. dimDate
4. dimStore
5. dimProduct
6. dimChannel
7. dimCustomer

List of fact tables:

1. factSales
2. factProductDaily
3. factCRSdaily

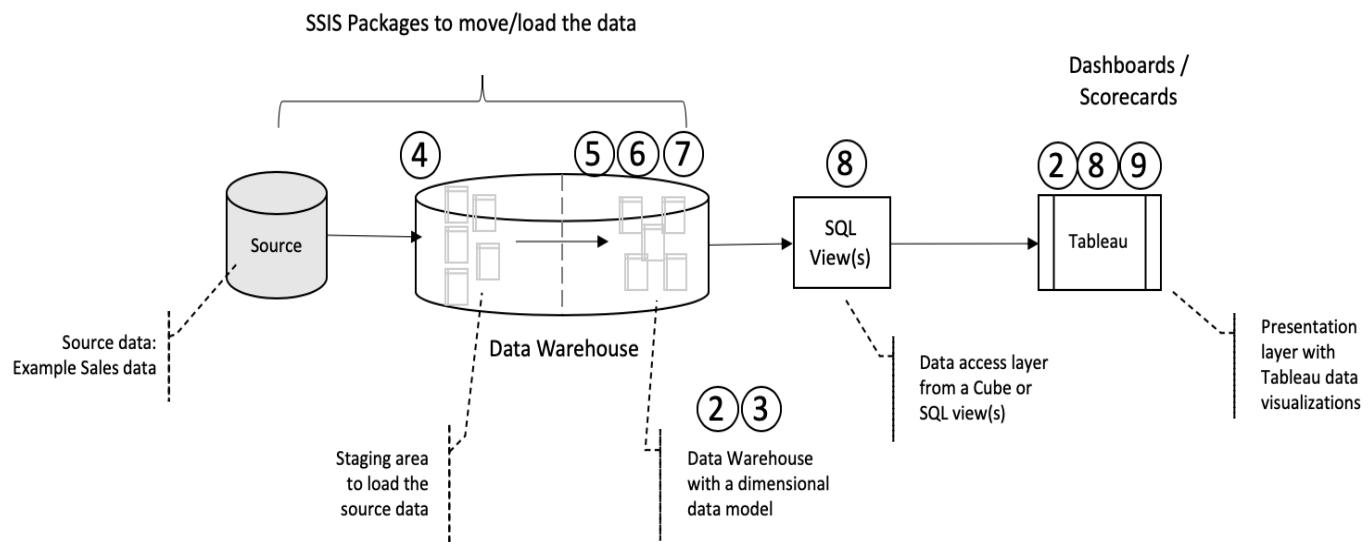
Use of Surrogate keys:

1. dimResellerKey
2. dimLocationKey
3. dimDateKey
4. dimStoreKey
5. dimProductKey
6. dimChannelKey
7. dimCustomerKey
8. factSalesKey (optional)
9. factCRSDailyKey (optional)
10. factProductDailyKey (optional)

Inclusion of Natural (source system) Keys:

1. ResellerID
2. StoreID
3. ProductID
4. ChannelID
5. CustomerID

Our flow according to the schedule:



Process Flow for SSIS:

I performed the staging and loading of our dimension and fact tables in the following manner:

Staging:

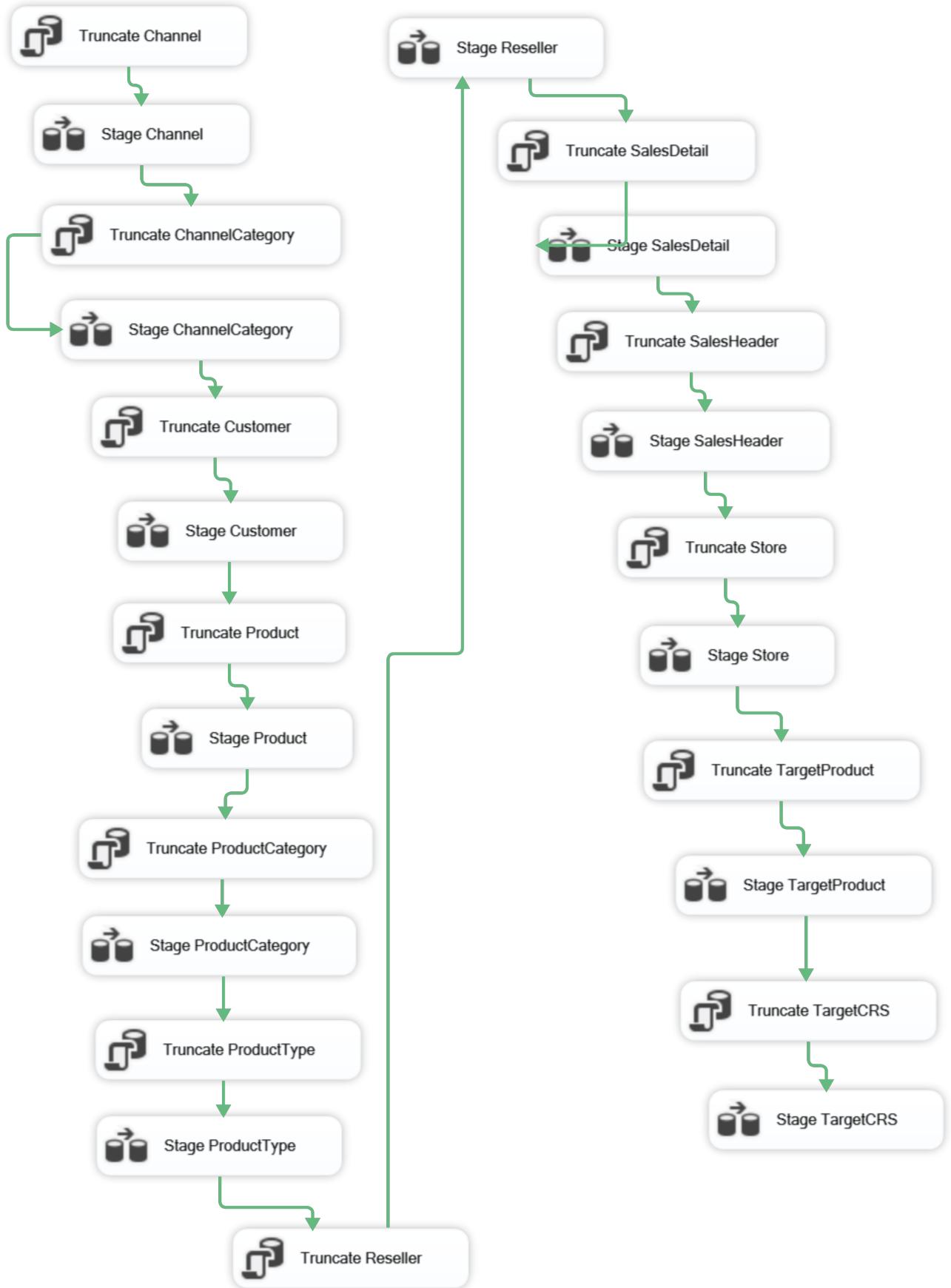
First staging process was done. In this step we load our tables from source data to staging area (data warehouse) for further processing of data.

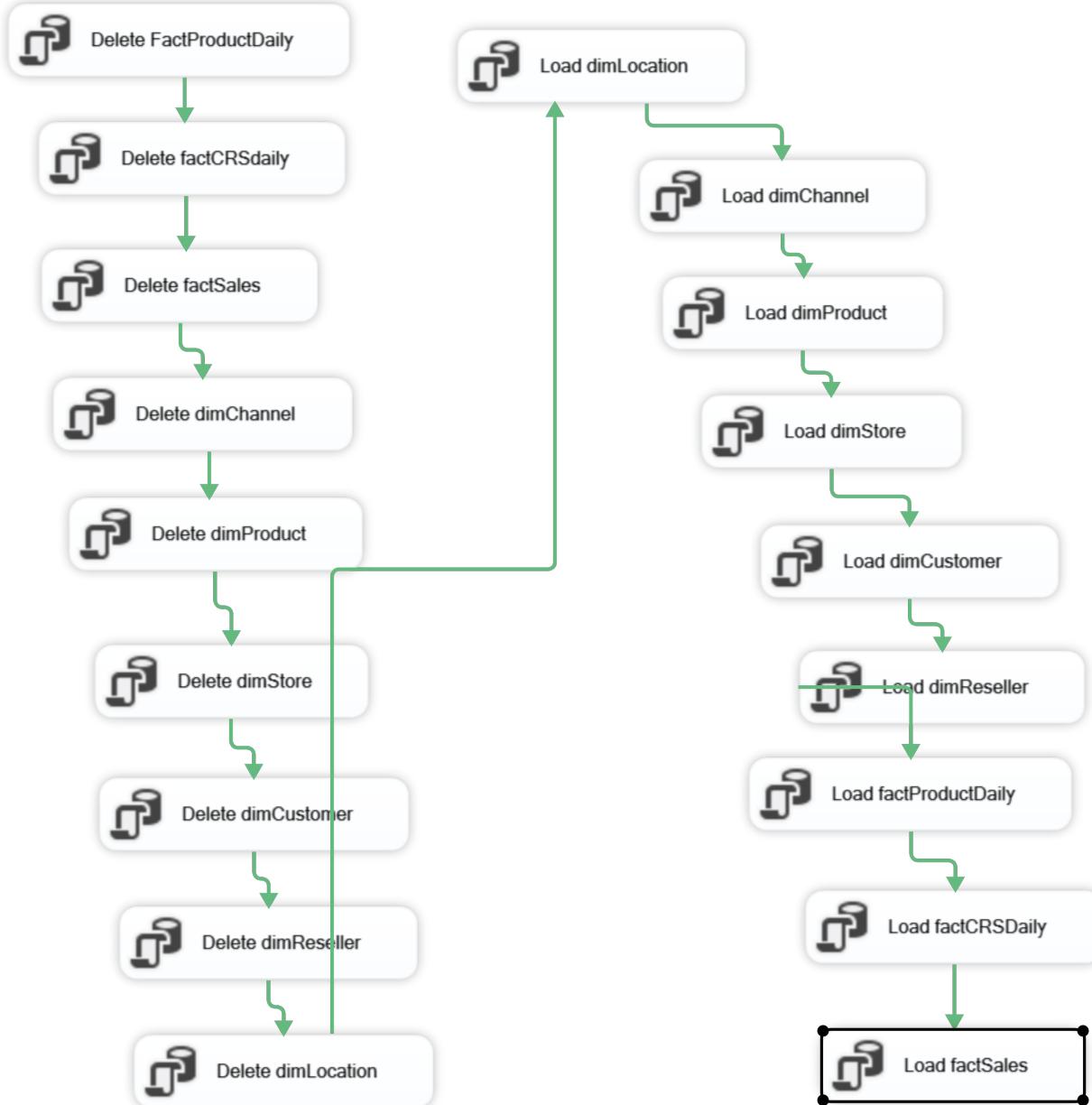
Dimension Load:

After staging, I loaded our dimensional tables into data-warehouse.

Fact Load:

Then I loaded our fact tables into the data warehouse.





Data Visualization using Tableau:

I first created views in SSMS which acted as data access layer for my tableau dashboards. Then we created the following graphs in tableau:

Unit Profit Margin

Sales Quantity vs Product

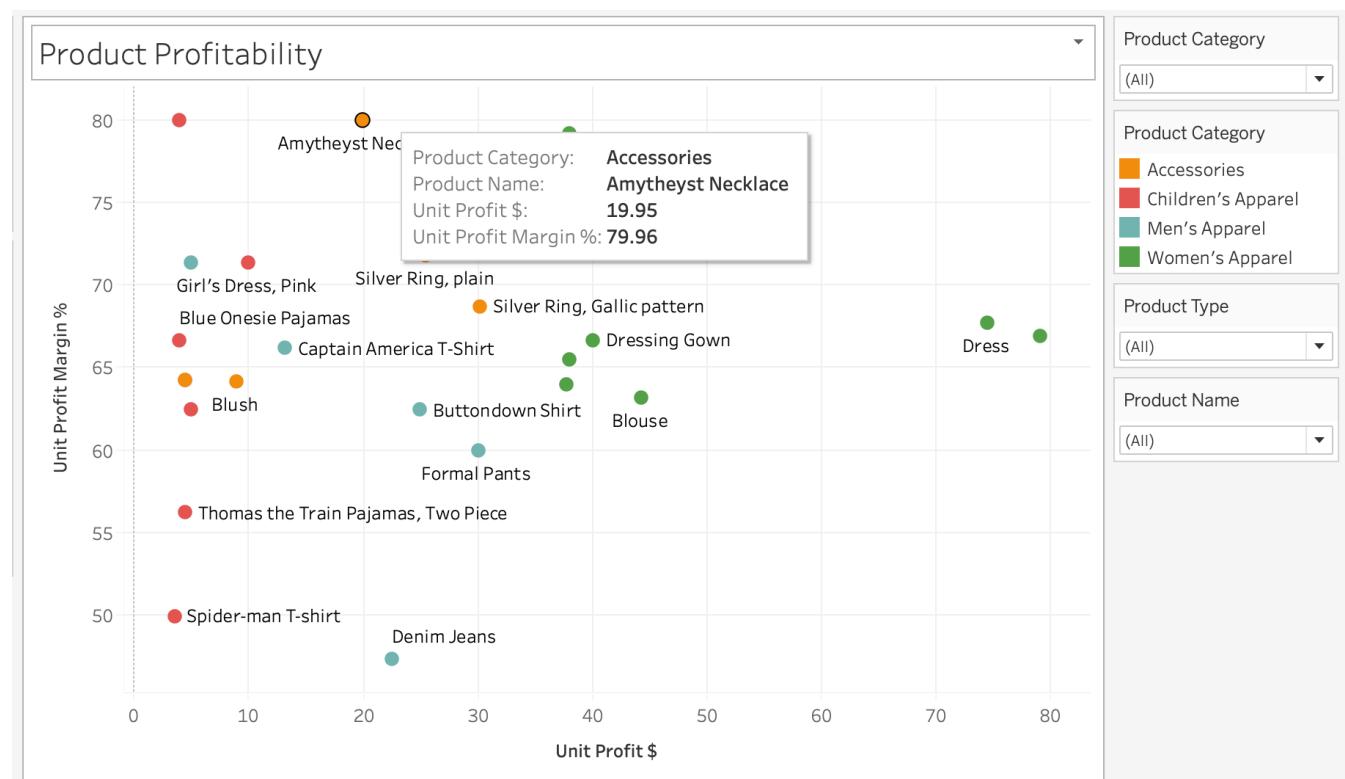
Monthly Sales Amount

Monthly Sales Profit

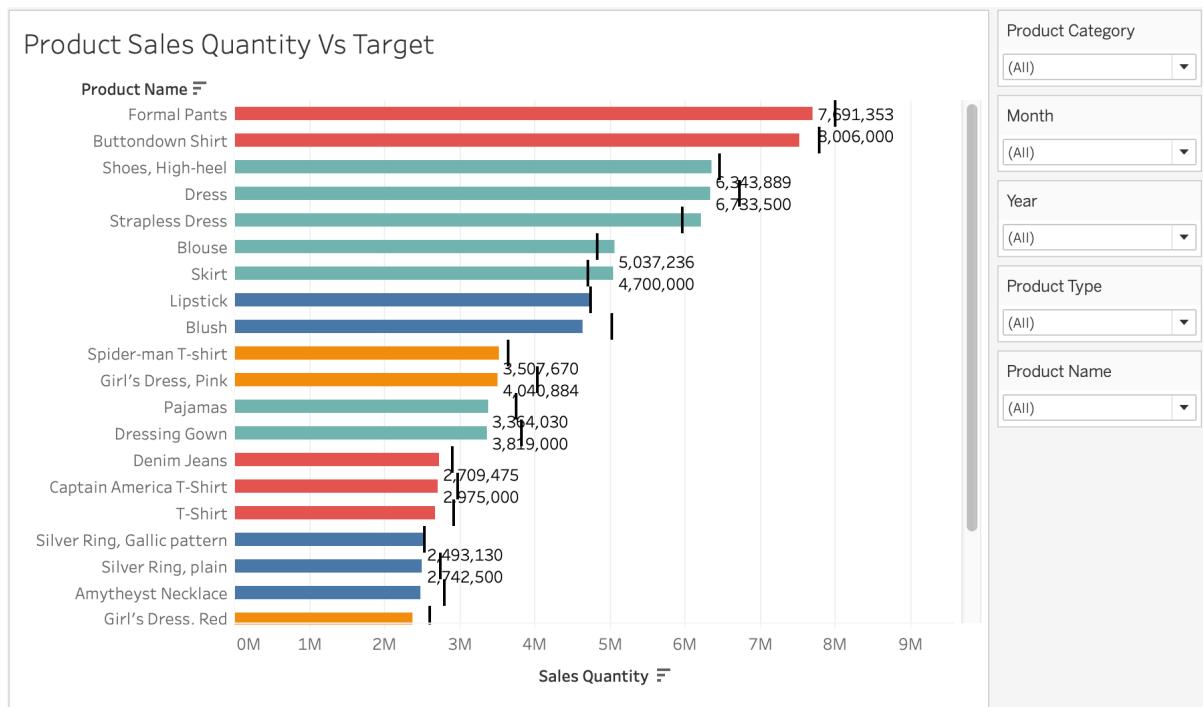
Sales Quantity By Month

Finally I created a tableau dashboard which included some of the above mentioned graphs. The dashboard also included global filters as well as interactions.

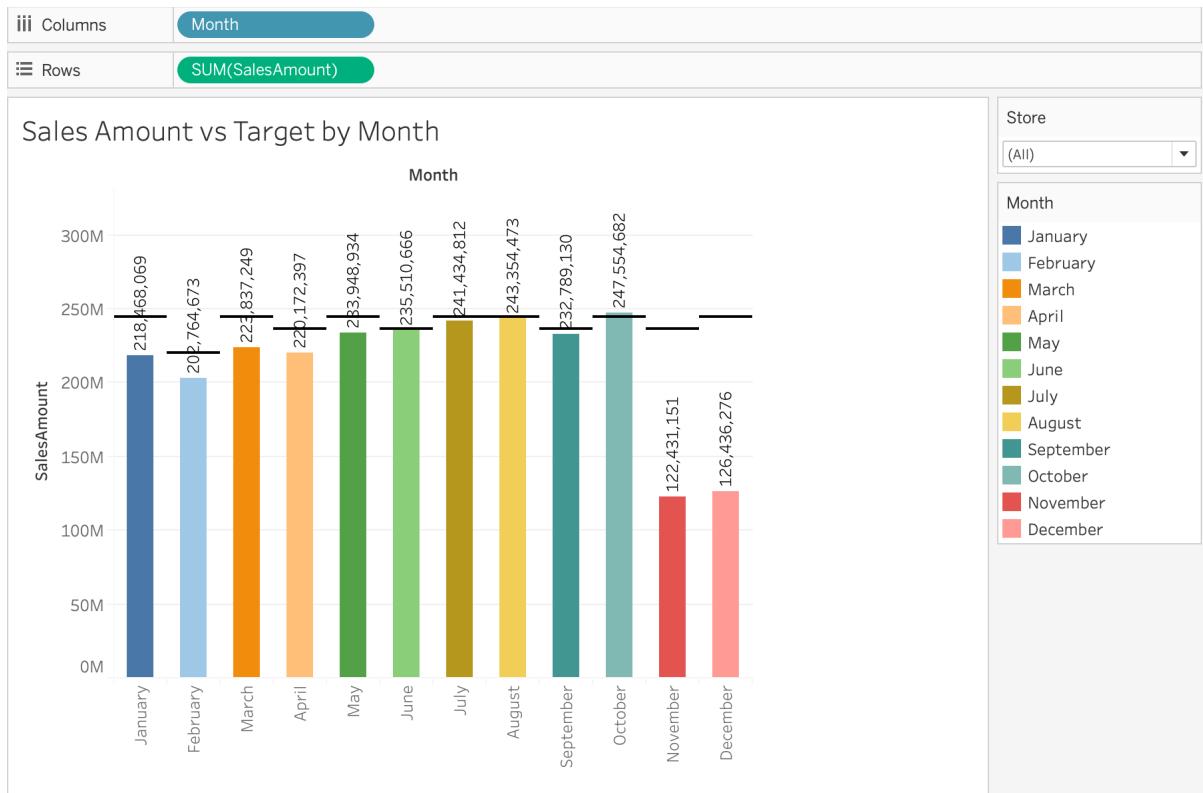
1.



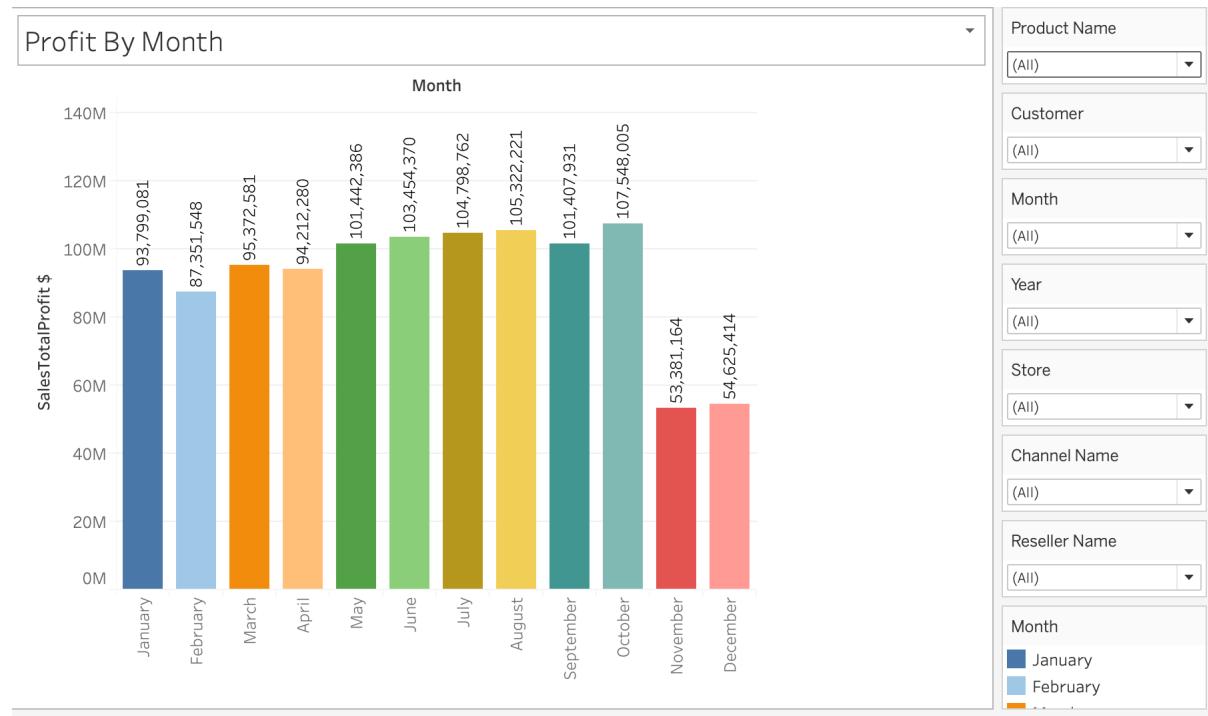
2.



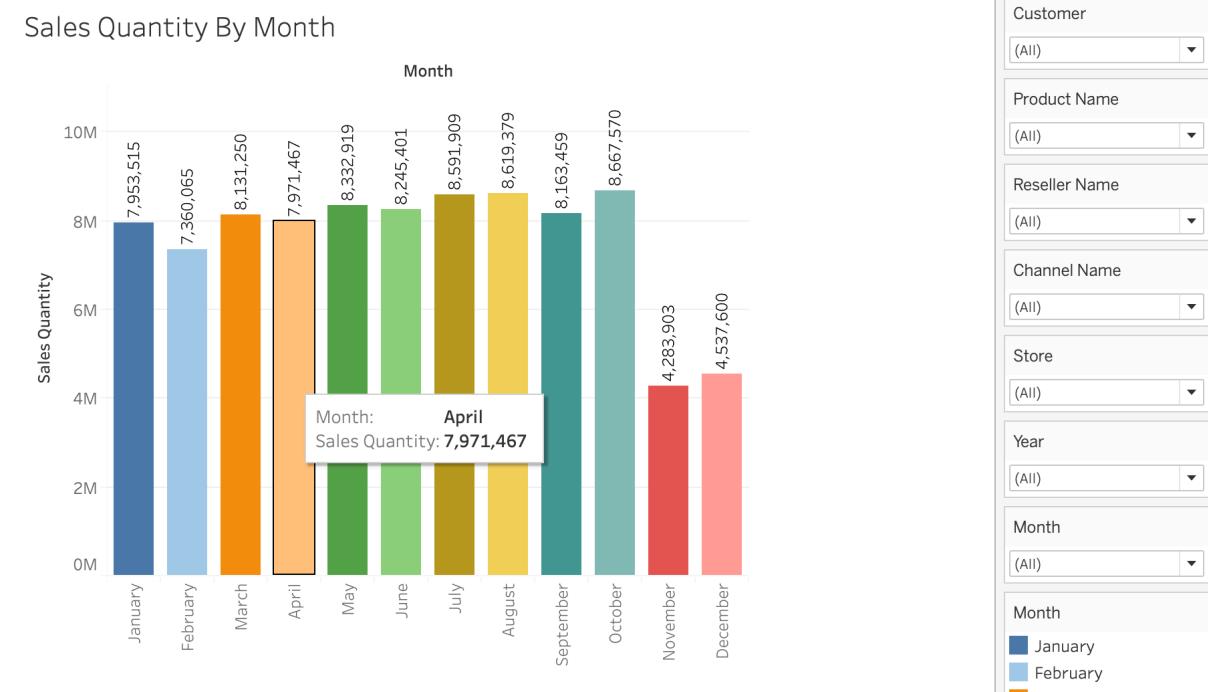
3.



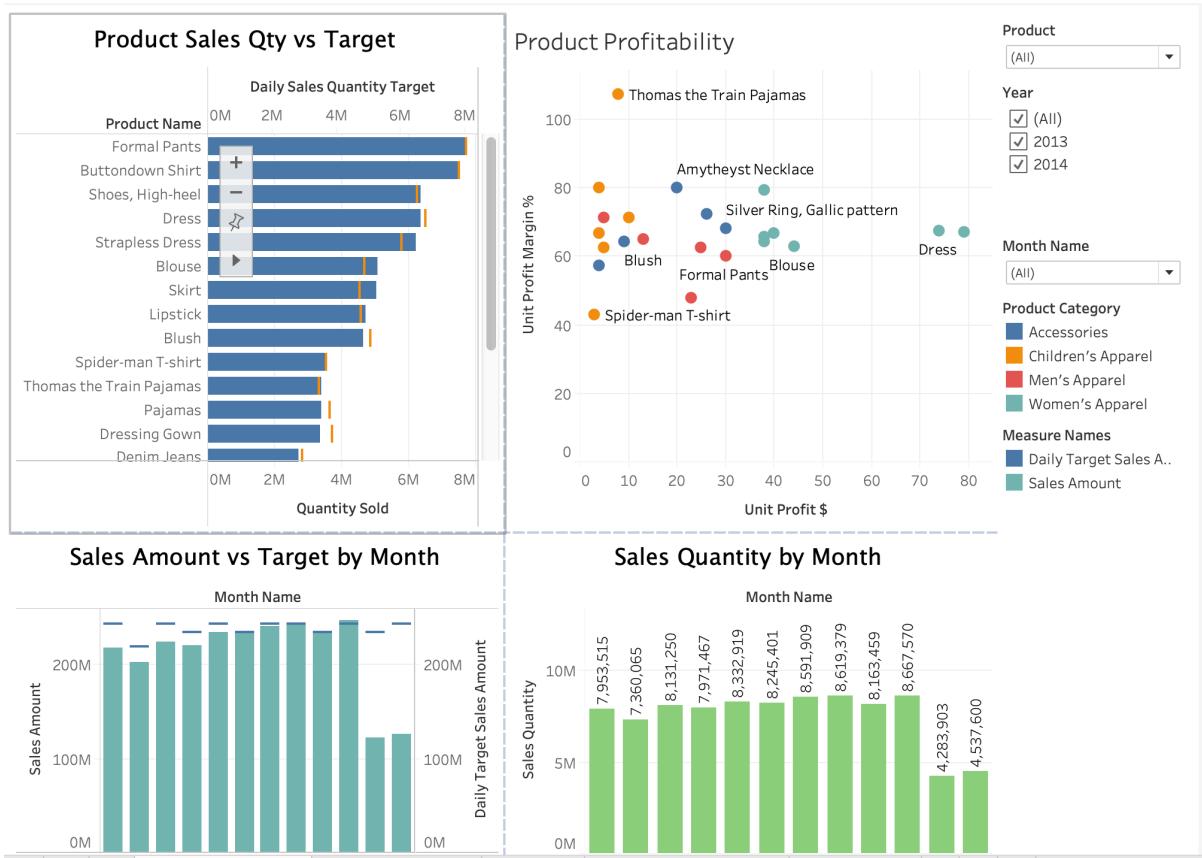
4.



5.

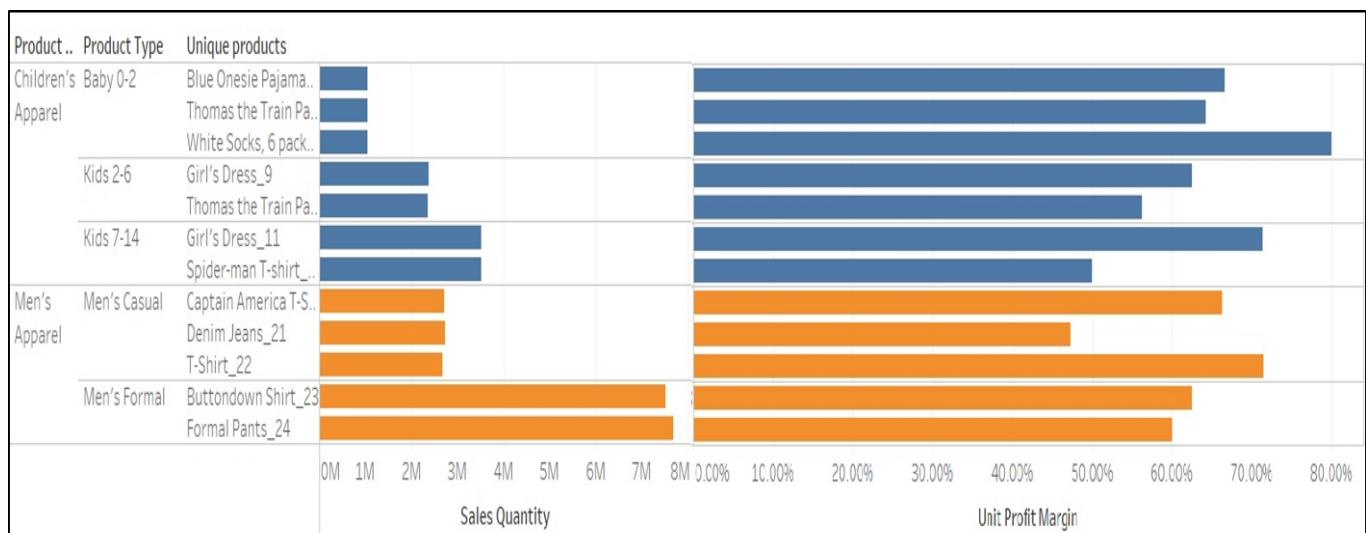


Dashboard:



Recommendations related to marketing strategy for men's and children's apparel:

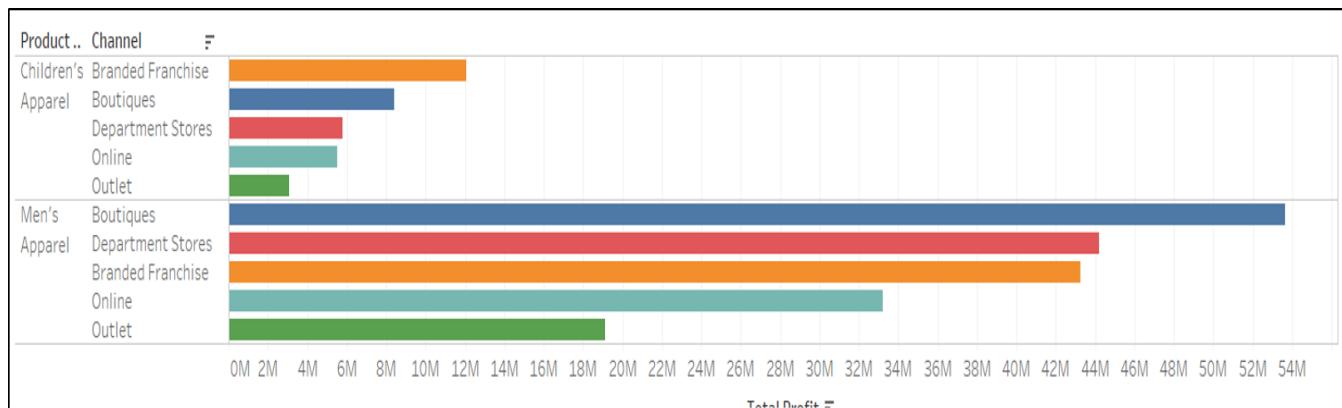
Featured Products



- Most selling products – Men's formals with >60% profit margin – Featured products
- Baby products, Captain America T-shirts, T-shirts – Low Sales but High Profit Margin
- Increase sales of such products – Featured products

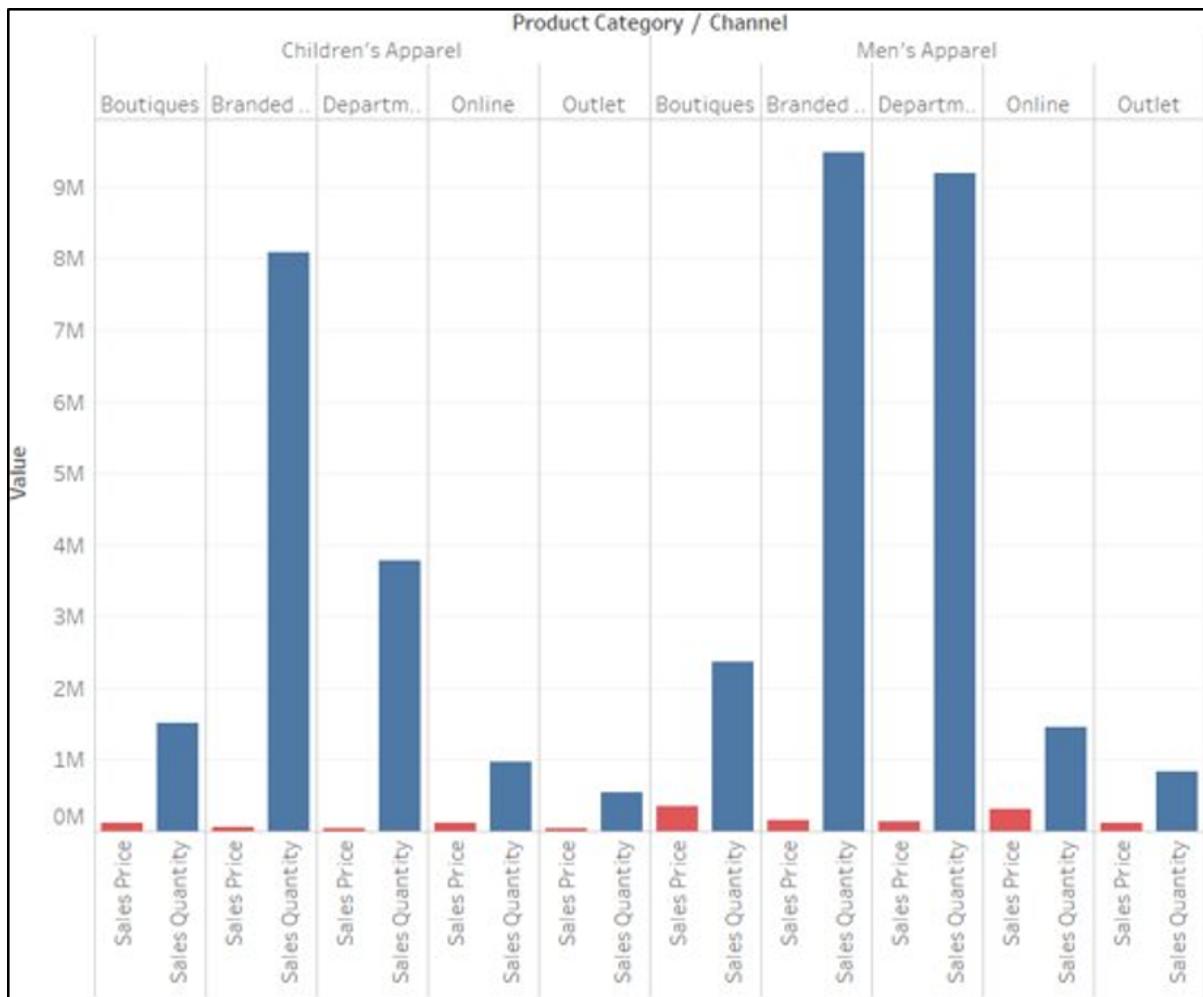
Maximizing Product Category profits

Recommendation 1:



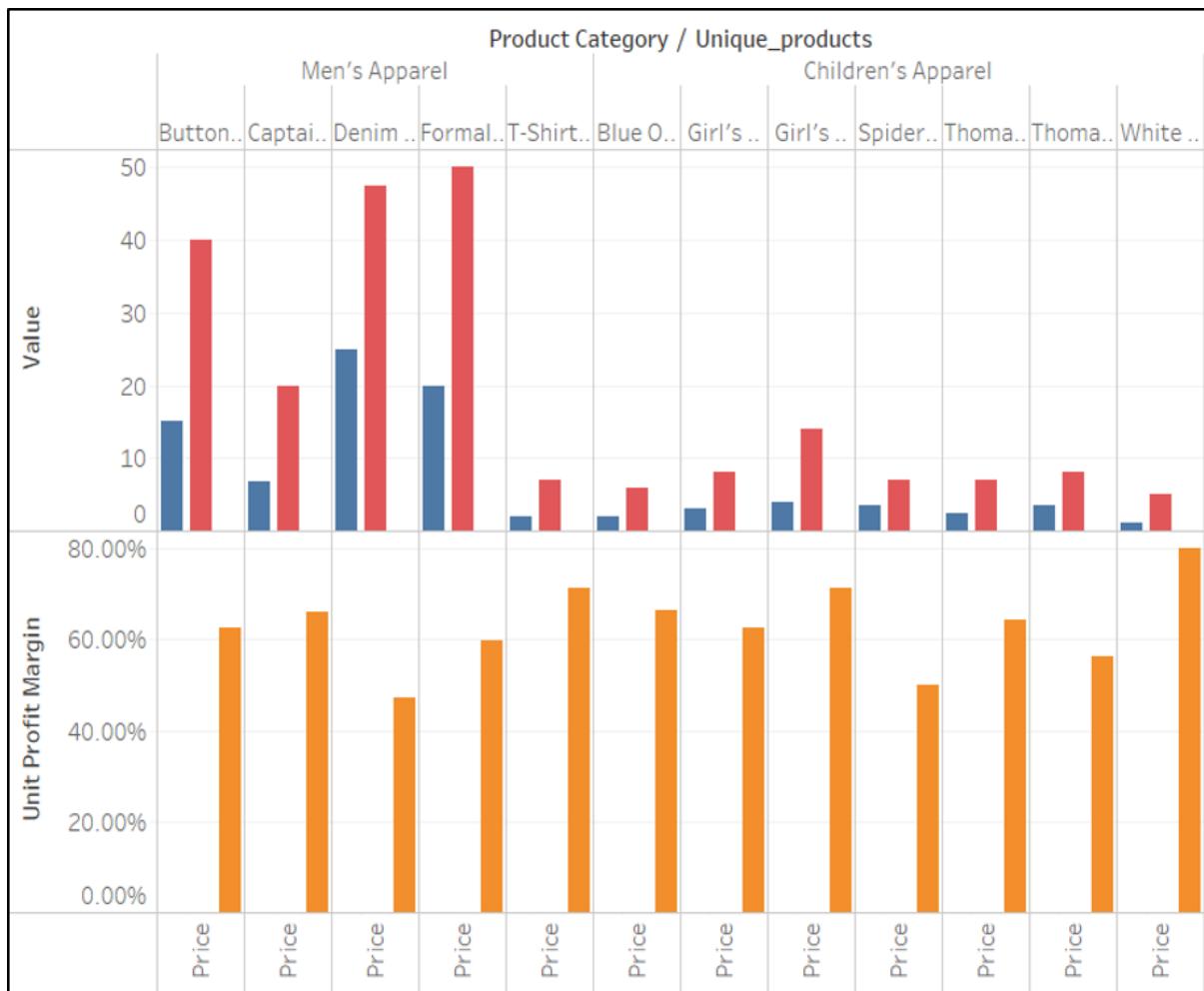
- Lowest profits through Online and Outlet Channels for both Product Categories
- Dive deeper to find if low profits are due to low sales, low selling price (more discount)

Recommendation 2:



- Comparing the sales price and sales quantity across all channels, sales quantity is low for these channels across both product categories
- Channels should increase sales through optimized schemes, advertising, increasing customer base
- Online channel can leverage product recommendation systems

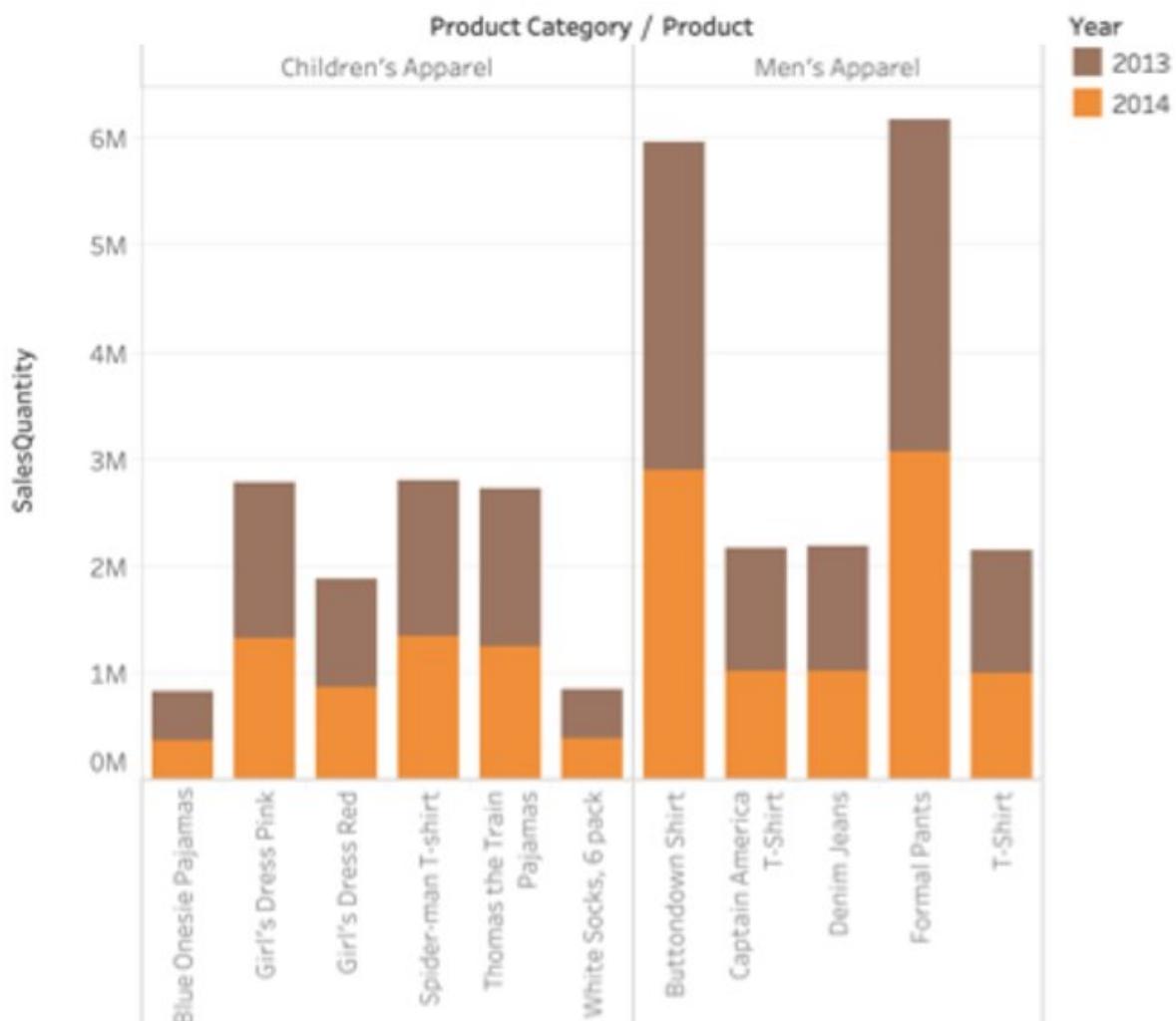
Recommendation 3:



- Denim Jeans & Spider-man T-shirt have lowest profit margins
- Decrease cost by:
 - 1) Better deals from supplier
 - 2) Cutdown production waste
- T-Shirt & White socks have highest profit margins
- Promote sales of these products

Assessment of trends with respect to Products

Trends in Products for Children and Men's Apparel



- Trends across many different products remains almost the same across the years 2013 and 2014.
- We observe that with respect to sales quantity Children's apparel do less business than Men's Apparel.

Maximizing the sales of Men's Apparel and Children's Apparel based on Channels



For maximizing Men's Apparel in Georgia Branded Franchise and in Mississippi Department stores channels should be leveraged the most.

Maximizing the sales of Men's Apparel and Children's Apparel based on Channels:

