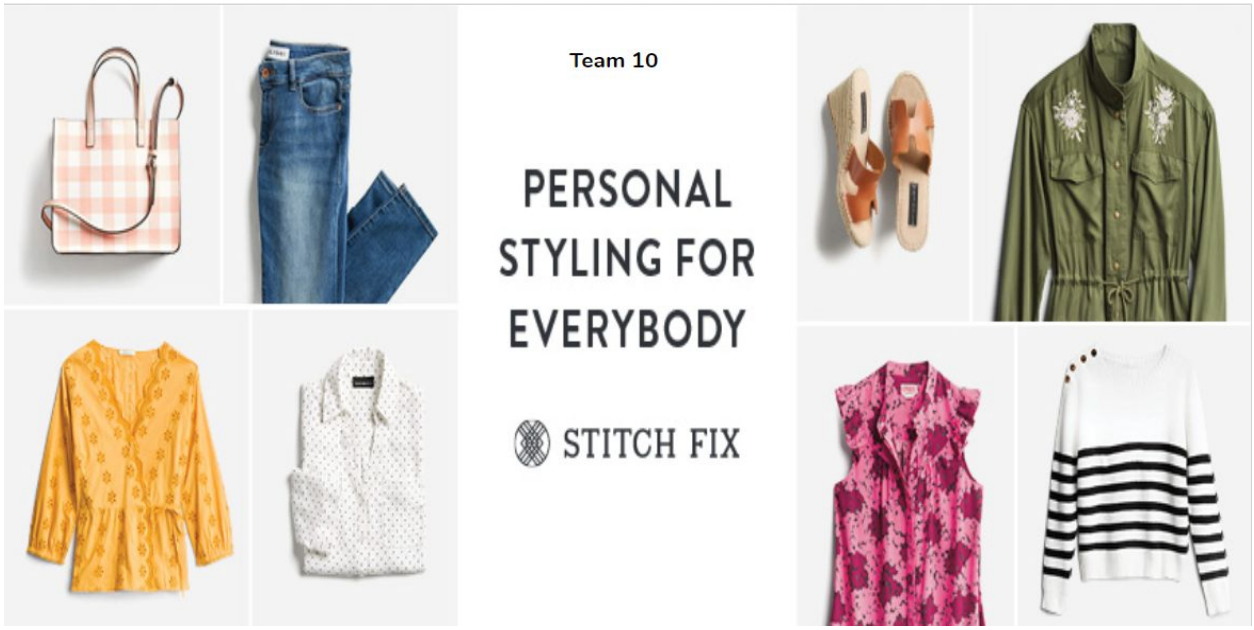


Stitch Fix

Team 10

Sayali Bhagat
Anuja Suryawanshi
Pooja Dharmadhikari

Overview

Stitch Fix is a personal styling service that sends individually picked clothing and accessories items for a one-time styling fee.

The personal styling company uses predictive algorithms to look at a number of attributes to determine who will make a successful client.

How it works:

To start using Stitch Fix, clients fill in a style profile questionnaire where clients provide personal preferences on style, fit, size, and price. Clients can then select the frequency, electing to either auto-ship at a set cadence or request the service on-demand as needed. Stitch Fix then sends each client a “fix”, or a box containing a personalized selection of apparel, shoes and accessories. Each fix contains five items, which she can try on in the convenience of her home. Within 3 days, she then identifies which items to keep, and returns the rest in a prepaid bag. There is a $20 styling fee attached to every box, which can be credited towards merchandise purchased. Customers who want to keep all the five items get a 25% discount on the order.

MySql Workbench:

Workbench was used to create Schema for the datasets.

Gliffy:

This was used to create the workflow of the datasets.

Tableau:

Tableau was used to create dashboards based on the data analysis.

Blenda:

This platform was used to integrate Salesforce with GCP BigQuery.

Salesforce:

Salesforce platform was used for analyzing the data and for creating dashboards based on the analysis of the data.

Dataflow based on the datasets was also built using this platform.

Google Cloud Pipeline:

The datasets form salesforce were then integrated on to the bigquery in Google cloud pipeline for other teams to access.

Goals

1. The goal of this project is to analyze the customer, sales, promotion data and to find a way to increase the sale and the manufacturing of the products accordingly.

Specifications

Datasets:

Datasets consists of Customer, Product, Review, Employee, Sales, Promotions and Social Media Influencer data.

1. Customer: customer data consists of the demographic information of the customers, the time they spend shopping on various platforms, the annual amount spent and status of the customer specifying whether they are active or inactive.
2. Product: Product dataset consist of the product specifications, price and discount applied to that product.
3. Reviews: This dataset consists of clothing id, reviews, ratings, recommendations and the category of the product.
4. Employee: This dataset consists of the demographic information of the employees along with their salary and position.
5. Sales data: Consists the particulars of sale i.e. id, shipment details, status of the package to be shipped.
6. Promotions: Contains the data about the promotions like expiry date of the promo

code, and the platform it is purchased from.

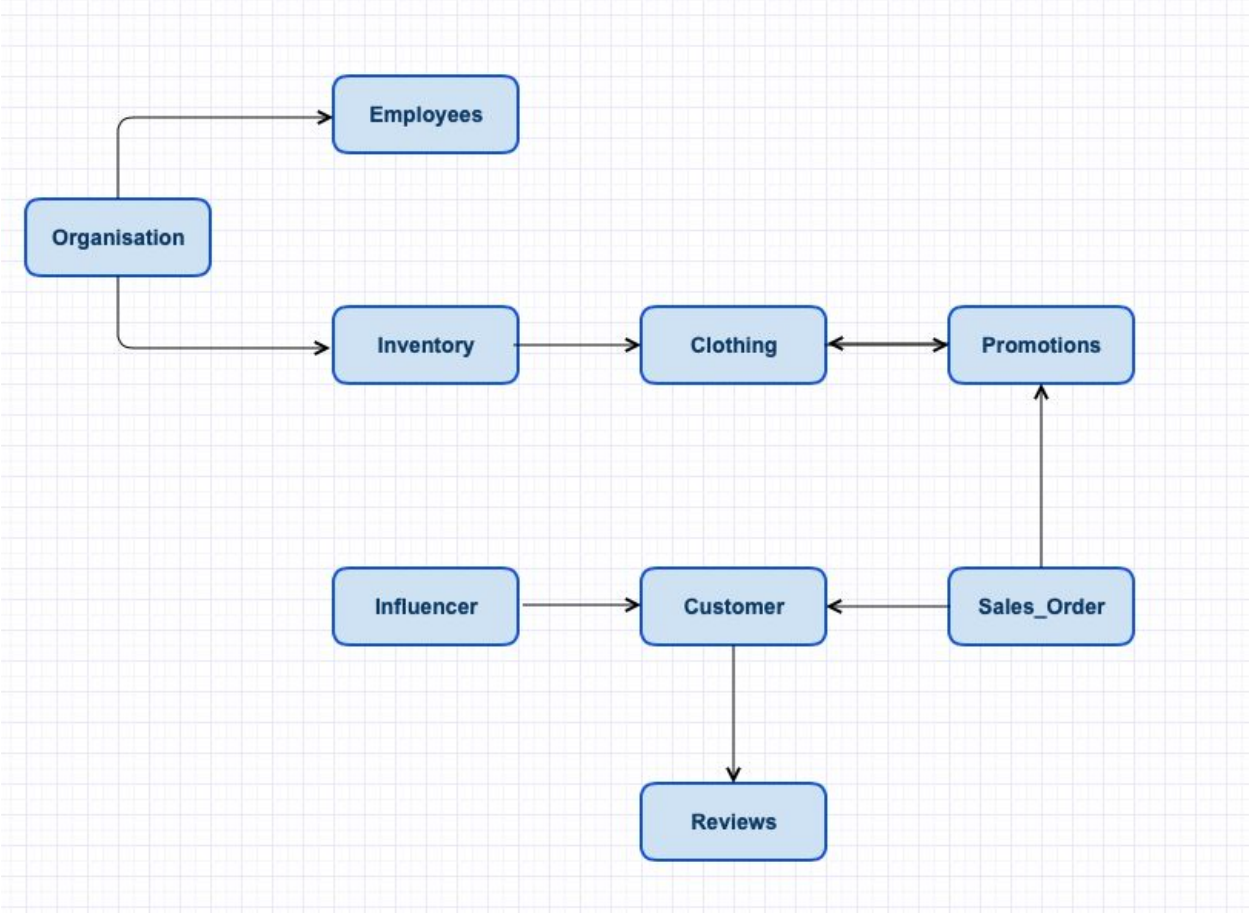
7. Social Media Influencer: This dataset consists of the social media influencer data i.e. which platform used and percentage of customers it attracted towards the product.

Milestones

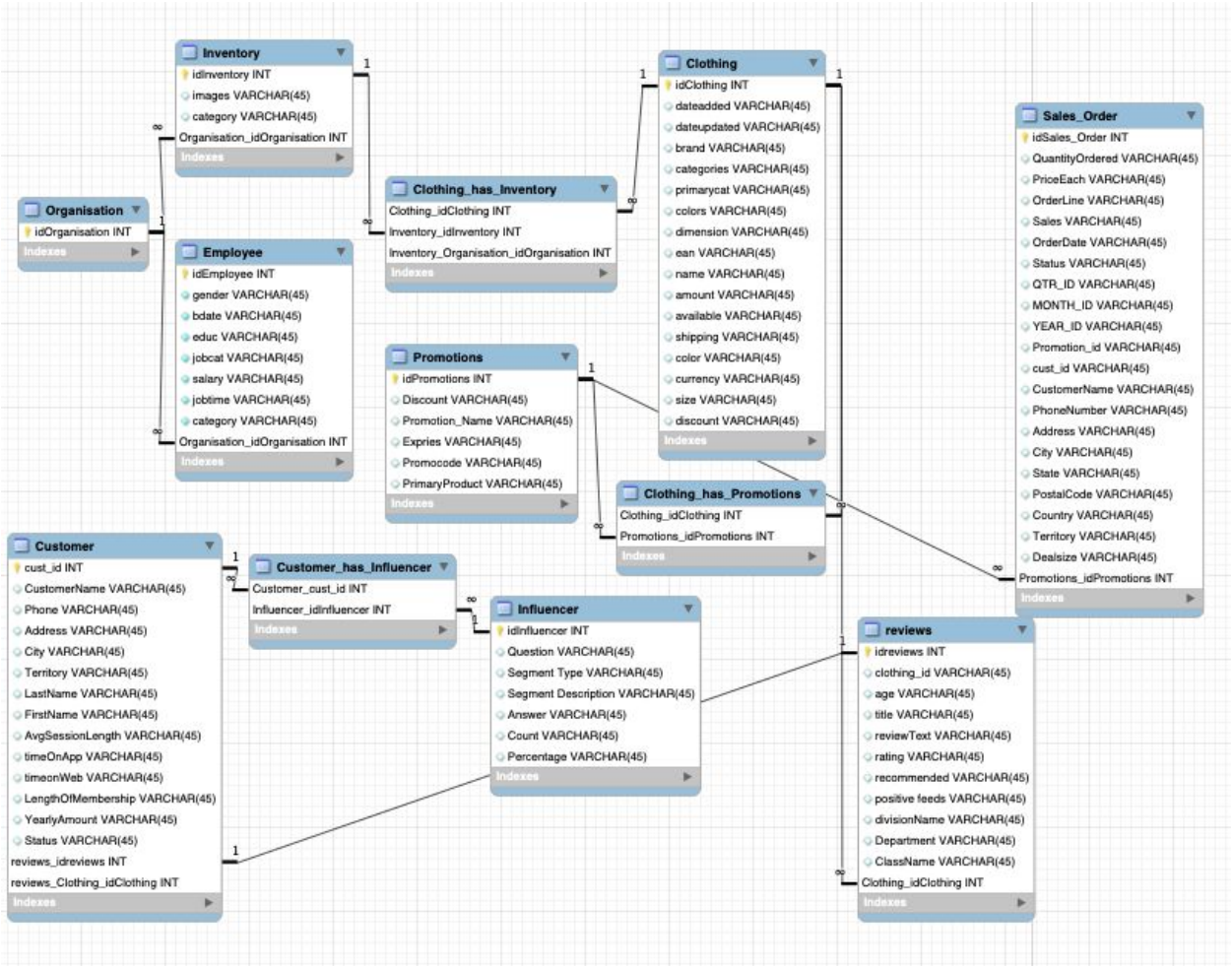
I. Schema and workflow generated using Workbench and Gliffy

Schema and workflow for the dataset was generated using MySql and Gliffy for further analysis of the data.

1. Workflow generated on Gliffy



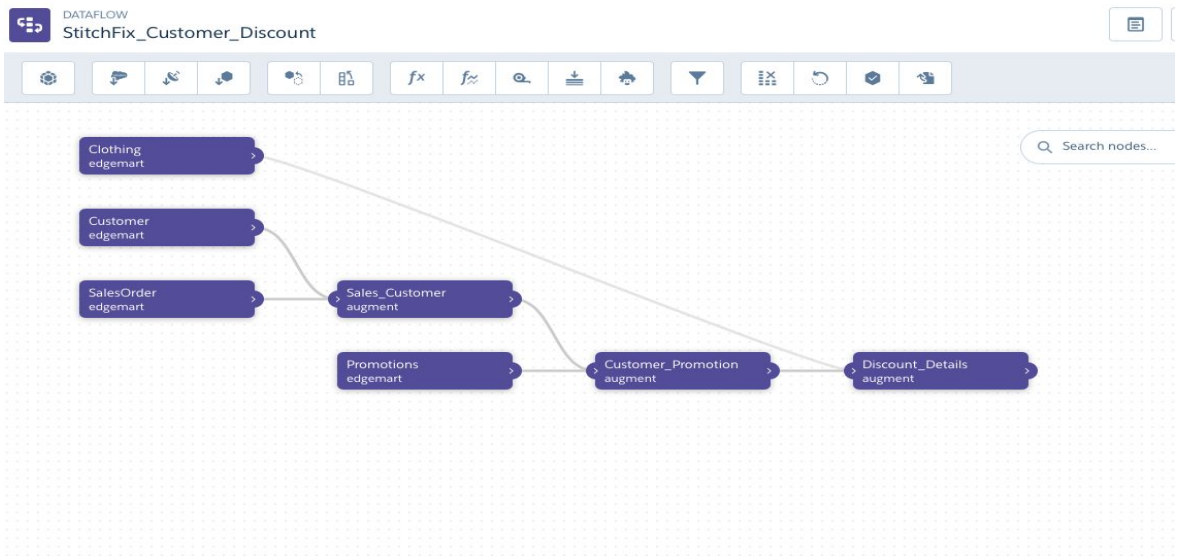
2. Schema built on MySql Workbench



II. Dataflow Generated in Salesforce

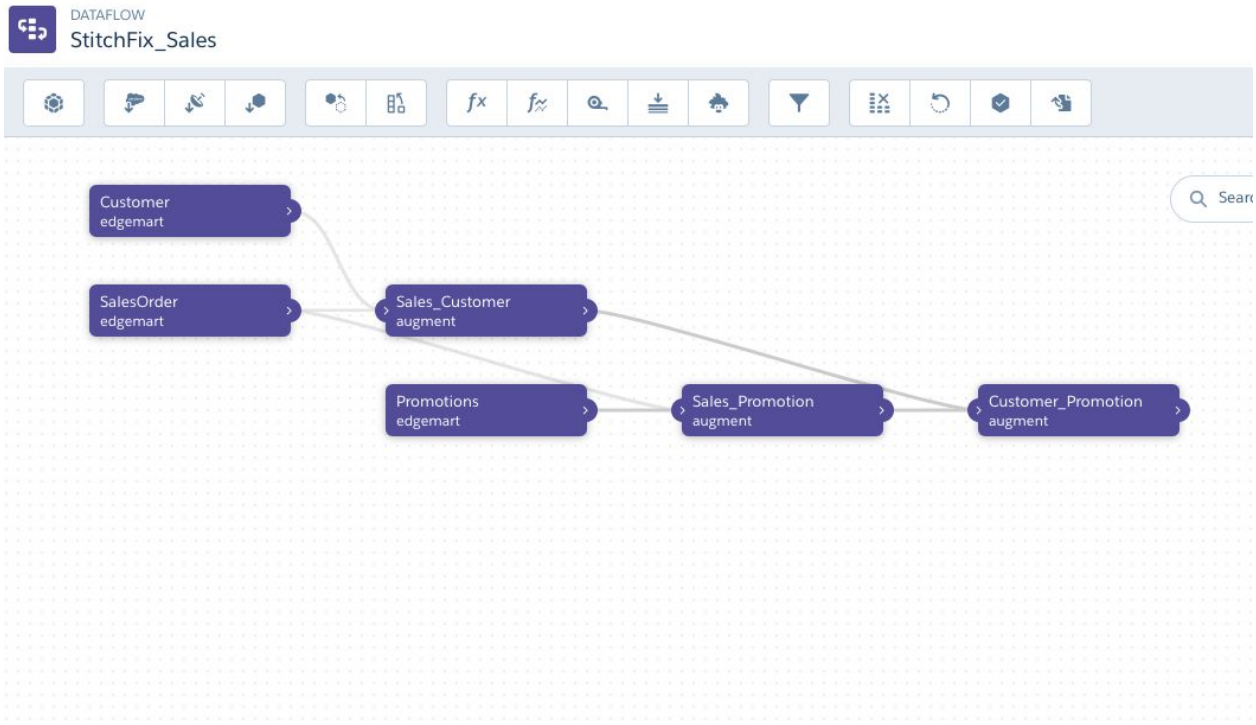
Customer discount, Sales and review dataflows were generated on salesforce which were then used to generate dashboards for analyzing the data.

1. StitchFix_Customer_Discount Dataflow



2. StitchFix_Reviews Dataflow

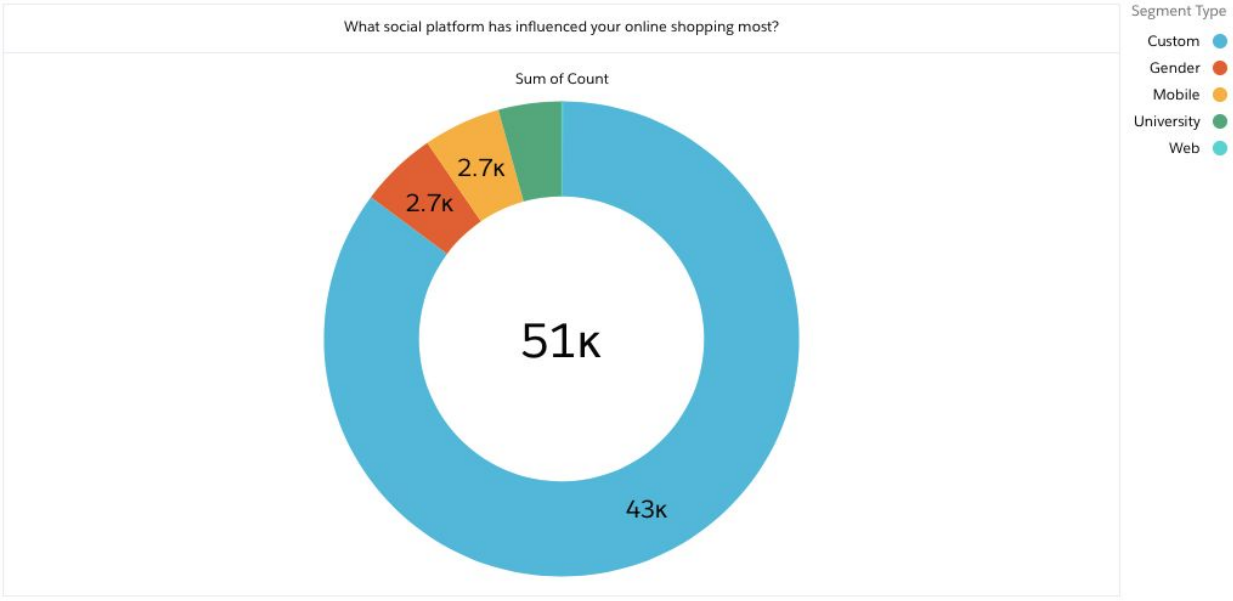
3. StitchFix_Sales



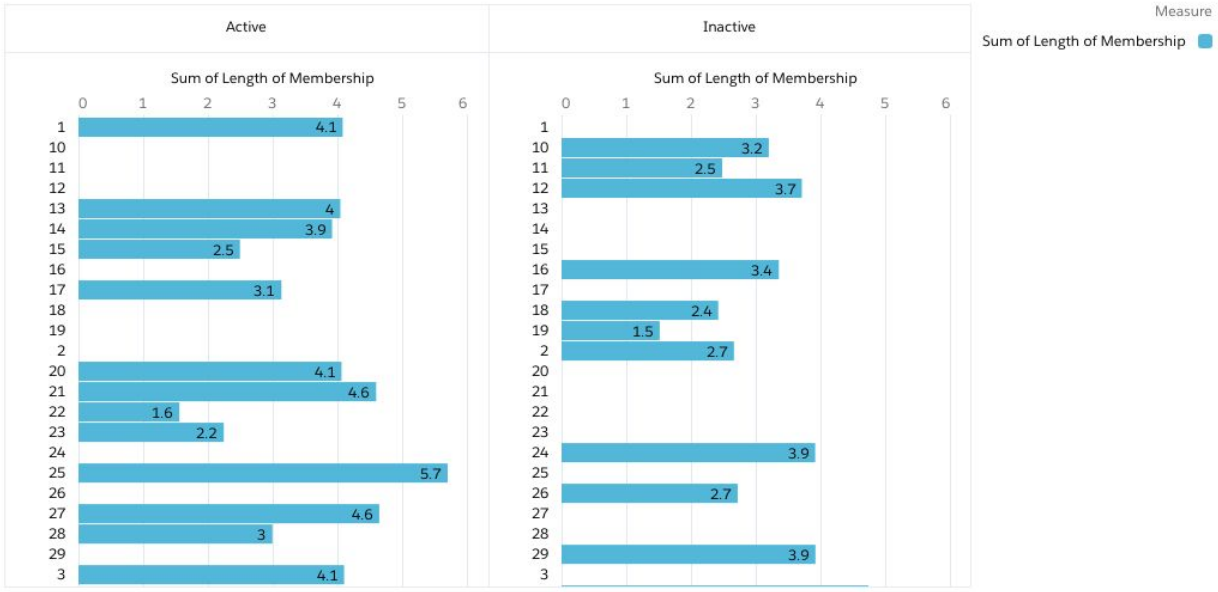
III. Dashboards generated on Tableau and Salesforce

Visualization based on the data provided was done based on the dataflows that were developed using Salesforce and Tableau

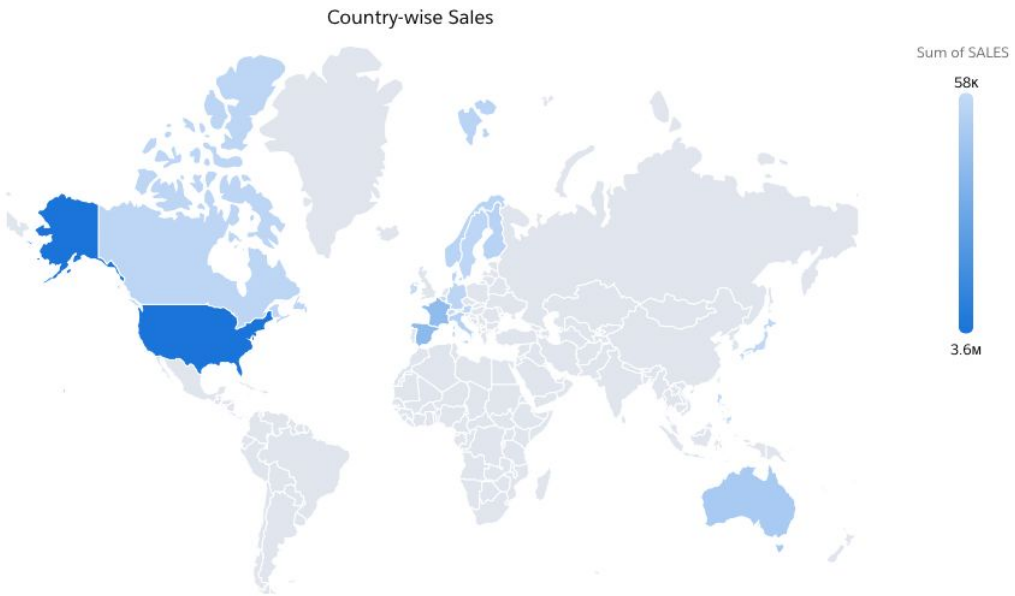
1. This dashboard was created to specify the platform influencing the online shopping the most



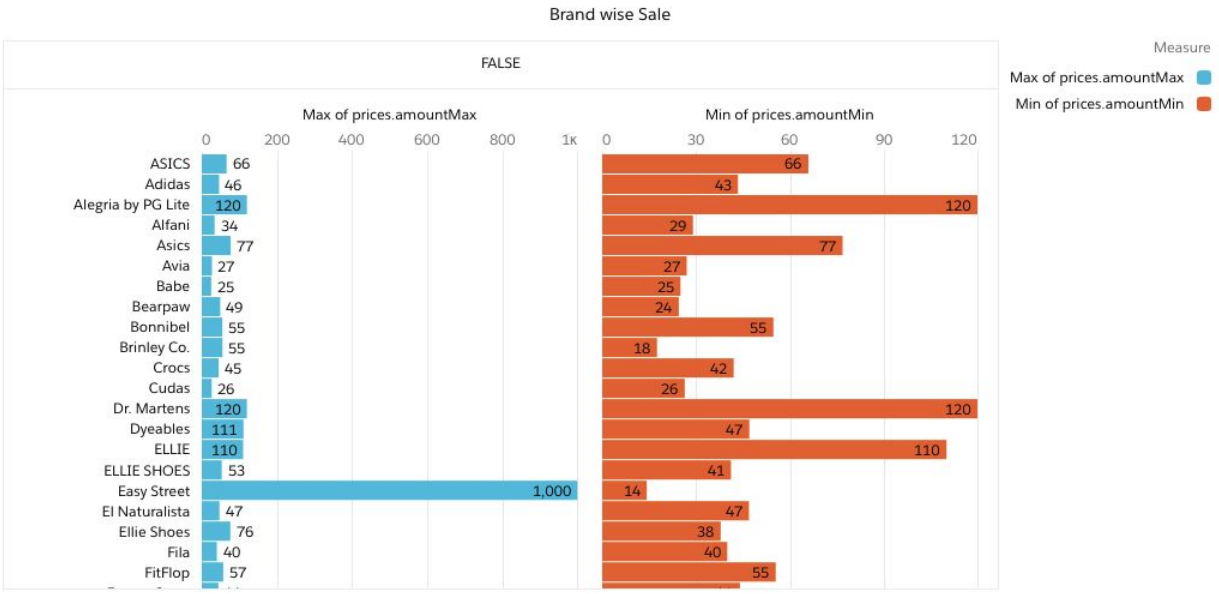
2. The following dashboard shows the sum of the length of membership and the status of the membership



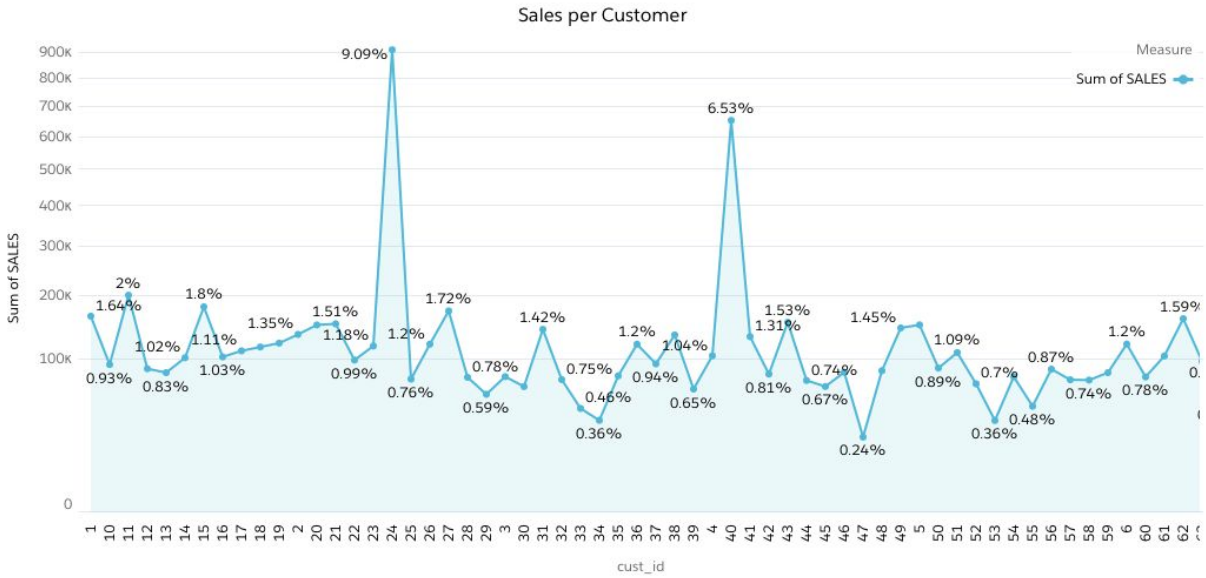
3. This visualization shows the country-wise sales of stitch fix



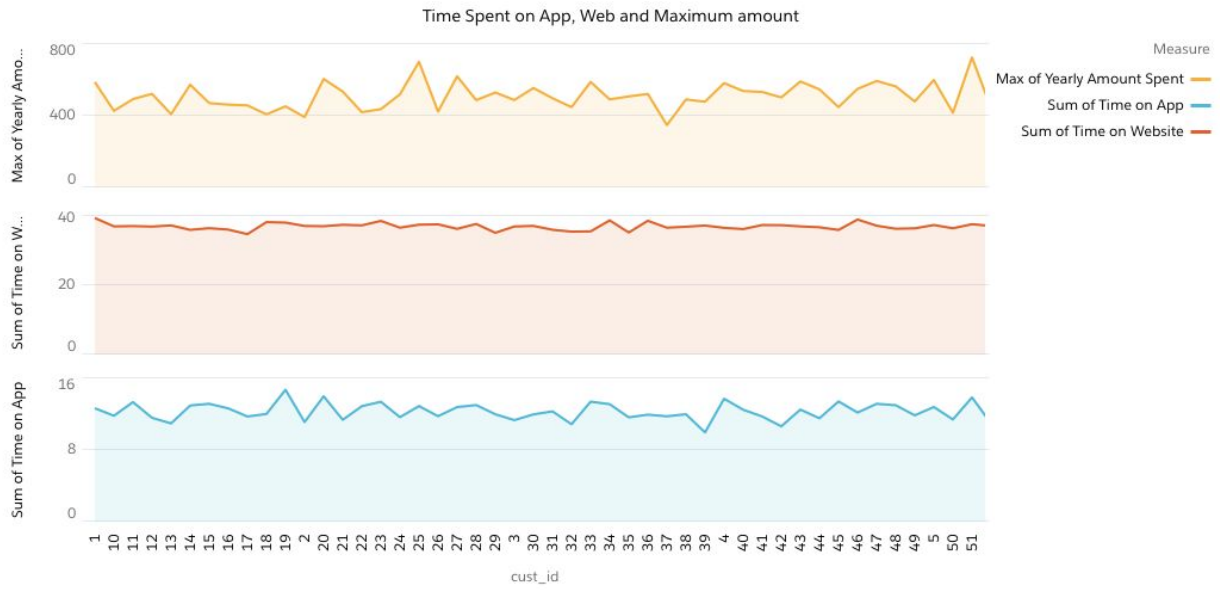
4. This Visualization was done to show the brandwise sale of the products.



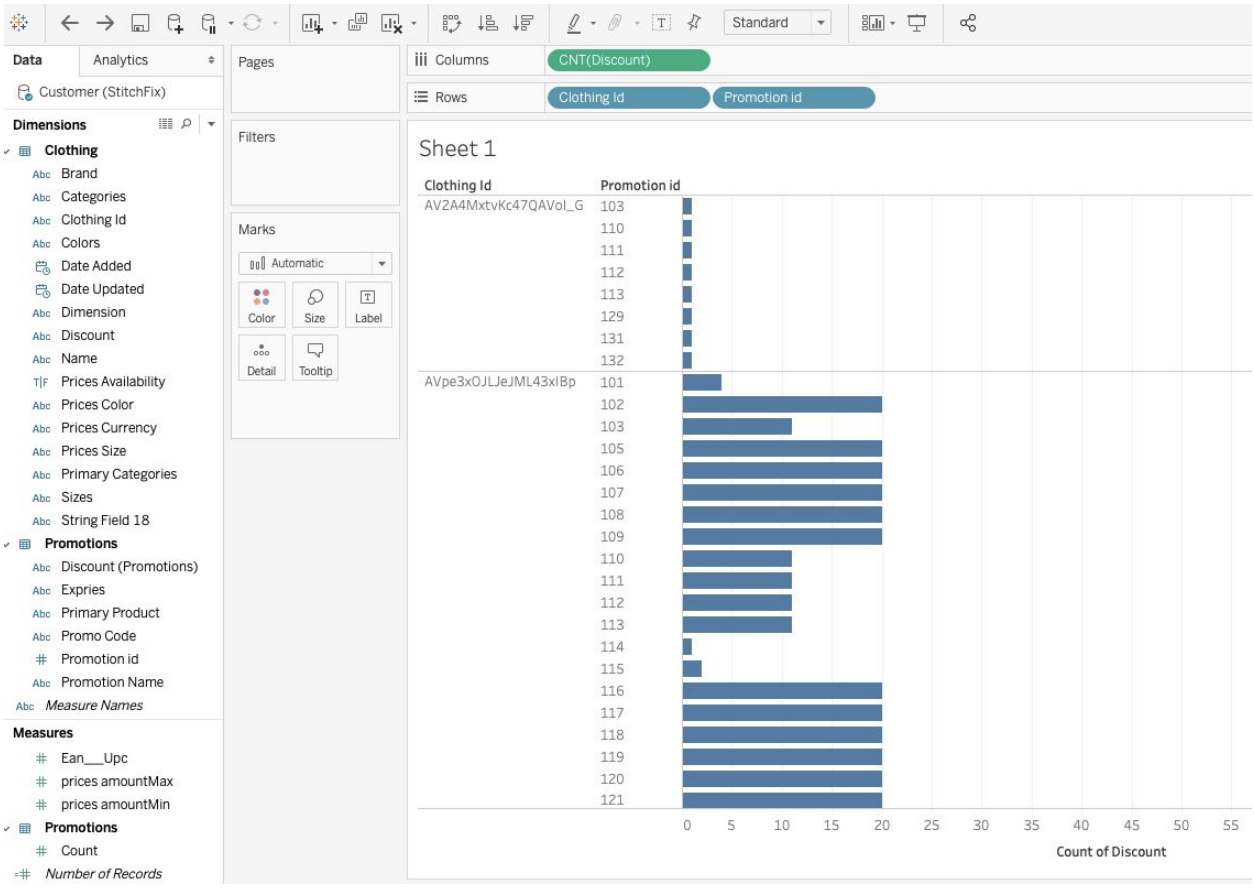
5. This dashboard shows the sales per customer visualization.



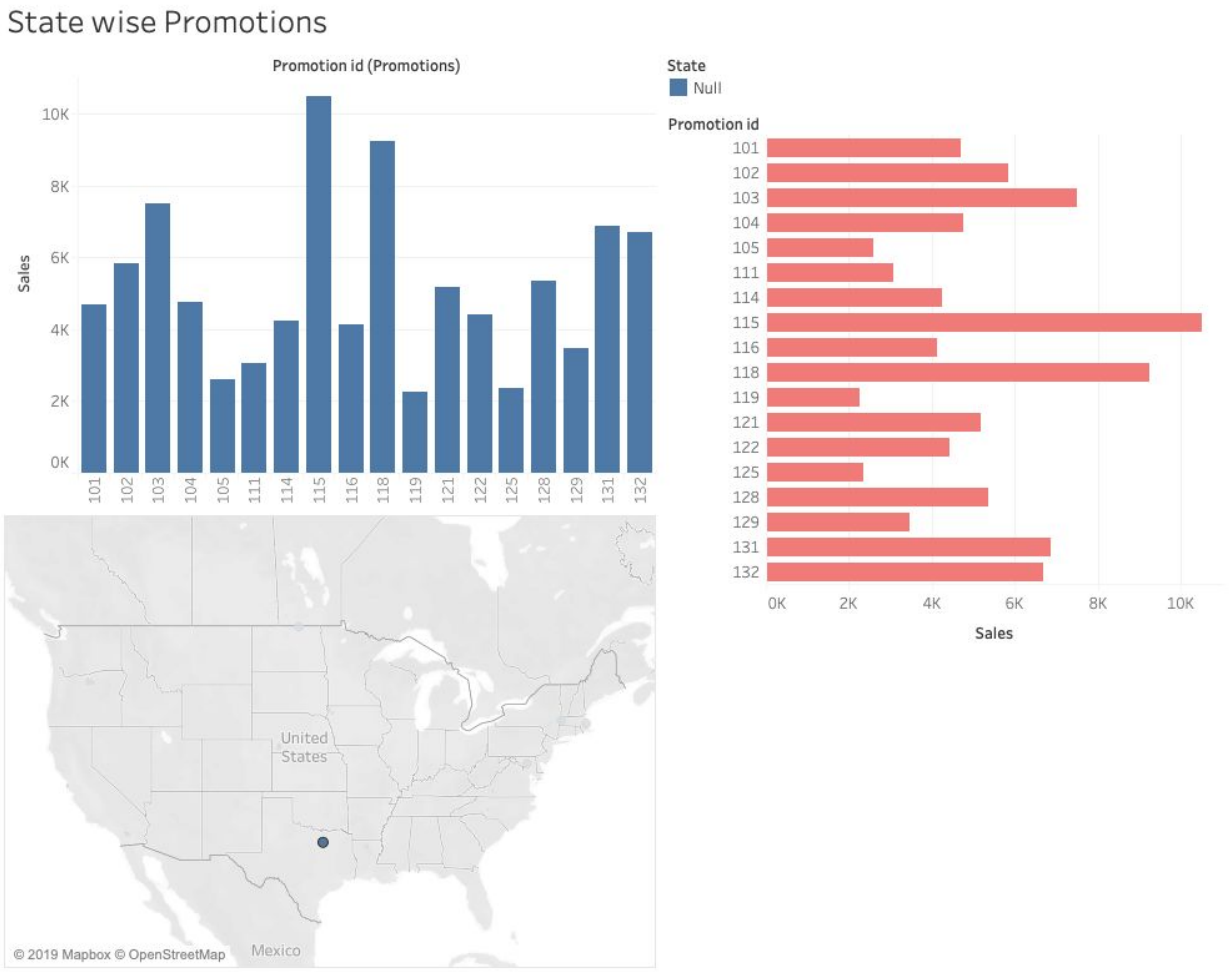
6. This shows the sum of time spent by the customers on the web and on the app as well as the maximum annual amount spent by the customers.



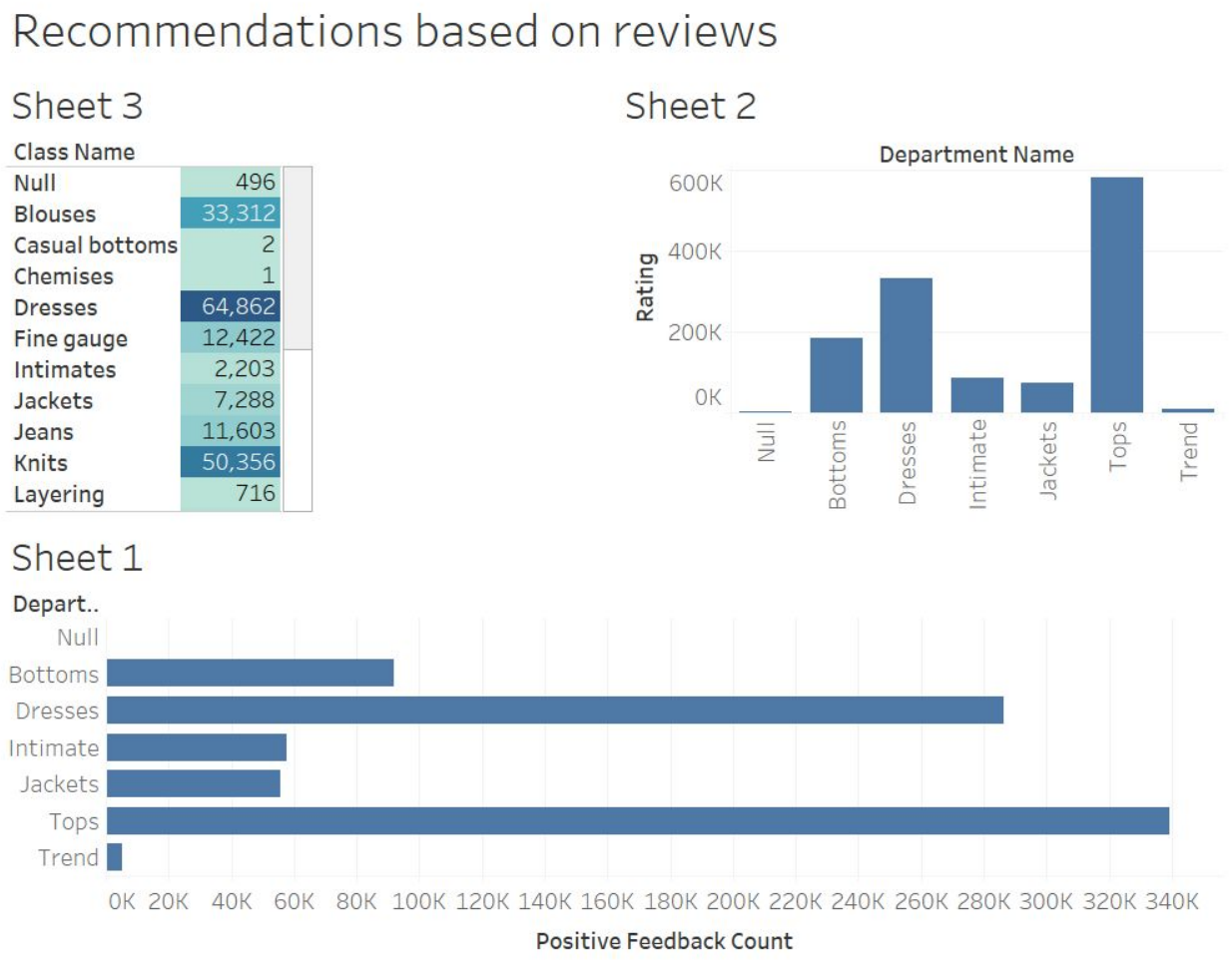
7. This Visualizations depicts the Products, promotions and the discount count on that product.



8. The following Dashboard shows the State-wise Promotions and the sales made through those promotions.

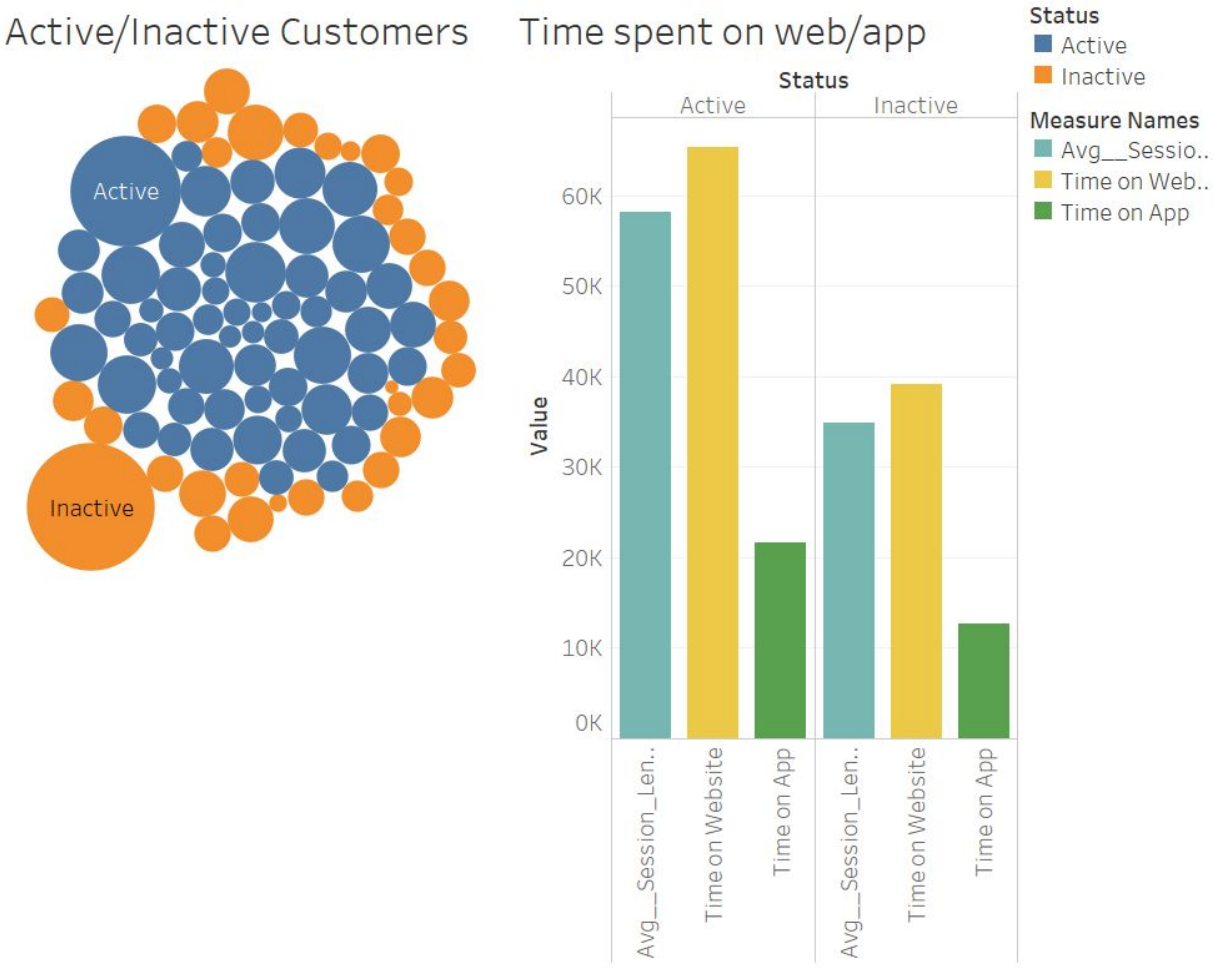


9. This Dashboard depicts the recommendations based on reviews



10. This dashboard shows the profit per month and the top most and lowest product of the month.

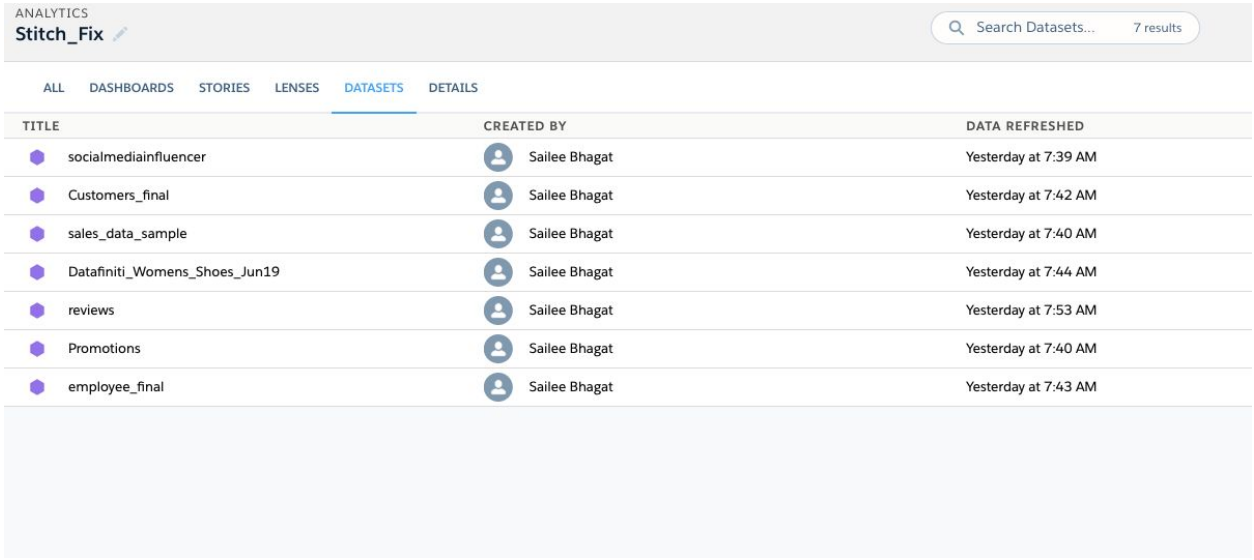
11. Customer Analysis Dashboard



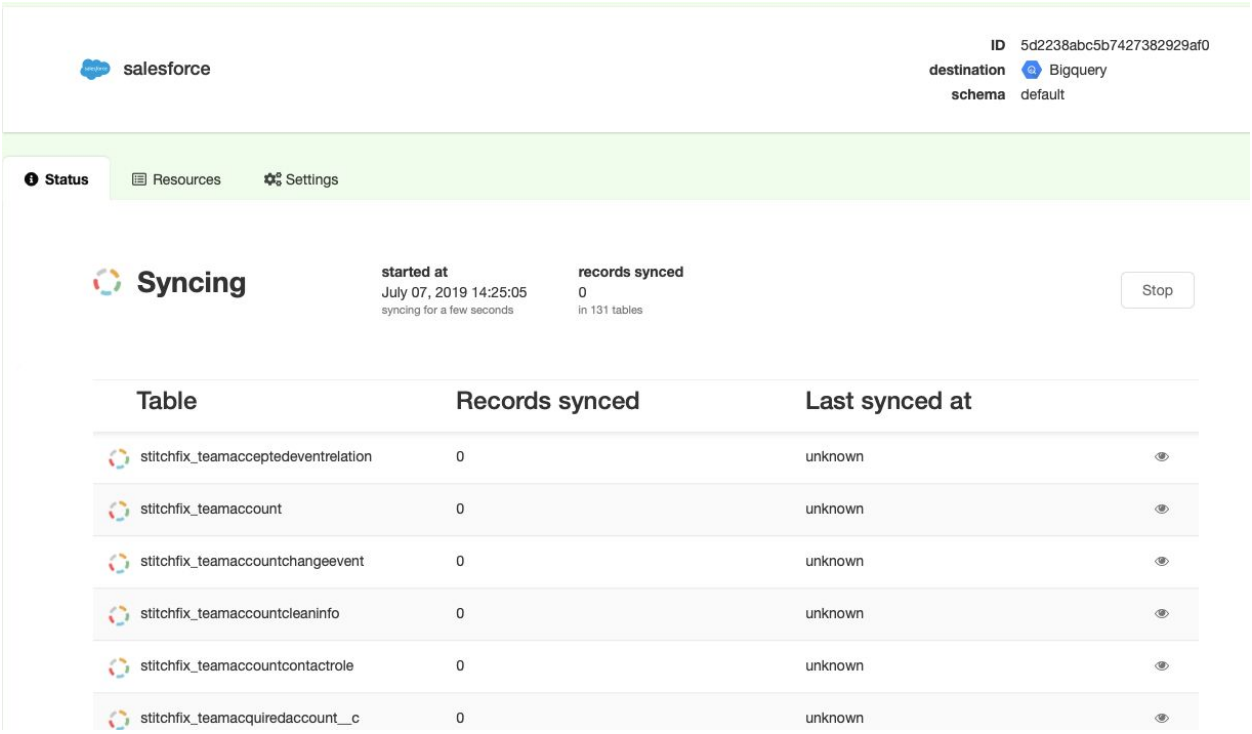
IV. Data Integration between Salesforce and BigQuery

Data was migrated from salesforce to BigQuery with the help of Blendo for the other teams to access.

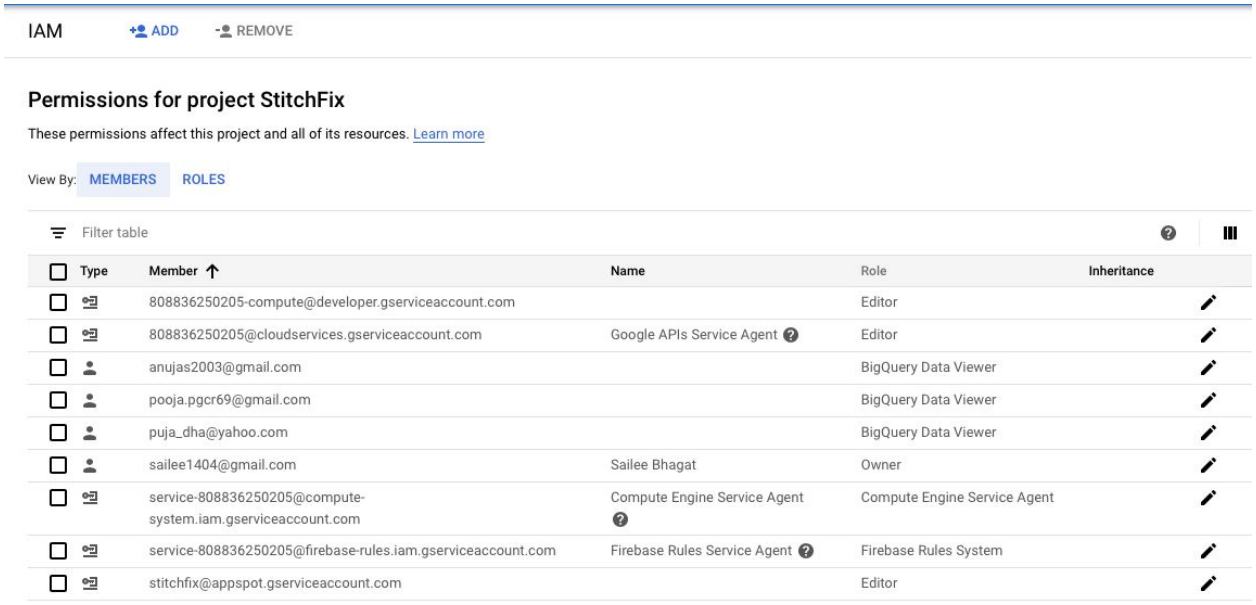
1. The Datasets in Salesforce:



2. Data from Salesforce was migrated to GCP BigQuery using Blenda

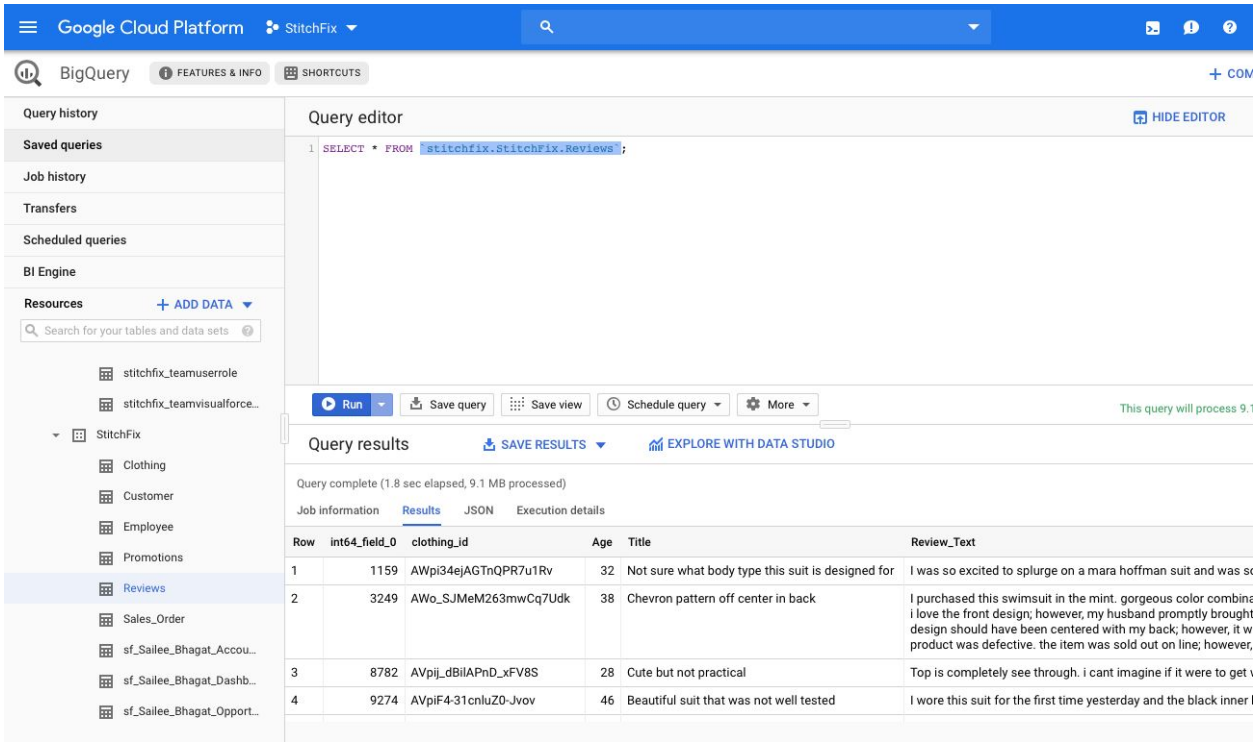


3. Permissions for accessing the data were given to the new users



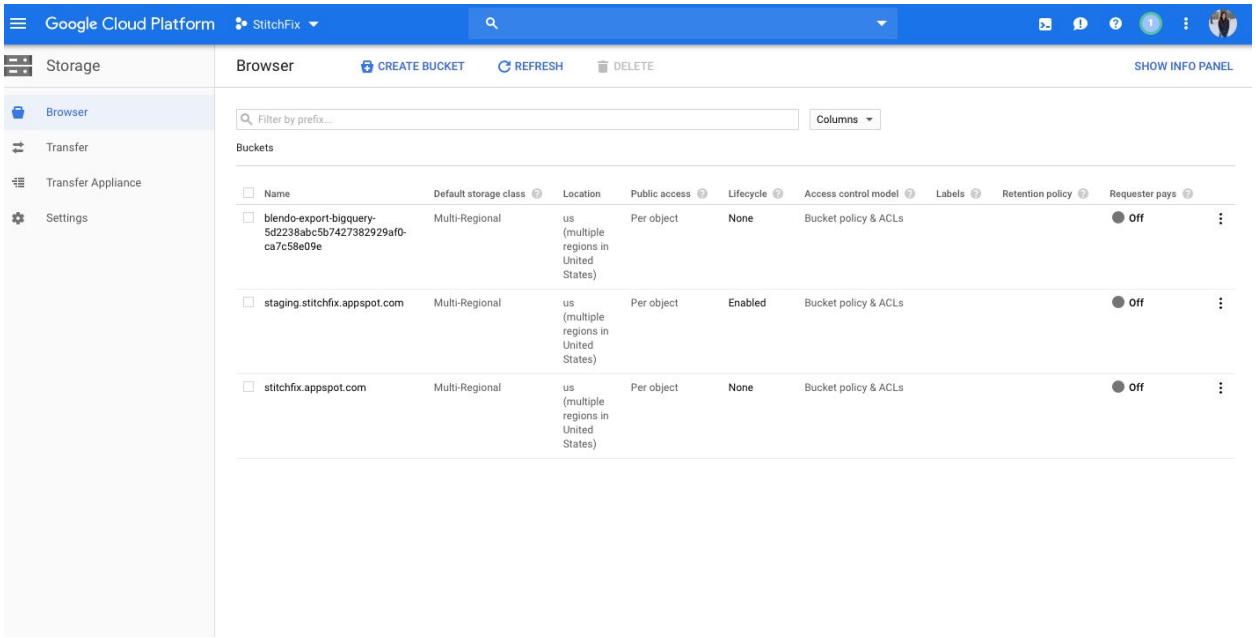
Type	Member	Name	Role	Inheritance
	808836250205-compute@developer.gserviceaccount.com		Editor	
	808836250205@cloudservices.gserviceaccount.com	Google APIs Service Agent	Editor	
	anujas2003@gmail.com		BigQuery Data Viewer	
	pooja.pgcr69@gmail.com		BigQuery Data Viewer	
	puja_dha@yahoo.com		BigQuery Data Viewer	
	sailee1404@gmail.com	Sailee Bhagat	Owner	
	service-808836250205@compute-system.iam.gserviceaccount.com	Compute Engine Service Agent	Compute Engine Service Agent	
	service-808836250205@firebase-rules.iam.gserviceaccount.com	Firebase Rules Service Agent	Firebase Rules System	
	stitchfix@appspot.gserviceaccount.com		Editor	

4. The data was checked in the BigQuery



Row	int64_field_0	clothing_id	Age	Title	Review_Text
1	1159	AWpi34ejAGTnQPR7u1Rv	32	Not sure what body type this suit is designed for	I was so excited to splurge on a mara hoffman suit and was so
2	3249	AWo_SJMeM263mwCq7Udk	38	Chevron pattern off center in back	I purchased this swimsuit in the mint. gorgeous color combina i love the front design; however, my husband promptly brought design should have been centered with my back; however, it w product was defective. the item was sold out on line; however,
3	8782	AVpij_dBilAPnD_xFV8S	28	Cute but not practical	Top is completely see through. i cant imagine if it were to get
4	9274	AVpiF4-31cnluZ0-Jvov	46	Beautiful suit that was not well tested	I wore this suit for the first time yesterday and the black inner

5. Checking if the data has entered the buckets.



Steps to Access the Data

Giving GCP BigQuery Access to other teams:

1. Login to GCP Cloud Console:

   Login to your Google Cloud Platform with your credentials
2. Select your working Project:

   Enter “stitchfix”
3. Goto IAM service
4. Select BigQuery
5. Connect the Dataset and Table in the TAbleau Connector