Market Analysis Report for National Clothing Chain

Final Project: Udacity nanodegree

SEPTEMBER 17, 2023

Yumna Mohammad Adnan Salaas

Table of Contents

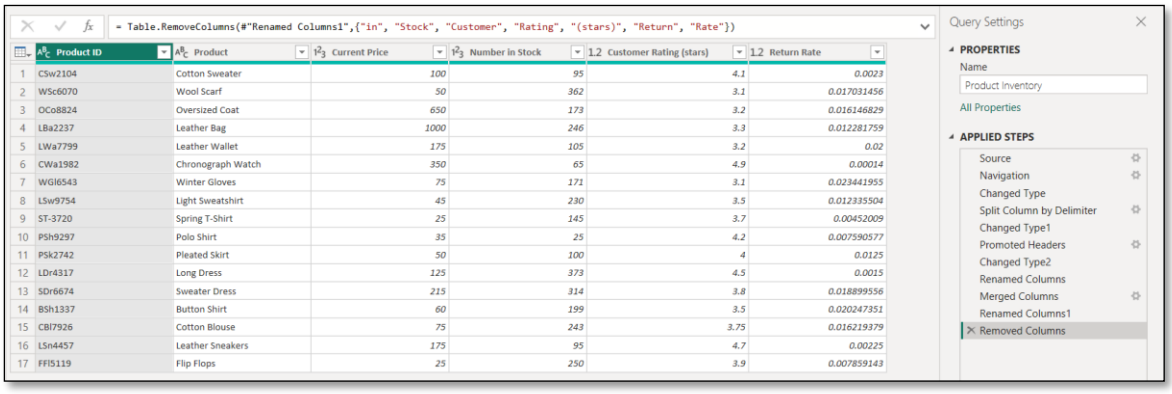
Introduction

This report is to help create a targeted marketing campaign to lure lost customers back. The analysis is based on 7 tables {Avg Income by State, Customer List, Incomes by State, Industries, Product Inventory, Purchase List, State List}. Steps done in the project are the following: ETL, Using DAX, and Creating visuals.

Power Query- Data Cleaning

Product Inventory table

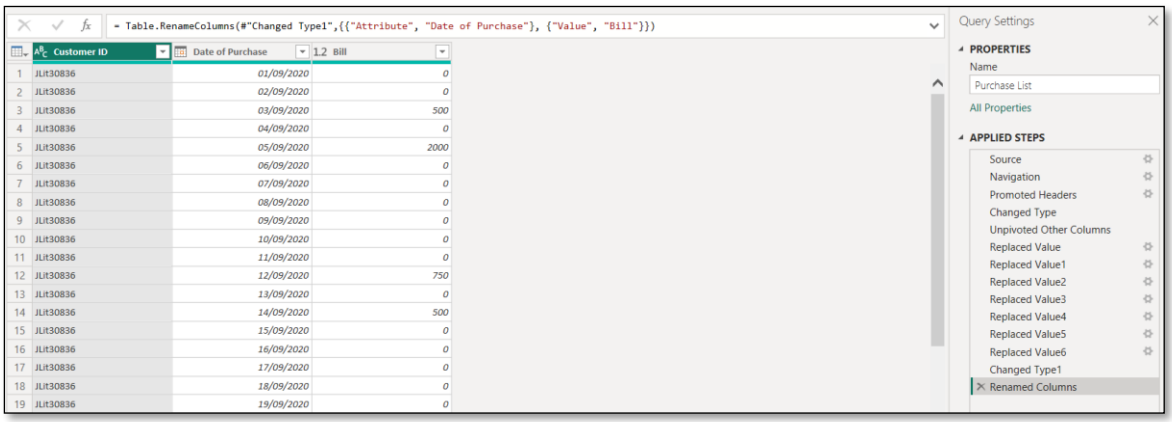
After cleaning



	Product ID	Product	Current Price	Number in Stock	Customer Rating (stars)	Return Rate
1	CSw2104	Cotton Sweater	100	95	4.1	0.0023
2	WSc6070	Wool Scarf	50	362	3.1	0.017031456
3	OCo8824	Oversized Coat	650	173	3.2	0.016146829
4	LBa2237	Leather Bag	1000	246	3.3	0.012281759
5	LWa7799	Leather Wallet	175	105	3.2	0.02
6	CWa1982	Chronograph Watch	350	65	4.9	0.00014
7	WGl6543	Winter Gloves	75	171	3.1	0.023441955
8	LSw9754	Light Sweatshirt	45	230	3.5	0.012335504
9	ST-3720	Spring T-Shirt	25	145	3.7	0.00452009
10	PSh9297	Polo Shirt	35	25	4.2	0.007590577
11	PSk2742	Pleated Skirt	50	100	4	0.0125
12	LDr4317	Long Dress	125	373	4.5	0.0015
13	SDr6674	Sweater Dress	215	314	3.8	0.018899556
14	BSh1337	Button Shirt	60	199	3.5	0.020247351
15	CBl7926	Cotton Blouse	75	243	3.75	0.016219379
16	LSn4457	Leather Sneakers	175	95	4.7	0.00225
17	FFl5119	Flip Flops	25	250	3.9	0.007859143

Purchase List table

After cleaning



	Customer ID	Date of Purchase	Bill
1	JLit30836	01/09/2020	0
2	JLit30836	02/09/2020	0
3	JLit30836	03/09/2020	500
4	JLit30836	04/09/2020	0
5	JLit30836	05/09/2020	2000
6	JLit30836	06/09/2020	0
7	JLit30836	07/09/2020	0
8	JLit30836	08/09/2020	0
9	JLit30836	09/09/2020	0
10	JLit30836	10/09/2020	0
11	JLit30836	11/09/2020	0
12	JLit30836	12/09/2020	750
13	JLit30836	13/09/2020	0
14	JLit30836	14/09/2020	500
15	JLit30836	15/09/2020	0
16	JLit30836	16/09/2020	0
17	JLit30836	17/09/2020	0
18	JLit30836	18/09/2020	0
19	JLit30836	19/09/2020	0

DAX

Histogram (Using buckets)

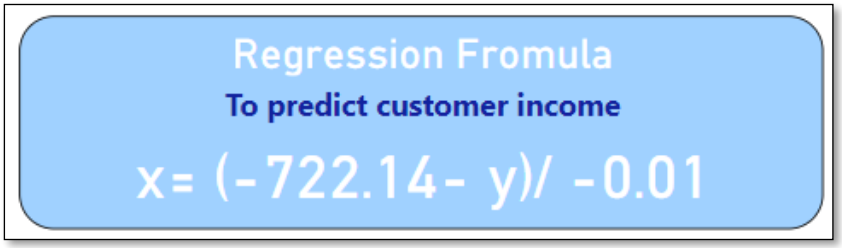
```
Bucket =
IF([Predicted income]<80000,"80k",
IF([Predicted income]>=80000&& [Predicted income]<100000,"100k",
IF([Predicted income]>=100000&& [Predicted income]<120000,"120k",
IF([Predicted income]>=120000&& [Predicted income]<140000,"140k",
IF([Predicted income]>=140000&& [Predicted income]<160000,"160k",
IF([Predicted income]>=160000&& [Predicted income]<180000,"180k",
IF([Predicted income]>=180000&& [Predicted income]<200000,"200k",
IF([Predicted income]>=200000,"above 200k"
))))))))
```

Column chart (Recommended products)

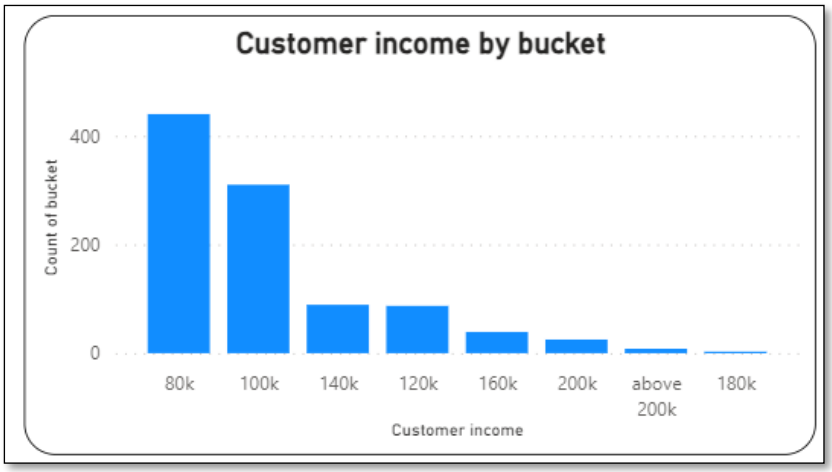
```
Recomended Product =
IF('Customer List'[Predicted income]>=60000 && 'Customer List'[Predicted income] <100000, "Shirt",
IF('Customer List'[Predicted income]>=100000 && 'Customer List'[Predicted income] <140000, "Sweater",
IF('Customer List'[Predicted income]>=140000, "Leather bag")))
```

Visualization and Analysis

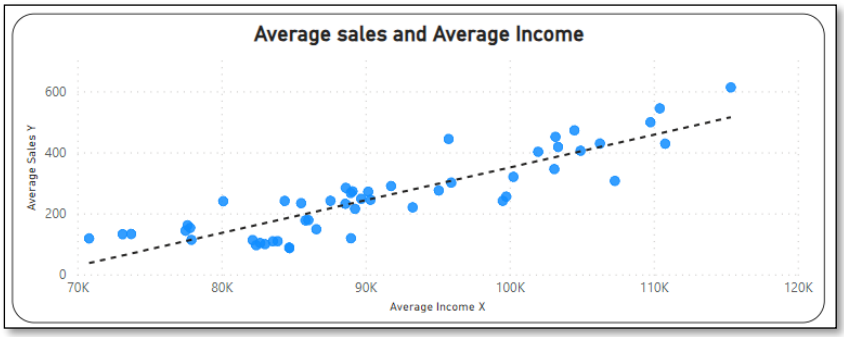
Linear Regression

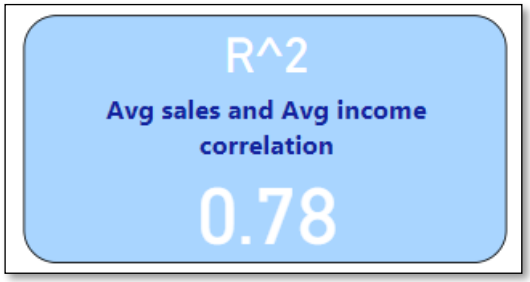


Histogram

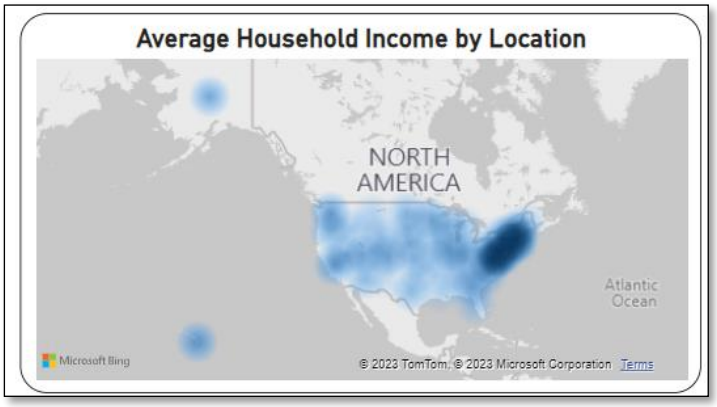


Scatter plot + Card



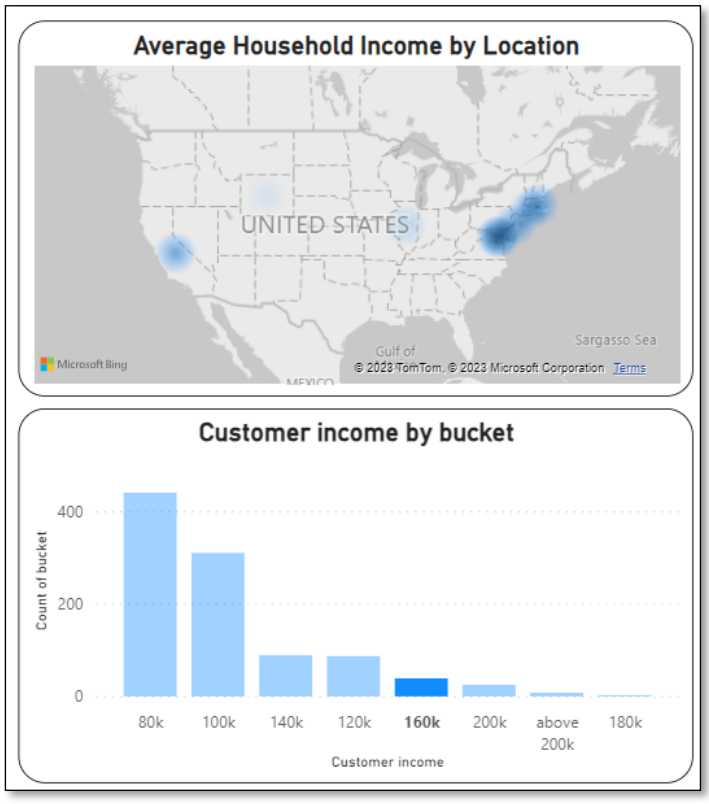


Heatmap

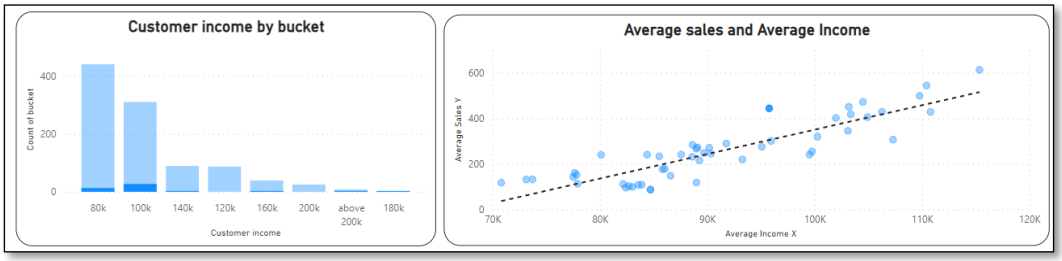


Cross Filtering

1. Heatmap and Histogram



2. Scatterplot and Histogram



3. Scatterplot and Heatmap

Findings and Recommendations

1. The correlation (R^2) between sales and income is 0.78. The number shows a strong correlation indicating the major dependency of sales on income. Illinois, New Jersey, and the District of Columbia are the top 3 states with the highest income as well as highest sales.

2. The correlation (R^2) between customer ratings and product return rate is 0.69. This shows that customers who give higher ratings tend to keep the purchased products, lowering the return rates.

3. The linear regression formula used to predict customer sales and customer income is X= b-Y/-m. As X represents the average income and Y represents the average sales.

4. Using DAX, the customer with the highest predicted income is Jon Little.

5. Using a histogram visual which counts the number of each recommended product, is shirts. This was observed through the analysis of the overall predicted income. However, in states like Illinois, New Jersey, and the District of Columbia, leather jackets should be advertised the most due to the higher income and higher sales.