**Market Analysis Report for National Clothing Chain**

Data Analysis and Visualization with Microsoft Power BI

**BY**

**Mai Salah Eldin – Data Analyst**

# SECTION 1: INTRODUCTION

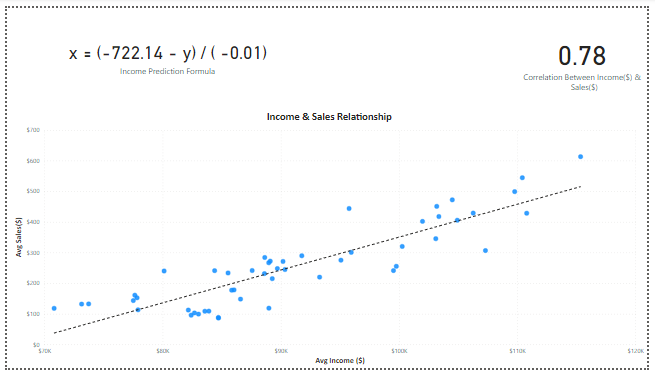
This report is prepared for National Clothing Chain to help them know where the greatest income exists around the country and whether there is a correlation between sales and income. Customers’ incomes are unknown at the moment, but it can be predicted by looking at the purchase history and locations.

The company is also interested in analyzing the inventory, specifically customer ratings and return rate and finding if there’s a correlation between the two.

The focus of this market analysis is to finding answers to the following questions;

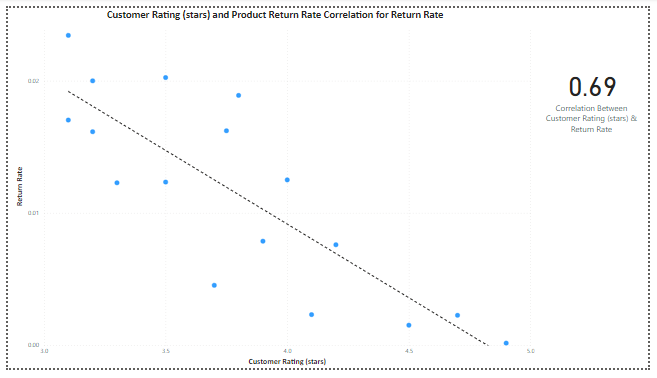
1. What is the correlation (R2 value) between sales and income?
2. What is the correlation (R2 value) between customer ratings and product return rate?
3. What are the linear regression formulas to predict customer sales and customer incomes?
4. Which customer do you predict has the highest income?
5. Which product will be advertised the most?

# SECTION 2: Insights and Analysis

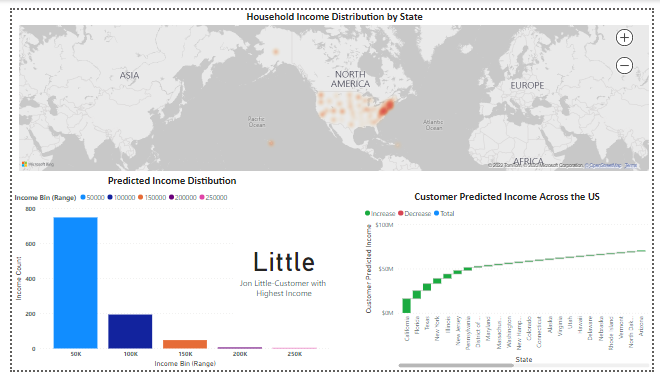
From the above scatterplot diagram we can conclude that there is a relatively strong correlation between income and Sales. As the customer income increases, their monthly purchases (Sales) increases as well. This relationship is directly proportional (Positive) with a **0.78** Coefficient of Determination(R^2) that indicates a strong correlation. It is a linear relationship since the trend line is linear.

It is for sure that National Clothing Chain does not know the income of all its customers however, it can be predicted using the following linear formula **y = mx + b.** This can be obtained by selecting a random point (x, y) from the above diagram then apply it to the formula

**x = (b – y) / -m ; Income Prediction Formula = x = (-722.14 – y) / (-0.01)**



We can also come to another conclusion regarding the relationship between the rate a customer gives to a product and the product return rate by applying linear regression analysis; we found that their relationship is Negative. That means the highest the product is rated; the less likely it is to be return. There is a relatively strong correlation between the two variables (**R^2≅ 0.7**).

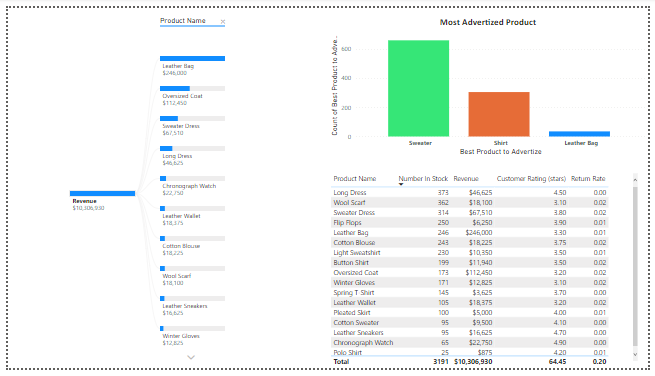


After confirming the strong positive relationship between purchases (sales) and customer’s income, we can use the LOOKUPVALUE function to find the customer with the highest predicted income.

**Customer with Highest Income =** LOOKUPVALUE('Customer List'[Last Name],'Customer List'[Customer Predicted Income], MAX('Customer List'[Customer Predicted Income]))



Jon Little from Illinois is the customer who has the highest predicted income of ($558,143.93). He has made the most purchases in the last 6 months ($5,250).



Based on the purchase amounts of our customers, we planned which product to advertise out of the three products:

* Shirt: $25
* Sweater: $100
* Leather Bag: $1,000

Sweater seems to be advertised the most since most customers made purchases of less than $1000 in the last 6 months.

# 

# SECTION 3: CONCLUSION & RECOMMENDATIONS

When analyzing the current stock we find that there are 246 Nos. of Leather bags that bring the highest revenue among all other products. Knowing this it is highly suggested to draw more attention to it during the next marketing campaign. Reaching more customers all over the states of a higher income range, assures that they can afford the relatively expensive leather bag, thus scores higher revenue to the company.