# Market Analysis Report for National Clothing Chain

The goal of this report is to create a targeted marketing campaign for specific products to specific customers in specific locations. The analysis in this report has been carried out using data from the US Census Bureau, Business Data and Customer Data.

What is the correlation (R2 value) between sales and income?
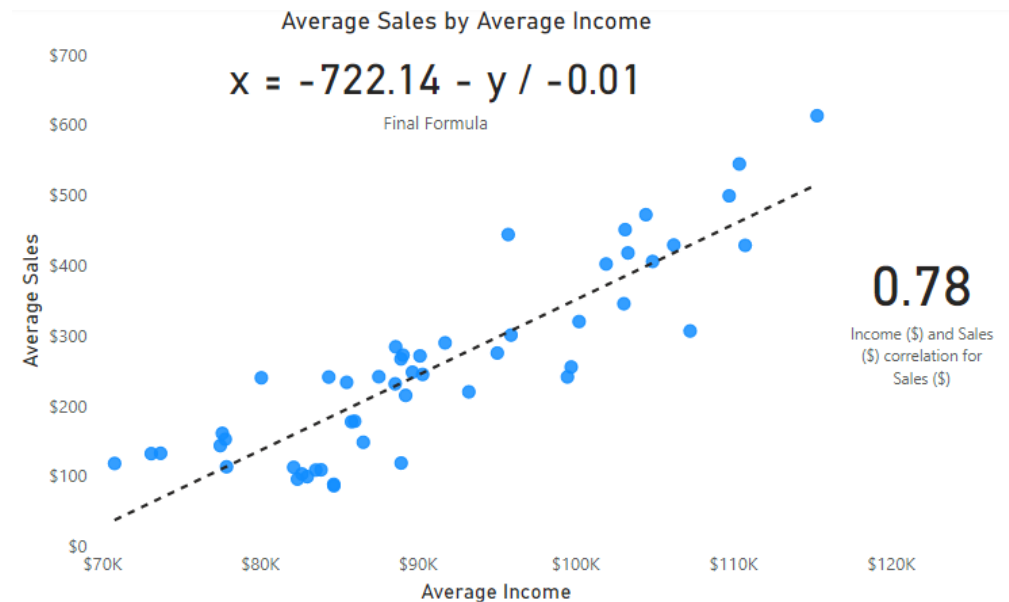
The figure above is a regression analysis showing the positive correlation between average sales and average income. It shows that as average income increases, the average sales is expected to go up. We have a strong R2 correlation of .78, and our regression formula is at the top. The formula also correctly calculates the expected salary based on population size.

What is the correlation (R2 value) between customer ratings and product return rate?

The figure above is a regression analysis showing the negative correlation between return rate of products and customer rating (stars). It shows that return rates decrease as the customer rating of products increase. We have a moderate R2 correlation of .69.

Return Rate by Customer Rating (stars)

0.69

Customer Rating (stars) and
Return Rate (%) correlation
for Return Rate (%)
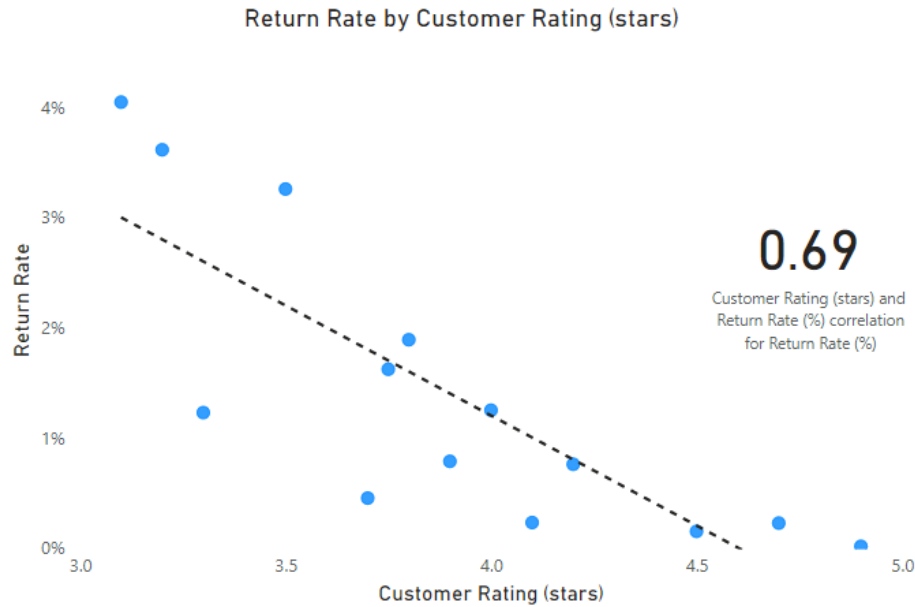
What are the linear regression formulas to predict customer income from customer sales?

$$x = -722.14 - y / -0.01$$

Final Formula

where x = customer income and y = customer sales

Which customer do you predict has the highest income?

My prediction is that 38 year old Jon Little from Illinois has the highest income. His customer ID is JLit30836

Which product will be advertised the most?

Income	Count of Customers	Average Sales	Recommended Product(s)
Less than $100,000	707	$ 123.47	Shirt
$100,000 - $150,000	216	$ 507.83	Sweater, Shirt
$150,000 - $200,000	56	$ 954	Sweater, Shirt
$200,000 - $250,000	18	$ 1,417.18	Leather Bag, Sweater, Shirt
Over $250,000	3	$ 3,366.67	Leather Bag, Shirt, Sweater

The Shirt $25 will be advertised the most. This is because it is perfectly priced for the segment of customers (707) with less than $100,000 income. The Sweater $100 should be the next most advertised product while the Leather Bag $1,000 would be the least advertised product.

Customers by Predicted Income

Average Sales based on Average Income Buckets
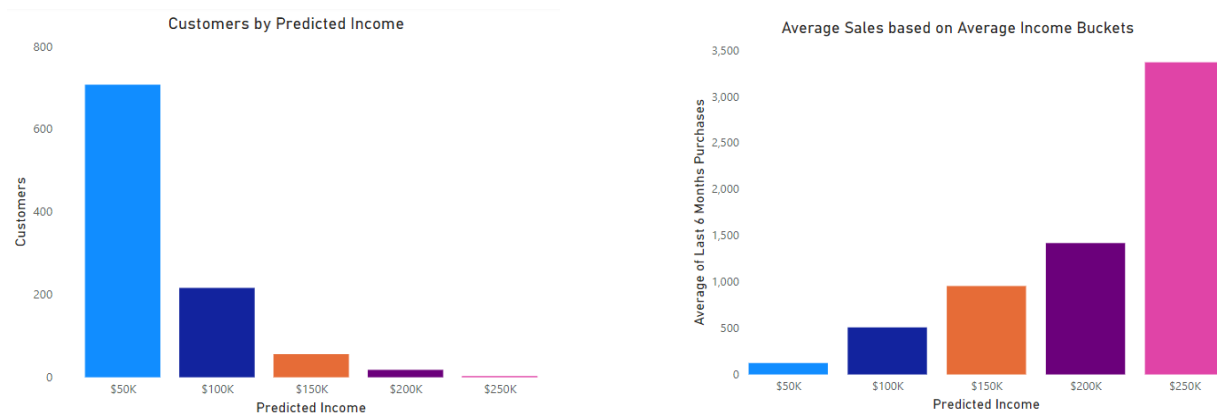
Customer Demographic Breakdown

Of the 1,000 total customers belonging to National Clothing Chain, California has the highest number of customers (150). In California, more than half (78) of the customers earn between less than $100,000 and the majority of these customers are between 40 - 49 age range.