

## ANALYSIS SUMMARY

In this report, we present the summary of the findings found with the statistical analysis and visualizations carried out in the attached Power BI file.

- The Average Income in the **total population** of US, shows a Histogram that is right skewed, with a Median in 85K.
- The average predicted Income of our **customers list** shows a Histogram also right skewed, with a bigger tail than the general population, but also a median in 85K
- The customers with **the 5 highest salaries** live in Illinois and New Jersey, with an average of purchases in the last 6 months between 5.250\$ and 1.736\$, among the highest also.
- The **Correlation (R2 Value) between sales and income** is positive, with a R2 of 0,78, a strong correlation
- The **Linear Regression formula** to predict customer income from customer sales, is as follows:

$$Y = mx + b$$

$$Y = 72,43 X + 72638,21$$

This formula allowed us to predict the Income of our customers, as shown in the visual of the 3rd page.

- Based on the 3 preselected items by the Clothing Chain, **the recommendation would be:**
  - Knowing the prediction of Income of our customers, as well as the correlation coefficient between Income and Sales, the suggestion is:
    - Promote *T-Shirts* to those clients with an average Income less than 80.000\$ (321 possible customers)
    - Promote *Sweaters* to those clients with an average Income between 80.000\$ and 120.000 \$ (549 possible customers)
    - Promote *Leather Bags* to those clients with an average Income superior to 120.000\$ (130 possible customers)
  - By State, California, Florida and Texas are the ones with more potential customers.
- The **correlation coefficient between customer ratings and product return** rate is negative, with a R2 of 0,69, a moderate correlation. Based on this:
  - From the 3 preselected product, 2 of them do not have a great rating:
    - Leather Bag = 3,3
    - Button Shirt = 3,5
  - The sweater has a good rating of 4,10.

If we consider that the lowest rating have the highest return rate, we should focus our marketing campaign in the sweater, rather than in the other 2 products. But we need to keep in mind that the stock for Leather bag and Shirt is still quite high (246 and 199), so maybe something should be done to improve the rating of those 2 product, in order to release them.
- In an effort to better improve the sales of the Company, we dare to suggest that the marketing campaign focus on **3 new products**, based on the correlation between “Income and sales” and “Return Rate and Rating”:
  - Promote products of *less than 100\$* to those clients with an average Income less than 80.000\$

- Promote products between 100-700\$ to those clients with an average Income between 80.000\$ and 120.000 \$
- Promote products of more than 700\$ to those clients with an average Income superior to 120.000\$

So, more specifically:

- To promote **Polo T-Shirt** (rating 4,2) to those customers with an income lower than 80K (321 possible customers). Note the stock is just 25
- To promote **Long Dress** (rating 4,5 and 373 units in stock) to those customers with an income between 80- 120K (549 possible customers)
- To promote **Chronograph Watch** (rating 4,9) to those customers with an income over 80K, with a focus on those with incomes over 120K (549 + 130 possible customers)