The dataset I have used is called “Sample – Superstore” from Tableau. After some brief exploration of the dataset, I found that no cleaning was necessary. I did however calculate a “Profit Margin” column by dividing “Profit” over “Sales”. I wanted to explore what parts of the data affected profit. After creating various small summary tables and graphs, I noticed that some customers had high levels of sales but also negative profit. Upon further investigation I found that excessive discounts given to some customers meant that some sales were made at a loss. I felt that this was an important fact that needed illustration using data visualization.

**Viz 1**

I wanted to find out why higher sales did not correspond with higher profits and see if I could effectively visualize the situation. The y-axis are the customer names and the x-axis is the amount of each customer’s combined sales. Normally I would orient the bars vertically and show sales on the y-axis, but it would be very difficult to display the customer’s names on the x-axis as a result.

By color-coding the bars based on the profit that a customer generated, I could clearly show that some customers represented money lost. I simplified the color palate of the bars to show the profit/loss on a 4-scale grey-to-red scheme. I tried to add labels regarding other variables, but I could only add one additional layer of labels on each bar. I needed to see additional levels in the data to further investigate this issue. Therefore, for this visualization, I only used simple coloring of the bars to sharply indicate that high sales do not necessarily mean higher profits.

**Viz 2**

Since “Viz 1” was not going to be of much help going forward, I created a scatterplot to see both individuals and overall trends. “Viz 2” is much better at showing the multivariate nature of the issue. Profit is along the y-axis and sales amount is on the x-axis. We can see there are sales of many different amounts that brought both positive and negative profits.

By using color and shape I could explore other aspects of the data. The most insight came from including the profit margin and total discount each customer received. The size of the bubbles show the percent discount applied to the customer while the grey and red colors indicate positive and negative profit margin respectively. The larger red bubbles pop out much more forcefully and draw attention to the discount --> profit-loss connection. I removed the grid lines and softened the axis labels so as not to distract from the overall trend in the chart.

From here, I could see a strong inverse correlation between the amount of discount a customer received and the profitability of their purchases. I could see that it was the discounts that were largely responsible for cutting into profits.

**Viz 3**

Since I could see that the least profitable orders had large discounts applied, I thought it would be good to simplify what was shown in the previous scatterplot. It isn’t necessary to show each customer. The important thing to see if the damage discounts do to profits. I aggregated all of the profit from all customers into one bar and split it by profit margin (profit/loss). I oriented the bar vertically and labeled the average discount from the orders on each bar. This way I’ve clearly shown that too many purchases with discounts cause losses (and vice-versa).

While there may be understandable reasons for selling at a loss (moving out inventory, etc.), the relationship between large discounts and profit loss is very important to understand. I hope that my analysis and visualizations help show this.

**Data Source:** “Sample – Superstore”https://community.tableau.com/docs/DOC-1236