Problem Set 3

ECON833, Prof. Jason DeBacker
Gyumin Kim

As I noted ProblemSet 1, my research interest lies in retail operations, specifically focusing on fitting room operations. To conduct this research, I obtained transaction data, fitting room data, and traffic data from Korean fast-fashion brand.

The transaction data record which items are sold or refunded in a specific store on a given day. They sell clothes, accessories, shoes and underwear, some of which are related to fitting rooms. For this study, I exclude accessories, shoes and underwear from the dataset. The traffic data report the number of visitors to each store on a given day. Finally, the fitting room data, which is unique, contain information on which items are brought into the fitting rooms in a specific store on a given day. These data are collected through RFID reader installed in front of fitting room entrance, and each item is tagged with RFID barcode. Thus, when customers bring their item into the fitting room, the RFID readers capture relevant information.

All datasets are recorded at the SKU level, meaning they contain detailed information such as size and color. For this practice, however, I use item-level data, which identify only broad categories of clothing, such as jeans, pullover, and shirts. I do so in order to demonstrate the logic behind my choices of research design and control variable. Accordingly, I merge the datasets at the daily item-store level.

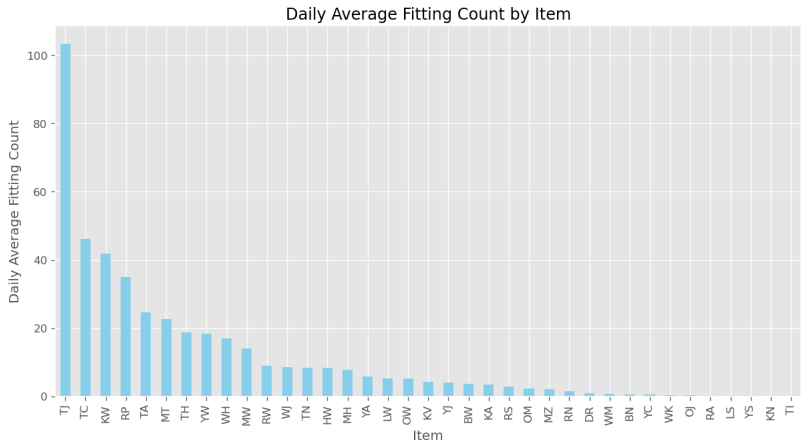

Figure 1

The first figure presents the daily average fitting count by item. Here, $fittingcount$ refers to the number of times a specific item is brought into the fitting room. For each item, I calculate the daily average number of fitting count and draw a histogram with the x-axis representing item codes and the y-axis representing the daily average

fitting count to show noticeable results.
As shown in the figure 1, The "TJ" item, which refers to jeans, stands out with an average daily fitting count of about 100, while all other items remain below 50. Following jeans, the next most frequently tried items are "TC" and "KW", which refers to cotton pants and knit sweater respectively.
This result indicates that jeans are the most frequently tried-on item in the store. Based on this, I selected jeans as the focal item to analyze the impact of fitting room usage on store performance.

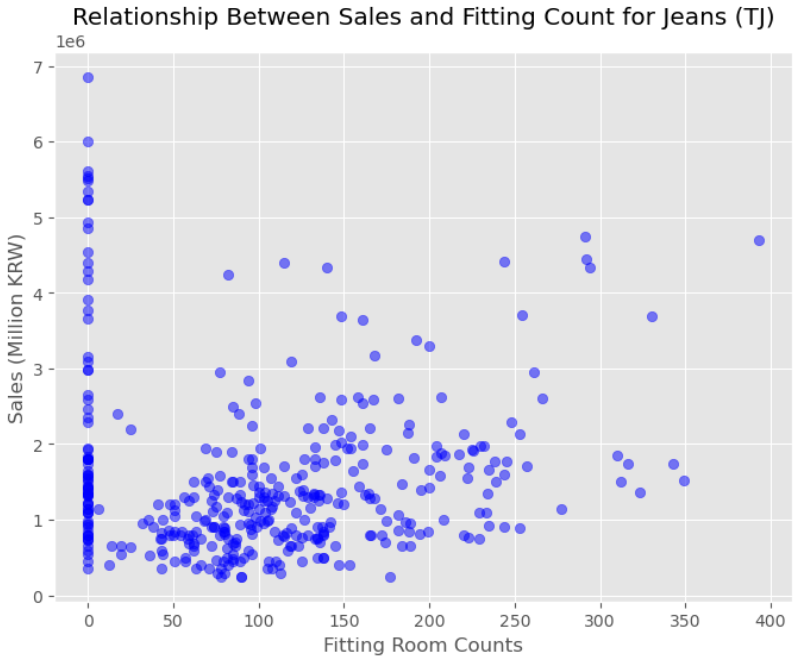

Figure 2

The second figure illustrates the relationship between daily sales (measure in in Korean won) and the daily fitting count for jeans. As shown in Figure 2, there is positive relationship between sales and fitting count. However, the data also reveal a large number of observations where the independent variable (fitting count) equals to zero, yet the dependent variable (sales) varies considerably.This suggests that many customers purchase jeans without trying them on in the fitting room. A likely explanation is that customers who already know their size may choose to avoid waiting in fitting room queues.

Additional evidence supports this interpretation. Using transaction data, I was able to identify which payment card customers used. Many of the high-sales transactions associated with zero fitting counts can be attributed to foreign customers. This makes sense in that the dataset includes two stores, Gangnam and Hongdae, that are popular among international visitors. Regardless of the specific explanation, theses observation highlight the importance of considering dataset charateristics when selecting a research design and modeling approach.

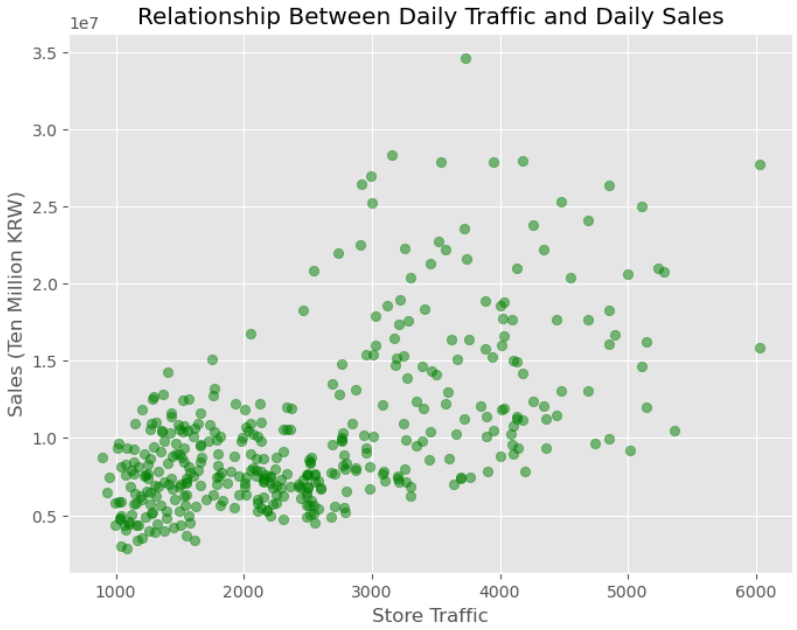

Figure 3

The third figure demonstrates the relationship between daily traffic and daily sales, not limited to jeans. I conduct this exploratory analysis to examine the impact of traffic on store performance and to find appropriate control variables in my research. Intuitively, higher store traffic should lead to increased sales, and this expectation is confirmed by the dataset. As shown in Figure 3, daily store traffic is positively correlated with daily sales. This result is consistent with prior empirical work. For instance, Olga et al.(2012) found a positive relationship between sales volume and store traffic, while also observing diminishing returns to scale with respect to traffic. Although my analysis measrues sales in Korean won rather than in sales volume, the insights provided by Figure 3 are qualitatively similar to those reported Olga et al. In conclusion, In conclusion, I include daily traffic as a control variable to address potential endogeneity concerns.