

Dear [CEO],

Thank you for your patience, and for entrusting me with this task! I'm happy to report that after analyzing the company's data from the past year, I was able to determine numerous avenues for improvement and growth. Below are the key takeaways; please see the attached file (analysis.html) for more details.

- There is a lot of variation in profit margin among “small” stores with low revenues and low costs. “Large” stores with higher revenues and costs, on the other hand, consistently achieve a high profit margin of around 30%. As such, it may be promising to expand small stores that are underperforming to obtain this profit margin boost.
- While high costs are generally necessary to achieve high revenues, the relationship is a lot murkier for rental cost (the correlation coefficient of gross revenue with rental cost is only about 0.66, as compared to 0.96 for fixed cost and 0.99 for variable cost). When planning a new store, finding a low-rent location is one of few opportunities to reduce costs without hurting revenue much. Similarly, when cutting costs at existing stores, we should first investigate unnecessary fixed costs before trying to reduce variable costs.
- Through both modeling and comparing top performing stores with the worst performing stores, I've found that having a greater number of products is associated with higher profit margins, and that certain states (New Jersey and Georgia) and months (January and November) are associated with lower profit margins. This can help inform where new stores should be opened, when marketing campaigns may be most impactful, and how many products stores need to have for the best odds of success.
- With slightly more detailed data, such as store addresses or the components of the fixed and variable costs, it would be possible to achieve even more powerful results. Additionally, now that we know of several potential adjustments, small amounts of experimentation would allow us to determine causality and achieve greater improvements to profit margin. Though my suggestions above will be a great start, I'd suggest exploring these additional data, and I would be happy to discuss the possibilities with you further!

Sincerely,

Michael Yin

Data Scientist, [Company]