NORTHWIND DATABASE

RAYMOND WILLEY

The Northwind Database project started with a simple question: *Do discounts have a statistically significant effect on the number of products customers order? If so, at what level(s) of discount?* It was also expected that additional queries and hypothesis tests be performed to gain further insight to the data. I wanted to take this a step further, and see how we can turn this task into an overall business strategy.

UNDERSTANDING THE PROBLEM

| Scenario	Questions
• Northwind is a global company that sells its products worldwide and is trying to figure out how to increase their sales volumes. Recently, they have experimented with discount offerings, and have provided access to a database of information. The firm wants to know how effective discounts are at increasing volume.	• At what levels do discounts provide a statistically significant impact on volume (items sold)? • Are there any other factors that have an impact on sales volume that can be used to improve the number of products sold per order? • The default assumption in all scenarios is that no factor significantly impacts volume.

As a global company that sells its products worldwide, Northwind wants to boost its sales volume, and using discounts is just one approach that can be used to do that. In addition to discounts, we will also explore other possible factors that may impact volume. Before looking at the specific questions we will be trying to answer, it should be noted that the default assumption in each scenario will be that the relevant factor has no significant impact on volume, and the statistical significance of this assumption will be measured. Our target variable in each scenario will be number of products sold.
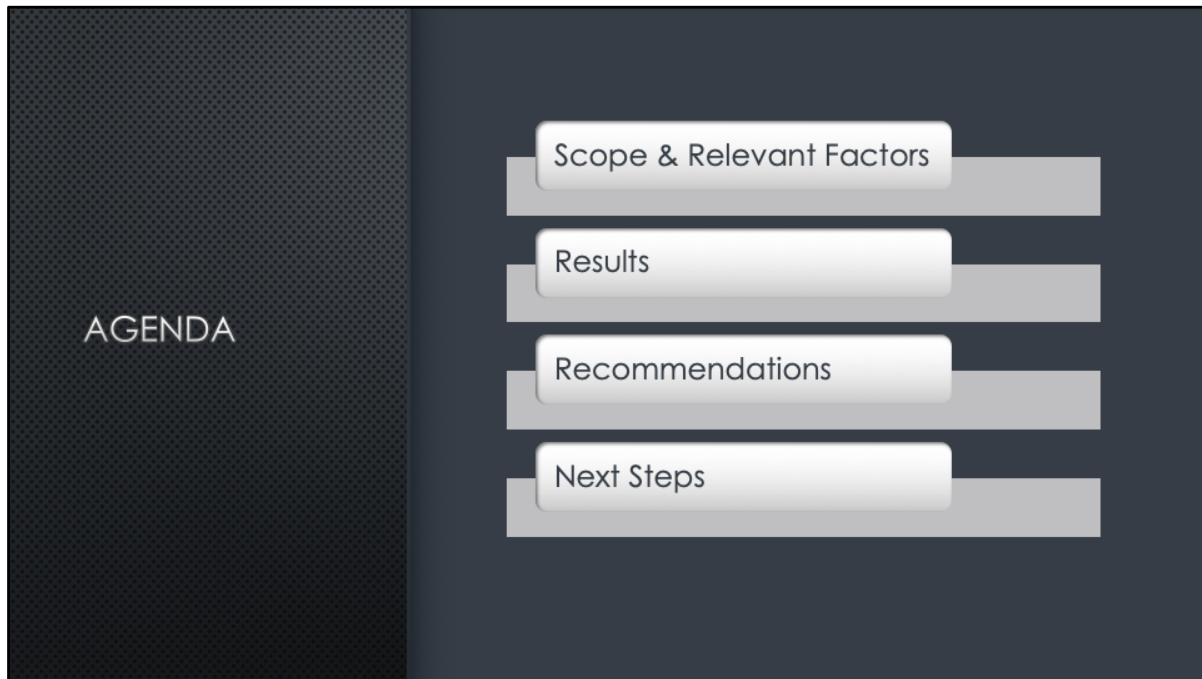
Here is a quick overview of the items we will be reviewing today. First, we will have a look at the scope relevant factors being considered. For each item, we'll have a look at the results, as well as insights we can take away from each. Then we'll put all of it together to make a few recommendations from a strategy perspective. And finally, we'll look at some options for next steps to further the analysis and create even more value.
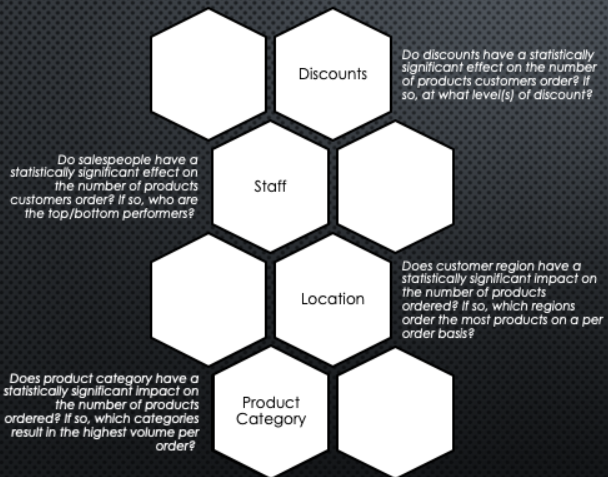
In addition to discounts, the following factors were considered for their potential impact on number of products sold per order: staff, location, and product category. With the exception of product category, all factors were found to have a statistically significant relationship with volume. Without such significance, nothing else remains to be said for product category, so we'll continue with the analysis of the other three, starting with discounts.
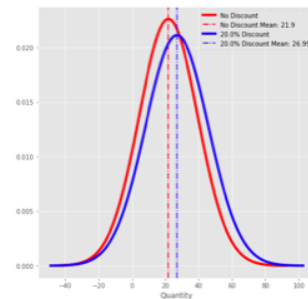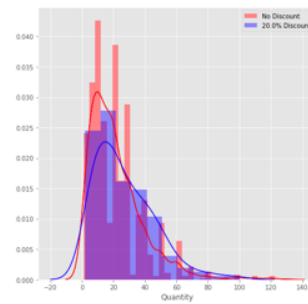
The first thing to be aware of with regard to discounts is that the level of impact can vary significantly depending on the level of discount being offered. Typically, we see the average number of products ordered increase by anywhere from 15% to 31%. However, the distribution around the average number of products ordered also widened, which is an indication that the consistency in effectiveness of discounts varied from order to order.

So, what we want to do is identify the level of discount which increases volume at a level more than the discount itself, and which yields the most consistent results. To do this, we want to look at the risk adjusted return on volume. By treating the level of variability as a measure of risk, and the volume percentage increase as return, we can create a measure that accounts for both volume and consistency. While discounts do have a meaningful impact at the 15% level and above, it is at the 20% level that the risk adjusted return is maximized.

The volume increase in sales at this point increases by 23%, while variability increases by a mere 7%, resulting in a risk adjusted volume return of 15%.
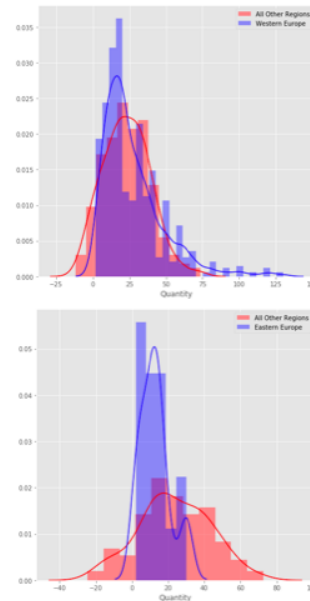
So, we know that discount effectiveness is maximized at the 20% level, but this doesn't necessarily mean such a discount should be applied globally. The factor that has the strongest relationship with volume is the region. Western Europe has the highest volume levels on a per order basis, but it also has the greatest level of variability in volume. On the other end of the spectrum are Central America and Eastern Europe, which have consistently lower volume per order.

By focusing discounts in Western Europe, overall variance in volume levels can be expected to increase, but if we assume that the effectiveness of discounts on volume is not affected by the existing levels, there is also the potential for the greatest, absolute level of volume per order. However, this may not be a safe assumption to make, and more data would be required to properly test it.

Increasing volume and variability in Central America and Eastern Europe by implementing discounts are likely to increase both variability and volume in these regions as well, without having to worry about the assumption mentioned earlier. In addition, there is the potential for offsetting business cycles with Western Europe, which could mitigate the overall increases in variability globally, despite regional increases. So here, we have the potential for greater volume, greater diversification of customer base, and more overall stability.
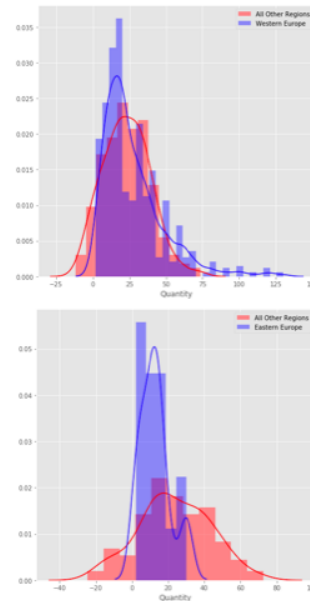
The last thing we're going to have a look at is volume performance by salesperson. It was found that the top performing volume salesperson was King7 (employee ID). His volume output is 18% greater than the rest of the company, though the results are not quite as consistent. On the other end of the spectrum, we have Suyama6, who has less volume, but less variability in the results. She is consistent.

There are many ways this can be interpreted, but if King is able to achieve greater levels of consistency by learning from Suyama, and Suyama can boost her volume by learning from King, then it may be possible to develop a training strategy that increases stability and volume across the firm.

When we put all of this together, a simple 2-part strategy becomes apparent.
1. Implement discounts at the 20% level in Eastern Europe and Central America to increase volume, while minimizing the impact on variability.
2. And partner King & Suyama together in an attempt to build a learning & development strategy that improves performance across the firm.

Now, before we finish, I would like to go over a couple next steps with this analysis that will help to refine the strategy.

1. The first thing we want to do is perform a regression analysis that explores the interaction of each of these factors. This would allow us to take into account the ways in which the factors interact with each other. In particular, do certain regions already receive a disproportionate share of discounts? If salespeople are assigned to specific regions, and if discounts have been provided disproportionately, performance may have less to do with the quality of the salesperson than it would seem at first glance. We need to account for these sorts of factors before implementing any strategy.
2. In addition, it would be a good idea to perform an analysis from not just a volume perspective, but a revenue perspective. Is the goal to maximize revenue or volume? Ultimately, we of course know the goal is to maximize profit, but without access to cost data, this will not be feasible.

Clearly, more work needs to be done, but we do have a good sense of direction to start building a strategic framework. And continued analysis on these points will allow us to more effectively guide the evolution and development of that strategy.

Thank you.