Question 19.1

Describe analytics models and data that could be used to make good recommendations to the retailer. How much shelf space should the company have, to maximize their sales or their profit?

Of course, there are some restrictions – for each product type, the retailer imposed a minimum amount of shelf space required, and a maximum amount that can be devoted; and of course, the physical size of each store means there's a total amount of shelf space that has to be used. But the key is the division of that shelf space among the product types.

For the purposes of this case, I want you to ignore other factors – for example, don't worry about promotions for certain products, and don't consider the fact that some companies pay stores to get more shelf space. Just think about the basic question asked by the retailer, and how you could use analytics to address it.

As part of your answer, I'd like you to think about how to measure the effects. How will you estimate the extra sales the company might get with different amounts of shelf space – and, for that matter, how will you determine whether the effect really exists at all? Maybe the retailer's hypotheses are not all true – can you use analytics to check?

Think about the problem and your approach. Then talk about it with other learners, and share and combine your ideas. And then, put your approaches up on the discussion forum, and give feedback and suggestions to each other.

You can use the {given, use, to} format to guide the discussions: Given {data}, use {model} to {result}.

One of the key issues in this case will be data – in this case, thinking about the data might be harder than thinking about the models.

Answer:

There are 3 main questions asked in the homework prompt here:

A) How much shelf space should the company have, to maximize their sales or their profit?

B) How will you estimate the extra sales the company might get with different amounts of shelf space?

C) How will you determine whether the effect really exists at all?

To look at how to solve these questions, we will split the assignment up into different parts and use different analytical models at each step.

Step 1: Remove seasonality and random variance to obtain the average units of a product sold on a weekly basis

Given: {time series data, units sold of a product}

Use : {Exponential smoothing}

To : {remove random variance and seasonality from a product's sale volume}

Our ultimate end goal will be to use an optimization model to determine how to use shelf space. For a year-round retailer, there will be products such as Halloween costumes or greeting cards that will have major demand spikes during certain seasons. If we do not use Exponential smoothing to remove seasonality and random variance from a product's sales volume, the high sales will skew the eventual optimization model to favor shelf space for these products even during times that they will not be in season and hence we will miss out on using valuable shelf space for better products during those seasons to sell.

Step 2: Optimize the shelf space allocation for each available product in the store

Given : { Exponentially smoothed sales volume from step 1, product price, area unit per shelf, total shelf

space in store, product name, product surface area per unit, product profit per unit sold }

Use : {Optimization Models}

To : { maximize the utilized shelf space and the profit }

In order to answer the first question "how much shelf space the company should have to maximize profit" using an Optimization Model is the obvious option. By feeding in exponentially smoothed data, we will have unbiased results that do not place unfair weights on seasonal items. Our objective function should consider the total number of products, number of units to be sold, and the areas (of both the products and the shelfs). We will also need to consider constraints such as the minimum and maximum Shelf space for each product (to be within a particular threshold) and the total shelf space used. The output of this step will show how many units of each product will be needed to stock the store considering all area restrictions.

We can further optimize our model by using a constraint of having the total shelf space to be used to be just 90% of the total shelf space available, and then again with 80% and so on down and compare how much total sales are affected when less than the total shelf space available in the store is used. This is important to know because if for example sales using 80% of total shelf space are equal to using 100%, then the profits would be less using 100% because we are purchasing inventory that is not being sold and hence we could eliminate this inefficiency. This helps give some insight into questions B and C from the main homework prompt questions above, but for further investigation those questions we will look at steps 3 and 4.

Step 3 : Identify complementary products for bundling in order to increase sales of both products

Given : {Products offered, products historically purchased}

Use : {K means clustering}

To : { identify cluster groups of high value paired items that benefit from shared space }

A clustering algorithm is a wonderful way to tackle the problem of finding pairs of products usually purchased together and labeling them as high value. By using K means clustering, specifically, we can define the distance criteria as based on if a certain number of customers have historically bought the same two products together in the past. Example: if a customer bought cereal and milk and peanut butter and jelly they can be placed in the same cluster as a high value pair and can then be bundled together to try to increase sales.

These answers the second question of "How will you estimate the extra sales the company might get with different amounts of shelf space?" more in depth. We can re-run our optimization algorithm from step 2, now using the clustered data we have created here in part 3. Since this clustered data uses different amounts of shelf space, through the use of bundled products, the new output can give us an estimate of the extra sales the company might get.

Step 4: Validation test to see if the combining products hypothesis leads to increased sales.

Given : { clustered high value pair items from step 3, optimized data from step 2 }

Use : {Greedy Multi-Arm Bandit Experimental Design (Bayesian Hypothesis Testing)}

To : {answer if the effect really exists and determine if sales were increased by the hypothesis}

Now that we have optimized data from step 2 and clusters of high value pairs from step 3, we can put together multiple stocking plans for stores. The first using the straight optimization plan from step 2 and subsequent ones making use of all or some of the high value pair bundling recommendations from our clustering algorithm in step 3. We can use a multi-arm bandit experiment and set up different stocking Options in stores and see how sales are impacted comparatively to test our hypothesis.
This has the advantage of providing value to the customer right away, while continuing to gather information about which type of stocking provides the best results.
For this reason I prefer this approach over doing a simulation where we compare all options theoretically and then select 1 to just implement as it causes a delay in selection where the retailer misses out on value. It is important to consider though that other factor such as demographics of customers, store location, etc may also be at play when doing a multi-arm bandit across multiple locations for a retailer (and not SOLELY the way we stock). However, for the purposes of this case it was given that we should forget other factors and focus solely on the analytics and hence I believe this approach would work best under these constraints. If we put those other factors back in play, I would advise that even after gaining some data from the multi-arm bandit, to tread cautiously before converting all stores to the stocking plan that is performing the best and be open to converting back if a plan in one store does not necessarily perform at the same level at another.