ISYE-6501 – HOMEWORK WEEK #14

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.

Solution:

Here is a list of potential useful data needed to solve the task:

1. Amount of shelf space required per product
    a. Minimum
    b. Maximum
2. Total daily revenue
3. Daily revenue (gross profit margin) generated by individual products
4. Daily revenue (gross profit margin) per square foot
5. Market basket (list of items in each customer's basket)
6. Retail space
7. Number of products available
8. Daily total sales
9. Food category
10. Inventory
11. Number of items per basket
12. Pay-to-stay fees
13. Weekday (Y,N)
14. Weekday (Monday through Sunday)

Hypotheses:

1. More self-space results into more sales
2. More sales result into more complementary sales
3. Larger effect if complementary products are placed in adjacent selves

Hypothesis testing -

- Using sales data to confirm Hypothesis 2 and 3
- As limited data was available for 1,
    - Used A/B testing
    - Change detection
        - Seasonality
        - Trend
        - External factors
    - Use exponential smoothing if multi-year data is available
        - This showed correlation, but not causation

New data set

- Tracking Cameras
    - Use logistic regression to match images (visual data) and then,
    - Use optimization for maximum probability matching based on logistic regression output
- Analytics VS Privacy
    - Camera tracking and associating that data with credit card etc. Will open privacy concerns
    - Keep ethical issues in mind

Approach:

1. Clustering Model - Distance between two products in the store is inversely proportional to their sales together (more distance = less sales)

Given...	Use...	To...
Sales and store Data	Clustering models	Correlate sale of complementary products

2. Community Finding Model (Louvain) - The first step is a "greedy" assignment of nodes to communities, favoring local optimizations of modularity between complementary products. The second step is the definition of a new coarse-grained network of products and shelf spaces, based on the communities found in the first step. These two steps are repeated until no further modularity-increasing reassignments of communities are possible.

Given...	Use...	To...
Cluster information from above	Louvain	Finding product and shelf space correlations and confirming the hypothesis earlier

3. Optimization model

Given...	Use...	To...
Sales and store Data	Discrete stochastic simulation model in ARENA	Maximize sale of pairs of complementary products under the given constraints

Once clusters are defined, use optimization to group products together on shelf space. Then locate the shelfs on the store floor for maximum sales benefit.

Conclusion

Once we complete building each of these components, we need to put the solution approach under test by piloting the end-to-end solution to a small set of stores that are diverse in size (Shelf space) and product selection (many known and potential pairs of complementary products). We should set target that must be met in order for the solution to be acceptable. Only upon successful completion of this trial/pilot that can be objectively measured and demonstrated, we should roll out the approach to the larger enterprise (all stores, or all stores in one country, as there may be additional variability in data for other locations). I suspect that during the pilot phase we will encounter many issues related to various aspects of data, assumptions and modeling that will need to be tweaked.

Overall, I enjoyed doing this assignment as its open-ended nature allowed me to think broadly in terms of all the knowledge I have gained during this course. It also posed a real-world challenge for the first time and I could see the various steps it takes to get to a final solution.