Homework 13: ISYE 6501 - Introduction to Analytics Modeling

Question 19.1

Prompt

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

A framework should be developed to structure the problem focusing on balancing constraints, objectives and real-world variability. The framework is a step-by-step process that includes:

- Linear Regression Modeling
- Optimize Shelf Space Allocation with Linear Programming (LP)
- Measure and Validate Recommendations

Regression for Shelf Space and Sales

A linear regression model can be utilized to determine the relationship between shelf space and sales. The goal of using linear regression is to determine how changes in shelf space for a product impact sales while controlling for other factors that may influence sales.

The data will include independent variables or predictors, which is the shelf space allocated to each product type. This could be either sq. feet or the number of shelves. The dependent or response variable would be the sales volume or revenue for each product type in either daily or weekly sales.

Data that must be considered are product attributes such as price, profit margin, and category. Store attributes such as location, total foot traffic and demographics could be considered. Additional external factors could also be included such as seasonality, holidays, or regional preferences. One must consider in the model that shelf space and sales data are available for the same time periods. Variables must also be normalized and scaled if needed. If there is any missing data, removal or imputation may be included.

The linear regression model can be formed considering:

$$Sales_i = a_0 + a_1(Space_{shelf,i}) + a_2(Price_i) + a_3(Category_i) + \cdots + \epsilon_i$$

where coefficients are determined for each predictor and error is included with the error term ϵ_i.

The linear regression model is then fit and the output should be checked. P-values or the coefficients for each predictor should be statistically significant, where $p < 0.05$. R^2 values could be check to determine how well the model explains the variability in sales. The results will quantify the marginal sales increase or decrease for each additional unit of shelf space. Check whether the retailer's hypotheses about space-sales relationships are valid and statistically significant.

Optimize Shelf Space Allocation

Shelf space allocation should be optimized with a linear programming (LP) model. Data would include historical sales and profit data per product type. LP can be used to maximize profit while considering constraints.

- Constraints:
  - Minimum and maximum shelf space per product type.
  - Total available shelf space.
  - Utility or profit objectives: maximizing total revenue or profit.

The decision variable is x_i to represent the shelf space allocated to product type i. The objective function is to maximize total profit across all product types, where the profit contribution of product i depends on its allocated shelf space.

$$Max : Z = \sum_{i=1}^{n} p_i \cdot f(x_i)$$

The max function includes p_i or the profit per units of sales per product i. The sales function based on shelf space for product i is defined by $f(x_i)$, which is modeled as a linear function. If linear regression is not optimal, the sales may have negative returns with shelf space. This could mean that $f(x_i)$ should be reconsidering, where maybe a piecewise linear function will solve the LP. Constraints should include shelf space for each product, total shelf space, and non-negativity constrains.

Results would include allocated shelf space, determining the optimal space for each product that maximizes profit. Sensitivity analysis can be performed to test how changes in parameters impact the solution. Iteration and validation will need to be performed. After implementing recommendations, collect data to validate whether the profit objectives were satisfied and adjust the model parameters as needed based on the observed results.

One challenge may be the relationship between sales and shelf space. Historical data might not always clearly reflect the impact of shelf space on sales due to confounding factors. For example, promotions or seasonal variations might skew data. Advanced methods like causal inference can help isolate effects.