In order to find an efficient and accurate solution to this problem, I will employ the use of collaborative filtering, clustering, and optimization to determine the best course of action for the retail company. First, many products are sold on a cyclic basis and thus we must account for this. I thought of using exponential smoothing however given that this was a forecasting method, I was not sure how to include it into the plan for optimization. So, I decided that by being given sales data over the past 3+ years, I will divide the datasets up into groups of three based on the month. For example, the first group would be all sales data within the months of Jan-Apr, then May-Aug, and finally Sept-Dec. Imagine for each of the below steps, they are run on each different subset of data.

The problem must be broken down into steps that can be solved with an analytical approach:

1. Determine the groups of products that are frequently bought together. This will be done based on the set of factors below:
    a. Given:
        i. Sales data over the past 3+ years, each data point detailing every item sold on that purchase.
    b. Use:
        i. Collaborative Filtering.[1]
    c. To:
        i. Obtain a response matrix per product that will detail 1 if the product is frequently purchased with the product that corresponds to the row label, and 0 if not.

Rationale:

The rationale behind this approach is that by using collaborative filtering, we can mathematically determine correlation between each product sold and the most common items sold alongside it. I have determined the collaborative filtering model to be the best model to use because of its simplicity to explain to the retail company, and its relevance to the problem at hand.[1]

2. Grouping of the products based on output matrix from collaborative filtering (step 1).
    a. Given:
        i. A product matrix as described in part c of step 1.
    b. Use:
        i. K-Means Clustering Model.
    c. To
        i. Group frequently bought products together to ensure that we keep products together to maximize profit in the optimization step.

Rationale:

The idea to use K-means clustering to geographically group similar customers will be very helpful for the later step of optimizing the routing problem[1]. In order to *k*, we will use the value of 2 for the maximum size of each cluster, which means that the number of clusters *k* will be equal to the number of products divided by 2. Grouping these products is important because it will help the optimization step factor in the distance between shelves of items frequently bought together, and thus maximize profits.

3. Optimize the shelf storage of the store
    a. Given:
        i. A dataset where each product has the predictors of quantity sold, profit*shelf space previously taken up (as a %), dimensions (size), total shelf space, and cluster labels from step 2.
    b. Use:
        i. An Optimization Model.
    c. To:
        i. Maximize the total profit.

Rationale:

The use of this optimization model would be to maximize the profit based on the hypothesis that items that are bought more frequently being placed closer together will result in increased profits. By using data from the clusters generated in the previous step, this step combines the frequently purchased products and optimization of the shelf space the retail

company should run to ensure the most profit maximizing solution. The following considerations are taken when performing the optimization model:

Objective Function:

- Maximize profits

Decision Variables:

- Quantity sold, profit*shelf space previously taken up (as a %), dimensions (size), total shelf space, and cluster labels from step 2.

Constraints:

- Minimum shelf space and maximum shelf space, as well as total space in store.

Discussion:

I chose this method of deducing which products to place where because of its interpretability and relevance to the problem. By this I mean that the use of collaborative filtering was chosen because it would be much easier to explain to the management at the retail company, and it has proven itself useful in other recommender systems.[1] I also used clustering before optimization as it would prioritize the products often sold together. Furthermore, in order to test the hypothesis of whether the store believes putting products closer together will result in extra sales, you can simply rerun the optimization step without the cluster labels predictor and compare the two profits.

References:

1. https://developers.google.com/machine-learning/recommendation/collaborative/basics