

Given:

- data on the sales of each product type in the store
- the amount of shelf space currently allocated to each product type
- We can also collect data on the minimum and maximum shelf space requirements for each product type
- Data on the shape and dimensions of all of the products
- All historical sales data to try and find similar products bought together
- Data about storage needs for particular types of products (freezer, refrigeration, etc.)

Use:

- **Clustering model:** analyze historical data to try and find which products cluster together
- **Regression Model:** build a model to estimate the extra sales that the company might get with different amounts of shelf space allocated to each product type. One possible model is a regression analysis, where we regress sales on the amount of shelf space allocated to each product type while controlling for other factors such as price, promotions, and seasonality. We can estimate the coefficients of the regression model to determine the marginal effect of each additional unit of shelf space on sales.
- **Optimization:** To estimate the optimal division of shelf space, we can use optimization to find the allocation that maximizes sales or profit subject to the constraints on minimum and maximum shelf space requirements for each product type, shape of each product type, the total amount of shelf space available in the store, and the maximum amount of each product available

To:

- **Validation:** To validate our model, I would run geo tests where I apply the recommendations of the models in some geos and do not in other geos. The control and test geos will be determined in a pre-analysis where we try and make the two cells as equal as possible. That way, if the recommendations are correct, we should see more sales in the test geos where we implement the recommendations.