**Question 19.1**

**Summary**

I propose a three-step solution to this question.

1. Hypothesis Testing to test whether the given hypothesis are correct
2. Clustering to divide products into groups of complementary products
3. Optimization for grouped values

**Description (general format)**

1. **Clustering**

**Given:**

1. market data on sales of individual data
2. current storage space distribution (amount)

**Use:** Clustering

**To:** Find complementary products groups

1. **Evaluating correlation within groups**

**Given:**

1. Output from clustering model
2. market data on sales of individual data

**Use:** Regression with threshold

**To:** evaluate strength of the correlation within groups of products, using a threshold to classify between “complementary” or not, resulting in -the “value” of a group

1. **A/B testing**

**Given:**

1. Complementary products output from previous step
2. market data on sales of individual data(grouped)

**Use:** A/B test by varying shelf spaces of grouped products

**To:** test hypothesis of causation from shelf space and sales

1. **Optimization for best value of groups**

**Given:**

1. Complementary Groups
2. Values of groups

**Use:** optimization

**To:** determine best space distribution

**Potential Problems**

**Data:**

-this model was modeled under circumstances of one store, ignoring demographic difference between stores.

-A/B testing may be hard to imply given the amount of products needed to change each day.

- Defining products as complimentary based on sales only instead of being bought together