Task 2: Store Format for New Stores

The grocery store chain is has 10 new stores opening up at the beginning of the year. The company wants to determine which store format each of the new stores should have. However, we don't have sales data for these new stores yet, so we'll have to determine the format using each of the new store's demographic data.

Pretty sweet grocery store, right?

Task 2: Determine the Store Format for New Stores

You've been asked to:

- Develop a model that predicts which segment a store falls into based on the demographic and socioeconomic characteristics of the population that resides in the area around each new store.
- Use a 20% validation sample with *Random Seed* = 3 when creating samples with which to compare the accuracy of the models. Make sure to compare a decision tree, forest, and boosted model.
- Use the model to predict the best store format for each of the 10 new stores.
- Use the StoreDemographicData.csv file, which contains the information for the area around each store.

variables in the model.

Task 2 Submission

- What methodology did you use to predict the best store format for the new stores? Why did you choose that methodology?
- What are the three most important variables that help explain the relationship between demographic indicators and store formats? Please include a visualization.
- What format do each of the 10 new stores fall into? Please provide a data table.

NEXT