US Beer Assortment Optimization

2018 Capstone Project with Anheuser-Busch InBev
New York, New York

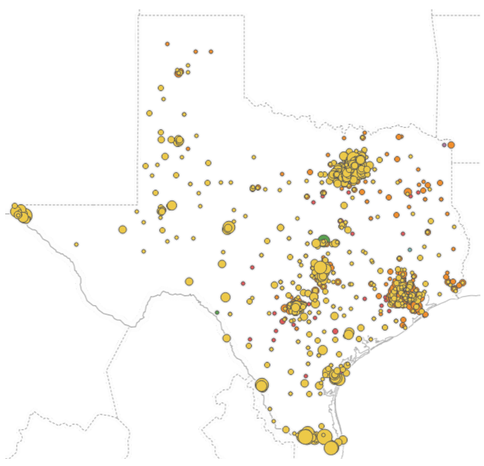
Team: Daria Brauner and Thomas Littrell
Advisor: Prof. Dimitris Bertsimas
PhD collaborators: Brad Sturt & Chris McCord
Thanks to: Brent Hayden, Emily Shapiro, Siyu Wang, and Debarshi Indra

1. Challenge

What products should we put into a store to make the most money while accounting for demand substitution?

2. Scope and Timeline

We worked with the U.S. business and focused on making assortment recommendations for chain accounts in Texas. These accounts are where we have the most data and can have the most immediate impact given that ABI representatives help chains design assortments every 6 months.



- Feb.: Project start
- Jun.: Full time in NYC
- Aug. 2nd: Presentation at RM conference
- Aug. 17th: Project end

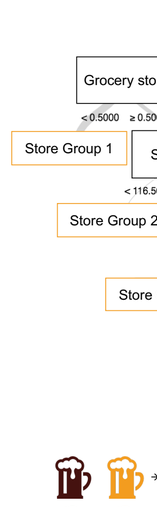
3. Data

Our dataset is a combination of internal and third party data. The main data source we used is the data provider JDA's beer retail data set, which provides sales information at the SKU level by store for around 3000 chain accounts in Texas. We supplemented the sales data with information on both stores and products. In particular, we used data provider IRI's sales data to impute price information for all products in our data.

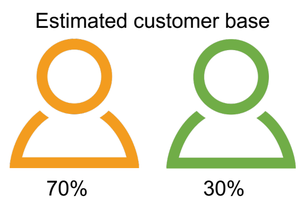
For computation reasons, we aggregated all products to product groups, providing the highest level of visibility for the largest and fastest growing brands.

4. Methodology

Our methodology consists of two pieces.[1] First, we estimate a consumer preference model to capture substitution patterns between products. Second, we use the consumer preference model as an input to a constrained optimization that searches over all the possible assortments and picks the one that gives the most expected revenue while respecting business constraints.



1. We assume that customers have a ranking, which they come to for any reason, of all products and buy their most preferred product from all the available products in a store.



2. Different rankings define different customer types and we solve a linear program to find the proportion of each type in a store's customer base that best explains sales data.

Beer	Share
Beer 1	45%
Beer 2	17%
Beer 3	38%

3. Given the customer base and rankings, we can estimate shares for new assortments. Changes in shares as assortment changes captures substitution.

4. Since stores have different customer bases, we train an optimal tree that learns from the data how to best cluster stores based on demographics for the choice model.

5. A mixed integer optimization problem looks at the expected value of all assortments (predicted share times price summed over products) and picks the best one.

Business rules:
- Assortment size
- Package size
- Disallowed products
- ABI only
- Number of changes

6. Working with subject matter experts, we add additional (optional) constraints to reflect business rules and improve the usefulness of our recommendations.

[1]Model taken from Bertsimas, Dimitris, and Velibor V. Mišic. "Data-driven assortment optimization." submitted to Management Science (2015).

5. Results

The model outputs recommendations for what to swap into and out of a chain store's current assortment. Below, we show the recommendations made for different numbers of swaps for a convenience store and highlight how demand substitution affects our recommendations. After letting the model make as many changes as it wants, we also observe that relatively few assortment changes can realize most of the overall benefit.

+1%

Estimated opportunity to increase unites moved for a typical store

The Miller pack pulls demand from ABI products

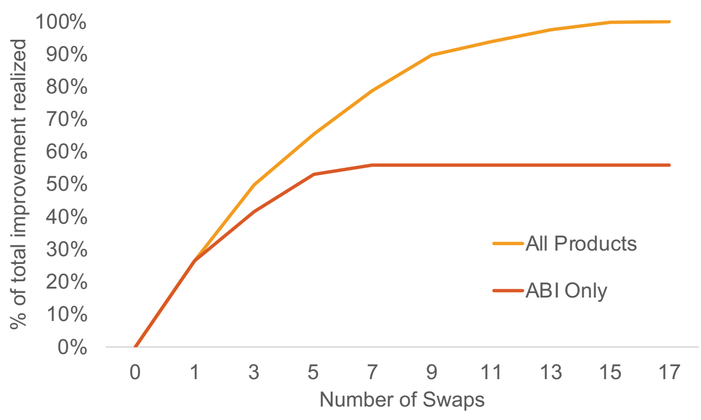
Michelob Ultra is a higher price brand

Swaps	Products to add	Products to remove
1	Bud Light large*	Miller Lite 18/12C
3	Bud Light large* Bud Light 15/16C Bud Light 24/12C	Bud Light 12/12B Miller Lite 18/12C Miller Lite 12/12B
5	Bud Light large* Michelob Ultra large* Bud Light 24/12C Bud Light 15/16C Bud Light 24/12BBB	Budweiser medium* Miller Lite 18/12C Budweiser small* Miller Lite small* Bud Light 12/12B

Demand substitution patterns show a willingness to buy larger SKUs

*Aggregate product

Number of changes to assortment vs. percent of total improvement



6. Next Steps

There are four areas in which the model can be extended:

1. Adding information on the market segment served by each chain
2. Using unit movement in the objective function
3. Using more refined space constraints
4. Accounting for inventory requirements

After refining the model further, we recommending running a field experiment with a partner chain in Texas.