Sophia Chung, Lana Huynh, Jamie Luna, Rachel Roggenkemper, James Rounthwaite
California Polytechnic State University - San Luis Obispo

Booze ‘R’ Us

13th October 2023

OVERVIEW

This proposal aims to analyze and forecast liquor store sales by creating two models: an aggregated inventory model that examines overall purchasing patterns and a more specific alcohol category inventory model to provide insights for informed decisions on sales forecasting and potential expansion.

DESCRIPTION OF DATA

For this project, we will be working with the Iowa Liquor Sales dataset.

Since there is no metric for how much sales the liquor stores have, we have to infer their sales. Our approach to accomplish this is to observe the trends in the amount of alcohol bought from the liquor stores. One way to do this is by aggregating the alcohol bought by each store and looking at the monthly trends. We will need to cross-reference other years to account for month-to-month variation. For example, an increase in alcohol bought from November to December does not necessarily mean the store is increasing their sales as a whole; the jump could just be attributed to seasonal fluctuations.

In order to make this dataset more manageable, there are several alterations we have made. We have added a Month and a Year column so that we have the option to use these as model predictors rather than only the date as a whole.

Additionally, for this report we chose to use one store in Iowa to act as a proof of concept for our analysis methods. The full model will use all of the data and stores available.

We removed certain columns from the dataset because they would not be useful to our analysis. Columns we removed included the computed region columns that contained miscellaneous information, the coordinates of the stores which we decided against keeping since we are already taking into account the name of the store and its city and county, the item description

since there are alternative product identifiers we believed would be more useful like category and item number, and state bottle cost which is the amount that the Alcoholic Beverages Division paid for each bottle and sale gallons which is the volume of liquor ordered because we felt like this information was not important to our analysis since we are keeping other variables that contain similar information so it would be redundant to keep both.

Table 1: Number of bottles purchased by Alcohol Category

Alcohol Category	Count
Straight Bourbon Whiskies	11092
Canadian Whiskies	10123
American Vodkas	7562
Imported Vodkas	7561
100% Agave Tequila	7423
Scotch Whiskies	5864
Vodka 80 Proof	5653
Imported Brandies	5582
American Brandies	5178
Single Malt Scotch	5148

Table 1 shows us the 10 most popular alcohol categories by bottle count from the years 2012 to 2023 for one particular store.

Figure 1. Bottles of Alcohol Bought from Vendor by Central City Liquors

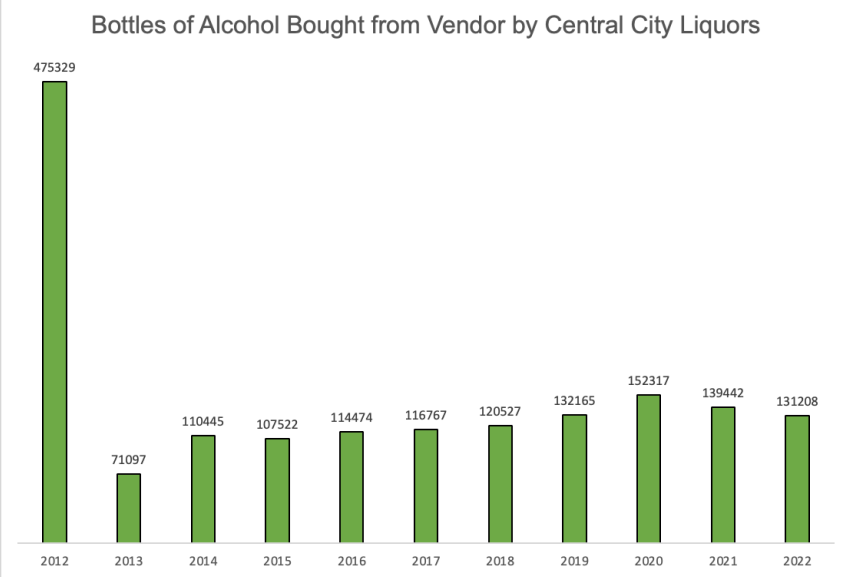

Figure 1 shows us a visual representation of the amount of bottles of alcohol bought by the liquor store. We do not currently know which stores Booze 'R' Us owns, but if they were to give us the names of their stores, we will easily be able to modify our model and graphs to adjust to the stores they own.

DESCRIPTION OF MODELS

Since we don't have sales as a feature on this data, to fit the models, we will be inventing a metric to represent the sales. Our strategy to achieve this involves monitoring the purchasing patterns of alcohol from liquor stores. We can do this by summing up the alcohol acquired by each store and examining the monthly patterns.

Aggregated Inventory Model

The first model we plan to fit is an aggregated inventory. When we talk about "aggregated inventory," we are referring to the collective sum of all inventory items, regardless of their type or category. This holistic view provides a broad perspective on the overall inventory status.
To gain insights into the financial aspects of inventory management, we will track the trend of the cost associated with the inventory. This means we'll be analyzing how much is being spent on inventory over time. By observing these costs on a monthly and yearly basis, we can identify patterns, such as seasonal fluctuations or long-term growth or decline. This data can be instrumental in making informed decisions about making accurate predictions on sales forecasting.

Alcohol Category Inventory Model

The second model we plan to fit is an alcohol category inventory. The "alcohol category inventory" is a more specific subset of the aggregated inventory. Here, we are focusing solely on items that fall under various alcohol categories, such as wines, beers, spirits, etc. Just like with the aggregated inventory, we aim to track the financial trend associated with the alcohol category. However, the distinction here is that we will be breaking down the costs by each specific alcohol category. This granular view allows us to see which categories might be more costly, which are more popular, and which might require more attention in terms of stock replenishment. By observing the trend of the cost of inventory for each alcohol category, ordered by month and year, we can pinpoint specific categories that might be driving up costs or those that are more profitable. This detailed analysis can guide decisions related to making accurate predictions on future sales.

In summary, while the aggregated inventory gives a bird's eye view of the entire inventory's financial trend, the alcohol category inventory dives deeper into specific categories, offering insights that can lead to more targeted and effective inventory management strategies. With both models, the model will take in various factors, and will output inventory predictions that can be used to decide whether there will be enough income to expand operations.

CONCLUSIONS

Overall, there is not an exact method to forecast liquor stores sales. We will work around this issue by examining the trends in the cost of inventory ordered by the liquor stores. We can use inventory/cost of inventory to infer sales, but we cannot make an exact claim. For example, if we observe that Liquor Store A has increased costs of inventory from November 2012 to December 2012, and there is also a jump from November 2012 to November 2013, we can reasonably infer that Liquor Store A is doing well in sales. Thus, “Booze ‘R’ Us” should consider expanding their operations.

We also want to take a closer look at the trends in different alcohol categories as well. Aggregated data does not always capture details at the granular level that we would like, so it is important to separate this out.