Project Plan

Anthony Rosamond[1], Madison Brown[1], Jade Towgood[1], and Jasmine Gomez[1]

[1]Arizona State University, Tempe, AZ 85281, USA

September 15, 2024

1 Background

Alcohol has a major role within social and economic issues across many regions. Within the State of Iowa, feelings towards alcohol and its sale has varied across the decades. However, the state's current high alcohol consumption rates have prompted this investigation into liquor sales across the different regions of the state. As of 2023, the state of Iowa has 100 breweries, 120 wineries, and 45 distilleries [6]. Alcohol is also readily available within retail locations, with individual sale transactions provided through the state of Iowa [4]. Reports from Penn State Extension provide an overview on changing consumer preferences within the alcoholic beverage market [5]. Generally, generational changes also reshape consumer behaviors but there is also the unique influence from the COVID-19 Pandemic which forced a change within consumer habits [2, 1].

Motivation

The motivation for this research stems from the need to understand how societal and lifestyle changes are reflected in consumer behaviors, particularly through liquor sales. In recent years, factors such as shifting social attitudes toward alcohol consumption, the rise of health-conscious trends like the "sober curious" movement [2], and significant events like the COVID-19 pandemic [1] have influenced the ways people purchase and consume alcohol. In Iowa, these changes are mirrored in liquor sales data, which can serve as a valuable indicator of broader societal trends. We aim to equip policymakers, public health professionals, and businesses with the insights they need to make informed decisions that reflect the evolving dynamics of Iowa's population.

Definition of the Research Problem

This research seeks to answer the following question: *How can analyzing liquor sales trends provide insights into broader lifestyle changes and shifts in Iowa's population?* Specifically, the study will focus on uncovering connections between liquor sales data and factors such as demographics, economic changes, and cultural movements [5], allowing us to better understand the underlying societal shifts in the state.

2 Research Question and Hypothesis

2.1 Sub-Question 1: How have external events, like the COVID-19 pandemic, affected the types and quantities of alcohol purchased?

The COVID-19 pandemic has significantly altered both the types and quantities of alcohol purchased, with a shift towards spirits and wine, and a decrease alcohol purchases in bars and restaurants due to social

restrictions. Consumers likely increased their purchases of alcohol for home consumption, reflecting an overall rise in retail liquor sales, as they shifted from social drinking in public spaces to private consumption at home.

2.2 Sub-Question 2: What role do demographic variables, such as age and income, play in influencing liquor consumption patterns?

Demographic factors, such as age and income, play a significant role in shaping liquor consumption patterns. Younger adults (ages 21-35) and individuals with higher disposable incomes are more likely to purchase premium liquors and experiment with new types of alcohol, such as craft spirits and organic wines, while older adults may show more loyalty to traditional categories like beer and mainstream liquor brands. Lower-income individuals, on the other hand, are likely to purchase lower-priced alcohol, such as inexpensive beer and spirits.

3 Datasets

3.1 Iowa Liquor Sales Data

This dataset [4] is published and maintained through the Iowa Department of Revenue, Alcoholic Beverages and contains data starting from January 1, 2012 to currently. It is typically updated the first of every month. As of September 1, 2024, there are 29.9M instances, individual product purchases, in this dataset with 24 different variables. The 24 variables are as follows: Invoice/Item Number, Date, Store Number, Store Name, Address, City, Zip Code, Store Location, County Number, County, Category, Category Name, Vendor Number, Vendor Name, Item Number, Item Description, Pack, Bottle Volume (ml), State Bottle Cost, State Bottle Retail, Bottles Sold, Sale (Dollars), and Volume Sold (Gallons).

The variables that we will be focusing on this dataset are Invoice/Item Number, Date, County, Category Name, Item Description, Sale (Dollars), and Volume Sold (Liters). The Invoice/Item Number is the concatenated invoice and line number associated with the liquor order to provide a unique identifier for each individual liquor products in the store order. The Date is just the date of the order, month, day, and year. The county is where the store is located. Category Name is for the type of the liquor that was ordered. The Item Description is the description/name of the individual liquor product ordered. The Sale (Dollars) encompasses the total cost of the liquor ordered, which is calculated using the number of bottles multiplied by the state bottle retail. Finally the Volume Sold (Liters) is the total volume of liquor ordered in liters, which is calculated by the bottle volume (ml) by bottles sold dived by 1,000.

3.2 Iowa Census Data

Using https://data.census.gov/ we are able to gather data from multiple datasets [3] from Iowa. For this project in particular we are going to gather information about population size, race and ethnicity, income, and possibly age. It is separated by year and can be filtered through individual county, which we will be doing for this project. We will be focusing on the population size, race and ethnicity, and average income. There are 99 counties in the state of Iowa and we will be looking at the data per county.

4 Methods

For the analysis of the datasets chosen we will employ a few different techniques. For the initial exploratory data analysis we will look at the average amount of sales across all counties in Iowa along with the total volume of alcohol sold. Then we will utilize the census data as well to see if there are any obvious relationships between alcohol sales and various demographic sub-groups. The statistical portion of this will be done in python and we will use heat-maps in tableau to visually see any differences that may come up.

While the initial EDA is critical of our question on alcohol sales and demographics, we will be diving into a deeper statistical analysis for that question. We will use clustering analysis to get a deeper understanding of how this data is grouped. The clustering will be done in two parts. First we will build and train an auto-encoder that will compress the data that hopefully retains patterns from the original data. We will then apply k-means clustering to this encoded data. With these clusters we will then run some of the data the decoder portion of the algorithm in order to analyze the demographics of these clusters. Then using the census data we will take counties with a higher proportion of demographic groups and perform t-tests against the data for the entire state in order to confirm whether any differences are statistically significant. Here the purpose of the clustering is to hopefully get a much deeper understanding on how demographics play a role in alcohol sales that is beyond the basic question of "Does group x purchase more alcohol than group y".

Finally in order to analyze the affects of state policies on alcohol sales we will look at the dates that laws and/or information campaigns started and then see if we can measure a difference in sales or a different trend when it comes to the type of alcohol sold. Here we will do a T-test (compared to the data before the start date) as well in order to see if any difference is statistically significant.

References

[1] Study shows uptick in u.s. alcohol beverage sales during covid-19 pandemic, 2021. Accessed: 2024-09-12.

[2] Can the sober curious trend change u.s. alcohol consumption?, 2023. Accessed: 2024-09-12.

[3] U. S. C. Bureau. State iowa, 2022. Accessed: 09-01-2024.

[4] A. B. D. Commerce. Iowa liquor sales, 2024. Accessed: 09-01-2024.

[5] K. Kelley. Alcoholic beverage consumption and purchasing trends, 2024. Accessed: 2024-09-12.

[6] A. Kutz. 'fascinating history': Iowa's relationship with alcohol, explained, 2023.