Iowa Liquor Retail Sales

Stefan Popov, B. Sc.

1 Introduction / Business problem

In this project, I will perform an analysis of a publicly available data set to solve some of the most common retail problems. Retailers often want to forecast their demand in order to predict when and how much goods should be stocked. Second, distributors need to have clear demand in order to optimize their delivery route.

In this work, I aim to analyze the data set and gain some initial insights into how liquor is being used in the state of Iowa, and then apply some machine learning methods to better predict the stock-out of stores and the retail demand, the results of which I hope to be practically used by liquor stores and retail companies (most probably located in Iowa).

Analysis of this data set can provide insights into the habits of alcohol consumption in Iowa. They might provide liquor stores with more accurate information about their demand and help them better control their stocks; and they might help distributors optimize their delivery routes by predicting the demand days ahead and thus preventing over- or under-stocking of different types of alcohols (supply-chain problems).

2 Data

The data is publicly available both on the Iowa's government web-page at *Iowa Liquor Retail Sales | data.iowa.gov* (2020) and is included in Google's publicly released data sets on the Google Cloud Platform at *Iowa Liquor Retail Sales | Marketplace* (2020). On the former page it is freely available to download in various formats (csv, Excel, etc.) while on the latter it is integrated in BigQuery - Google's analytic warehouse tool where users can process up to 1 TB of data free of charge.

This data set contains every wholesale purchase of liquor in the State of Iowa by retailers for sale to individuals since January 1, 2012. The State of Iowa controls the wholesale distribution of liquor intended for retail sale, which means this data set offers a complete view of retail liquor sales in the entire state. The data set contains every wholesale order of liquor by all grocery stores, liquor stores, convenience stores, etc., with details about the store and location, the exact liquor brand and size, and the number of bottles ordered.

The data set is updated monthly, usually on the first day of the month, therefore it is up to date. Currently, there are 18.8 million rows with 24 columns.

Some of the attributes are geographical, such as store location, ZIP code, address, city name, county, etc. and these will be used to join or query Foursquare data to visualize some interesting maps.

Other attributes include information about:

– Vendors: name and number
– Items: number, description, bottle volume, category
– Order: store cost, retail cost, number of bottles sold, volume sold (litters and gallons),

These fields enable us to analyze the data and drill down alongside multiple dimensions, for example: (1) group sales by vendors and analyze their financial value, (2) analyze sales by date and type of alcohol consumed - this can be useful if we find some dates (holidays) when people consume some alcohol more and retailers can offer discount or special price so their sales increase.

Bibliography

Iowa Liquor Retail Sales | data.iowa.gov (2020). Last accessed 06 June 2020.
   URL: *https://data.iowa.gov/Sales-Distribution/Iowa-Liquor-Sales/m3tr-qhgy*

Iowa Liquor Retail Sales | Marketplace (2020). Last accessed 06 June 2020.
   URL: *https://console.cloud.google.com/marketplace/details/iowa-department-of-commerce/iowa-liquor-sales?filter=solution-type%3Adataset—&filter=category%3Aanalytics—&id=18f0a495-8e20-4124-a349-0c4c167b60ab*