

Regional Liquor Sales in Iowa

Christophe Hunt^{*,a}, Senthil Dhanapal^a, Yadu Chittampalli^a

^a*CUNY School of Professional Studies, Data Analytics, New York, NY*

Abstract

This is the abstract.

It consists of two paragraphs

Keywords: *Liquor, Liquor Sales.*

Problem

Liquor sales are highly variable and the objective of this report is to create a statistical model for the volume sold of liquor in gallons by region within the state of Iowa. This will help us predict inventory and assist wholesale distributors to plan for predicted volume of distribution.

Introduction

In February, the Distilled Spirits Council (DISCUS), announced that spirits had an estimated retail sales of nearly \$72 billion in 2015. Additionally, DISCUS credits the continuous growth of the distilled spirits industry to several key factors - continuous fascination with American Whiskeys in the United States and abroad, innovations in flavors, premiumization across all spirits categories leading to consumer interest, improved regulatory and tax environment resulting in expanded market access and a relatively low number of state tax threats, and the growth of small distillers, which expanded grassroots and overall interest in the spirits category Del Buono (2016).

This establishes that spirit sales in the United States is a valuable market worth exploring for a more detailed and statistical understanding of sales and volume. We hope to more thoroughly understand what impact regional and season impacts might have on liquor sales. We will limit the analysis to Iowa which has also reported sales at a record pace during the last half of 2000 Boshart (2001). While this older information we do have data up to 2016 to review.

Research Background (Literature Review)

Our goal is inventory prediction.

^{*} Author

Email addresses: christophe.hunt@spsmail.cuny.edu (Christophe Hunt), senthil.dhanapal@spsmail.cuny.edu (Senthil Dhanapal), yadu.chittampalli@spsmail.cuny.edu (Yadu Chittampalli)

Methodology

The data set contains 5 number of variables, County, Category name, Bottles Sold, Sale Dollars, Average Price Per Bottle and the dependent variable is Volume Sold in Gallons. The initial data set is very large as it includes sales by location and is very granular. The size of the initial data set has every liquor transaction from 2012 to present so it approaches ~2gb. For the purposes of this report, to analyze a data set this large is not feasible. Therefore, we reduced the number of variables and summarized to a higher level regional aggregate.

We first looked at the top 5 liquor categories for each year by number of bottles sold. In 2015, the top categories were "X" and interestingly straight burbon appears to have more sales in 2015 than 2014 which coincides with the literature of growing whiskey sales "ADD REFERENCE".

discuss the key aspects of your problem, data set and regression model(s). Given that you are working on real-world data, explain at a high-level your exploratory data analysis, how you prepared the data for regression modeling, your process for building regression models, and your model selection.

Experimentation and Results

Data Aquisition

The dataset contains the spirits purchase information of Iowa Class "E" liquor licensees by product and date of purchase from January 1, 2012 to current. The data set is provided by the Iowa Department of Commerce, Alcoholic Beverages Division, [click here](#) to view the data set at Data.Iowa.Gov.

describe the specifics of what you did (data exploration, data preparation, model building, model selection, model evaluation, etc.), and what you found out (statistical analyses, interpretation and discussion of the results, etc.).

Discussion and Conclusions

In another study conducted in 2012 in Idaho, the monthly revenue generated was examined rather than the yearly revenue generated. The continued growth was rather owed to the number of weekends a month has (five instead of four) and to the higher prices in neighboring states. In Washington, the voters approved an initiative that led the state to sell its liquor stores and add new distributor and retail fees, making prices in the neighboring states (Idaho and Oregon) look better. There were no changes made in marketing or pricing in response to the regulatory shift in Washington (???). Further research into the proximity of our counties to states and towns with higher prices and regulation may provide more insight into sales and volume of liquor sold. Additionally, reviewing the data by identifying months that has 5 weekends instead of four could provide further insights.

conclude your findings, limitations, and suggest areas for future work.

Appendices

Supplemental tables and/or figures.

Session Info

- R version 3.3.2 (2016-10-31), x86_64-w64-mingw32
- Locale: LC_COLLATE=English_United_States.1252, LC_CTYPE=English_United_States.1252, LC_MONETARY=English_United_States.1252, LC_NUMERIC=C, LC_TIME=English_United_States.1252
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils
- Other packages: ggplot2 2.2.0, pacman 0.4.1
- Loaded via a namespace (and not attached): assertthat 0.1, backports 1.0.4, colorspace 1.3-1, digest 0.6.10, evaluate 0.10, grid 3.3.2, gtable 0.2.0, htmltools 0.3.5, knitr 1.15.1, lazyeval 0.2.0, magrittr 1.5, munsell 0.4.3, plyr 1.8.4, Rcpp 0.12.8, rmarkdown 1.2, rprojroot 1.1, rticles 0.2, scales 0.4.1, stringi 1.1.2, stringr 1.1.0, tibble 1.2, tools 3.3.2, yaml 2.1.14

R statistical programming code.

Please see [Final Project.rmd](#) on GitHub for source code.

<https://github.com/ChristopheHunt/DATA-621-Group-1/blob/master/Final%20Project/Final%20Project.Rmd>

References

Boshart, Rod. 2001. "Liquor Sales in Iowa Set Record." *Gazette*.

Del Buono, Amanda. 2016. "Keeping Spirits High." *Beverage Industry* 107.4: 14–16, 18.