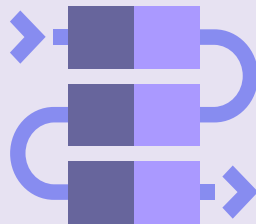


Booze 'R' Us

Business Solutions



Data Overview



Data Modeling Objective

- Predict future growth/decline in sales for liquor stores in Iowa

Data Source

- Data was provided by Iowa Department of Revenue, Alcoholic Beverages subdivision

Variables

- The data tracks the purchasing of Class “E” liquor by individual liquor stores in Iowa

Data Engineering and Modeling

Yearly Arrangement of Data

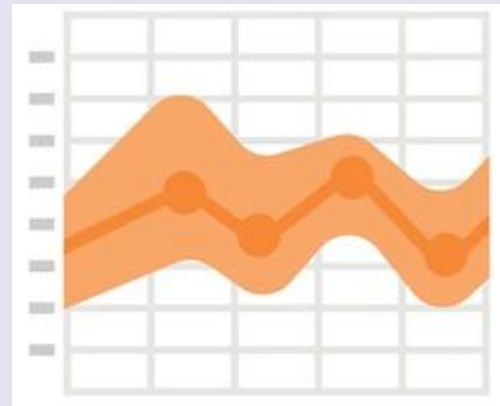
- Aggregate variables for liquor store for each year
- Use data from previous years to predict the next

Response Variable

- Predict the amount of Alcohol Purchased by a liquor store in dollars for a single year

Modeling Intuition

- The purchasing of more alcohol by a store is likely tied to how much alcohol a store is selling
- Predict a store to purchase 10% more alcohol → project a store to sell 10% more alcohol



Model Fitting, Tuning, and Validation

Definition of “Best” Model

- Model that best predicts next years total Alcohol purchased (\$)
 - Minimizes prediction error (RMSE & MAE)

Model Variable Tuning

- How many previous years of data are needed to best predict response for next year?
 - What variables should be included for each previous year of data?
- What additional characteristics for a store should be considered

Best Model

- Total alcohol purchased for the previous two years was sufficient and optimal to predict total alcohol purchased for the next year

Model Performance on Individual Stores



0.89

- Our model is able to account for **89%** of the variation in the amount of alcohol stores purchase over a year
- Standard Deviation for the yearly amount of alcohol purchased by liquor store in Iowa is \$530,000

Model Performance on Individual Stores

RMSE

\$168,000

- Measure of prediction error that punishes larger errors more
- Larger errors tend to be associated with stores with major drops in spending
 - \$3,000,000 spent to \$800,000 spent in the span of 1 year
- \$168,000 roughly equates to 0.30 standard deviations from the true value

Model Performance on Individual Stores

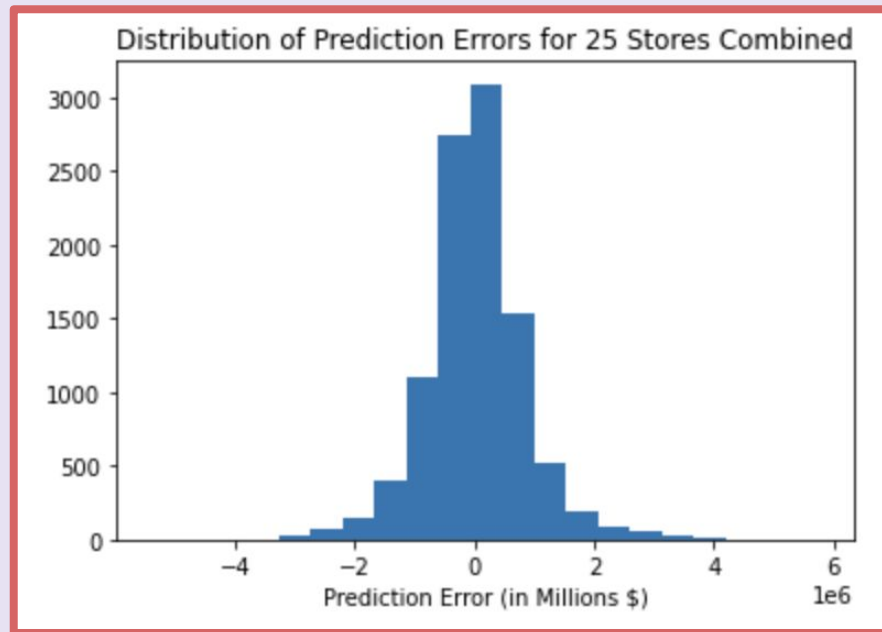
MAE

\$65,000

- Measure of prediction error that equally weighs all errors
- \$65,000 equates to 0.13 standard deviations away from the true value

Model Performance on Collections of Stores

- Predictions errors for collections of stores both peak and are centered at 0
- Model doesn't tend to either overestimate or underestimate



Conclusions

- Our model is very effective at accounting for a large portion of the variation in alcohol purchased by liquor stores
- Our model is very effective at predicting next years purchase amount for individual stores and collections of stores that are stable
- Our model can quantify the growth/decline in sales by comparing next years projection to previous years numbers

R^2 0.89

MAE \$65,000

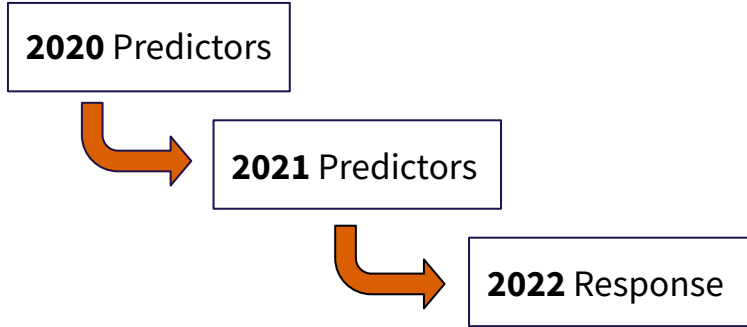


Thank You





Store 4102 Data **Pre Join**



Store 4102 Data **Post Join**

