Iowa liquor data: Findings and recommendations

Joshua Sung

Data Scientist, Iowa State

Tax Board

Guiding questions

- From the liquor data, what will be the predicted total sales for 2016?
- Determine which factors can greatly affect sales

Goals of the Study

- Overarching goal: Use the Iowa liquor data to be able to create a model to predict total sales for 2016.
- What components of the data will be able to perform the "best"?

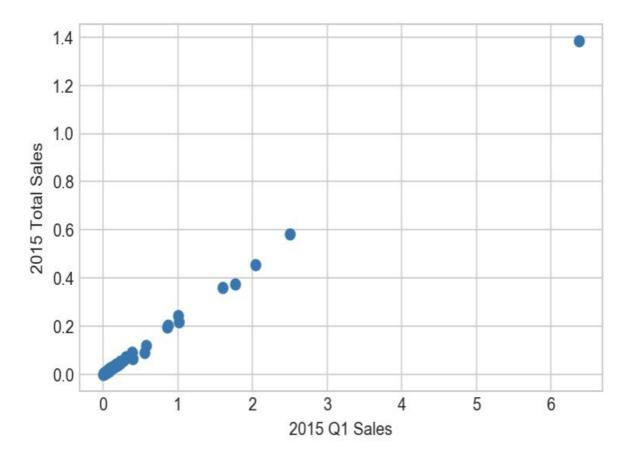
Predictors → Model → Projection of 2015 Total Sales

2016 Q1 Sales → Model → Projection of 2016 Total Sales

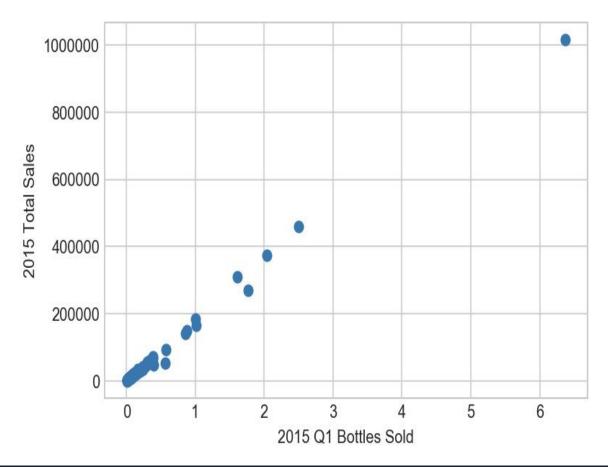
Process, Methodology

- Filtering data that I actually want to analyze → what is in the data that can "best" project sales? (i.e. creating new data, removing data)
- Examine relationships between the aspects of the data
- Creating the models and testing them

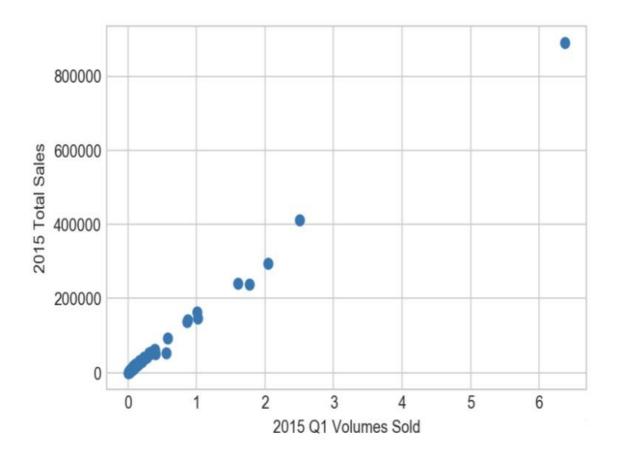
2015 Q1 Sales: A Strong Relationship with 2015 Total Sales!



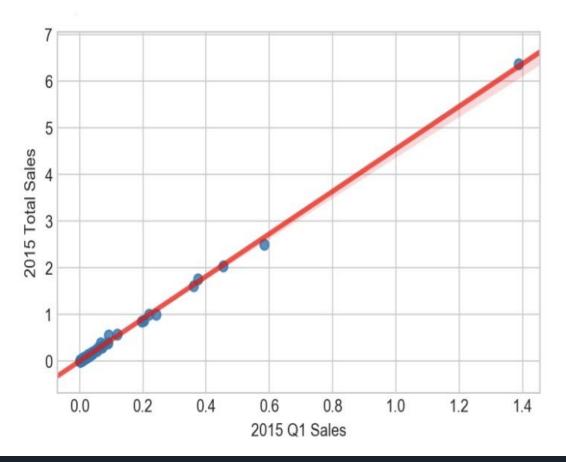
2015 Q1 Bottles Sold: A Strong Relationship with 2015 Total Sales!



2015 Q1 Volumes Sold: A Strong Relationship with 2015 Total Sales!



Quarter One Sales Has a Strong Positive Relationship with Total Sales, thus a Strong Predictor!



Findings and Results

• The model correctly projected the 2015 total sales.

• The model projects a 3.8% increase in total sales for 2016 or from \$284 million to \$295 million.

 All my models created showed a very strong relationship with 2015 Total Sales

Recommendations

- Future data collection:
 - Some data not included (i.e. demographics, consumer preference)
 - Assumptions were (i.e. static economy, steady prices)
 - Additional variables that could better predict sales

Conclusion

- The model performed well projecting 2015 total sales → 2016 total sales
- Many components of the data I could use to project sales
- 2015 Q1 Sales model performed the best
- \$295 million project total sales for 2016