



Iowa Liquor Store Proposal

Michael Costa

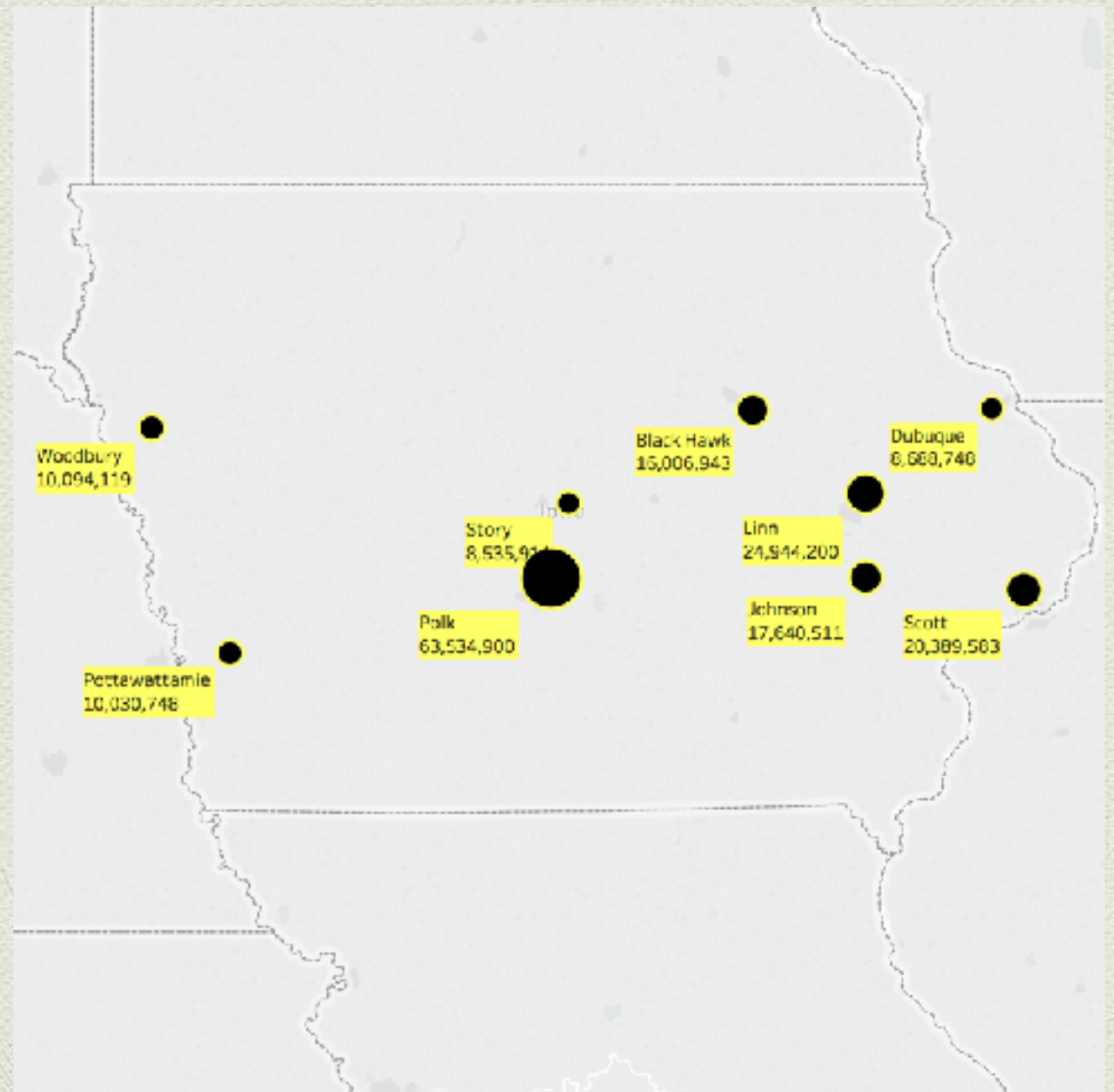
DSI-5

Project Objectives

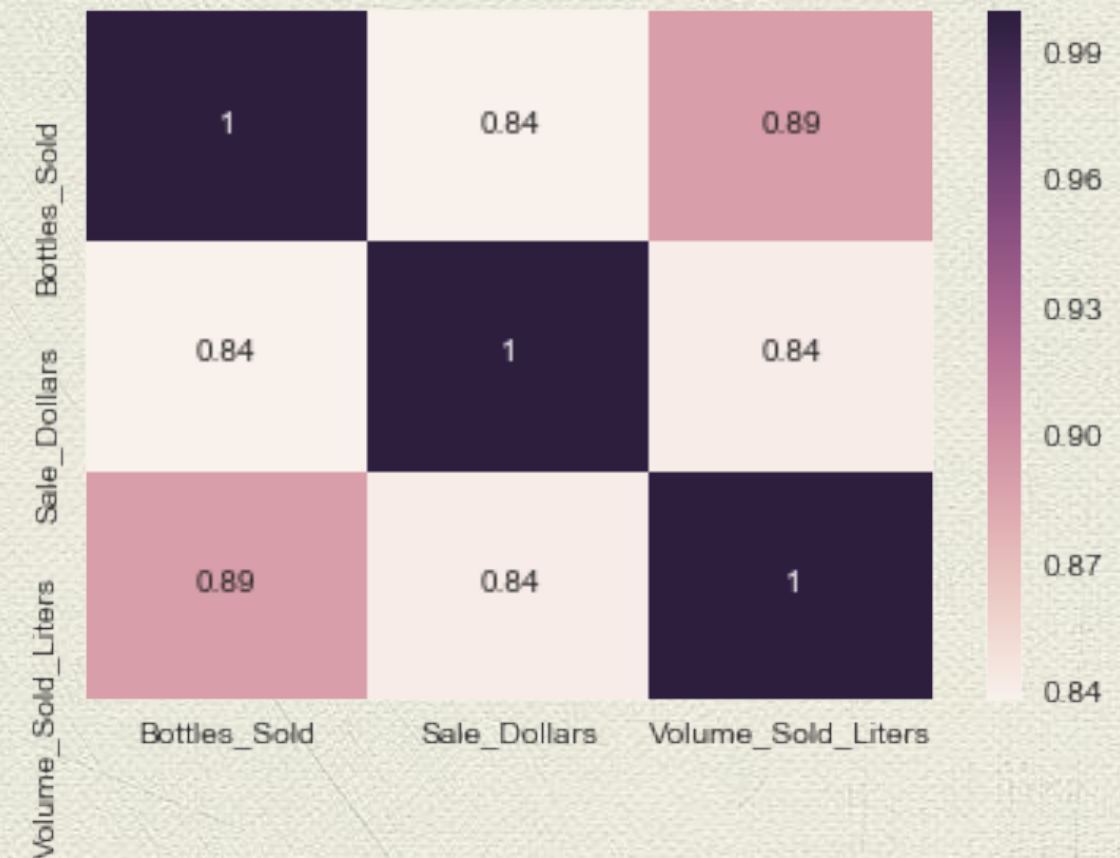
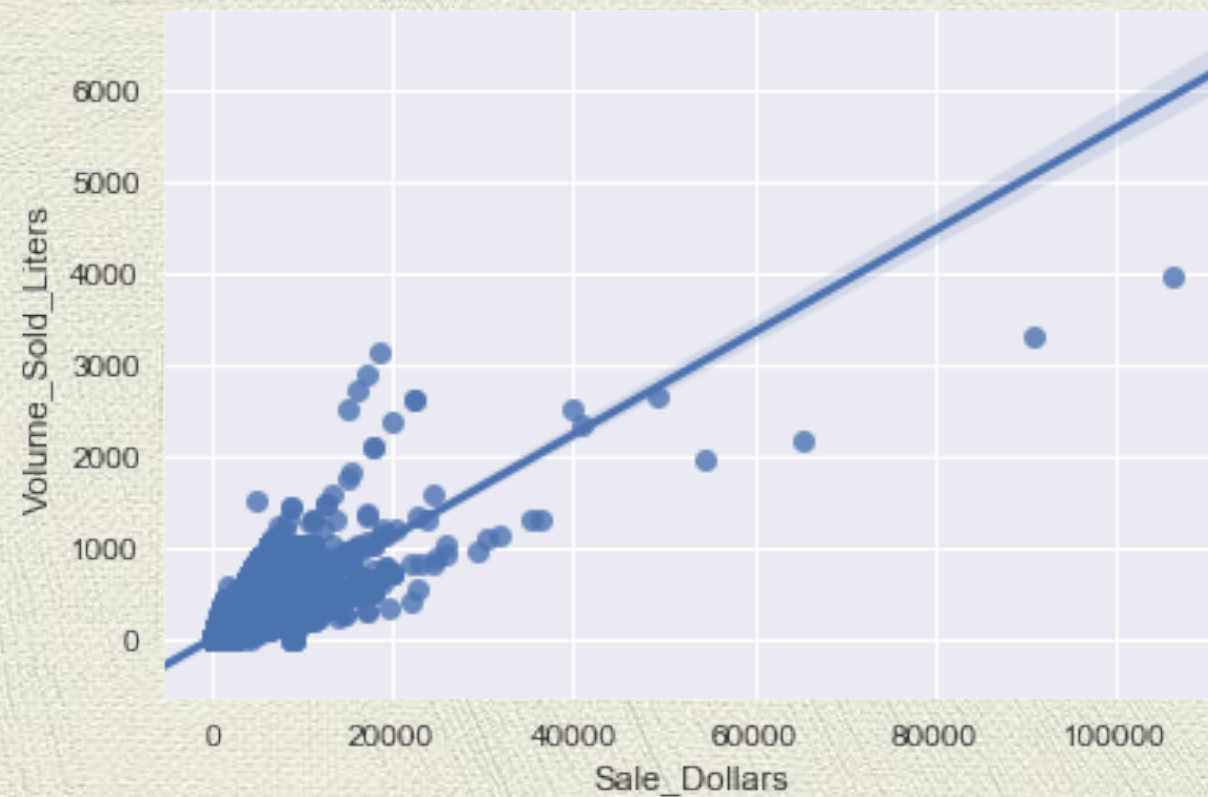
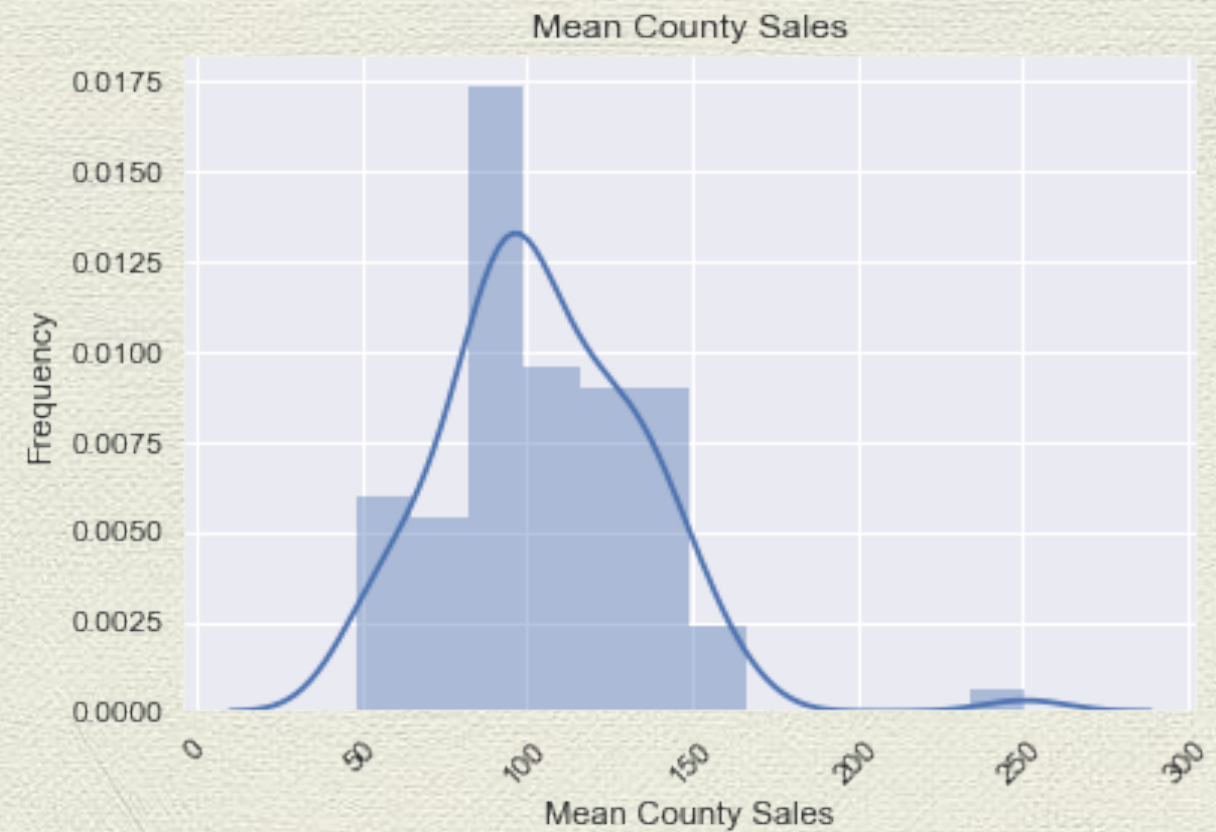
- ◆ Build models to understand Iowa's best performing counties for sales.
- ◆ Use data provided by state regulatory body and consideration of Iowa's laws.
- ◆ Use models to help suggest locations for new stores.

Wrangling The Data

- ◆ Data contains 2015 and the Q1 of 2016.
- ◆ Sale numbers represent state sales.
- ◆ Missing values, duplicates.
- ◆ Location data is our most important indicator.
- ◆ Deciding to focus on 9 high volume counties.



- Examining the mean sales
- Checking for correlation between.
- Linearity of sales and volume



Modeling

- ◆ Isolating data by counties.
- ◆ Standard Linear Regression
 - ◆ Simple and effective.
- ◆ Providing inference on future store locations from looking at volumes sold effect on price in each county.

- ◆ Results of our modeling.
- ◆ Comparing results of Polk, Scott and Johnson.
- ◆ What coefficients mean.
- ◆ How our models did in predicting sales.

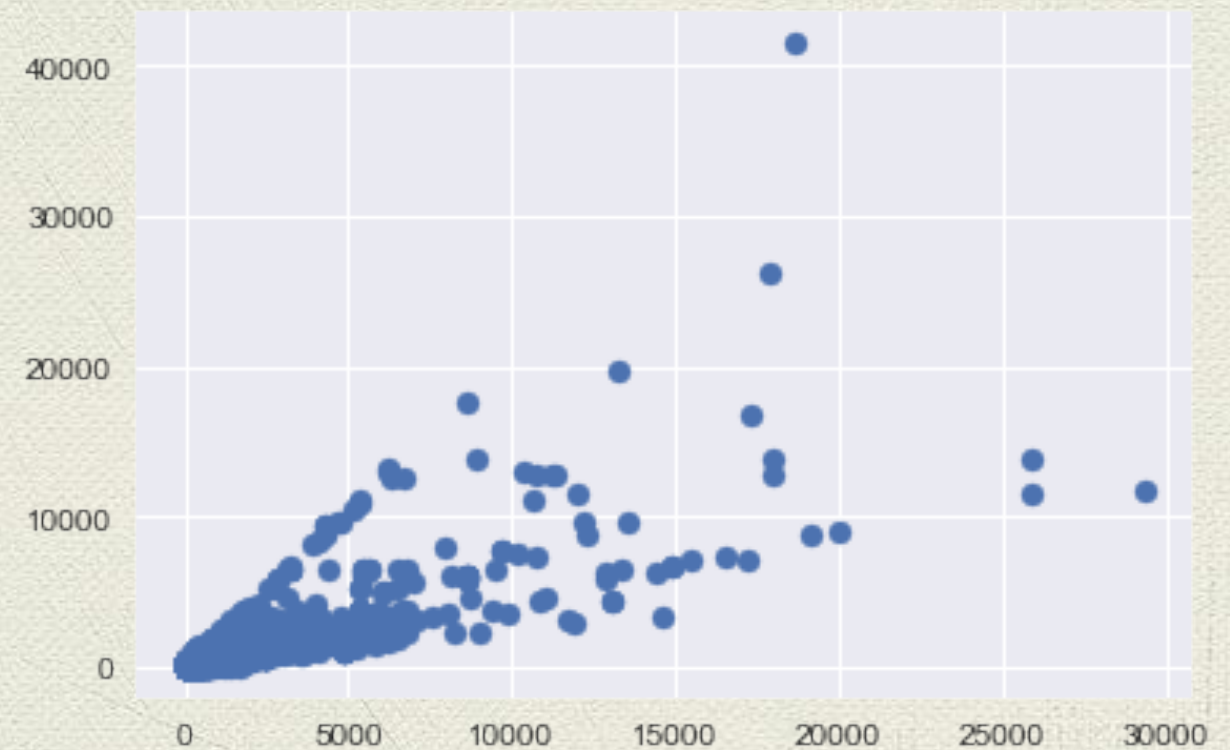
Cross-Predicted Accuracy: 0.663677643922

OLS Regression Results

```
=====
Dep. Variable:          Sale_Dollars      R-squared:
Model:                  OLS              Adj. R-squared:
Method:                 Least Squares    F-statistic:
Date:                  Fri, 14 Jul 2017  Prob (F-statistic):
Time:                  00:16:57         Log-Likelihood:
No. Observations:      76047            AIC:
Df Residuals:          76045            BIC:
Df Model:              1
Covariance Type:       nonrobust
=====
```

	coef	std err	t
const	35.8384	1.055	33.662
Volume_Sold_Liters	12.3816	0.032	392.991

```
=====
Omnibus:              119383.365      Durbin-Watson:
Prob(Omnibus):        0.000          Jarque-Bera (JB)
=====
```



Where to build

- ◆ Strengths of building in Johnson county, big slice of a small pie.
- ◆ Strengths of building in Polk. High volume high competition.
- ◆ Scott county the middle ground and my recommendation.

Additional Research

- ◆ Limitations of data in predicting store location.
- ◆ Demographic data and precise geographic data.
- ◆ Comments, questions and concerns?

