

Sale predictions on keyboards

Sale rate predicted by a Poisson
Regression model

Jung-a Kim

Sale Volume of Keyboards on e-platform

- Client
 - Wants to sell a used or brand-new keyboard on a platform like eBay.
- Goal
 - Passing the bar of minimum seller performance standards
 - Increase a selling limit in the next month

Data Collection

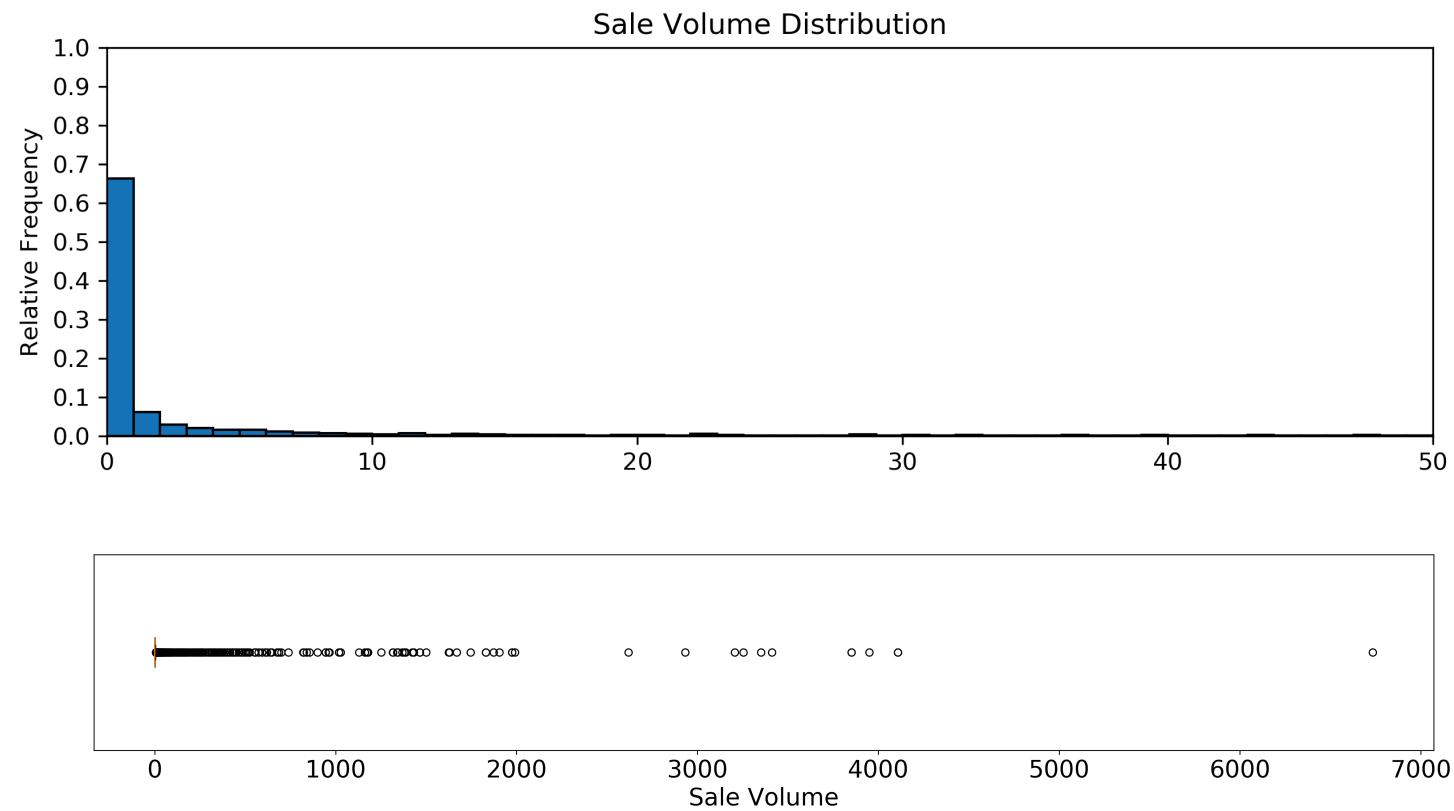
- 7000+ query results collected on 7/9/2020
- Explanatory variables
 - Price
 - Number of reviews
 - Average rating: 1.0 ~ 5.0
 - View count
 - Shipping cost
 - Added-to-watcher's-list count
 - Product's status: brand-new / open-box / refurbished / pre-owned
 - Free return
 - Benefits-charity
- Target variable
 - Sale volume collected on 7/13/2020
 - Sale in 4+ days

Data Cleaning

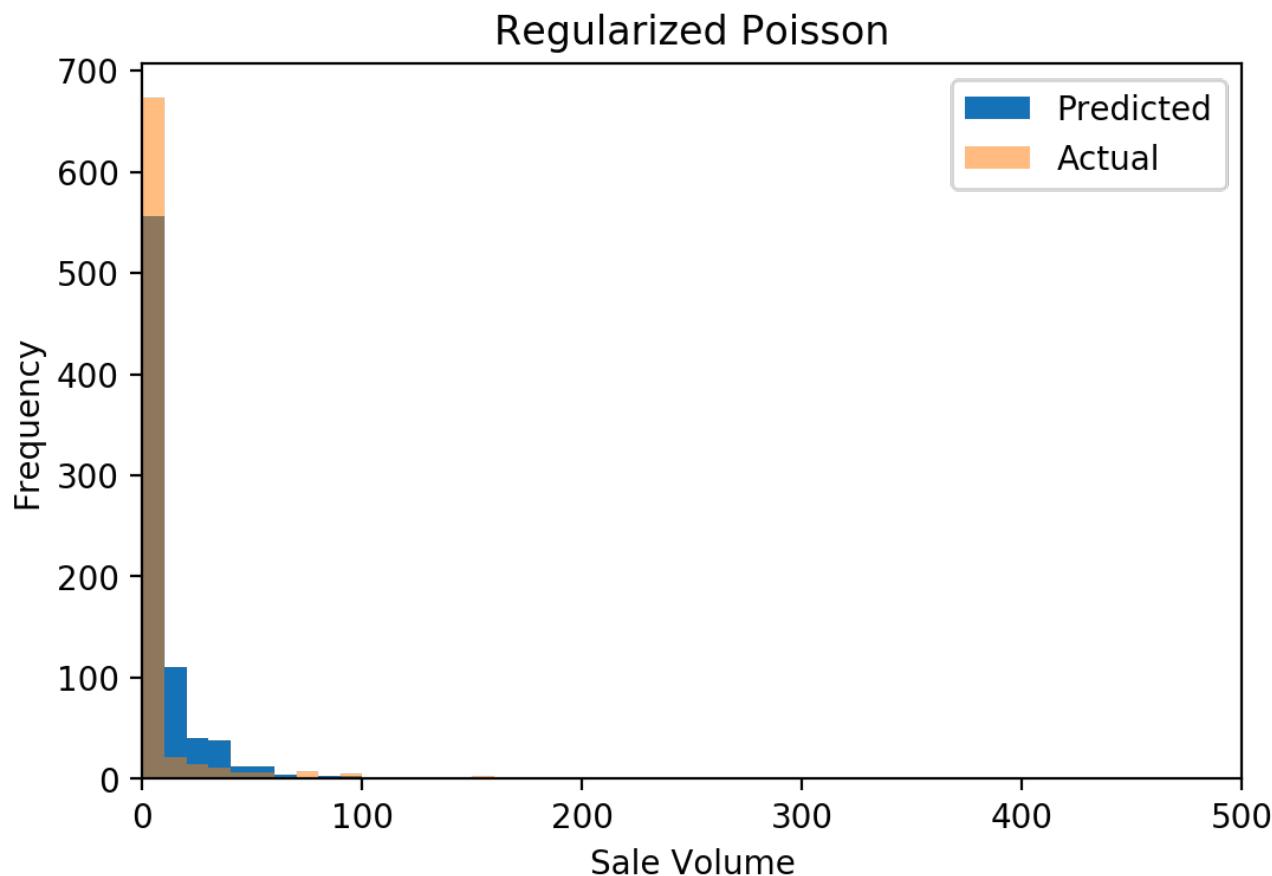
- Observations with duplicate information removed
 - 7016 results -> 5794 results
- Missing values
 - Average rating (92% missing), Price, Shipping cost, Status
 - Imputed by Multiple Imputation by Chained Equations
- Outlier
 - Keyboard-related item

Sale Volume Distribution

- Candidate models:
- 1. Poisson regression
- 2. Linear regression with $\log(\text{Sale Volume})$



Predicted Sale Volume Distribution by the model



$$NMAE = 0.58$$

- $MAE = 42.64$
- $MAD = 73.70$

Model Equation

- $E(\text{Sale Volume} | \text{Price}, \text{Shipping Cost}, \text{Free return}, \text{Rating present}) = \lambda = C \cdot \exp\left\{\frac{\beta_1}{\sqrt{\text{Price} + 1}} + \frac{\beta_2}{\sqrt{\text{Shipping Cost} + 1}} + \beta_3 \text{Free return} + \beta_4 \text{Rating present}\right\}$

where Sale Volume \sim Poisson(λ)

- $C \approx 0.239$
- $\beta_1 \approx 9.731$
- $\beta_2 \approx 1.413$
- $\beta_3 \approx 1.307$
- $\beta_4 \approx 1.167$

Model Equation – Weights Interpretation

- $E(\text{Sale Volume} | \text{Price}, \text{Shipping Cost}, \text{Free return}, \text{Rating present}) = \lambda = C \cdot \exp\left\{\frac{\beta_1}{\sqrt{\text{Price} + 1}} + \frac{\beta_2}{\sqrt{\text{Shipping Cost} + 1}} + \beta_3 \text{Free return} + \beta_4 \text{Rating present}\right\}$ where Sale Volume $\sim \text{Poisson}(\lambda)$

- $C \approx 0.239$
- $\beta_1 \approx 9.731$
 - Increasing price of your keyboard from \$29.00 to \$39.00 can reduce Sale Volume by 21.2% on average.

$$1 - \exp\left\{\beta_1 \left(\frac{1}{\sqrt{39 + 1}} - \frac{1}{\sqrt{29 + 1}}\right)\right\} \approx 0.212$$

- $\beta_2 \approx 1.413$
- $\beta_3 \approx 1.307$
- $\beta_4 \approx 1.167$

Model Equation – Weights Interpretation

- $E(\text{Sale Volume} | \text{Price}, \text{Shipping Cost}, \text{Free return}, \text{Rating present}) = \lambda = C \cdot \exp\left\{\frac{\beta_1}{\sqrt{\text{Price} + 1}} + \frac{\beta_2}{\sqrt{\text{Shipping Cost} + 1}} + \beta_3 \text{Free return} + \beta_4 \text{Rating present}\right\}$ where Sale Volume $\sim \text{Poisson}(\lambda)$

- $C \approx 0.239$
- $\beta_1 \approx 9.731$
- $\beta_2 \approx 1.413$
 - Increasing shipping cost from free-of-charge to just by \$1.00 can reduce Sale Volume by 33.9% on average.

$$1 - \exp\left\{\beta_2 \left(\frac{1}{\sqrt{1+1}} - \frac{1}{\sqrt{0+1}}\right)\right\} \approx 0.339$$

- $\beta_3 \approx 1.307$
- $\beta_4 \approx 1.167$

Model Equation – Weights Interpretation

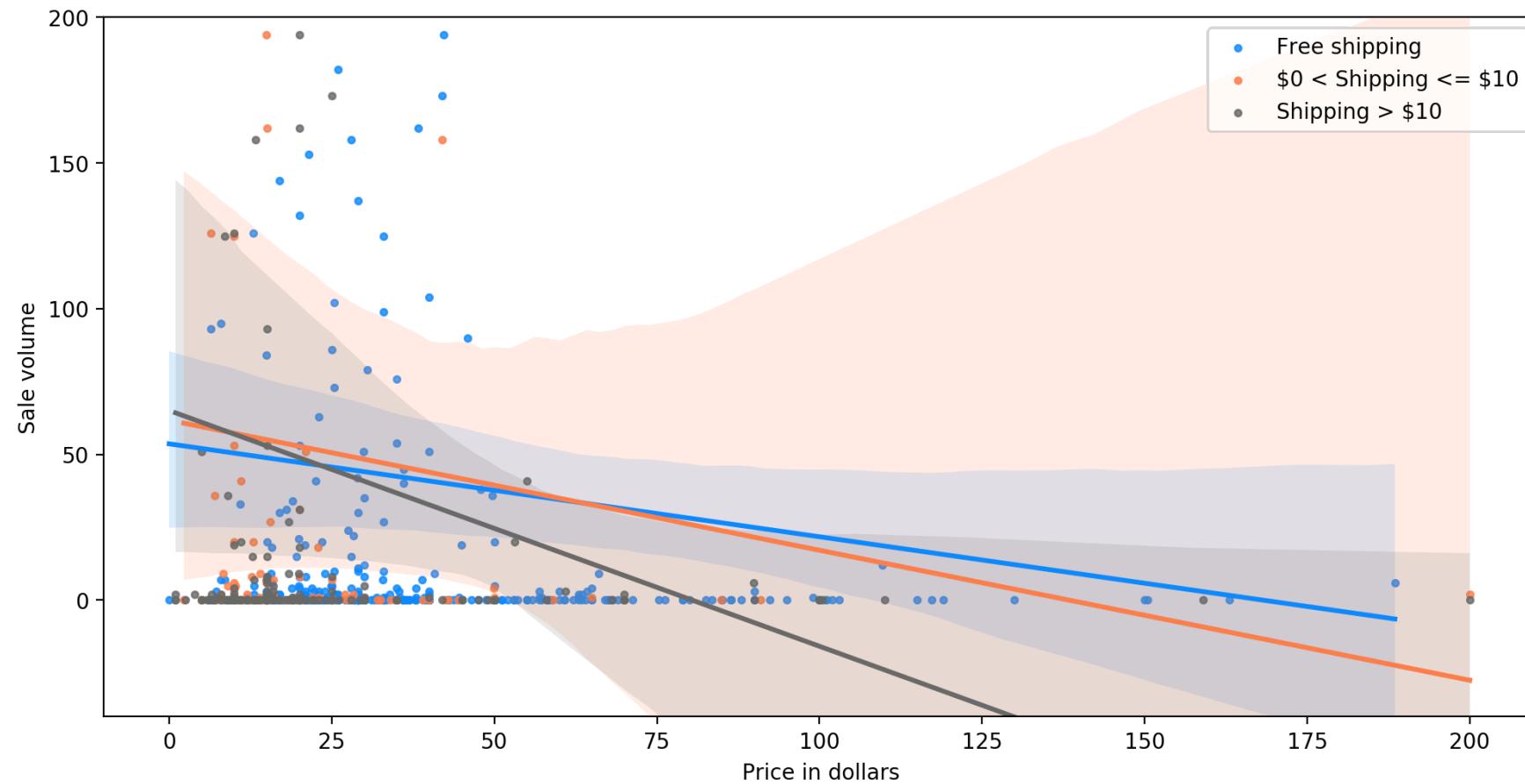
- $E(\text{Sale Volume} | \text{Price}, \text{Shipping Cost}, \text{Free return}, \text{Rating present}) = \lambda = C \cdot \exp\left\{\frac{\beta_1}{\sqrt{\text{Price} + 1}} + \frac{\beta_2}{\sqrt{\text{Shipping Cost} + 1}} + \beta_3 \text{Free return} + \beta_4 \text{Rating present}\right\}$ where Sale Volume $\sim \text{Poisson}(\lambda)$
 - $C \approx 0.239$
 - $\beta_1 \approx 9.731$
 - $\beta_2 \approx 1.413$
 - $\beta_3 \approx 1.307$
 - If you remove the cost of returning, it can increase Sale Volume by 3.7 times on average.
 $e^{\beta_3} \approx 3.70$
 - $\beta_4 \approx 1.167$

Model Equation – Weights Interpretation

- $E(\text{Sale Volume} | \text{Price}, \text{Shipping Cost}, \text{Free return}, \text{Rating present}) = \lambda = C \cdot \exp\left\{\frac{\beta_1}{\sqrt{\text{Price} + 1}} + \frac{\beta_2}{\sqrt{\text{Shipping Cost} + 1}} + \beta_3 \text{Free return} + \beta_4 \text{Rating present}\right\}$ where Sale Volume $\sim \text{Poisson}(\lambda)$
 - $C \approx 0.239$
 - $\beta_1 \approx 9.731$
 - $\beta_2 \approx 1.413$
 - $\beta_3 \approx 1.307$
 - $\beta_4 \approx 1.167$
 - If there is at least one review with the rating, it can increase Sale Volume by 3.2 times on average

$$e^{\beta_4} \approx 3.21$$

Interaction between Price and Shipping Cost



Recommendations

- Be cautious about increasing price under \$50 as it can reduce sale rate by at least 10% on average
- People are sensitive about shipping cost more than \$10
 - Increasing price will suffer large decrease in sale rate
- Free return is highly encouraged!
- Reward writing product reviews!

Caution

- Lack of Confidence Interval for the weights
- Likelihood of bias in the weights
 - Strong predictors missing. e.g) Content of the reviews, Rating of the seller

Future work

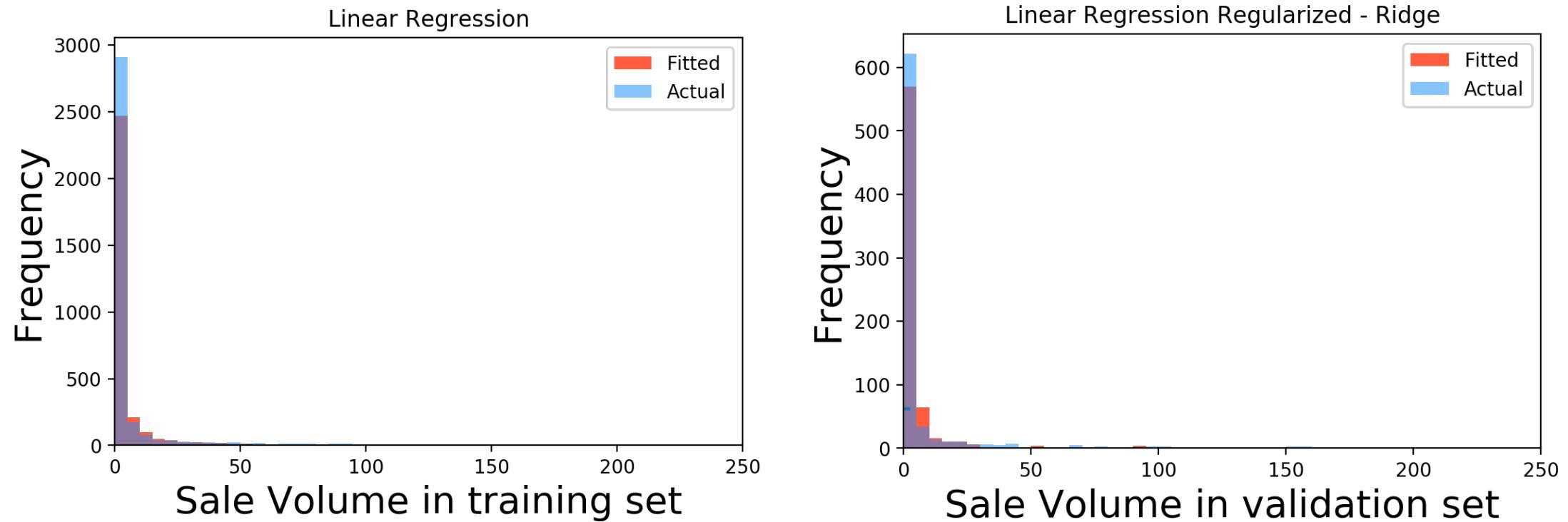
- Reduce the variance of the weights
 - e.g) Dimensionality reduction
- Increase the accuracy of the weights
 - e.g) Collect more explanatory variables
- Increase predictive power of the model
 - e.g) Impute missing values in the test set with the training set
- Negative Binomial Regression to overcome dispersion



Thank you for your attention!

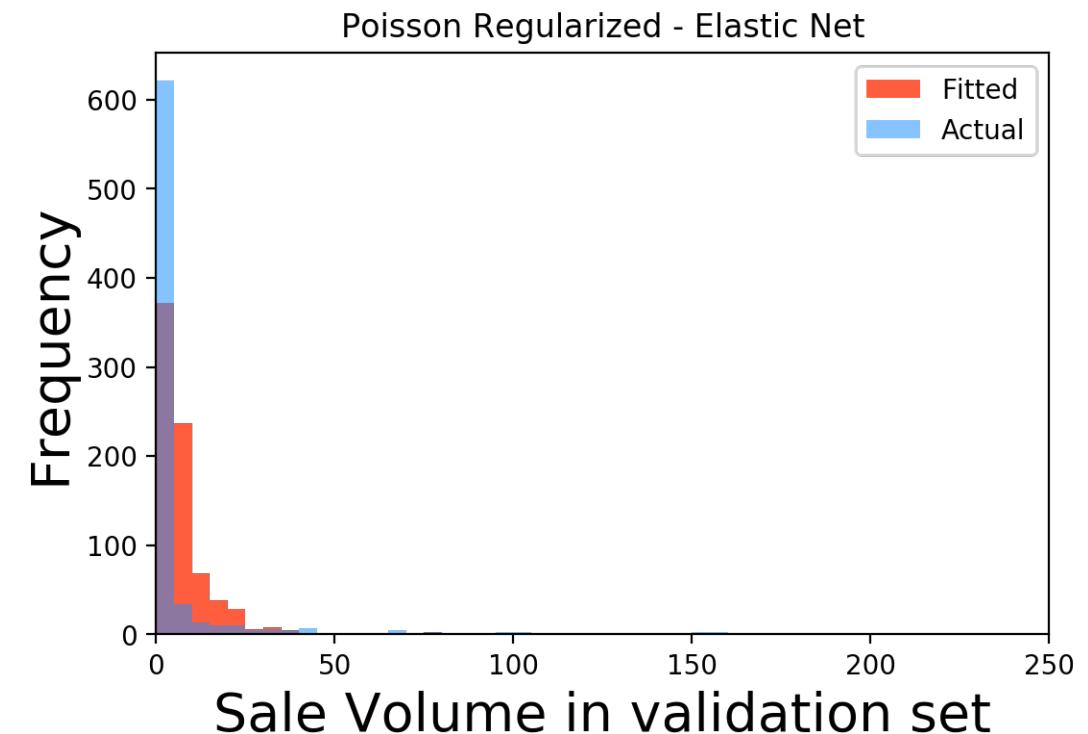
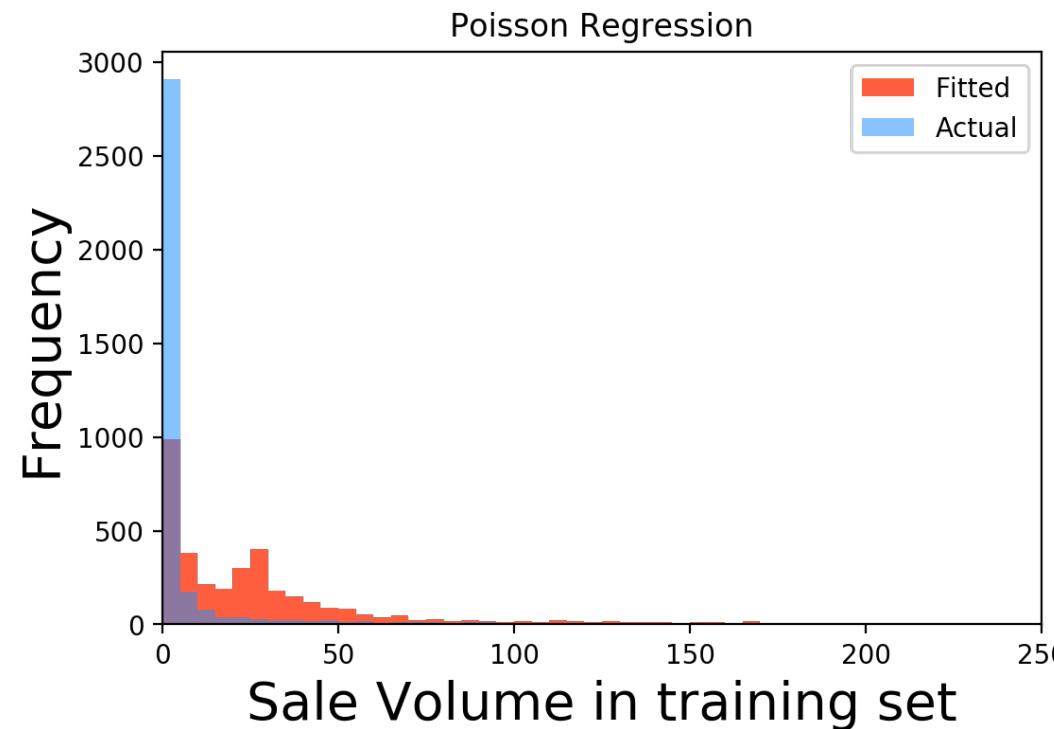
Q & A

Linear Regression - Training vs. Validation



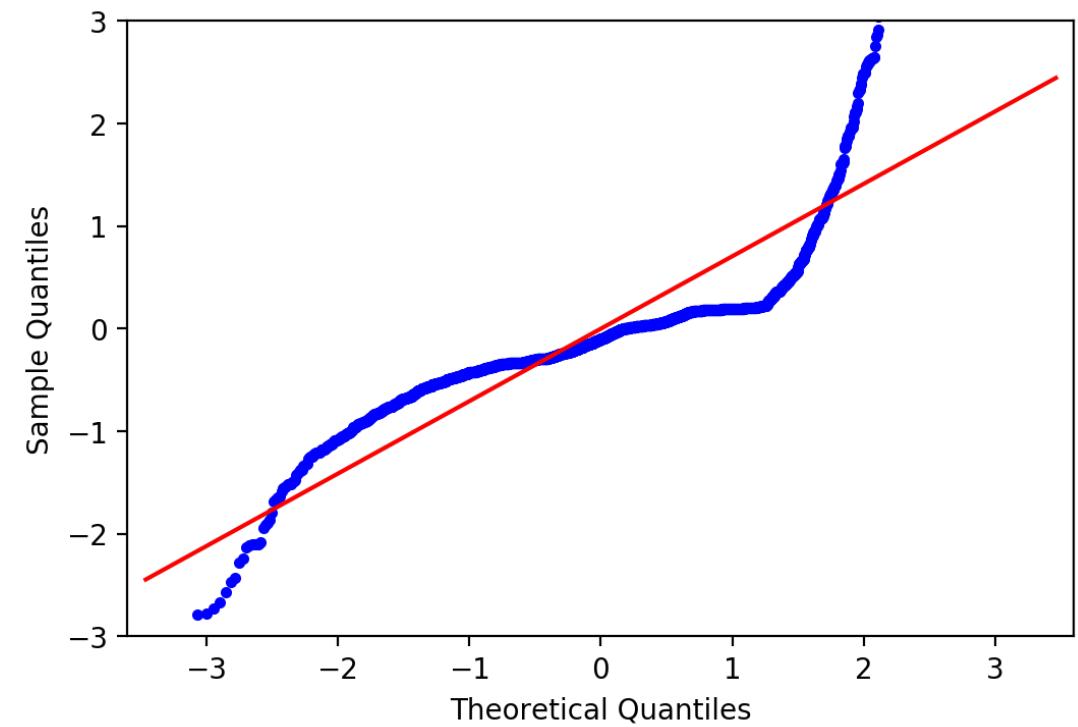
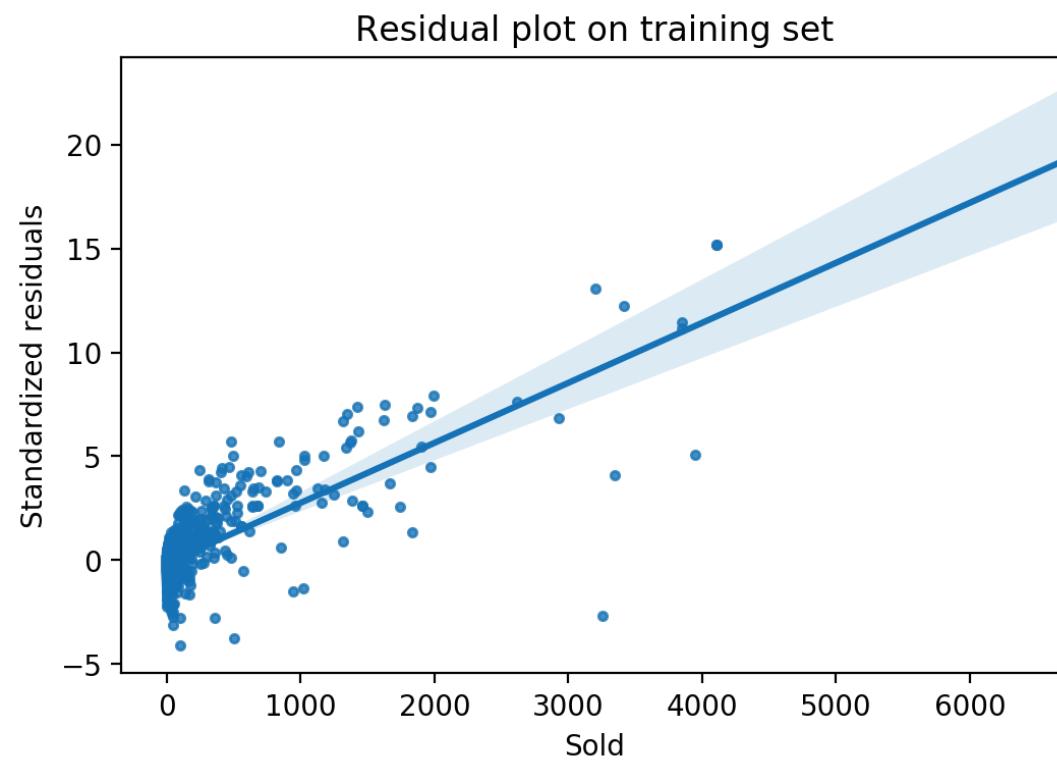
Metric \ Cross validation	Fit to the training set	Cross validation Before regularization	Cross validation After regularization
MAE	43.8	62.9	62.1
R ²	0.398	0.01	0.006

Poisson Regression – Training vs. Validation

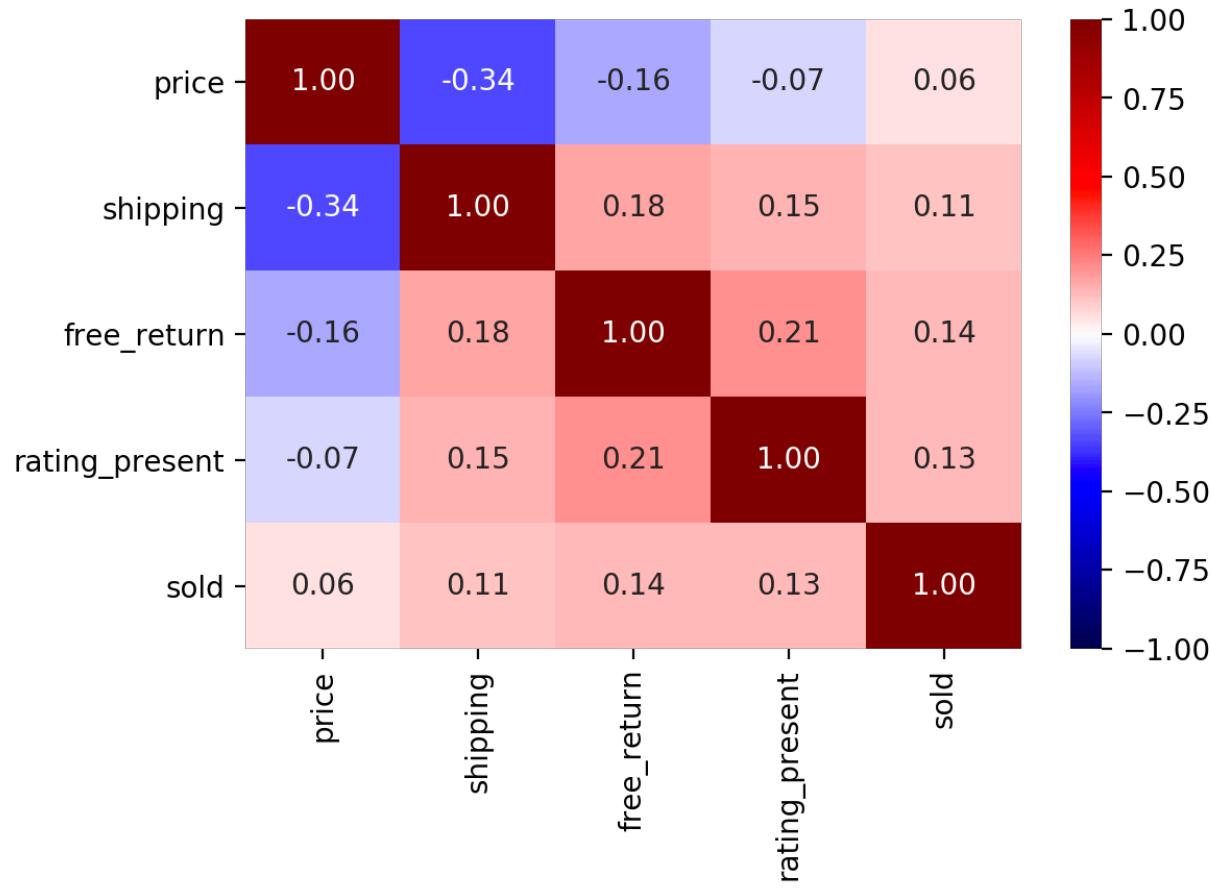


Metric \ Cross validation	Fit to the training set	Cross validation Before regularization	Cross validation After regularization
MAE	60.54	69.9	62.1
Pseudo R ²	0.40		

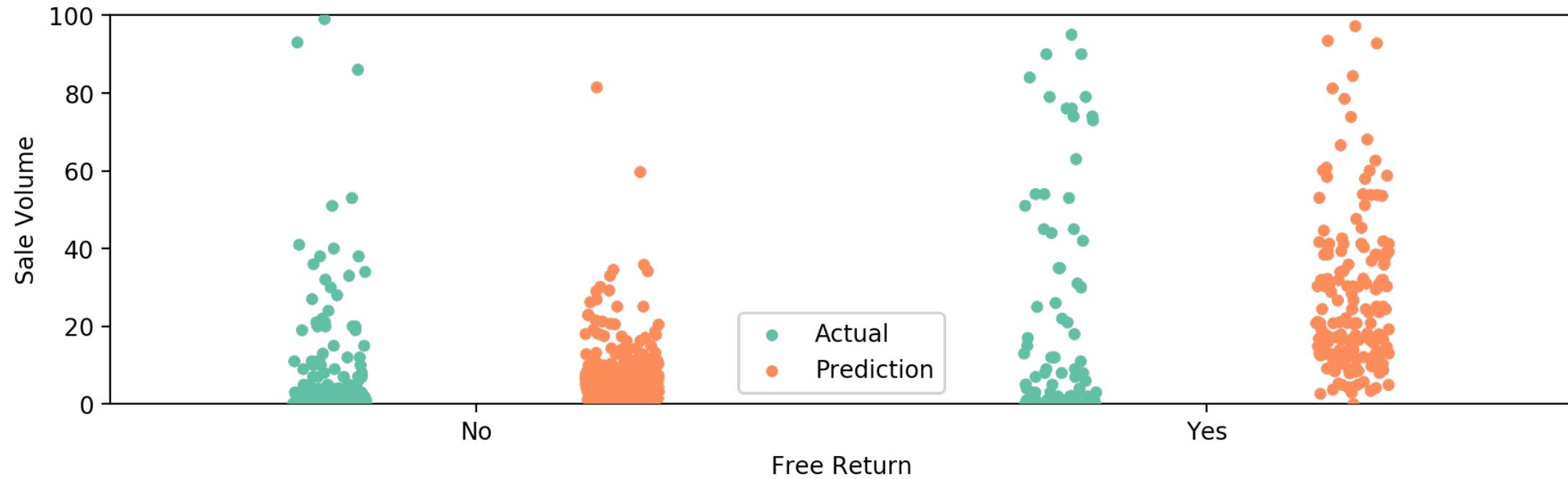
Deviance Residual plot – Poisson Regression



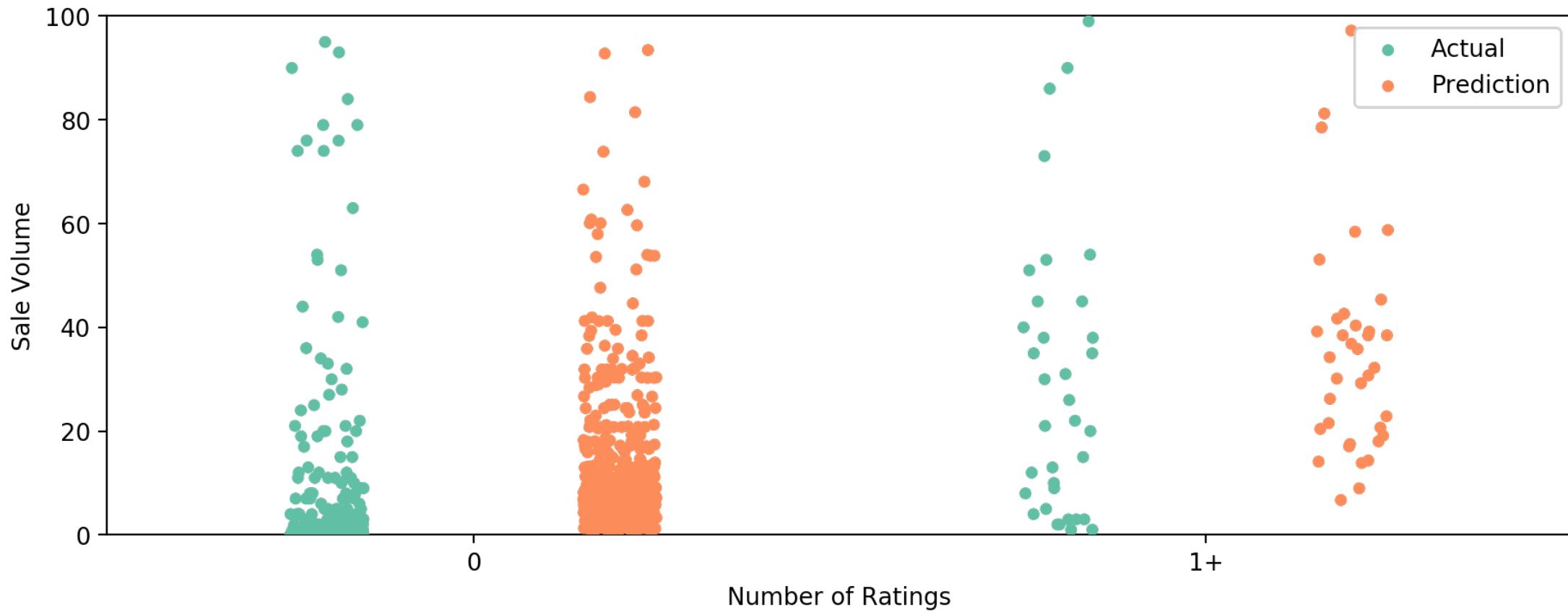
Correlation heatmap



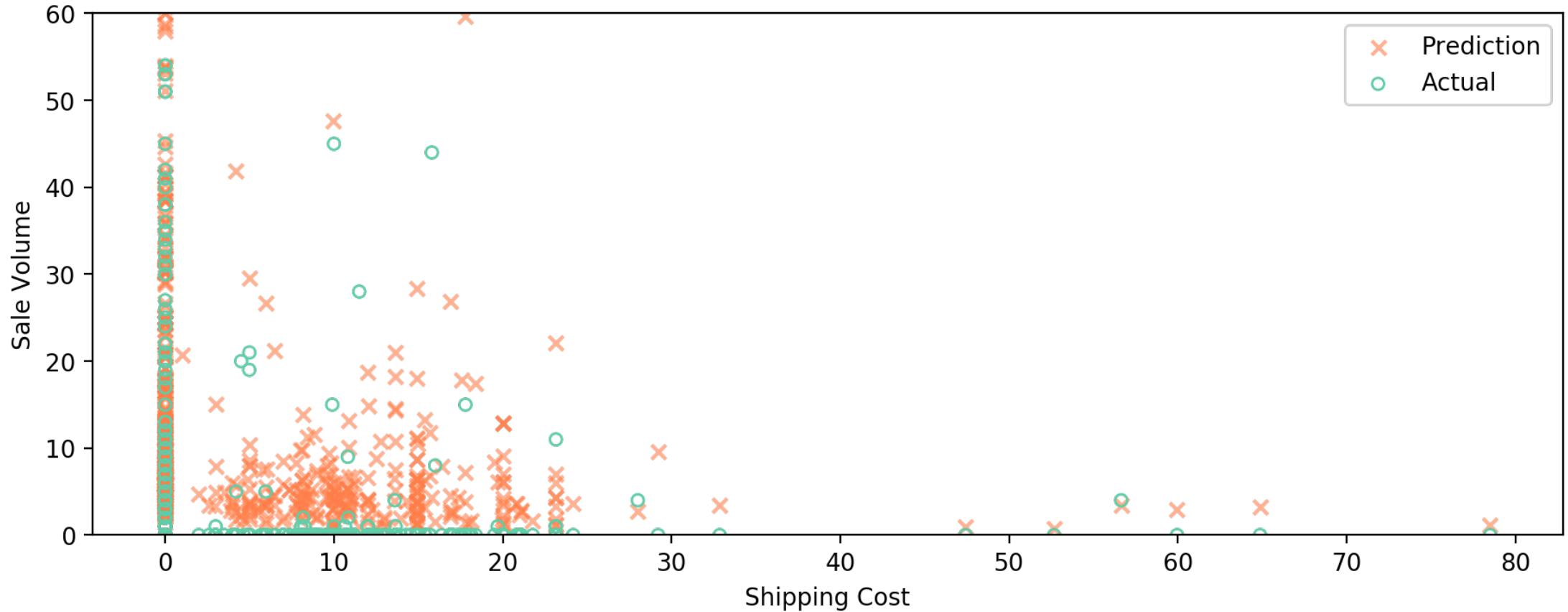
Sale Volume vs. Free Return



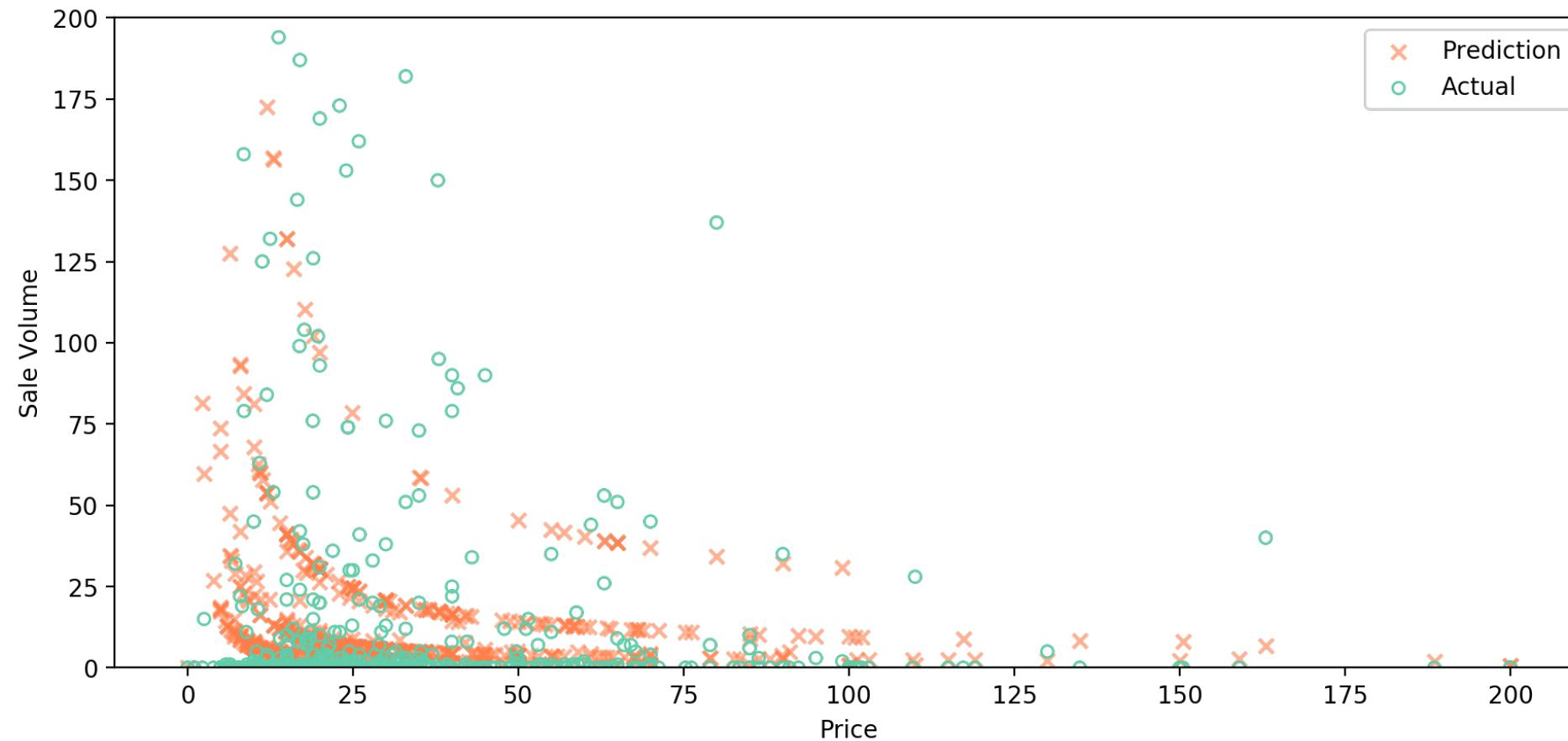
Sale Volume vs. Rating Present



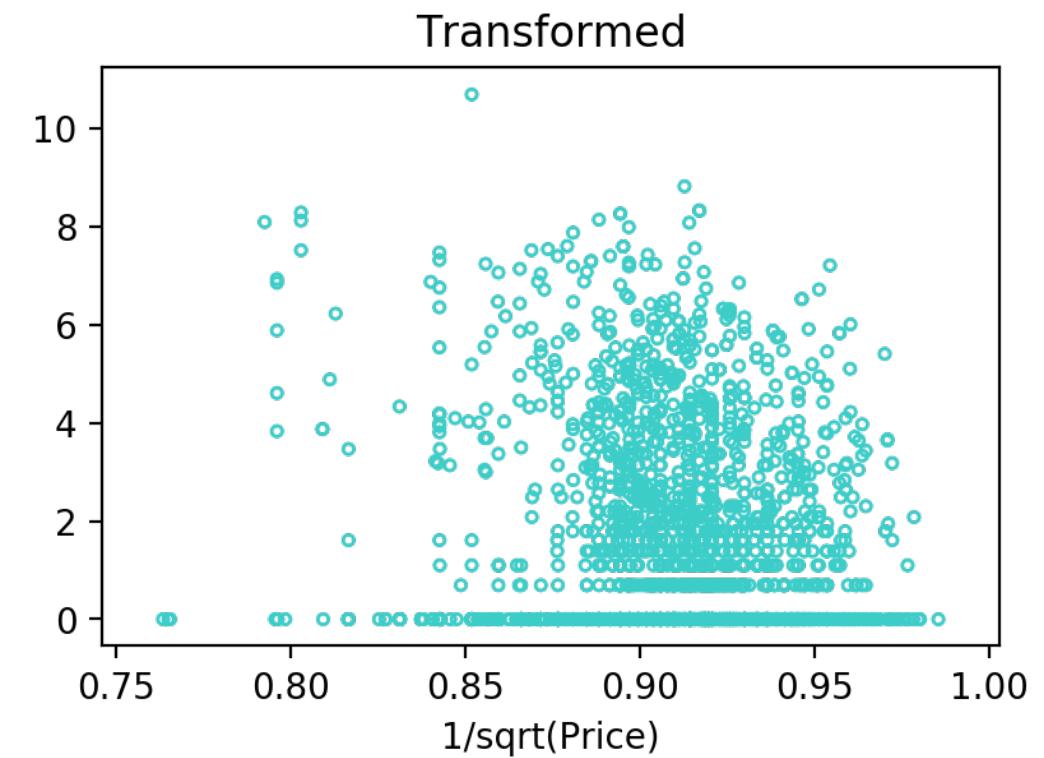
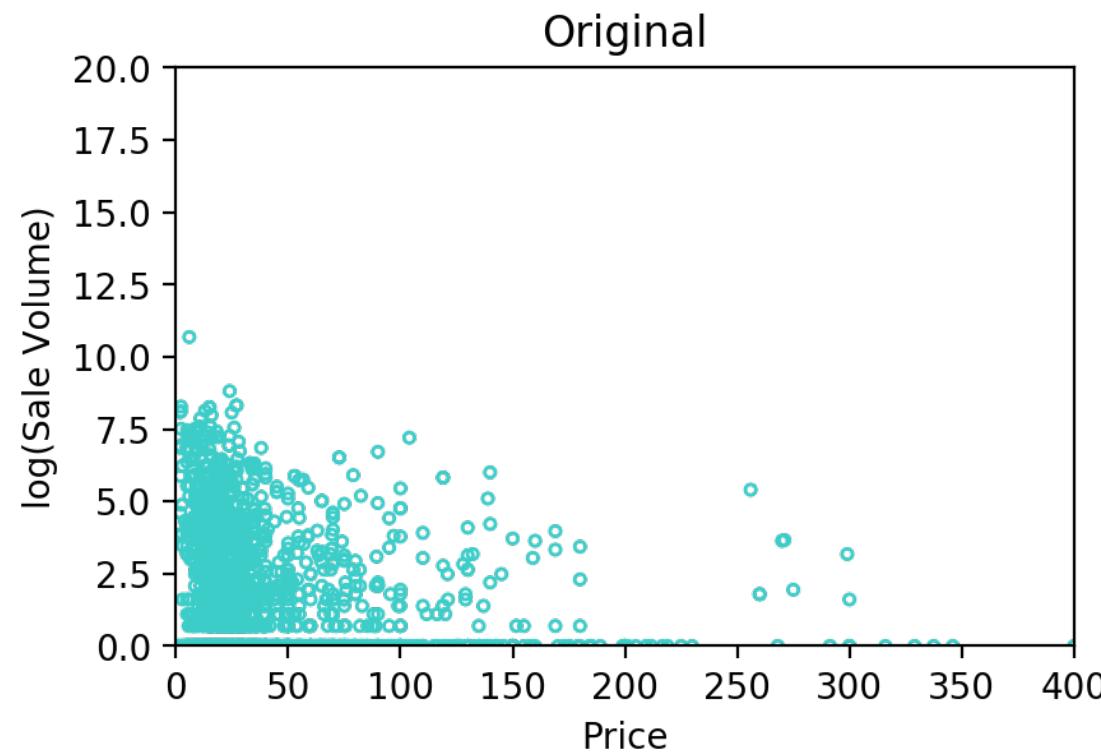
Sale Volume vs. Shipping Cost



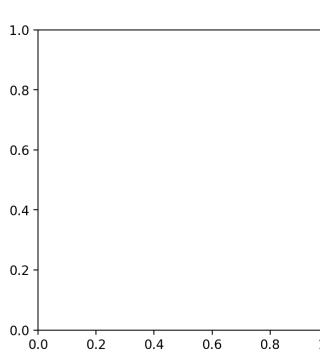
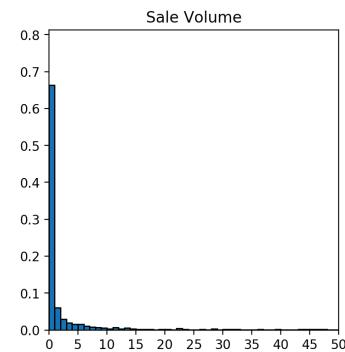
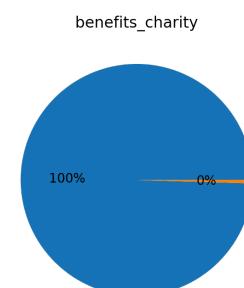
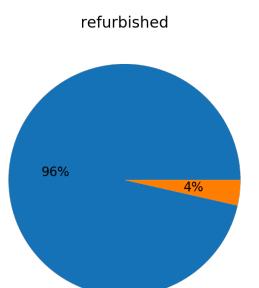
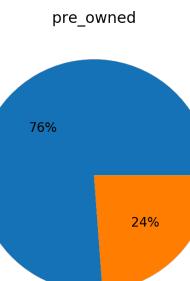
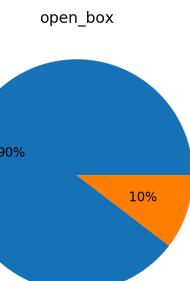
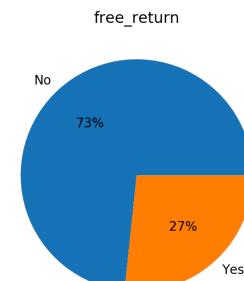
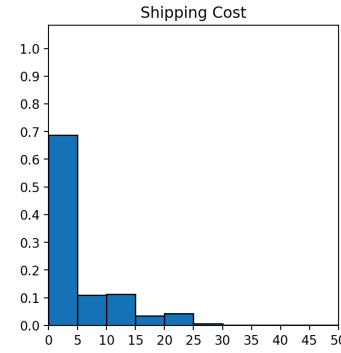
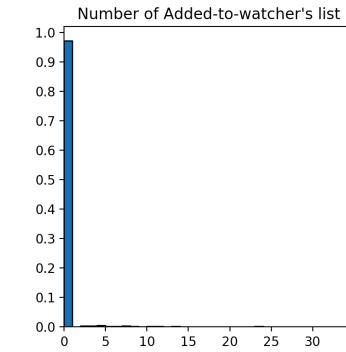
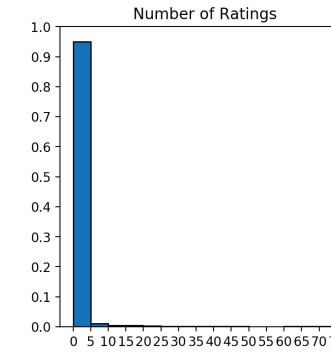
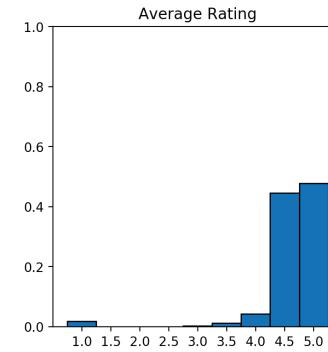
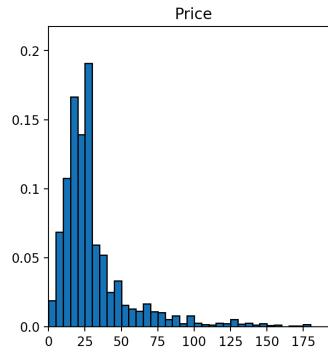
Sale Volume vs. Price



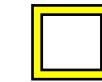
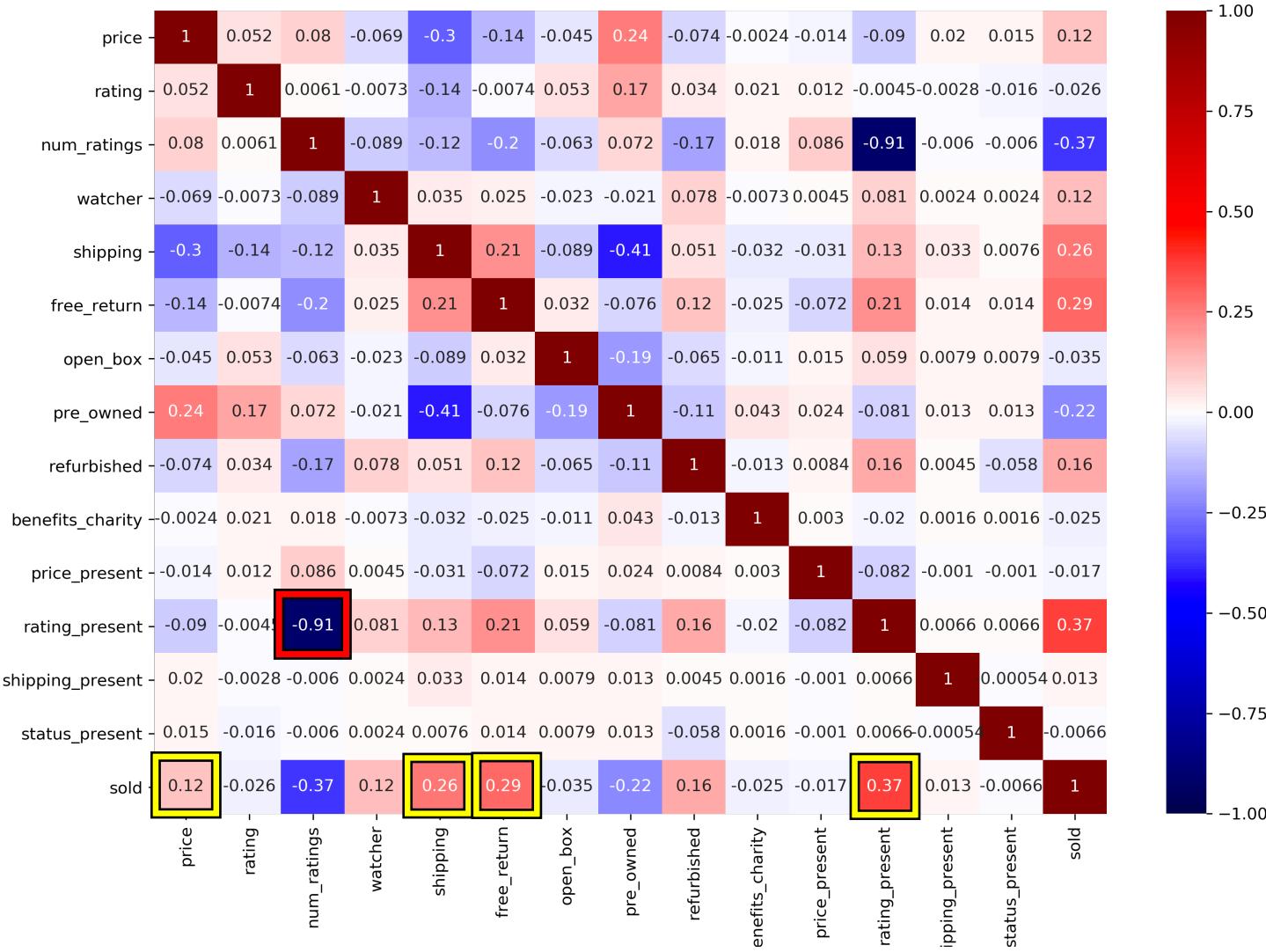
Linearity between $\log(Sale\ Volume)$ and $\frac{1}{\sqrt{Price}}$



Distribution of variables



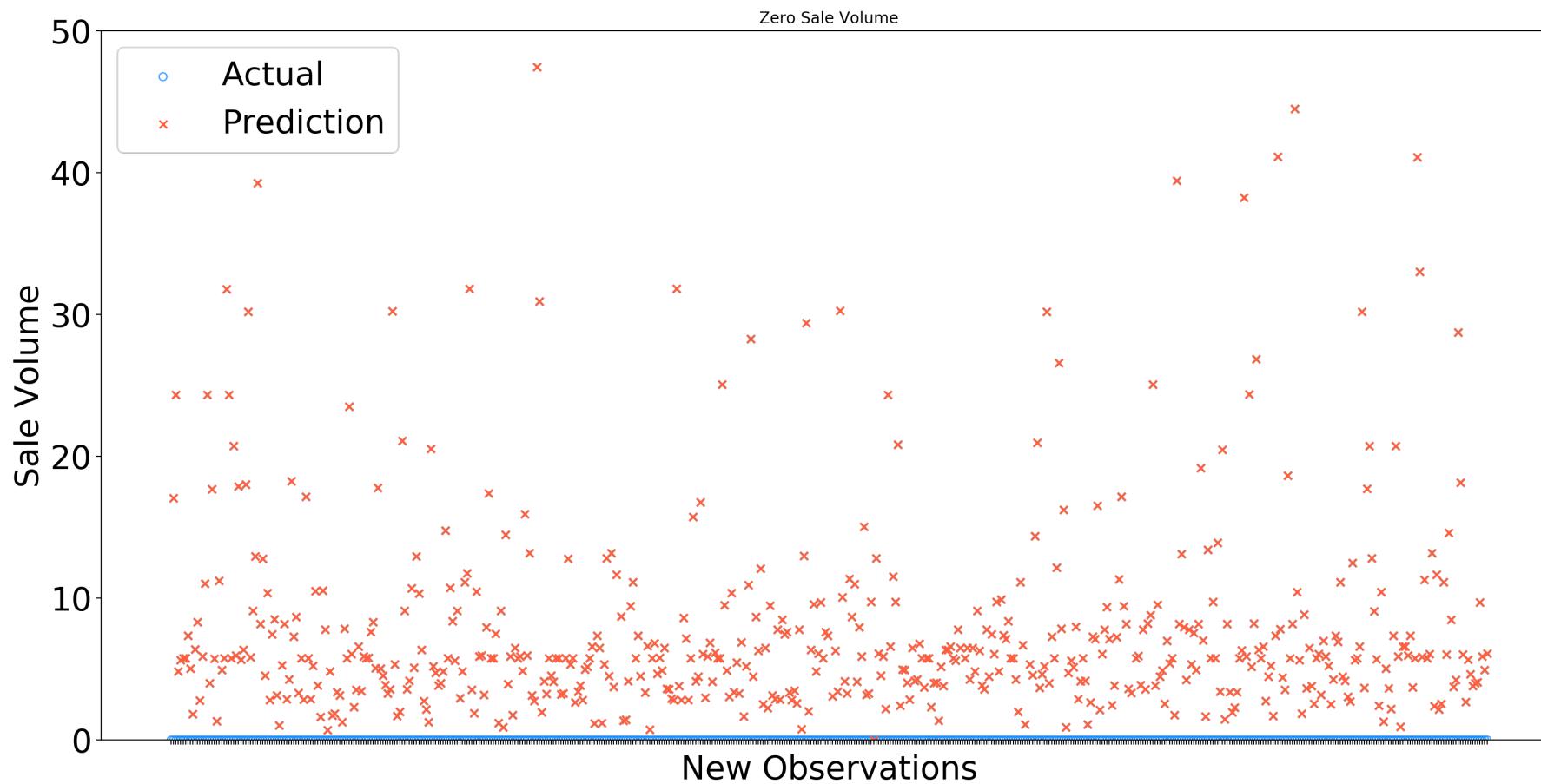
Correlation with Sale Volume



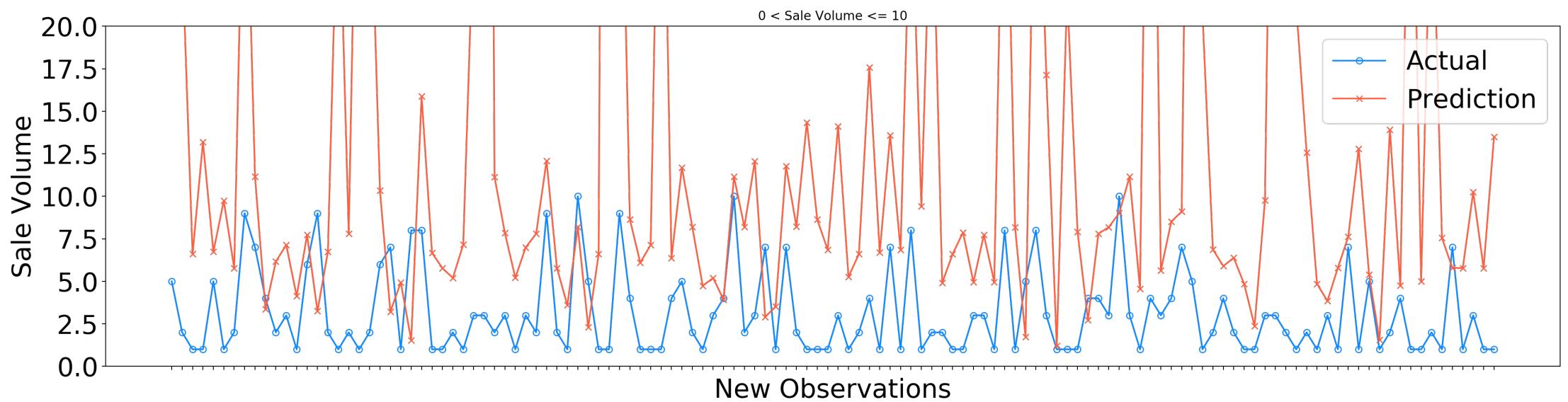
: Features in the final model

- Price
- Shipping Cost
- Free Return
- At least one review

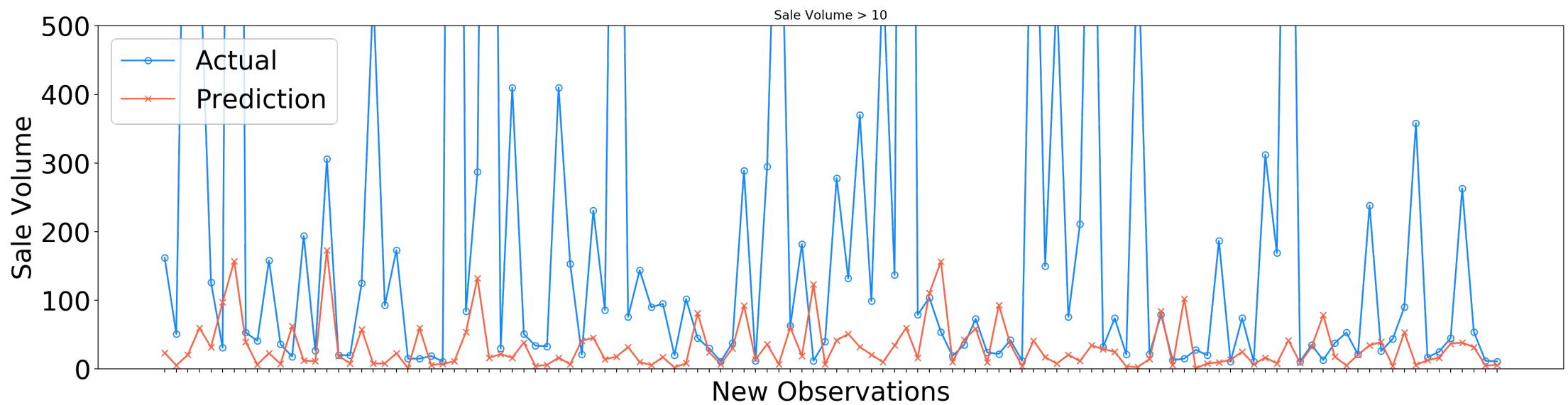
Zero Sale Volume Prediction



$1 \leq \text{Sale Volume} \leq 10$ Prediction



10 < Sale Volume Prediction



Outlier



Hard Case Shell for Macbook Air 13 / 11 Pro 13 / 15 Retina 12 + Keyboard Cover

1 sold in last hour

Condition: New

Color: - Select -

Model Type: - Select -

Quantity: 1 More than 10 available
43,749 sold / See feedback

Price: US \$5.99

Buy It Now

Add to cart

Add to Watchlist

Free delivery in 4 days

43,749 sold

More than 80% sold

Shipping: FAST 'N FREE

Guaranteed by Wed. Jul. 22 | See details

Item location: Dixon, California, United States

Ships to: United States | See exclusions

Payments: PayPal VISA Mastercard Discover

PayPal CREDIT

Earn up to 5x points when you use your eBay Mastercard. Learn more

Returns: 30 day returns. Buyer pays for return shipping | See details

\$ Have one to sell? Sell now

Shop with confidence

eBay Money Back Guarantee

Get the item you ordered or get your money back. [Learn more](#)

Seller information

emmasilmon (204831) ★

99.4% Positive feedback

Save this Seller

Contact seller

Visit store

See other items