

Amazon Price War

General Assembly: Data Science Remote

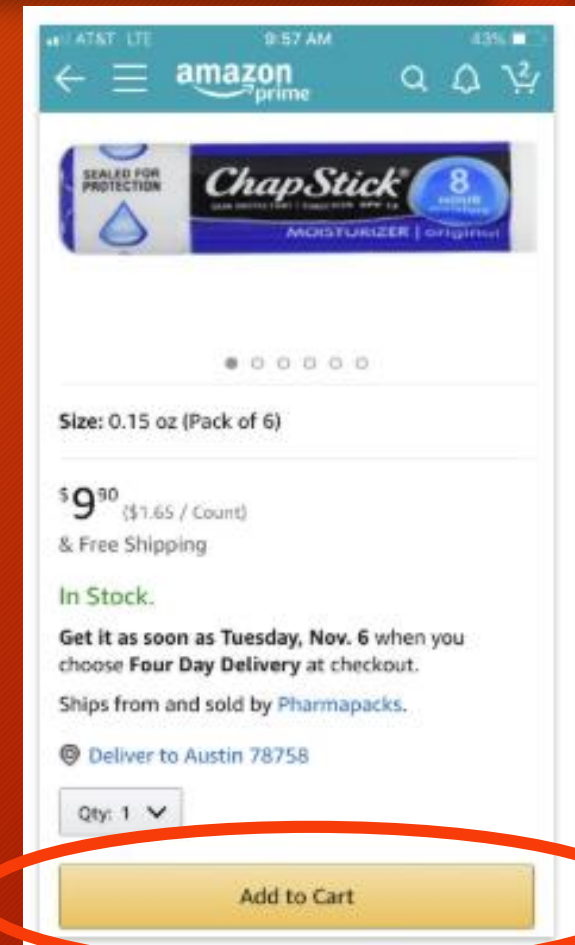
Hypothesis, Assumptions, and Goals

How win the **Buy Box** among multiple type of Amazon Reseller?

- Lower the price?
- Higher the ranking?
- Use Amazon platform?



82% of Amazon sales go through the **Buy Box**, and the percentage is even higher for mobile purchases.



Data Set

Data: Amazon Ranking Analysis during summer 2016

ProPublica reporters examined Amazon's shopping algorithm. We scraped data from the company's website to examine listings for **250 best-selling products** across a wide range of categories, from electronics to household supplies, over a period of several weeks during summer 2016. We compared pricing and shipping costs for products offered by multiple vendors, including those sold by Amazon and sellers in the "Fulfilled by Amazon" program. In total, we examined **6,973** vendor listings.

Models

- The data source has been split to training set (70%) and testing set (30%).
- Tried out four different models, including:
 - Logistic Regression
 - K-Nearest Neighbors
 - Decision Tree Classifier
 - Random Forest

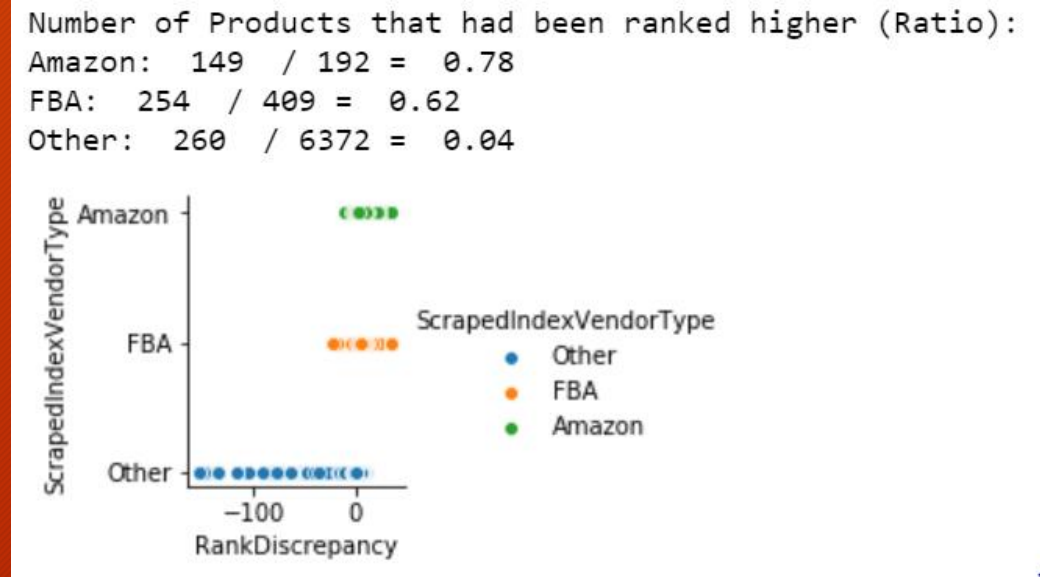
Results

- Logistic Regression
 - Accuracy Score: 98%
- K-Nearest Neighbors (neighbors=7)
 - Accuracy Score: 97%
- Decision Tree (depth=6)
 - Accuracy Score: 97%
- Random Forest (features=3)
 - Mean Accuracy Score: 80%

Current Statistic

The seller who use Amazon Platform has a higher chance to win the buy box.

- The items that lacked a shipping price were sold by Amazon or sellers in the “Fulfilled by Amazon” program
- Amazon-linked products **higher rankings** in more than 80 percent of cases.

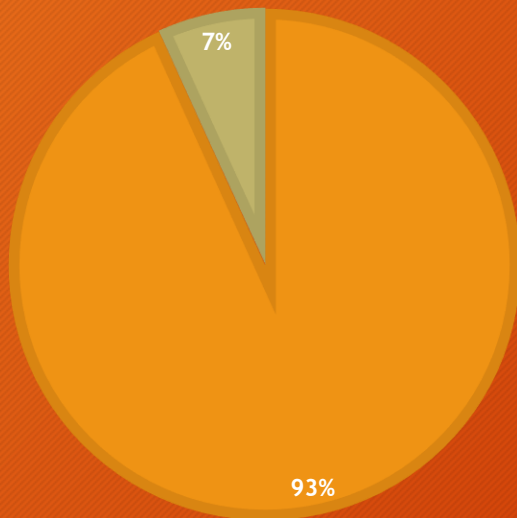


Current Statistic

Amazon = $179 / 192 = 93\%$

AMAZON

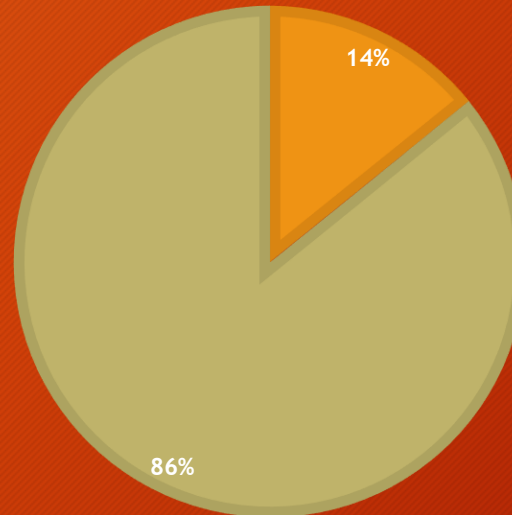
Win Buy Box Didn't Win Buy Box



FBA = $58 / 409 = 14\%$

FBA

Win Buy Box Didn't Win Buy Box



Others = $105 / 6372 = 1\%$

OTHERS

Win Buy Box Didn't Win Buy Box

