



# Improving Maryson's Sales

Cordelia Aitkin

Slides: FPPT.com

# Maryson's

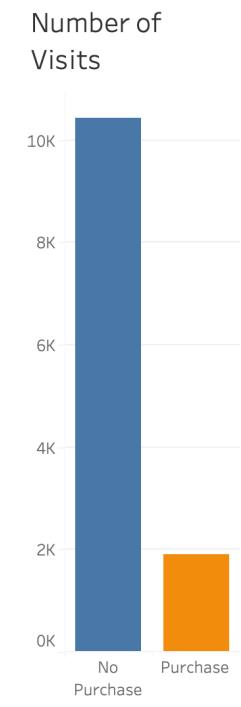
- You get a lot of visitors but not a lot of sales
- Improve sales rate
- Publicly available data

# Data Available



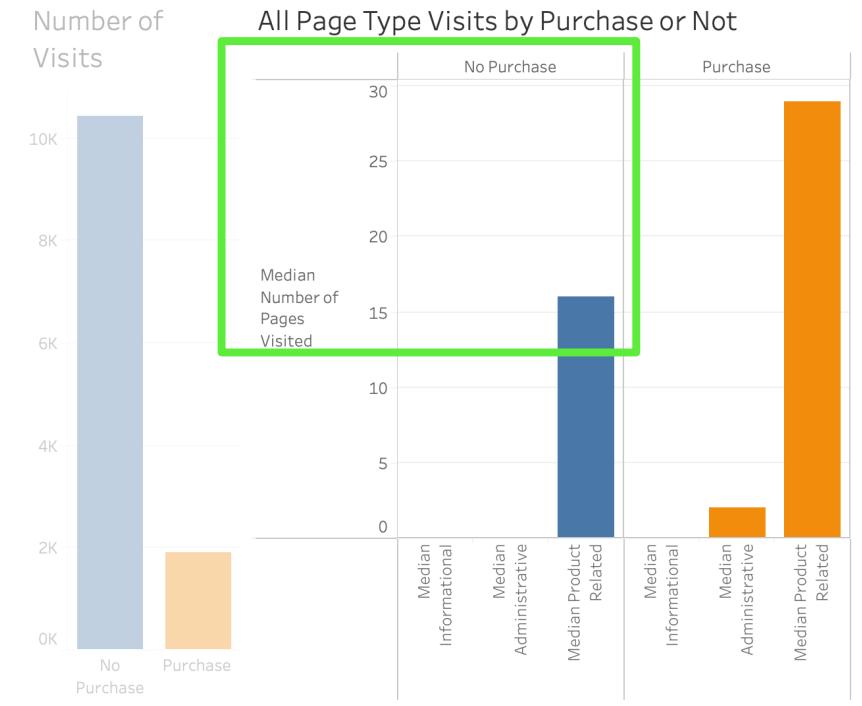
# Data

- 12,330 visits over the course of approximately a year
  - Missing: January, April
- 10,442 visits resulted in NO purchase



# Data

- Visitors who don't purchase anything do spend time on multiple pages



# Opportunity



# Proposal

- Convert more visits to sales
- Impacts
  - Increase revenue
  - Increase reputation

# Measures of Success

- Increase current sales by 5% in six months
  - 100 additional sales
- Understand
  - Predict users' behavior
  - Understand website

# Method

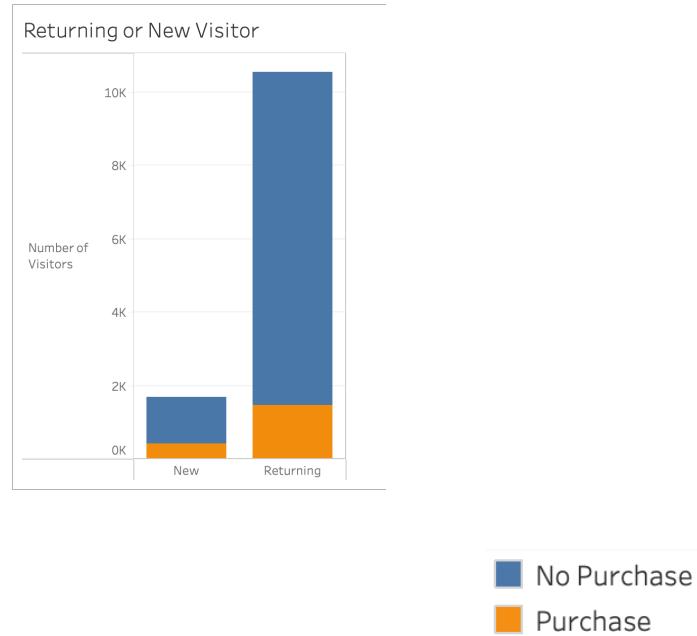


# Method

- Convert more returning visitors to sales

# Returning Vs. New Visitors

**Most of the site visitors are return visitors, and more absolute purchases come from return visitors...**

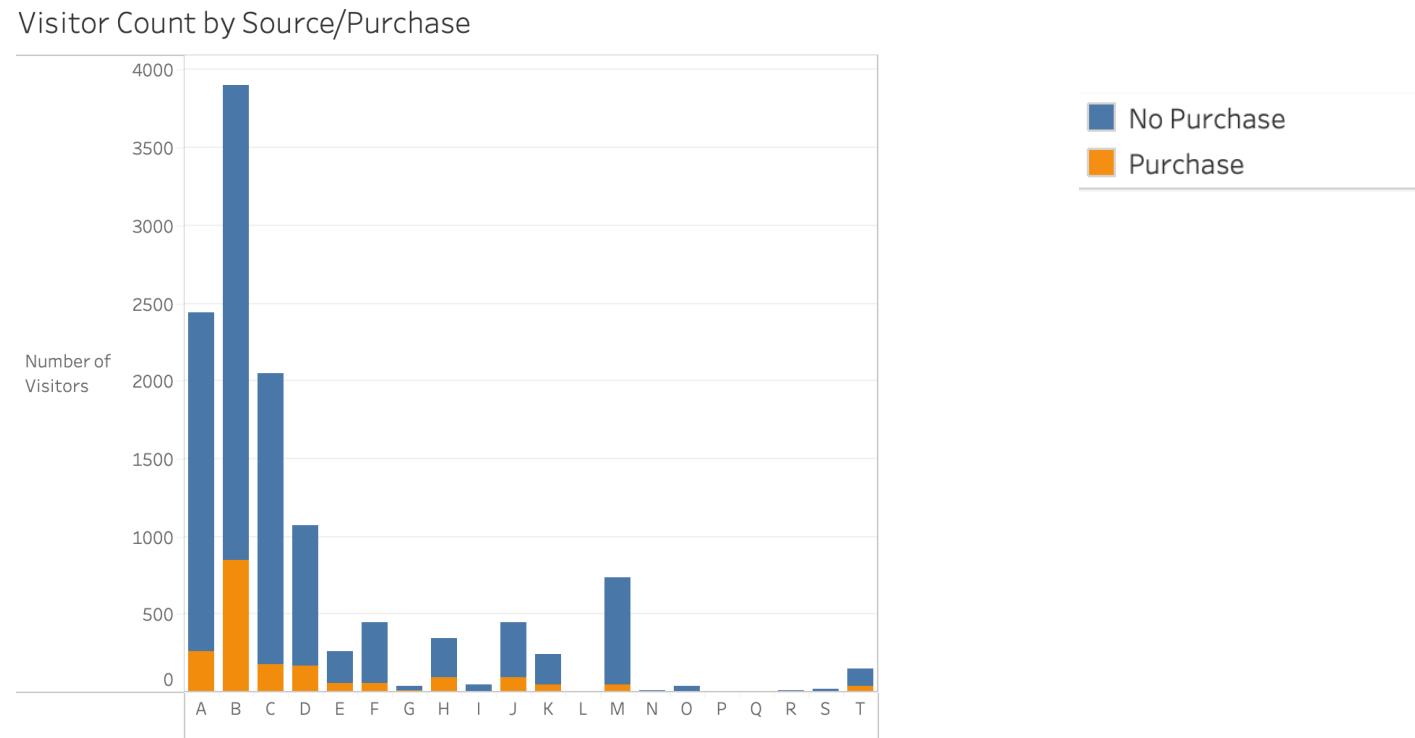


# Method

- Convert more returning visitors to sales
- Some origins of traffic produce better results

# “Source”

“Source”: where the visitors come to the site from, such as search engines, banner ads, etc.



# Design



Photo by [Priscilla Du Preez](#) on [Unsplash](#)

# Design

- Analyze the pages which current visitors visit to find exit points

# Design

- Analyze the pages which current visitors visit to find exit points
- Analyze “click paths” through product pages to see if there are any paths that turn off potential customers

Markov Model

# Design

- Analyze the pages which current visitors visit to find exit points
- Analyze “click paths” through product pages to see if there are any paths that turn off potential customers
- Analyze “Source” to increase presence on generative sources

Logistic Regression

# Data Needed



Photo by [Tamara Bellis](#) on [Unsplash](#)

# Data Needed

- Order of clicks within the website

# Data Needed

- Order of clicks within the website
- Specific exit pages

# Data Needed

- Order of clicks within the website
- Specific exit pages
- What letters in “Source” mean

# Future Directions



# Future Directions – Mine the Details

- Analyze which product pages produced best results
  - Is there something different about the product pages that result in sales?
- See if can determine why return visitors are not purchasing as much as expected
  - Are there return purchasers?
- Google Analytics “Page Rank”
  - Find which pages most effective



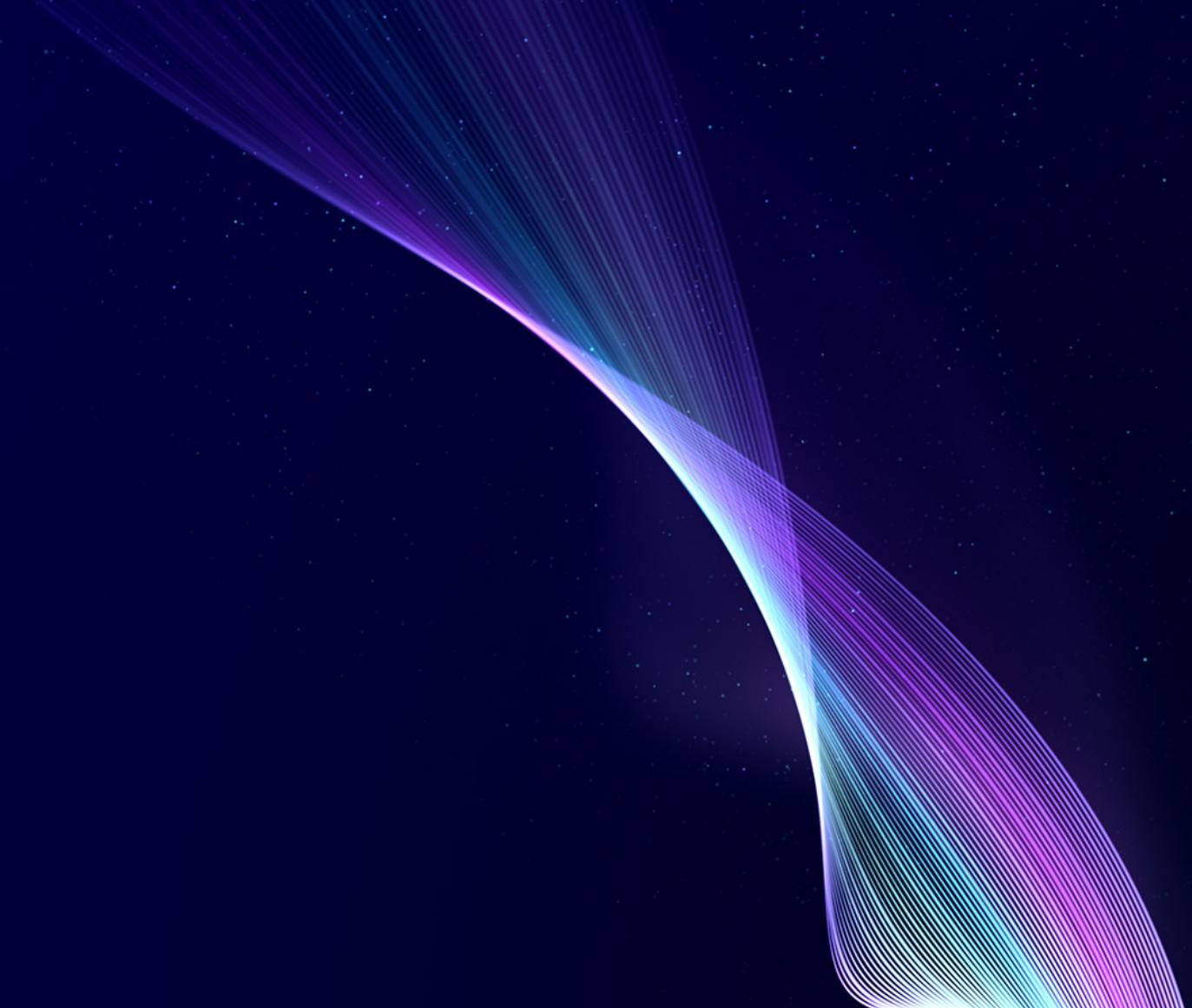
# Thank you!

Photo by [Alexi Romano](#) on [Unsplash](#)

# Credits

- <https://archive.ics.uci.edu/ml/datasets/Online+Shoppers+Purchasing+Intention+Dataset>
  - Sakar, C. Okan, Polat, S. Olcay, Katircioglu, Mete, & Castro, Yomi (2018) Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. Neural Computing and Applications <https://doi.org/10.1007/s00521-018-3523-0>
- FFTP.com
- <https://www.pixelstalk.net/fashion-wallpaper-high-resolution/>
- <https://media.istockphoto.com>
- <https://unsplash.com/>

# Appendix



# Data

## Continuous Used

- Product page duration (minutes)
- Administrative duration (seconds)
- Informational duration (seconds)
  - Product page count
  - Administrative page count
  - Informational page count
    - Page Value
    - Bounce rate
    - Exit rate

## Categorical Used

- Month
- Revenue \*\*

\*\*This is our Purchase/ No-purchase variable!

# Analysis

- Order of clicks within the website
  - Markov Model
  - Random forest
- Specific exit pages
  - Logistic Regression: individual pages on purchase
  - Grouping: something similar predicting exit pages
- “Source”
  - Logistic Regression: source on purchase
- Which pages produce best results
  - Logistic regression: individual pages on purchase
  - Grouping: something similar predicting purchases
- Why return visitors aren’t buying as much
  - Logistic Regression: click patterns on purchase
  - Markov Model
- Google Analytics
  - Logistic regression: page rank on purchase

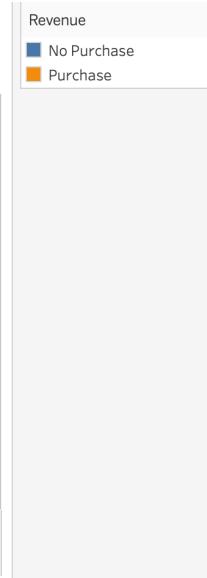
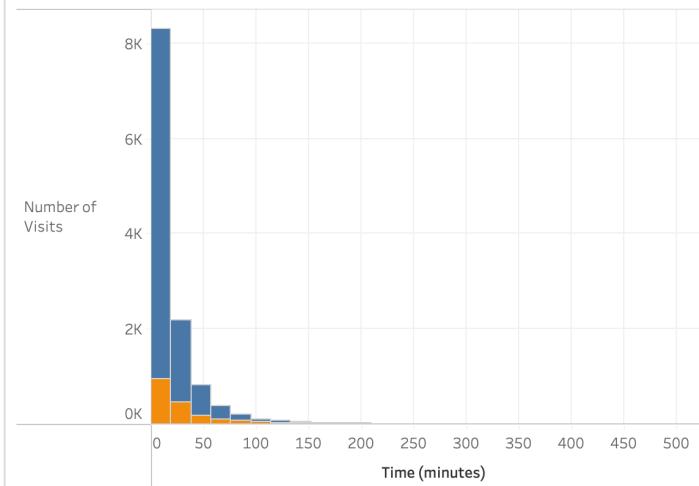
# Expected Vs. Observed New Vs. Old

		Observed		
		Bought	Didn't	Total
Visitor Type	New	422	1272	1694
	Old	1470	9081	10551
	Total	1892	10353	12245
		Expected		
		Bought	Didn't	Total
Visitor Type	New	261.7434	1432.2566	1694
	Old	1630.257	8920.7434	10551
	Total	1892	10353	12245
p value:		3.87E-31		

# Data

- Visitors who don't purchase do spend time on the website

Number of Visitors by Time Spent on Product Pages (minutes)

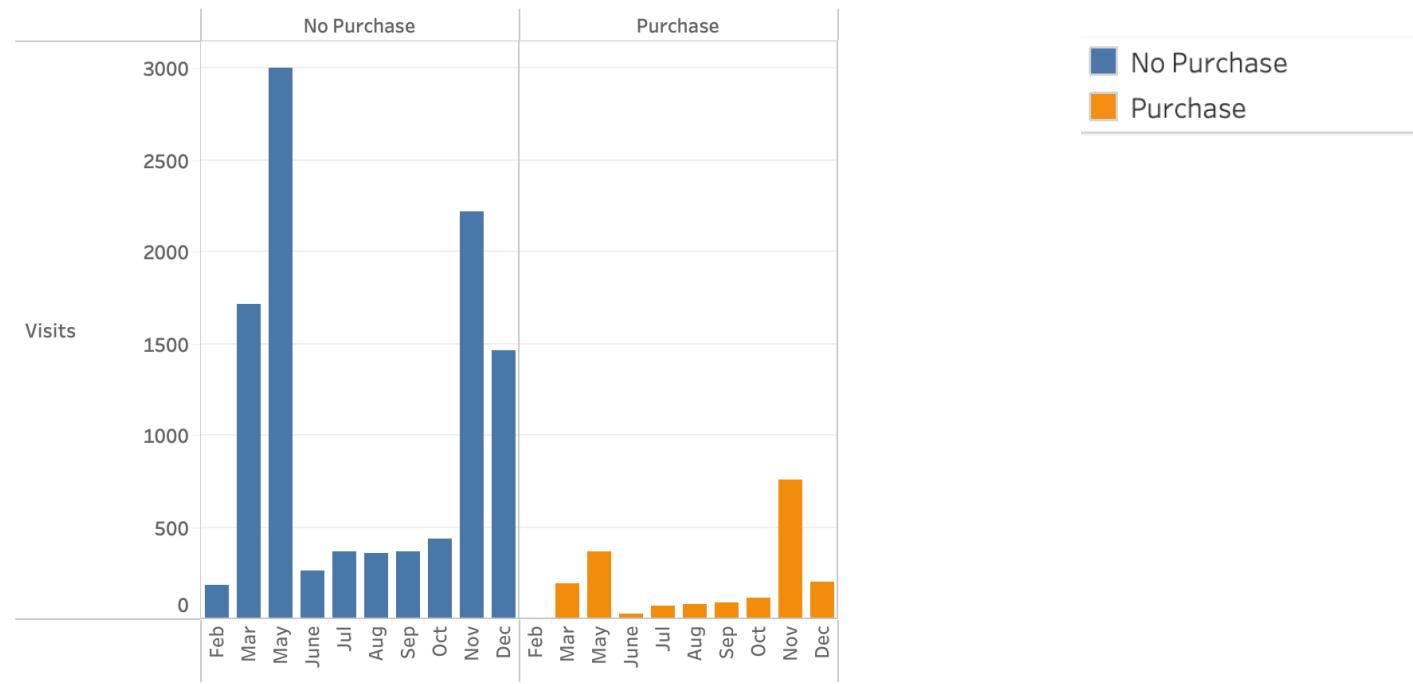


# Method

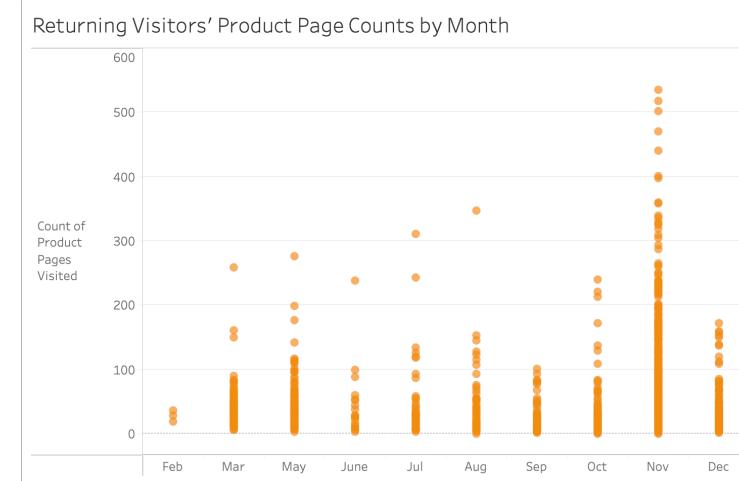
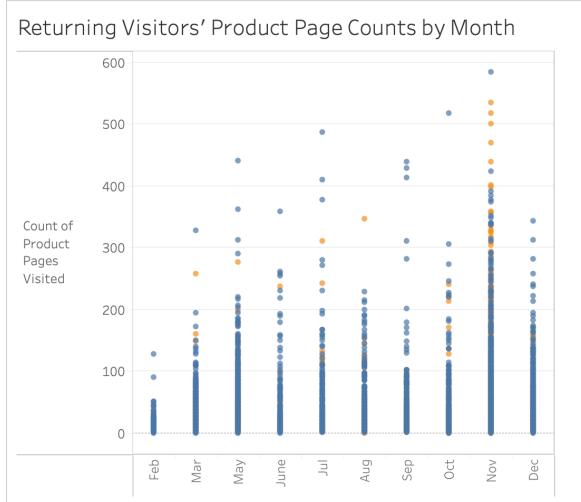
- Convert more returning visitors to sales
- Some months are more popular than others

# Monthly Visits

Number of Visits Each Month by Purchase or Not



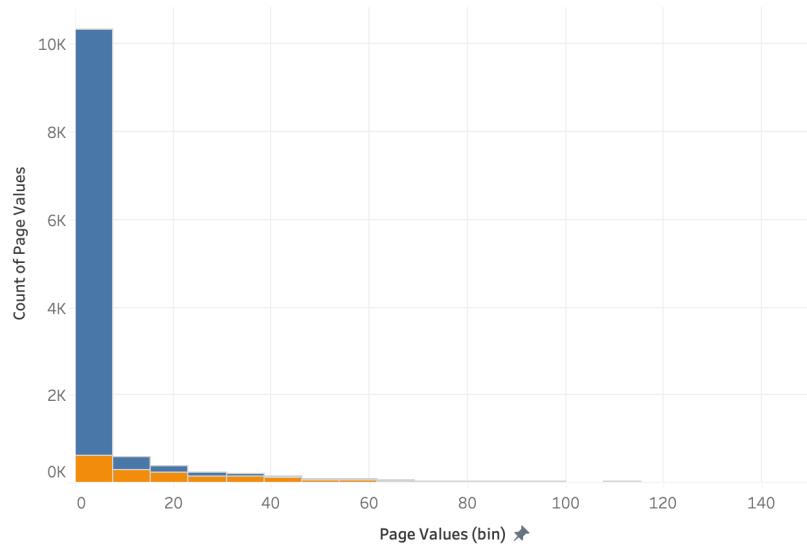
# Month by Purchase or Not



■ No Purchase  
■ Purchase

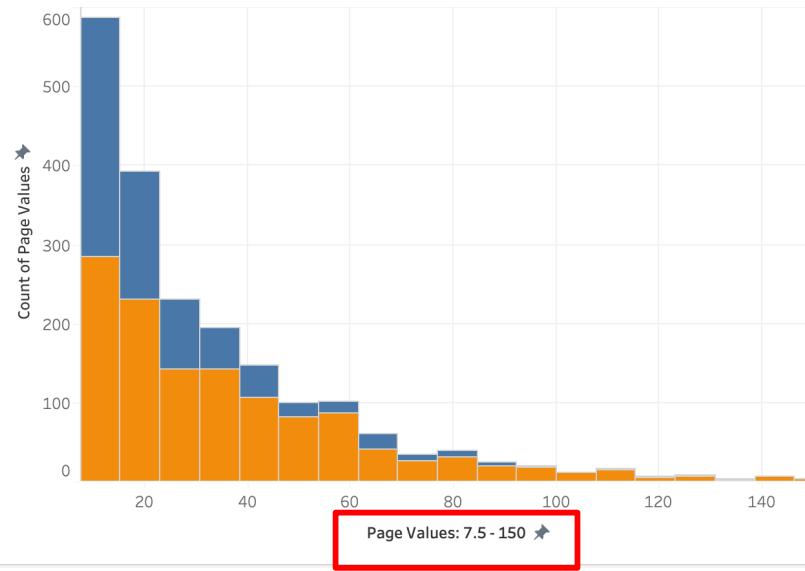


Visits by Page Value Only up to Value = 150



■ No Purchase  
■ Purchase

Visits by Page Values - Zoom In

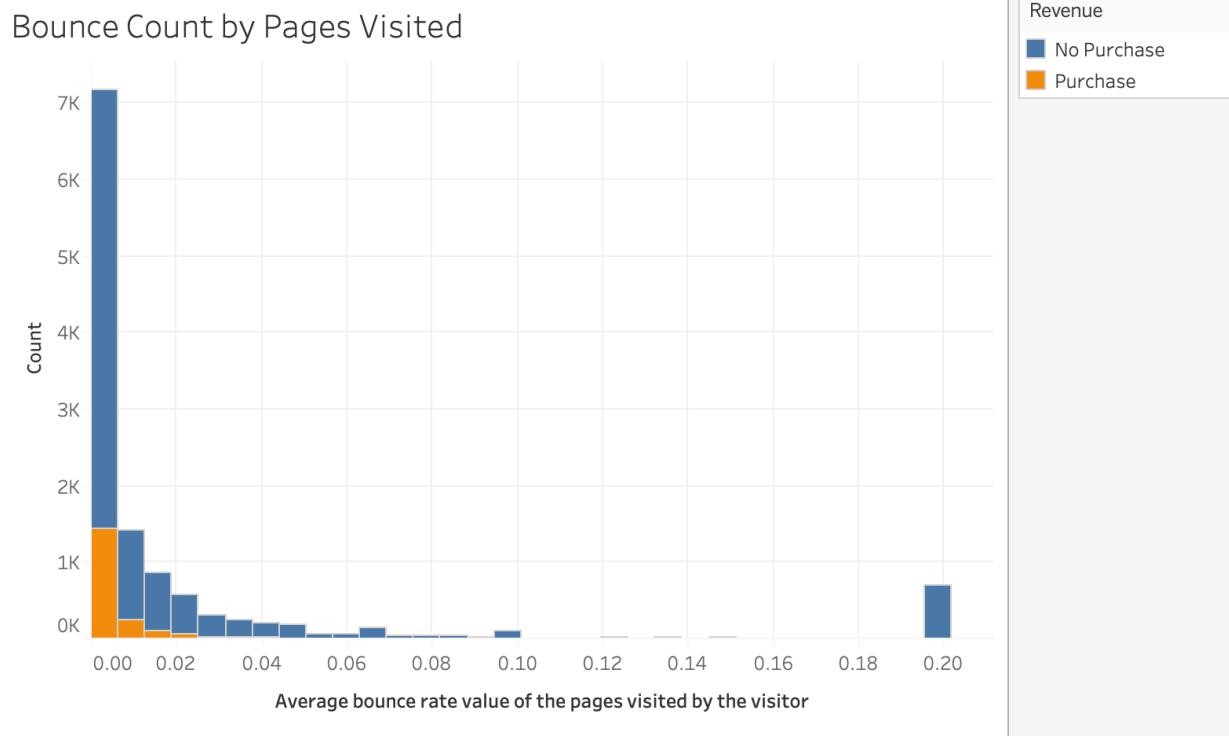


Logistic regression score: 0.88

# Some Jargon

- Exit rate: the percentage of times a single webpage is where a customer leaves the website
- Bounce rate: the percentage of times a single webpage is visited and the customer leaves without doing anything else on the website
  - E.g. didn't move to another webpage
- Exit Rate and Bounce Rate indicate which from webpages people in general leave the website
  - **In this data set, they are not the webpages where the visitor left the website**
- Page Value: a metric that analyzes how much a webpage contributes to revenue

# Bounce Count



# Exit Count

