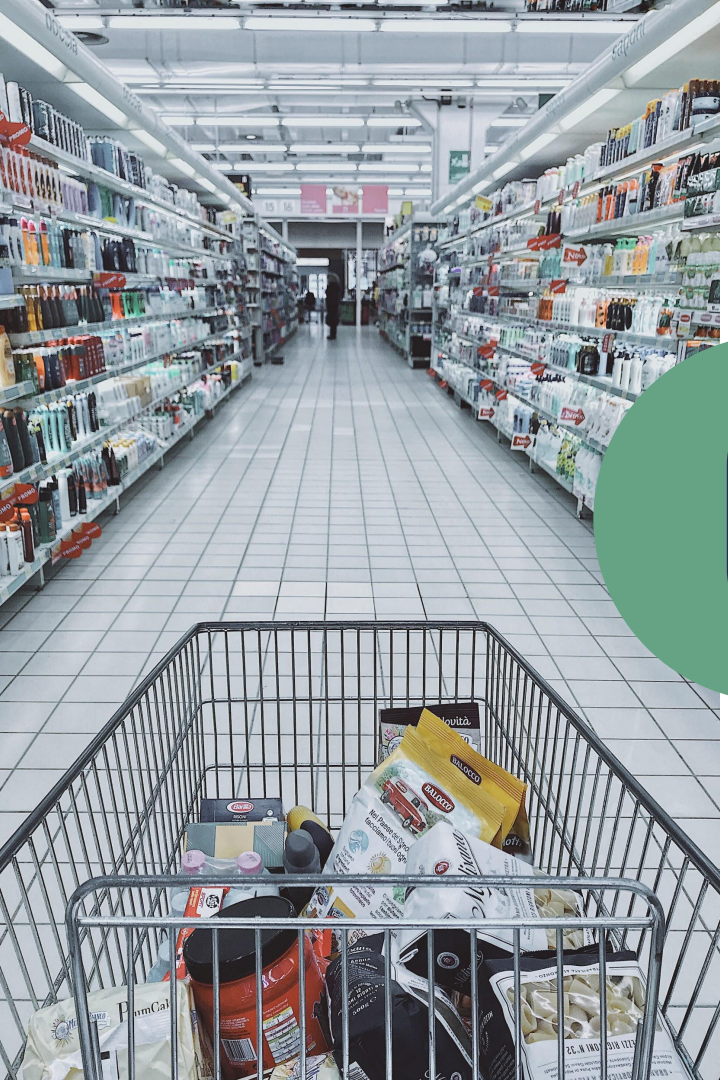
In Search of Spending

Jim Fay

Table of Contents



01

The Problem & Context

The E-commerce Market

$602B

Online Spending (2019)

16%

% of Total Retail (2019)

75%

Frequent Shoppers

Our Goal

- Offer **personalized marketing**.
- Increase **revenue** from online store.
- Determine **which factors** influence consumer spending.

02 The Data

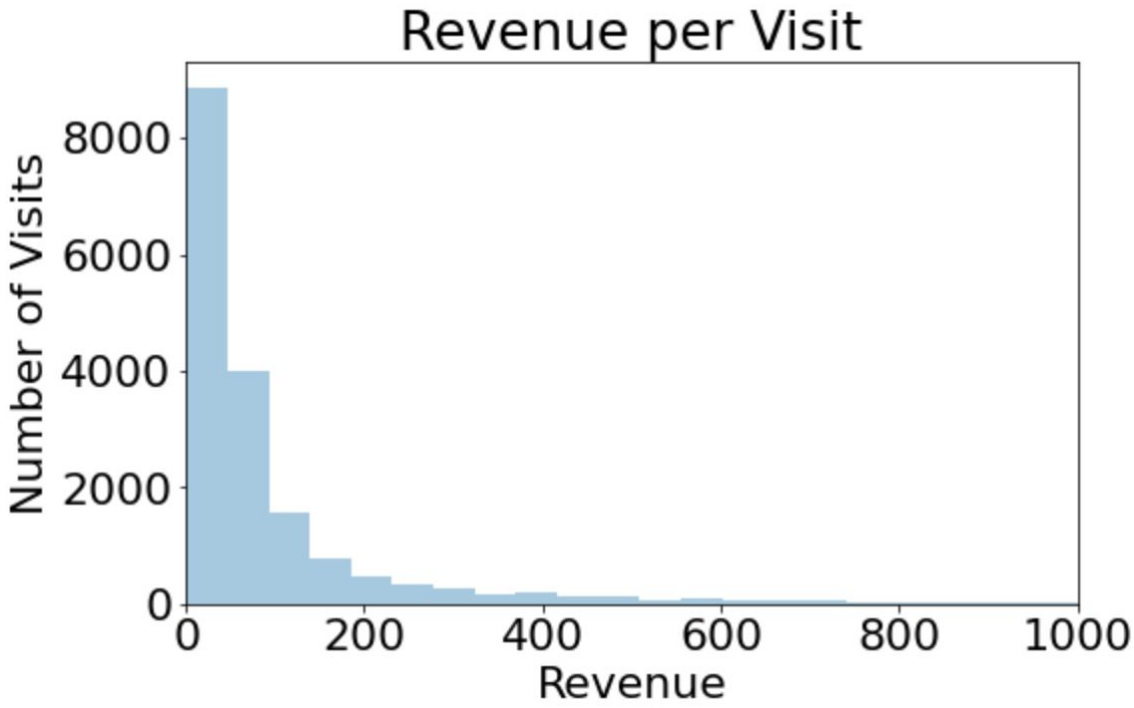
The Data

Breakdown

- 717k visits to Google's online store
- 2016-2018
- Variable Types:
  - Geographical
  - Economic
  - Device
  - User Activity
  - Price
- Merged economic data with Google data from Kaggle.*

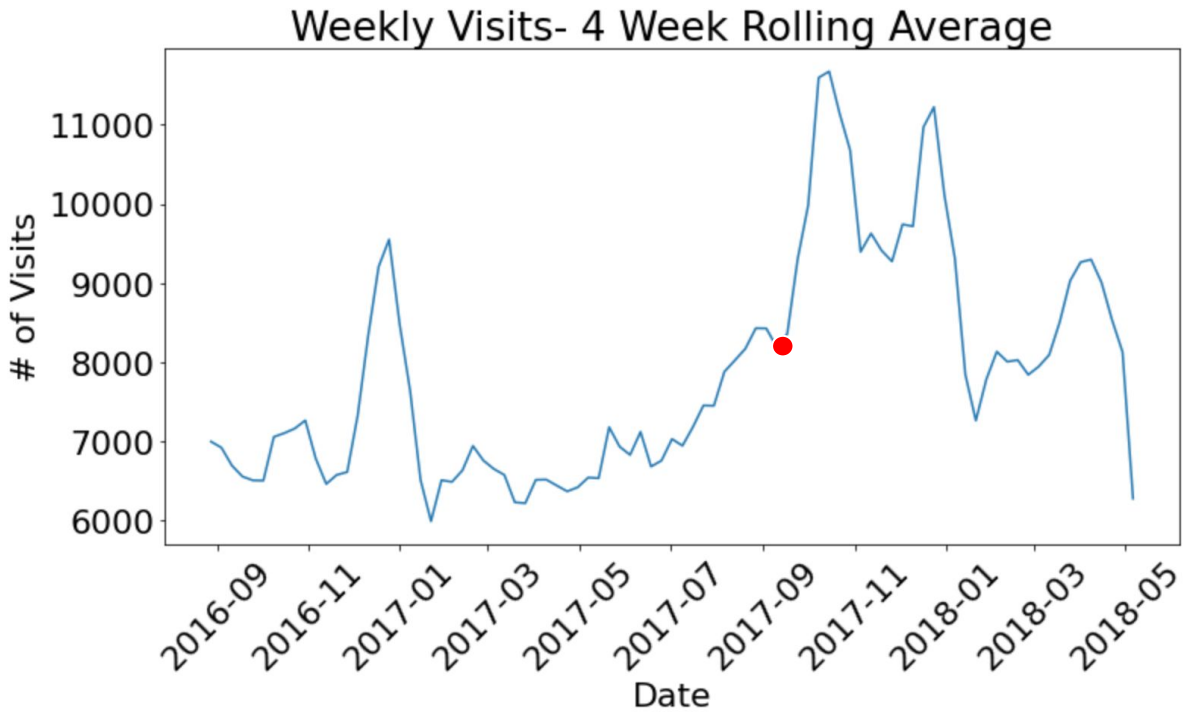
Breakdown: Purchases

- 2.46% of visits result in a purchase.
- Average Purchase: $124



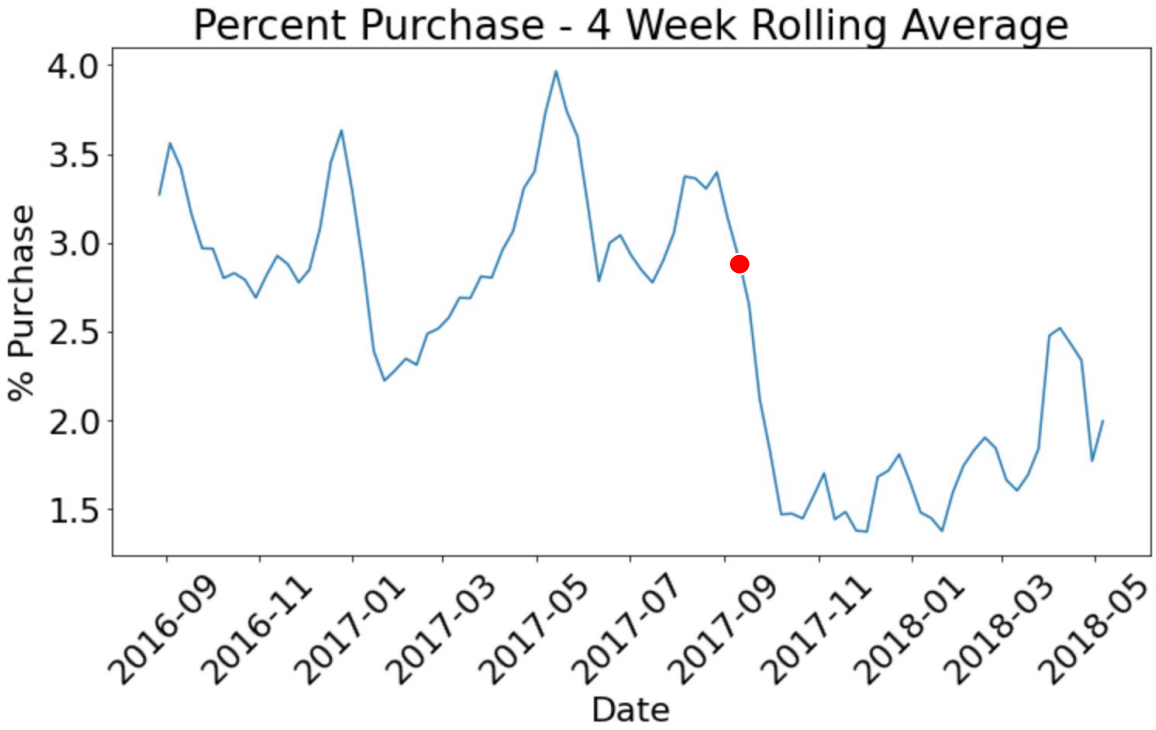
Purchases Over Time

- Visits increase with launch of new products.*



Purchases Over Time

- New visitors don't make purchases at the same rate.



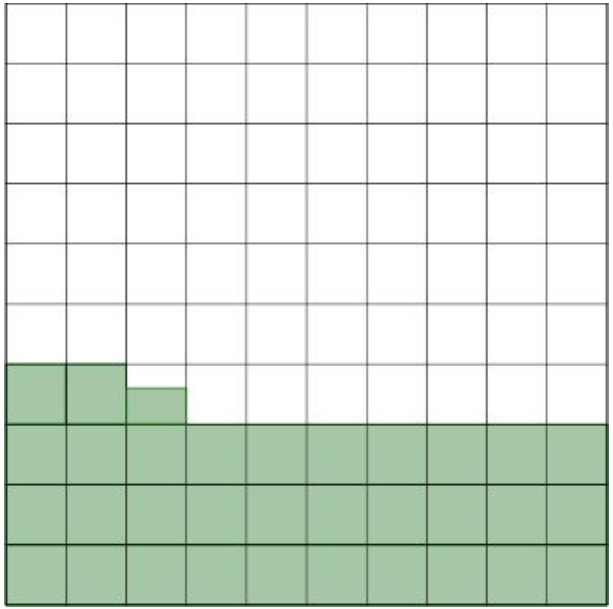
03 Data Analysis

The Model

Random Forest Performance

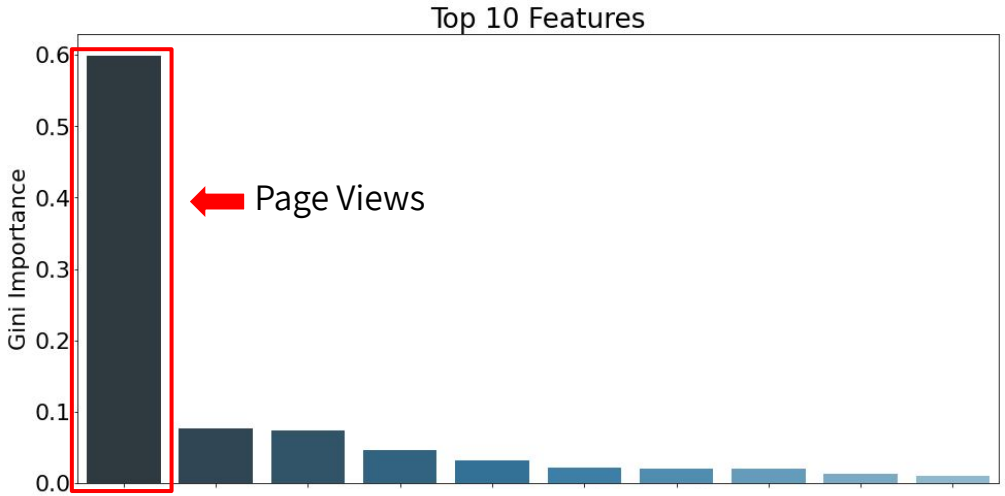
- Our model can explain 32.6% of variance in revenue.

- Model Features
- Other Explanations
  - Customer Income
  - Previous Purchases
  - Cart Data



Predicting Spending

Model Results



Interpretation:

- **Activity** and **time** spent on the site are key factors.

Predicting Spending

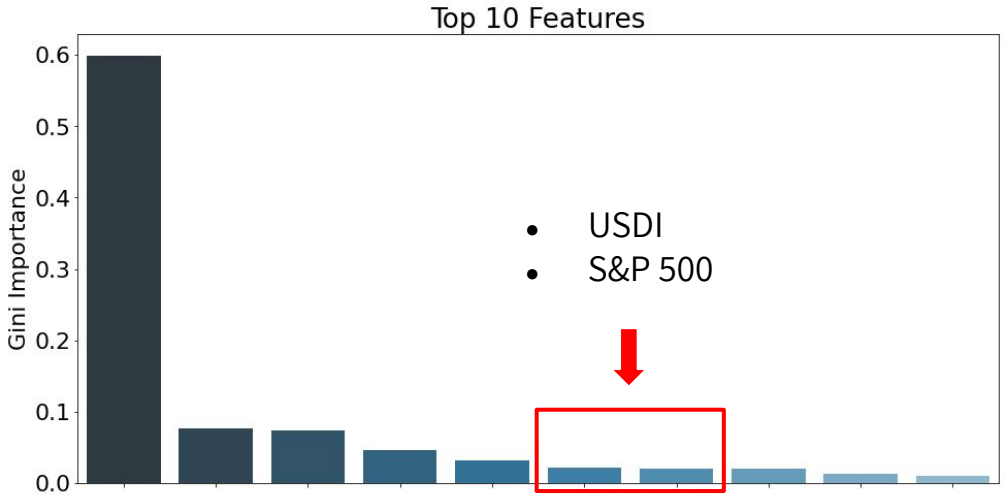
Model Results

Interpretation:

- **Activity** and **time** spent on the site are key factors.

Predicting Spending

Model Results



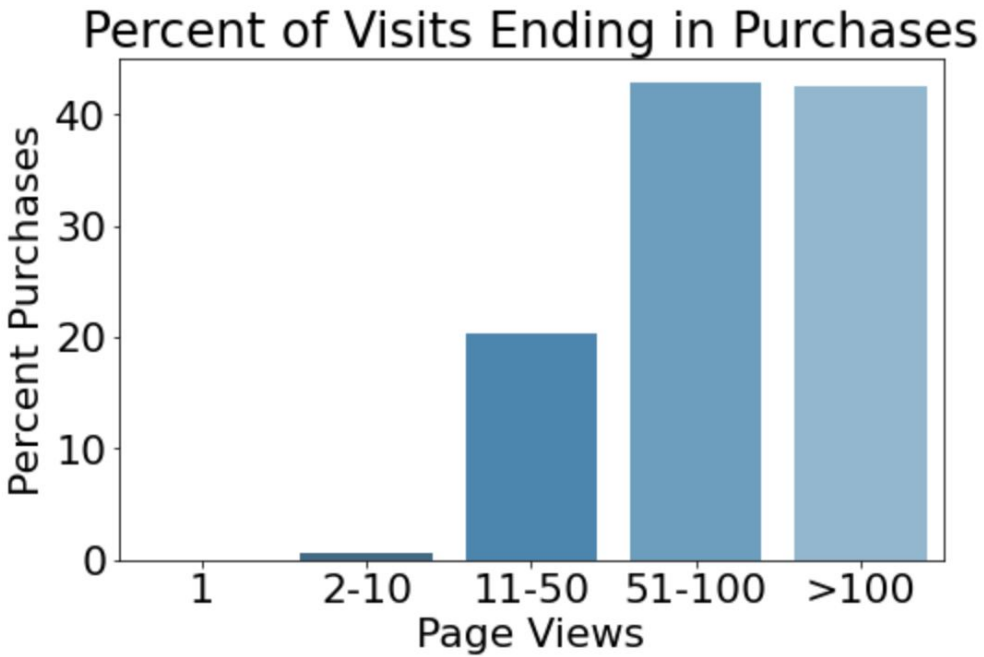
Interpretation:

- **Economic** environment is also important.

Page Views

Influence on Spending

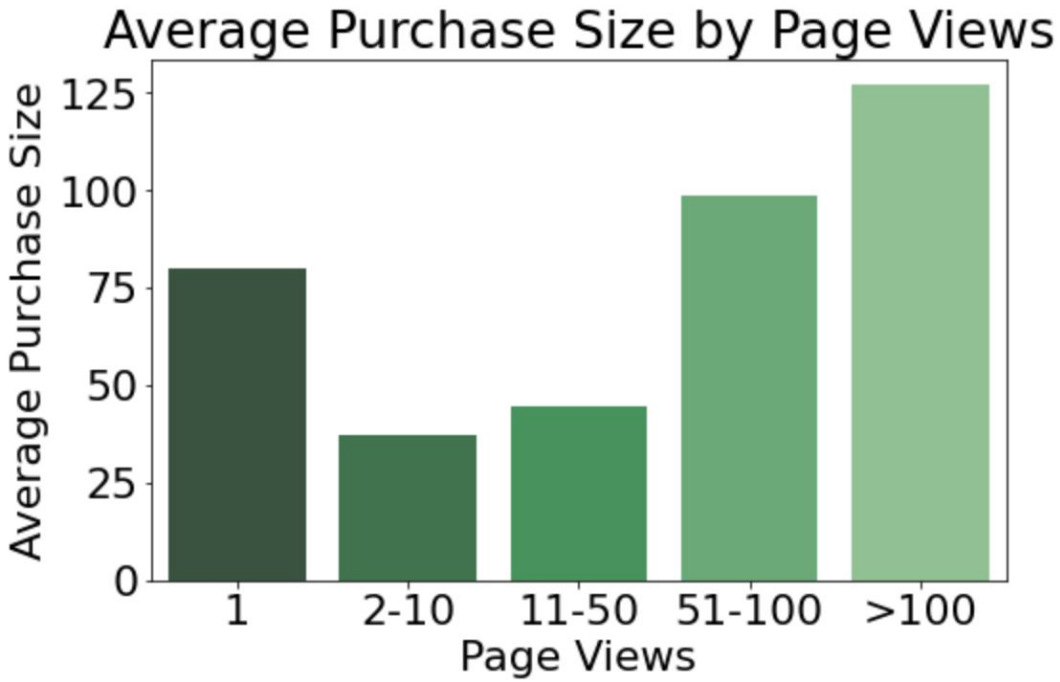
- More page views → Higher purchase **rate**.



Page Views

Influence on Spending

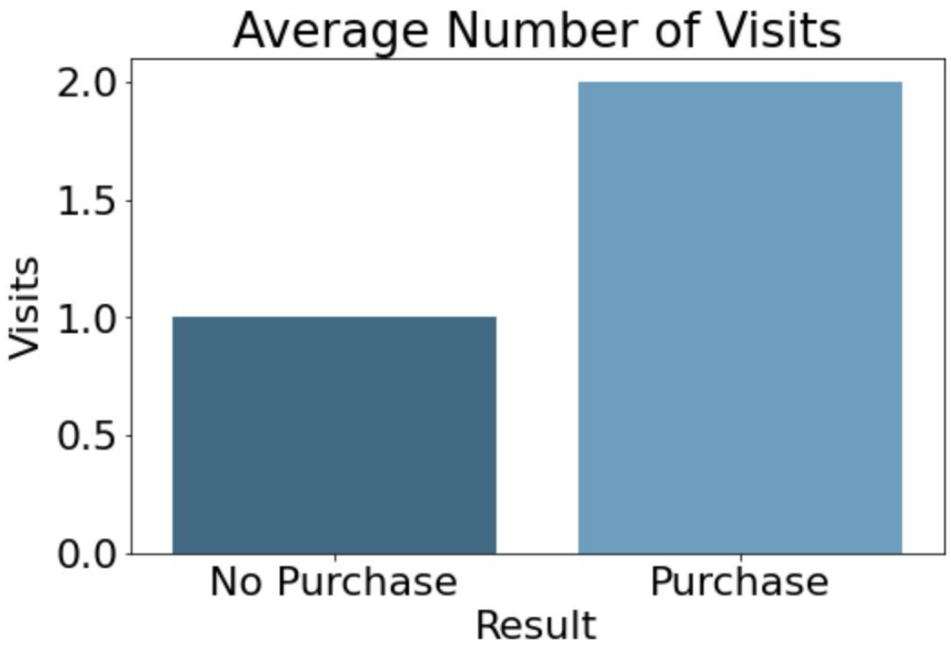
- More page views → Higher purchase **size**.



Number of Visits

Influence on Spending

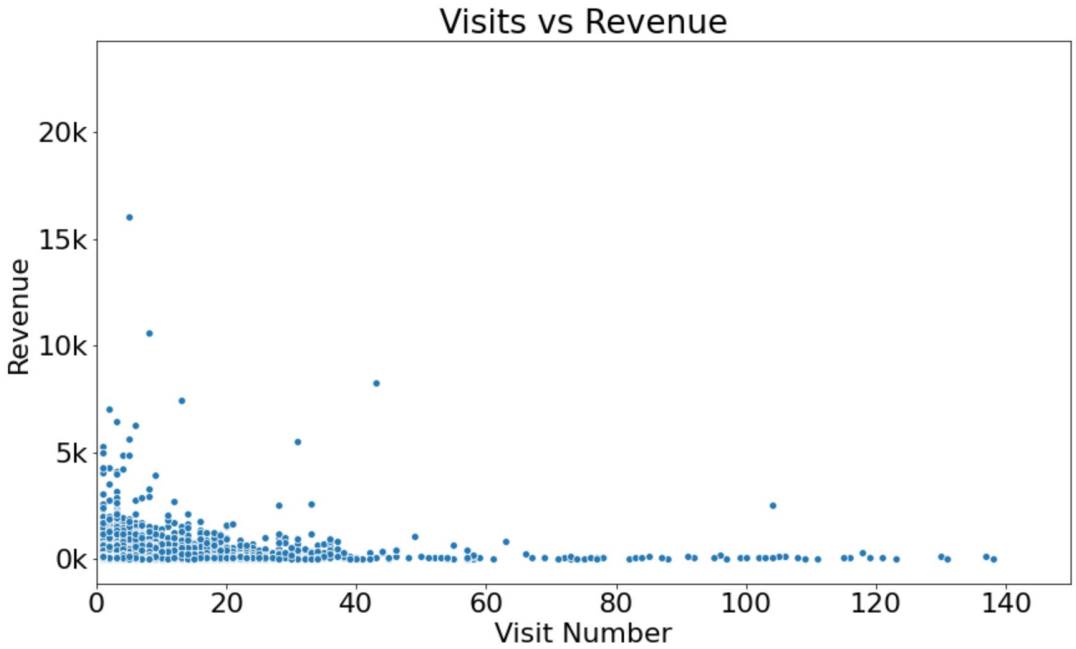
- Repeated visits → Purchases



Number of Visits

Influence on Spending

- Most purchases within first 25 visits.
- Purchases drop after 40 visits.



Recommendations

Keep Customers Coming Back

- Keep customers on the site.
  - Better product recommendations
  - Improved UI

- Offer discounts to buy within 25 visits.

04

Future Improvements

Limitations

The Data

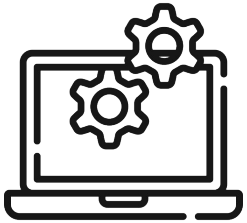
Unavailable data

Generalizability

Sampling

Was there a **pattern** to data collection?

Is it **exhaustive**?



The Model

Limited **predictive** power

Computation **time**

Improvements

Additional Data

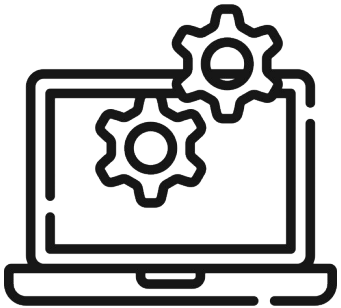
Consumer **Income**

Past **Purchases**

Cart Info

Deployment

Deploy the model as a web application



Improve Model

Try **different** ML models

Run on data from **other stores**

Questions?

jrf6xh@virginia.edu

linkedin.com/in/james-fay/

github.com/jrf6xh

* All images from pexels