Supply Chain Modeling and Analysis

Nikhil Anand, Serena Lakhiani

June 7, 2023

Problem Overview

- Inventory Management and Supply Chain Analysis are vital to any business, as well as general nationwide and global markets
- Customer and inventory segmentation are key methods in marketing analytics that drive business decisions using information on customer behavior and company resources
    - Status-quo methods of analysis, such as RFM Analysis and ABC Classification, can be improved with machine learning
- Possible lack of use of ensembles so the team hopes to generate high(er) performing models in a few areas of analysis

Status of Project Deliverables

- EDA
- ABC Segmentation
  - Rule-based
  - Multi-criterion decision-making with Machine Learning
- RFM Segmentation
  - Rule-based (IP)
  - Clustering
- Shipping Time Classifier
- Multi-Approach Customer Segmentation with Machine Learning (IP)
- Statistical Hypothesis Testing

Using ML to Extracting Meaningful Information

Shipping Time Classification

- Estimate shipping time:
  - < 2 days
  - 2-3 days
  - > 3 days
- Based on order details:
  - Shipping origin and destination
  - Product information (item, quantity, price)
  - Shipping Mode selected

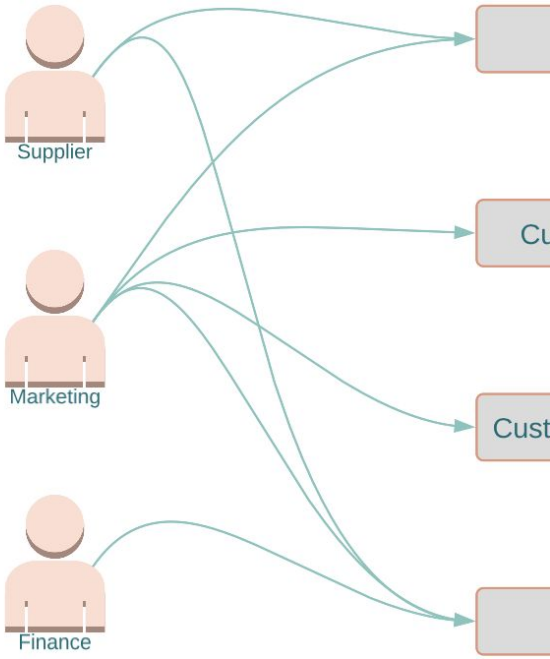
Customer Segmentation (RFM)

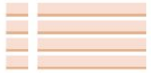
- Similar basis to RFM Analysis, but using ML to optimize clustering and minimize total groups
  - K-means
  - DBSCAN
  - Agglomerative

Customer Segmentation (ext)

- Use gender-guesser to guess customers' genders based on first name
- Use (limited) demographic data in addition to purchasing history to cluster customers

Informal Use Case Diagram





- Aggregation of sales history of items
- Run rule-based calculations
  - A = 75% of revenue
  - B = 20% of revenue
  - C = 5% of revenue

- Preprocessing
- Aggregation of purchase history
- Apply clustering algorithms
- Hypertune to find optimal parameters
- Evaluate clusters

- Preprocessing
- Aggregation of purchase history (less intensive)
- Guess gender based on first name
- Apply ML algorithms
- Hypertune to find optimal parameters
- Evaluate model performance

- Establish Null Hypotheses
- Create groups
- Run t-test/ANOVA
- Determine statistical significance to [not] reject null hypotheses

To the codewalk!