IE 492: GRADUATION PROJECT

FINAL REPORT

Measuring the Effects of Free Shipping on Sales and Price

Hüseyin Emre BACAK

Tevfik Buğra TÜRKER

Fatmanur YAMAN

Project Advisor: Taner BİLGİÇ

Abstract

The number of people using online shopping platforms is increasing gradually in the world. As a result, huge and competitive markets are formed. To prosper in these harsh market conditions, sellers offer customers various benefits. One of the commonly used methods is the contingent free shipping application. For sellers to be successful in this application, the threshold value for the contingent free shipping option should be adjusted accurately. In this project, the optimization of the threshold value for contingent free shipping is discussed. Throughout the project, methods such as effect formulation, statistical distribution fitting, and scenario simulations were utilized. As a result, using the discussed method, it was possible to obtain a %5.36 increase in revenue with the analyzed data.

Key Words: Contingent free shipping, threshold value, shopping basket, e-commerce, revenue maximization, predictive analysis, order padding

Özet

Dünyada her geçen gün, online alışveriş platformlarını kullananların sayısı artmaktadır. Bu artış sonucunda devasa ve rekabetçi marketler oluşmaktadır. Satıcılar, şartların zorlu olduğu bu ortamlarda satışlarını arttırmak için müşterilere belirli avantajlar sunmaktadır. Bu avantajlardan bir tanesi de koşullu bedava kargo uygulamasıdır. Satıcıların bu uygulamada başarılı olabilmesi için bedava kargo eşiğini doğru ayarlaması gereklidir. Bu projede, koşullu bedava kargo uygulaması için eşik değerinin optimize edilmesi konusu tartışılmaktadır. Efekt formülasyonu, istatistiksel dağılım uygunlaştırması ve senaryo simülasyonları gibi yöntemler kullanılarak yürütülen projenin sonunda önerilen yeni eşik değeriyle birlikte, üzerinde çalışılan modelde %5,36 oranında gelir artışı elde etmek mümkün olmuştur.

Anahtar Kelimeler: Koşullu ücretsiz gönderim, eşik değer, alışveriş sepeti, e-ticaret, gelir maksimizasyonu, tahmine dayalı analiz, sipariş doldurma

Table of Contents

1. Introduction

In recent decades, e-commerce and online platforms have gained significant importance due to the advancements in digital technology and the proliferation of the internet. Especially during the coronavirus pandemic, the share of online shopping in everyday life has spiked, almost replacing the traditional brick-and-mortar shopping style. This is no surprise since online platforms offer substantial advantages over their conventional counterparts. With online shopping, customers can browse hundreds of thousands of products in seconds and order anything they would like to without being physically in a store. This shift in consumer behavior towards online means has led to a change in market structure. Furthermore, due to the effect of globalization, the competition between market players is higher than ever. To survive in this turmoil, all players, whether small or large, should develop reliable and innovative strategies.

One of the fundamental differences between online shopping and conventional shopping is the delivery of the products after purchase. In contrast to conventional shopping, in online shopping, the ordered products must be shipped to the customer, which creates an additional cost. In the literature, there are various implementations regarding who will bear the shipping cost and how. The most basic and common types of shipping are *free shipping* and *paid shipping,* where in the former, the shipping cost is covered by the seller and in the latter by the customer. Apart from those, in order to gain a competitive advantage, market players have come up with more complex shipping models which resulted in the advent of *contingent free shipping*. In this type of shipping, customers need to satisfy some pre-conditions to obtain the opportunity of free shipping. Depending on the market structure and customer expectations, different types of pre-conditions for contingent shipping can be utilized. However, the focus of this study will be mostly on the threshold approach, in which customers need to spend more than a predetermined threshold to qualify for free shipping. With this approach, sellers aim to incentivize customers to spend more, consequently increasing average order value and total sales. Nevertheless, finding the optimal threshold value is not straightforward and depends on multiple complex factors.

In this project, the effects of shipping options are discussed in the light of real sales-transaction data from an online platform. One of the major aims of this project is to find an effective method to decide on the threshold for contingent free shipping. For the identification and analysis of the data, various machine learning algorithms such as Apriori and K-Nearest-Neighbor are used. After that, customer behavior is captured with the combination of two different statistical distributions. Then utilizing these distributions in various simulations, different scenarios are tested. According to the results of these scenarios, the optimal threshold value is found to provide an additional expected increase in revenues by %5.36 compared to the current level used in the data. Furthermore, the analysis is not limited to this data alone. The method suggested in this project to determine threshold values can be applied to a wide variety of products, categories, and subsets.

2. Problem Definition, Requirements and Limitations

2.1. The Problem Definition

The online platform owner and the sellers in the platform need to decide on the optimal threshold level for contingent free shipping to be successful in the competitive structure of the market. Finding the optimal level is vital for revenue maximization by convincing the users to purchase more products with contingent free shipping. The problem to be solved is adjusting the threshold value for the contingent free shipping in a way that the gained revenue from those can cover and exceed the shipping cost spent. Assuming that the market size stays constant, what should be the threshold value is the main question.

2.2. Requirements

There are some key requirements for feasibility and success. From the technical view, a proper statistical model that fits very well with the customer behavior and also predicts the future behavior of the customers regarding the changing threshold value is the main requirement. After that, an accurate way for the simulations is also required. From the business point of view, the optimal threshold value suggested by the mathematical model should be adaptable to the changing market conditions and shipping cost dynamics. In addition to those, robust software tools that can handle large datasets and perform detailed statistical analysis and simulation are needed.

2.3. Limitations

There are several limitations to finding the optimal threshold value in the problem. First of all, there are some deficiencies in the data quality and availability. There is no Customer ID information to track the individual customer behavior. Thus, the individual demand is not observable and that’s why the market size is assumed to be constant throughout the project. In addition to this, there are some missing data that could not be handled by popular data-generating techniques. Moreover, the shipping cost is dynamic and is affected by a lot of variables in the data which makes the problem difficult to interpret and predict the approximate future costs. Because of that, the profit could not be calculated and be the objective of the problem.

2.4. The Data Dimensions and Terminology

The dataset consists of sales transaction data of an online e-commerce shopping platform between 1 March 2017 and 30 April 2017. There are 698.679 single transactions in the data. Here are some of the important terms that are used throughout the project:

- **Basket:** A complete order from a customer, which may consist of multiple items.
- **Shipping Types:** There exist three different types of shipping in the data: *Free Shipping*, *Paid Shipping,* and *Contingent Free Shipping*.
- **Threshold:** The minimum monetary value required in a shopping basket to qualify for contingent free shipping.
- **Categories:** Every sold product in the data belongs to one of the five main groups. Here are the main groups in the data with their abbreviations:
    - *BC* = Basic Consumption Goods,
    - *HEA* = Household Electronic Appliances,
    - *PET* = Pet Shop Products,
    - *TEL1* = Telephone Accessories,
    - *TEL2* = Telephones.
- **SKU:** A unique product I.D. which is used to identify products.
- **Basket Revenue:** The total value of the products in a basket.

CARGO_ID	[2444494, 2444494, 2444494]
DATE	[2017-03-12 00:00:00, 2017-03-12 00:00:00, 201...
SKU	[ZYSAN914883, ZYKRISTAL259168, ZYGUR930225]
PRODUCT_NAME	[Sarelle Şekersiz Kakaolu Fındık Ezmesi 350 gr...
SUB_CATEGORY	[Gıda Ürünü, Organik ve Dogal Ürünler, Gıda Ür...
SUB_CATEGORY_CODE	[3580, 3039, 3580]
QTY	[1, 1, 1]
Seller	[PO, PO, PO]
DESI	[1.0, 5.0, 1.0]
SHIPPING_COMPANY	[ARS, ARS, ARS]
CAMPAIGN	[100 TL üzeri çatı, 100 TL üzeri çatı, 100 TL ...
FreeSH	[False, False, False]
Weekend	[True, True, True]
Total_Price	[12.01, 98.1, 12.76]
Unit_Price	[12.01, 98.1, 12.76]
Revenue	[12.01, 98.1, 12.76]
Unit_SHF	[0.0, 0.0, 0.0]
Real_SHC	[0.478410972, 2.392054861, 0.478410972]
Category	[BC, BC, BC]
CFreeSH	[True, True, True]
Basket Revenue	122.87
RSC/VW	[0.478410972, 0.4784109722, 0.478410972]
R/RSC	[25.10393929677683, 41.01076509549168, 26.6716...
Sum_Real_SHC	3.348877
Sum_DESI	7.0
Sum_Real_SHC/Sum_DESI	0.478411
Basket_Revenue/Sum_Real_SHC	36.689913
Remove Order	False
Shipping_Category	contingent_free_shipping

Figure 1. *An example basket that displays all the dimensions*

An example basket is given above in **Figure 1**. The most important features are SKU, Basket Revenue, Shipping Category, and Seller. In the data, there are two different sellers in the platform, the marketplaces (MP or individual sellers) and the platform owner (PO) which compete with each other. It is important to note that there are no additional descriptions for each marketplace i.e. all marketplaces are aggregated and treated as a single MP. In addition to this, another important dimension in the data is the shipping type. Three different shipping strategies have been applied to the data. The first is Free Shipping, where the shipping cost is paid by the seller. The second is Paid Shipping, where the users pay the total amount of the shipping cost. Finally, Contingent Free Shipping is the strategy where the customers are exempt from the shipping cost if their total amount of purchased basket exceeds a certain threshold value. This threshold value is determined by the PO as 100 TL and the total basket amount is calculated regardless of the category, seller, and SKU of the items in the basket. Other than those, the Category feature helps in the interpretation of the effects of different shipping types on customer purchase behavior.

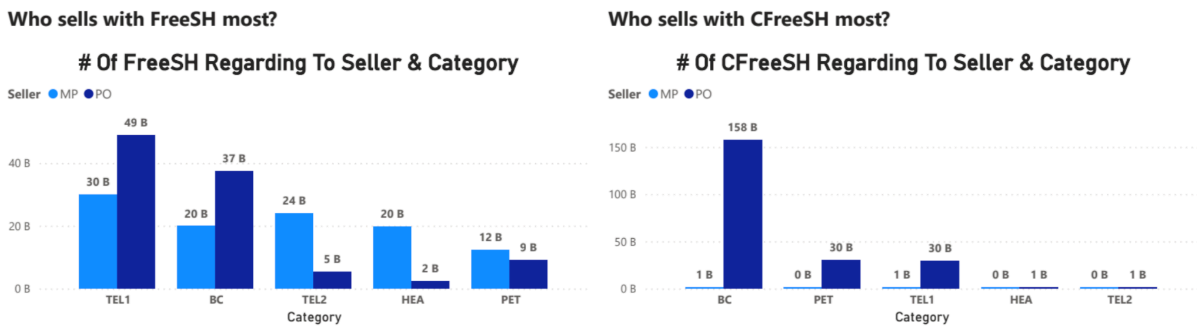


Figure 2. *Statistics regarding shipping types and categories*

According to **Figure 2**, while Free Shipping is dominant in TEL1 and TEL2 categories, Contingent Free Shipping is dominant in BC and PET categories. The meaningful categories for Contingent Free Shipping with average SKU prices under the initial threshold value are TEL1, PET, and BC categories. When a campaign is planned to promote Contingent Free Shipping, the customers who buy products from those categories should be considered as the main candidates. In addition to the category, when the Seller filter was added, while PO sold more with Free Shipping in TEL1 and BC categories, the MPs offered free shipping more in TEL2 and HEA categories. The distribution is approximately equal in the PET category. When it comes to Contingent Free Shipping, the PO is dominant in selling with this option. The share of the MPs is almost none.

2.4.1. Analysis of Seller Shipping Behavior: *Do the basket revenue and shipping cost affect the seller's decision on the shipping type?*

KNN is a lazy-clustering algorithm that searches for k number of neighbors and decides the cluster based on the majority of those k points. Since a huge dataset with numerous points exists, the KNN algorithm can draw the boundaries for different categories regarding shipping cost and basket revenue.

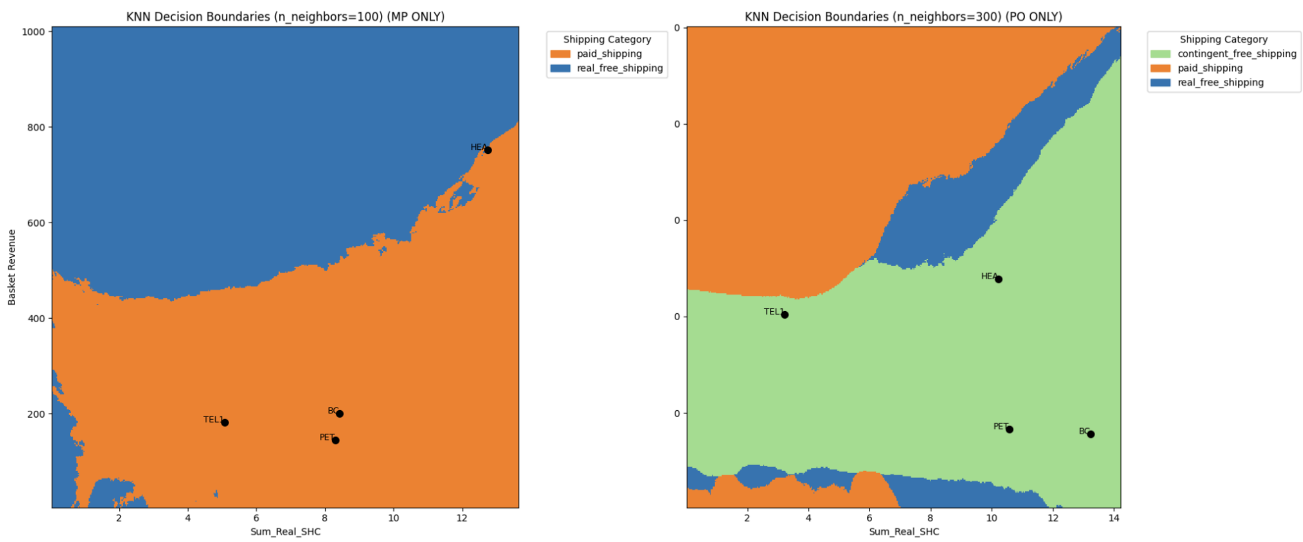


Figure 3. *Plots of Basket Revenue versus Total Shipping Cost of the basket for individual sellers (MP) and the platform owner (PO) after the KNN algorithm is executed*

As it can be seen from **Figure 3.**, the KNN analysis indicates that the PO is dominant in Contingent Free Shipping. However, the behavior of PO and MP is reversed when it comes to Free Shipping and Paid Shipping. While PO prefers to offer Free Shipping when both basket revenue and the shipping are high, MP follows a strategy to give Free Shipping to all baskets above a value. It can be said that the MPs are trying to catch the PO in the number of sales by offering Free Shipping. Lastly, when the shipping costs are low, the PO prefers to make the customer pay the whole amount of the cost.

2.4.2. Data Mining for Predictive Analysis of Baskets: *Which products are being sold together frequently and what are their effects on the total revenue?*

To understand what is being sold together and the overall content of the baskets, the market basket analysis has been made. The antecedent is the main basket content and the consequent is the tested item whether it is in the baskets with the antecedent. There are 3 metrics used in the market basket analysis. Support is the ratio of the occurrence of the itemset over the whole basket. Confidence is the ratio of occurrence of antecedent and consequent together over the baskets with the antecedent. The lift is the ratio of confidence over the support. The baskets with a support value above a threshold are selected. Those baskets can be used for suggesting dynamic shipping costs, strategies, and campaigns. In the below **Table 1.**, the most frequent baskets among all shipping types can be found.

Table 1. *Most frequent baskets in the data, regardless of shipping type*

support	itemsets
0,040	'Prima Bebek Bezi Aktif Bebek 5 Beden Junior Aylık Fırsat Paketi 150 Adet'
0,032	'Fairy Platinum Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 81 Yıkama 2li Paket + Platinum Sıvı Bulaşık Deterjanı 870 ml'
0,029	'Ariel Toz Çamaşır Deterjanı Dağ Esintisi 9 kg + Parlak Renkler 7 kg'
0,021	'Prima Bebek Bezi Aktif Bebek 4 Beden Maxi Aylık Fırsat Paketi 174 Adet'
0,020	'Finish Powerball Quantum Bulaşık Makinesi Deterjanı 144 Tablet 72x2 + Cillit Bang Kireç Sökücü Sprey'

Additionally, **Table 2.** provides the list of the most frequent baskets with contingent free shipping.

Table 2. *Most frequent baskets in the data, for contingent free shipping*

support	itemsets
0,119	'Prima Bebek Bezi Aktif Bebek 5 Beden Junior Aylık Fırsat Paketi 150 Adet'
0,102	'Ariel Toz Çamaşır Deterjanı Dağ Esintisi 9 kg + Parlak Renkler 7 kg'
0,101	'Fairy Platinum Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 81 Yıkama 2li Paket + Platinum Sıvı Bulaşık Deterjanı 870 ml'
0,069	'Prima Bebek Bezi Aktif Bebek 4 Beden Maxi Aylık Fırsat Paketi 174 Adet'
0,059	'Finish Powerball Quantum Bulaşık Makinesi Deterjanı 144 Tablet 72x2 + Cillit Bang Kireç Sökücü Sprey'

Also, **Table 3.** and **Table 4.** are constructed in the same manner, with different shipping types.

Table 3. *Most frequent baskets in the data, for free shipping*

support	itemsets
0,026	Omo Toz Çamaşır Deterjanı Active Fresh 40 Yıkama 6 Kg'
0,023	Fairy Platinum Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 81 Yıkama 2li Paket + Platinum Sıvı Bulaşık Deterjanı 870 ml'
0,022	Xiaomi Mi Band 2 Akıllı Bileklik Siyah'
0,021	Spigen Araç Tutacağı Manyetik Evrensel Tüm Cihazlarla Uyumlu Araç Tutucu - SGP11583'
0,019	Omo Toz Çamaşır Deterjanı Color 40 Yıkama 6 Kg'

Table 4. *Most frequent baskets in the data, for paid shipping*

support	itemsets
0,011	Case 4U U8 Siyah iOS ve Android Uyumlu Akıllı Saat'
0,010	Vancat Quardo Kokulu İnce Taneli Kedi Kumu 10 Kg'
0,009	Samsung Galaxy C7 İthalatçı Garantili'
0,008	Case 4U Manyetik Mıknatıslı Araç İçi Telefon Tutucu'
0,008	Eurogold Kedi Çimi'

After gathering these tables, association rules generation – a key step in market basket analysis – has been made. These association rules are sorted by descending order of confidence, and they are tabulated for a minimum confidence level of 0.1 in **Table 5.**, **Table 6.**, **Table 7.**, and **Table 8**.

Table 5. *Association rules with the highest confidence levels, regardless of shipping type*

antecedents	consequents	antecedent support	consequent support	support	confidence	lift
'Omo Toz Çamaşır Deterjanı Color 40 Yıkama 6 Kg'	'Omo Toz Çamaşır Deterjanı Active Fresh 40 Yıkama 6 Kg'	0,010	0,012	0,006	0,622	52,356
'Omo Toz Çamaşır Deterjanı Active Fresh 40 Yıkama 6 Kg'	'Omo Toz Çamaşır Deterjanı Color 40 Yıkama 6 Kg'	0,012	0,010	0,006	0,515	52,356
'Uni Baby Çamaşır Deterjanı 1500 ml'	'Uni Baby Çamaşır Yumuşatıcı 1500 ml'	0,008	0,010	0,004	0,492	49,229
'Solo Tuvalet Kağıdı 32 li'	'Solo Ultra Kağıt Havlu 12 li'	0,008	0,018	0,003	0,425	24,070
'Uni Baby Çamaşır Yumuşatıcı 1500 ml'	'Uni Baby Çamaşır Deterjanı 1500 ml'	0,010	0,008	0,004	0,401	49,229

Table 6. *Association rules with the highest confidence levels, for contingent free shipping*

antecedents	consequents	antecedent support	consequent support	support	confidence	lift
Fairy Hepsi Bir Arada Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 96 Yıkama + Platinum Sıvı Bulaşık Deterjanı 870 ml 2li Paket', 'Ariel Toz Çamaşır Deterjanı Dağ Esintisi 9 kg + Parlak Renkler 7 kg', 'Fairy Hepsi Bir Arada Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 108 Yıkama 2li Paket + Platinum Sıvı Bulaşık Deterjanı 870 ml'	Fairy Platinum Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 81 Yıkama 2li Paket + Platinum Sıvı Bulaşık Deterjanı 870 ml'	0,003	0,101	0,003	0,986	9,769
Xiaomi 20000 mAh Taşınabilir Şarj Cihazı Silikon Kılıf Siyah'	Xiaomi 20000 mAh Taşınabilir Şarj Cihazı'	0,006	0,012	0,006	0,977	84,079
Xiaomi Mi Band 2 Akıllı Bileklik Kordonu Askeri Kamuflaj Desenli'	Xiaomi Mi Band 2 Akıllı Bileklik Siyah'	0,006	0,037	0,006	0,969	26,432
Fairy Platinum Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 81 Yıkama 2li Paket + Platinum Sıvı Bulaşık Deterjanı 870 ml', 'Fairy Hepsi Bir Arada Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 96 Yıkama + Platinum Sıvı Bulaşık Deterjanı 870 ml 2li Paket', 'Ariel Toz Çamaşır Deterjanı Dağ Esintisi 9 kg + Parlak Renkler 7 kg'	Fairy Hepsi Bir Arada Bulaşık Makinesi Deterjanı Kapsülü Limon Kokulu 108 Yıkama 2li Paket + Platinum Sıvı Bulaşık Deterjanı 870 ml'	0,003	0,044	0,003	0,958	21,622
Xiaomi Mi Band 2 Akıllı Bileklik Kordonu Yeşil'	Xiaomi Mi Band 2 Akıllı Bileklik Siyah'	0,003	0,037	0,003	0,948	25,867

Table 7. *Association rules with the highest confidence levels, for free shipping*

antecedents	consequents	antecedent support	consequent support	support	confidence	lift
Xiaomi 20000 mAh Taşınabilir Şarj Cihazı Silikon Kılıf Mavi'	Xiaomi 20000 mAh Taşınabilir Şarj Cihazı'	0,003	0,007	0,003	0,973	140,329
Xiaomi Mi Band 2 Akıllı Bileklik Kordonu Mavi'	Xiaomi Mi Band 2 Akıllı Bileklik Siyah'	0,006	0,022	0,005	0,967	44,521
Xiaomi 10000 mAh Taşınabilir Şarj Cihazı Siyah Kılıf'	Xiaomi 10000 mAh Taşınabilir Şarj Cihazı'	0,004	0,010	0,003	0,921	93,112
Omo Toz Çamaşır Deterjanı Color 40 Yıkama 6 Kg'	Omo Toz Çamaşır Deterjanı Active Fresh 40 Yıkama 6 Kg'	0,019	0,026	0,015	0,798	31,051
Molfix Bebek Bezi Comfort Fix Süper Fırsat Paketi 4 Beden 74 Adet x 2 Paket 148 Adet'	Molfix Bebek Bezi Comfort Fix Süper Fırsat Paketi 5 Beden 66 adet x 2 Paket 132 Adet'	0,006	0,007	0,004	0,762	107,033

Table 8. *Association rules with the highest confidence levels, for paid shipping*

antecedents	consequents	antecedent support	consequent support	support	confidence	lift
'Case 4U Samsung Galaxy C7 Kılıf Ultra İnce Silikon Şeffaf'	'Samsung Galaxy C7 İthalatçı Garantili'	0,004	0,009	0,003	0,746	82,259

2.5. Performance Criteria and Improvements

There are some KPIs to measure the success of the suggested threshold. Revenue impact, number of total sales, customer retention rate, shipping cost coverage, net profit margin, order processing time, and the market share are the main metrics to measure the performance. Also, there is room for further improvements. For example, since the customer information is not available in the data, the personalized recommendations for shipping campaigns could not be applied. The upcoming improvement could be the implementation of customer-based promotions. Furthermore, this threshold adjustment process should be dynamic with a model that runs with new data to adjust the threshold value regarding changing trends and customer behavior.

2.6. The Context Diagram

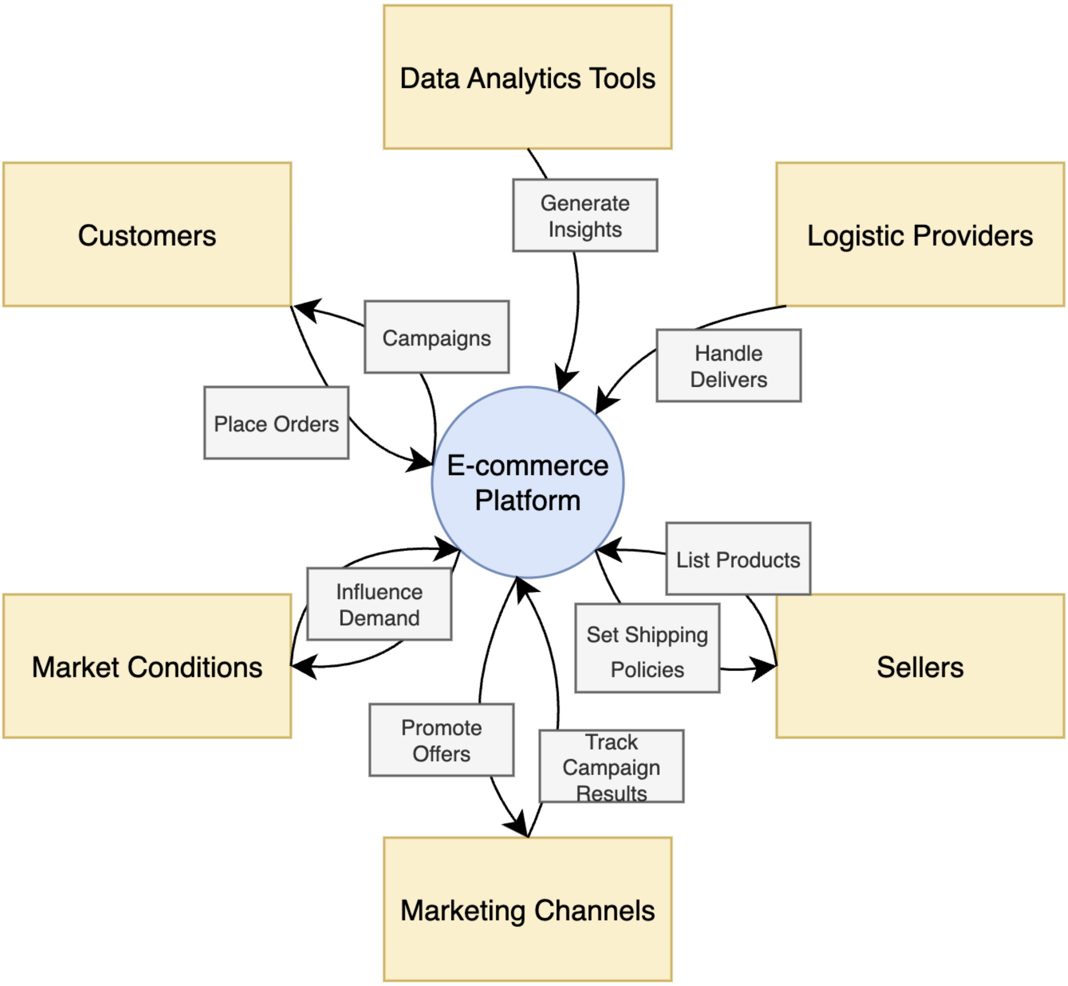


Figure 4. *Context diagram of the model*

The key components in this study and their interactions are demonstrated in the context diagram above.

- **E-commerce Platform:** Central to this study, the e-commerce platform integrates various components to facilitate online sales. It acts as a hub where data is generated and processed.
- **Customers:** Customers interact with the platform by placing orders. Their purchasing behavior, influenced by shipping policies, is crucial for this study. The platform gathers data on customer behaviors, order ingredients, and how different shipping options affect their purchasing decisions.
- **Sellers:** Sellers list their products on the e-commerce platform. They also set shipping policies that can include free shipping, paid shipping, or contingent free shipping based on order value. Understanding the impact of these shipping policies on their sales performance is a key aspect of this study.
- **Logistic Providers:** Logistic providers handle the delivery of products. The efficiency and cost of different shipping options provided by logistic providers can significantly impact customer satisfaction and overall sales. Data from these interactions help in assessing the effectiveness of shipping strategies.
- **Marketing Channels:** Marketing channels promote offers and track campaign results. They play a role in driving traffic to the platform and influencing customer demand. Analyzing marketing campaigns in conjunction with shipping policies can provide insights into how promotional strategies affect sales.
- **Data Analytics Tools:** These tools are composed of software packages and generate insights by analyzing the data collected from various interactions on the platform.

They help in quantifying the effects of different shipping types on total revenue and sales. Advanced analytics can reveal patterns and trends that inform decision-making.

- **Market Conditions:** Market conditions, including competition, economic factors, and seasonal trends, influence demand. Understanding how these external factors interact with shipping policies can provide a comprehensive view of their impact on sales performance.

3. Analysis for Solution/Design Methodology

3.1. Literature Overview

In recent years, the impact of free shipping on sales, customer behavior, and price has been the subject of academic research conducted by Marketing, Economics, Management, and also other similar interested departments. Huang, Shen, and Liang (2019) addressed the issue from the effects of threshold (contingent) free shipping angle by using retrospective experience sampling on participants who were asked to recall recent online shopping experiences and then asked to answer open- and closed-ended questions about this experience. Their research aimed to build a model to explain how consumer perceptions of the characteristics of the TFS (threshold free shipping) policy (i.e., threshold quantity, shipping charge, and delivery timing) may influence their inferred motive and value perceptions, and ultimately, their perceptions of policy fairness as well as intentions to pay for delivery when a threshold is out of reach. They also believed that this would not only assist e-retailers to build better communication strategies for the implementation of TFS but would also make a broader contribution to the online retailing literature, by adding insight into the psychology that governs how online shoppers evaluate and respond to TFS policies. [1]

In another study, Wang, Li, Jiang, and Zhang (2022) tried to fill the void in scholarly research by investigating and empirically analyzing retailers' decisions regarding offering free shipping within the competitive online marketplace, where there are always multiple retailers selling identical products while carrying numerous substitutable products. Their research was centered on answering 2 essential questions: (1) What factors determine the market structure concerning the percentage of retailers offering free shipping? Especially how competition affects the market structure. (2) How does the structure of the market, concerning the percentage of retailers providing free shipping, impact the sales of products? In order to answer these questions, they developed a conceptual framework to characterize retailers' free shipping decisions, consumers' purchase decisions, and their relationships. Also, they came up with an empirical model based on simultaneous equations to investigate research questions. [2]

In addition to these findings, Etumnu (2023) came up with important questions: (1) How does free shipping affect market outcomes from the retailer's perspective? (2) Do freely shipped products receive less or more sales? The goal of his study was to answer these questions and thus shed light on the free shipping business strategy and how it affects the daily dealings of online retailers and the public. His study was focused on the free shipping strategy of Amazon, the biggest online retailer in the world. Because Amazon has various product subcategories and items, and handling the whole sales data of the biggest online retailer couldn't be feasible, he narrowed down his study to focus on ground coffee, which is one of the most popular grocery items on Amazon. In order to estimate the effect of free shipping on sales, he specified a regression model, where the "Best Seller Rank" of each product is regressed against a dummy variable that represents whether the product is eligible for free shipping, price of the product, number of ratings of the product, and average rating of the product. Since this approach may raise endogeneity concerns because the price is a choice variable that consumers set in response to product sales, he collected ground coffee data from Amazon US and matched their prices to ground coffee products from Amazon Canada. This is justified because prices from Amazon US will be exogenous to the estimated sales model in Amazon Canada. After using several regression techniques such as OLS, IV, and quantiles regression, he concluded that products that are freely shipped received significantly higher sales and this effect persisted at different sales levels of the products. However, one must note that this is a case study of a single product — ground coffee, further research must be conducted to generalize the free shipping idea. [3]

The study of Cachon, Gallino, and Xu (2018) explores the economic impact of free shipping threshold policies in online retail. The authors develop a data-driven model to evaluate and optimize these policies, factoring in customer behavior changes like order padding (adding items to meet the free shipping threshold) and product returns. The study suggests that profitable free shipping thresholds are achievable under specific conditions, such as setting the threshold slightly above the average basket size and ensuring shipping revenue is a minor part of total revenue. The findings are based on real-world data from an online apparel retailer, highlighting the complex dynamics between shipping policies, customer purchasing behavior, and profitability. [4]

3.2. Alternative Solution/Design Approaches

Before the development of the proposed mathematical model where the combination of 2 different statistical distributions is presented, different solution and design approaches were also considered. For example, according to results obtained from market basket analysis with the Apriori algorithm, the idea of determining different dynamic threshold values based on categories (or subcategories) of current ingredients of a basket of the customer emerged. However, setting a dynamic value for an arbitrary customer was not considered as a feasible option since it would not correctly work in cases where "unusual" couples of different categories (or subcategories) are observed in the available data and demand is not observable – since there is no customer information, ID, etc. –. The meaning of "unusual" in the previous sentence can be defined as categories (or subcategories) in the content of a basket which are very unlikely to be bought together, i.e. categories (or subcategories) with the lowest collaborative purchase frequencies.

The proposed mathematical model divides our data into two meaningful parts: *Pure Customer Behavior* (without observing contingent free shipping) and *Attracted Customers with Contingent Free Shipping*. The main purpose of this distinction is to derive conclusions about whether different contingent free shipping threshold values affect total revenue. Following this separation, two different statistical distributions are fitted to these data sets based on statistical model comparison metrics, such as AIC and BIC. Then, using certain mathematical calculations, these two different statistical distributions are combined. After this combination, a ratio calculation is made regarding how these two statistical distributions will be combined for different threshold values, and using the existing ₺100 threshold model and this ratio, new statistical distributions (or density functions) for any threshold value are obtained. With the help of these distributions, new basket data can be simulated, and by analyzing these data sets, actions can be taken to solve the problem.

3.3. Assumptions

The proposed mathematical model and the whole process of analyzing and fitting distributions are based on specific assumptions. Firstly, the total market size is assumed to be constant when different thresholds for contingent free shipping are used. Therefore, there is an assumption of the fact that the threshold value does not have any influence on the size of the current market. In addition to that assumption, only aggregated basket revenues are used when modeling, which means there are not any detailed pieces of information regarding the basket content included. Finally, there might be new problems related to higher threshold values for contingent free shipping such as increasing shipping costs, and –potentially– increased volumetric weight of packages which affects order fulfillment and costs. These potential problems related to the structure of the new orders under higher threshold values (DESI, shipping costs, ...) are neglected.

3.4. Brief Overview of the Selected Approach

The chosen approach is based on dividing the dataset into two different meaningful parts: *Pure Customer Behavior* and *Attracted Customers with Contingent Free Shipping*, and fitting different statistical distributions to these parts. These different statistical distributions are combined according to a formula which will be explained in detail in the next chapter. Then, a parameter called “ratio” is defined using different threshold values and a reference value (determined as ₺100 in the current dataset). Using the calculated “ratio”, new statistical distributions can be obtained for any threshold value. These obtained different statistical distributions have been used to perform simulations and derive results related to these simulations.

3.5. IE Tools & Methods Integration

Integration of several industrial engineering tools and techniques has been essential while performing exploratory data analysis, understanding the data and therefore making deductions, and developing a comprehensive mathematical model for a solution. In the preliminary steps of the project, to address whether basket revenue and shipping costs influence the seller's shipping type decision, the K-Nearest-Neighbor (KNN) algorithm was used to analyze the decision boundaries. In addition, information on frequently sold products (based on category or subcategory) was sought using the Apriori algorithm, which is commonly used for market basket analysis in the field of data mining. The aim was to derive insights into how these frequently sold products impact total revenue.

In the later stages of the project, various industrial engineering techniques were employed, such as fitting different statistical distributions using maximum likelihood estimation with the help of available samples, comparing different statistical models based on various metrics like AIC and BIC, and performing simulations using the acceptance-rejection algorithm from the obtained model. Through these techniques, the process of preparing a solution appropriate to the specific discipline of interest was completed.

In summary, throughout the project, many techniques from various sub-disciplines of industrial engineering were utilized. The process was not limited to these techniques alone; software languages like Python and R, along with their existing packages for data visualization, data preprocessing, data manipulation, statistical modeling, mathematical modeling, and computation, were extensively used during the project.

4. Development Process of Solutions

The development of reliable representations of customer behavior under different contingent shipping threshold values is a rather complex process. The way that customers react to distinct thresholds depends on numerous factors, some of which are quite beyond the limitations of the available data. Nevertheless, with mathematical and statistical models, it is possible to represent customer behavior to a satisfactory level. In the following sections, the details of predicting customer behavior are explained step by step.

4.1. Effect Formulation

To enable ease of modeling, the available data is divided into two meaningful parts: *Pure Customer Behavior* and *Attracted Customers with Contingent Free Shipping*.

4.1.1. Pure Customer Behavior

Pure Customer Behavior represents the customer base of the online platform without the effect of contingent free shipping. The data of this subset only consists of transactions with free shipping and paid shipping. The main idea behind this subset is to understand how customers behave in an undistorted environment. Pure customer behavior will play a vital role in the modeling of the base condition. In this subset, there are 315,433 individual baskets. In **Figure 5.**, the distribution of the pure customer behavior is given.

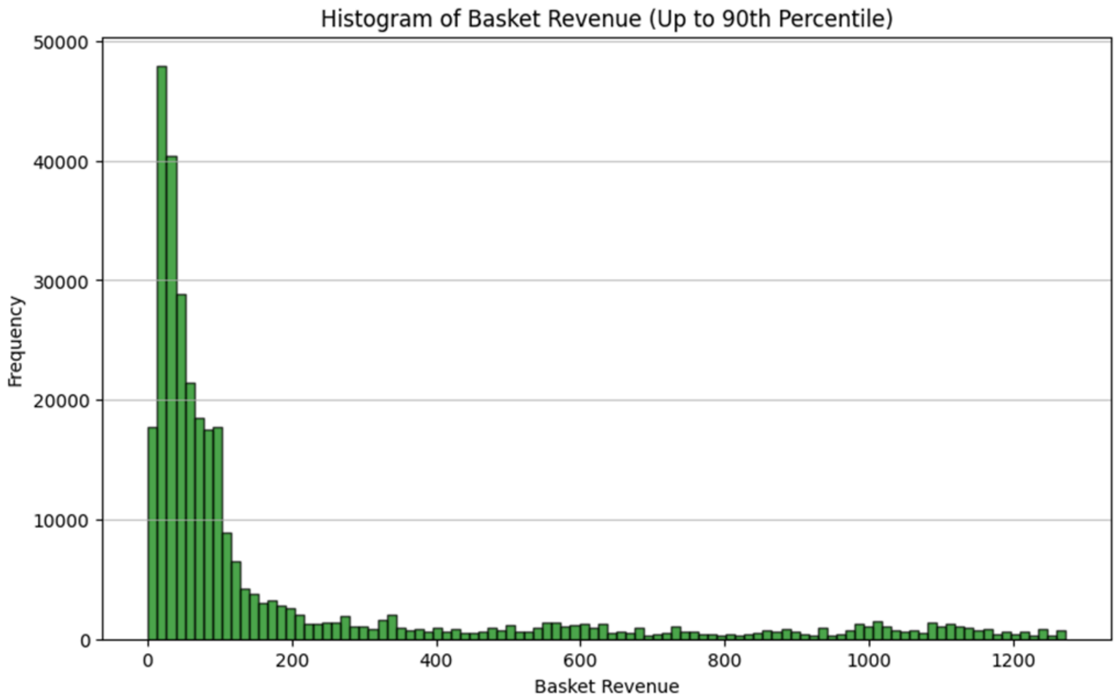


Figure 5. *Histogram of Pure Customer Behavior (90th percentile cut-off)*

Since there are only free shipping and paid shipping options in this subset, the corresponding histogram behaves approximately smoothly as expected. The frequency of the customer baskets increases until around ₺30, and then decays exponentially, with a heavy tail towards the higher values. As seen from the histogram, there exists some noise in the data that is caused by the heterogeneity of the products and baskets. This noise will be handled in the following sections. Now, let’s define the other important effect.

4.1.2. Attracted Customers with Contingent Free Shipping

The other important effect to be formulated is the *Attracted Customers with Contingent Free Shipping*. This subset includes the baskets that are formed with contingent free shipping only. According to the condition of the campaign, the customers do not pay any shipping fee for their transactions larger than ₺100. The main effect to be observed here is how customers react to the offered level of contingent free shipping. In this subset, there are 34,251 individual baskets. In **Figure 6.**, the distribution of the attracted customers with contingent shipping is given.

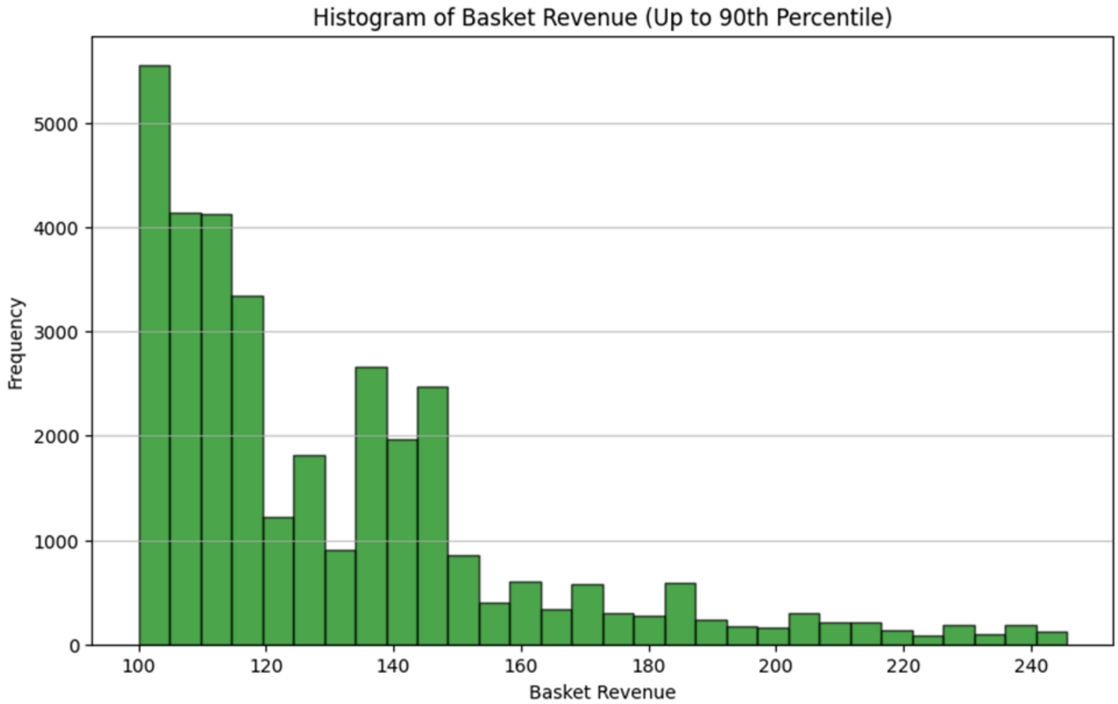


Figure 6. *Histogram of Attracted Customers with Contingent Free Shipping (90th percentile cut-off)*

Since there are only baskets with contingent free shipping in this subset, the basket size in the histogram starts with ₺100, which is the value of the current threshold for free shipping. From the data, it seems like the contingent free shipping could attract some customers and this effect is decaying as basket sizes increase. It appears that customers with basket sizes close to the promotion threshold have added extra products to obtain free shipping. This behavior is called order padding and as the basket size moves far away from the threshold, the effect diminishes. Similar to the previous data, there exists some noise due to product and customer diversity. This noise will be handled in the following section.

4.2. Representing Effects with Statistical Distributions

After the formulation of effects, it is now time to represent these effects with statistical distributions. Fitting statistical distributions provides several benefits such as understanding data characteristics, identifying patterns, and generalizing data behavior. Furthermore, these statistical distributions will be used in simulations and predictions in the subsequent processes.

In this respect, many statistical distributions have been reviewed to find the best-fitting model. For each effect, the following distributions were reviewed: Chi-Square, Lognormal, Exponential, Gamma, Weibull, Pareto, Gumbel_R, Gumbel_L, and GEV (Generalized Extreme Value). These statistical distributions are fitted one by one, and the quality of the fits is decided with AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion). The statistical distribution with the smallest AIC is selected to be the best fit and, therefore, the representative model for the corresponding effect.

4.2.1. Statistical Distribution for Pure Customer Behavior

In the below **Table 9.**, the results of fitted distributions and corresponding parameters are given. The pure customer behavior subset is used to find the distributions.

Table 9. *Related results for distribution fitting for Pure Customer Behavior*

Dist.	AIC	BIC	Parameters
GEV	4433136.1	4148808.67	(-1.02, 44.51, 47.75)
Lognormal	4449631.51	4165304.08	(1.32, 1.02, 77.81)
Pareto	4480349.91	4196022.48	(1.50, -125.03, 126.05)
Weibull	4523672	4239344.57	(0.73, 1.02, 155.08)
Gamma	4548815.26	4264487.83	(0.66, 1.02, 297.69)
Chi-square	4548824.18	4264496.75	(1.33, 1.02, 147.80)
Exponential	4591738.31	4179513.80	(1.02, 196.07)
Gumbel_R	4879862.60	4467638.10	(86.81, 143.70)
Gumbel_L	5328534.44	4916309.93	(373.08, 410.70)

According to the results, the best fitting statistical distribution to represent pure consumer behavior is the GEV (Generalized Extreme Value) distribution with the shape parameter -1.02, the location parameter 44.51, and the scale parameter 47.75. (**Figure 7.**)

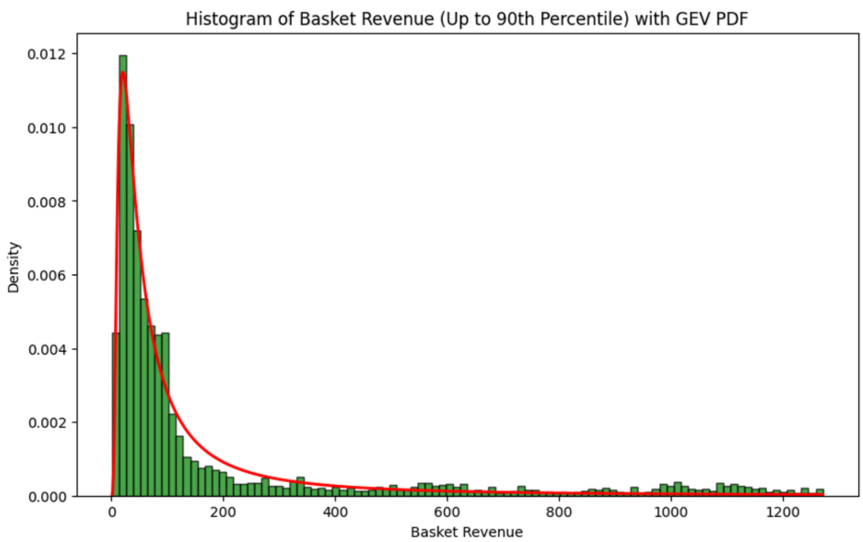


Figure 7. *Fitted distribution with the corresponding data of Pure Customer Behavior*

4.2.2. Statistical Distribution for Attracted Customers with Contingent Free Shipping

In the below **Table 10.**, the results of fitted distributions and corresponding parameters are given. The attracted customers with the contingent free shipping subset are used to find the distributions.

Table 10. *Related results for distribution fitting for Attracted Customers with Contingent Free Shipping*

Dist.	AIC	BIC	Parameters
Gamma	370743.61	339870.18	(0.96, 100, 31.60)
Chi-square	370744.95	339871.52	(1.93, 100, 15.67)
Weibull	370764.28	339890.85	(0.98, 100, 30.22)
Exponential	370791.59	326030.57	(100, 30.35)
Pareto	370791.59	339918.16	(17688237.53, -536870812, 536870912)
Lognormal	374339.13	343465.70	(1.06, 98.45, 19.93)
GEV	377216.49	346343.06	(-0.50, 113.46, 14.20)
Gumbel_R	383719.20	338958.19	(117.92, 19.15)
Gumbel_L	417289.05	372528.03	(147.09, 38.92)

According to the results, the best fitting statistical distribution to represent attracted customers with contingent free shipping is the Gamma distribution with the shape parameter 0.96, the location parameter 100, and the scale parameter 31.60. The threshold value for contingent free shipping in the data is ₺100, therefore the location parameter is fitted to 100. (**Figure 8.**)

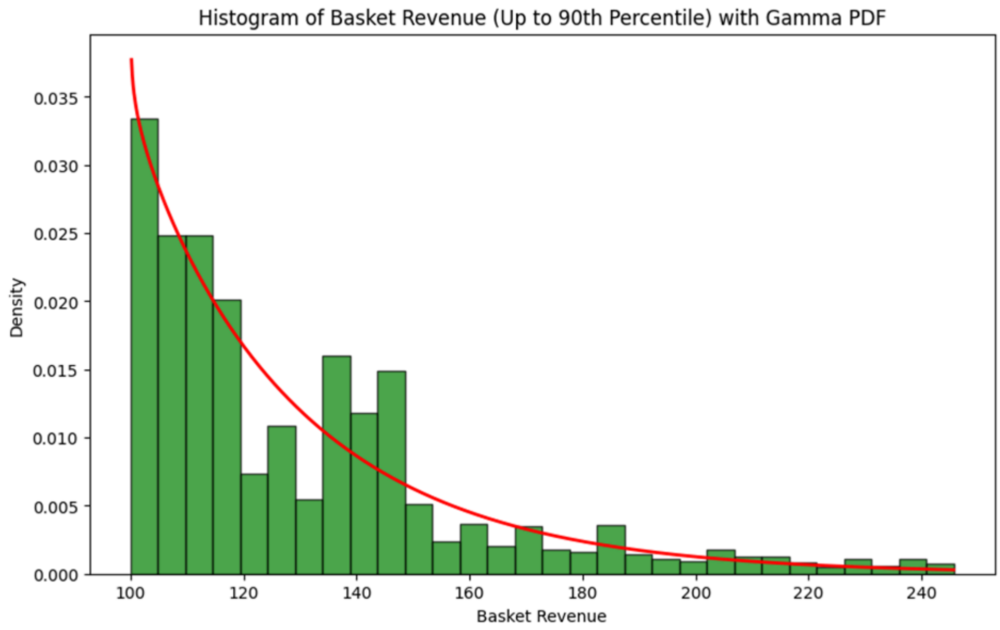


Figure 8. *Fitted distribution with the corresponding data of Attracted Customers with Contingent Free Shipping*

4.3. Combination of Effects

The accurate combination of the previously constructed two effects could lead to a meaningful explanation of the behavior of customers in a market with contingent free shipping options. In the following parts, the process of effect combination is described.

4.3.1. Combined PDF

The combination of these two effects means the combination of the statistical distributions that they represent. In this direction, the probability distribution functions of these two effects are combined with the formula given below. (**Figure 9.)**

$$\text{Combined PDF} = \boxed{\frac{(\mathbf{a} * \text{pdf1} + \text{pdf2})}{(\mathbf{a}+1)}}$$

Figure 9. *The formula for combining 2 probability distribution functions*

a: *combination factor*

pdf1 *= PDF of GEV distribution*

pdf2 *= PDF of Gamma distribution*

Here in this combined pdf formula, the weighted average of two PDFs is obtained with the combination factor "**a**" which decides the weight of the first PDF, relative to the second PDF. Then, for normalization, the weighted sum is divided by the sum of the weights. In the end, this formula yields another PDF which is the combination of two effects.

4.3.2. Finding the Optimal Combination Factor

To find the optimal combination factor "**a**", the negative log-likelihood of the combined pdf is minimized. The result of this minimization problem yielded 4.46 as the value of "**a**". In **Figure 10.**, the GEV distribution, the Gamma distribution, and their combination with the optimal "**a**" are given.

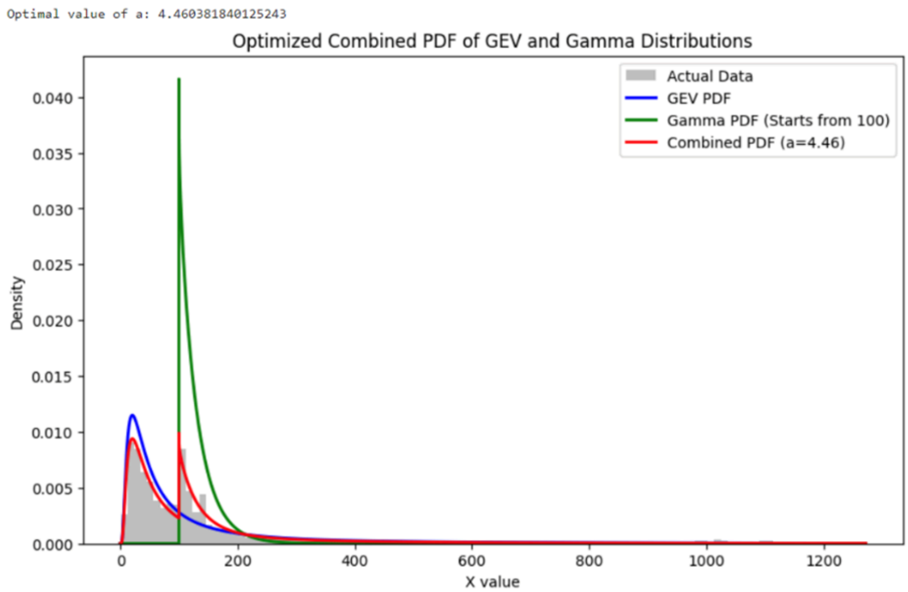


Figure 10. *Actual data, GEV distribution, gamma distribution with combination factor "a"*

4.3.3. The Resulting Model for Threshold at ₺100

The resulting optimal combined PDF is calculated as described in the previous section. In **Figure 11.**, the resulting optimal combined PDF is presented with the actual data that it is supposed to fit.

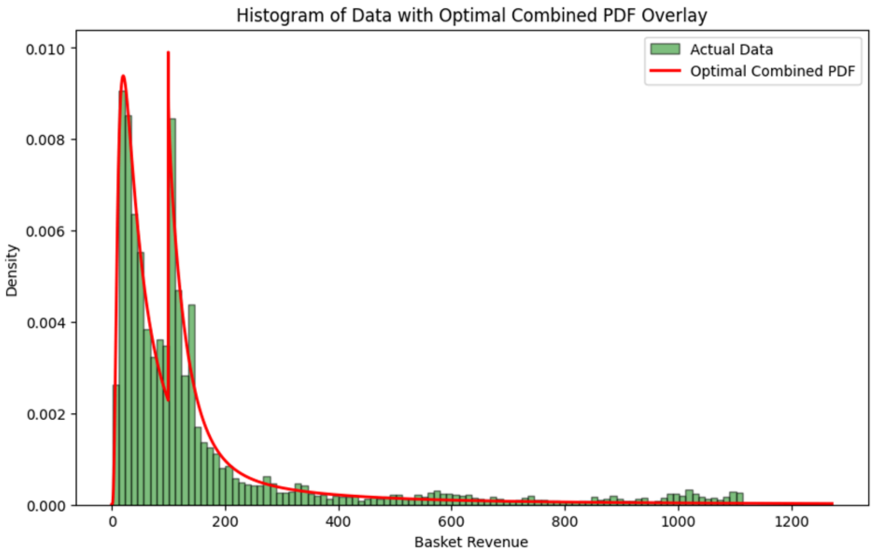


Figure 11. *Existing data and 2 different distributions with their combinations*

According to **Figure 11.,** the model appears to capture the padding behavior of customers around the contingent free shipping threshold value of ₺100. Having found this model, raised some further questions:

- What would be the customer's behavior if the threshold is moved to another value?
- How to find the optimal threshold value?

These questions will be answered in the following chapters.

4.4. Calculating Models for Different Thresholds

The available data is limited to contingent free shipping with a threshold ₺100. In the previous parts, the combined PDF is found to reflect the customer behavior for contingent free shipping with the available threshold. However, when the value of this threshold changes, the weights of the two effects (or two PDFs) will surely change. This means that the combination factor "**a**" should be changed to reflect the new weights of the PDFs. In that manner, a new parameter called "**ratio**" is introduced to readjust the weights.

4.4.1. Ratio Calculation

The "**Ratio**" for any threshold (T) value can be calculated with the following formula:

$$\mathrm{Ratio}_{\mathrm{T}} = \frac{\text{PDF of GEV at 100}}{\text{PDF of GEV at } T}$$

Figure 12. *The formula for ratio calculation*

The "**Ratio**" at threshold T is obtained by dividing the PDF of GEV at 100 by the PDF of GEV at the intended threshold T. The 100 level serves as a reference point for other values.

4.4.2. Combined PDF for Different Thresholds

The combined PDF for any threshold value T is given by the following formula:

$$(\text{Combined PDF})_T = \frac{a \cdot \mathrm{Ratio}_T \cdot \mathrm{GEV} + \mathrm{Gamma}_T}{a \cdot \mathrm{Ratio}_T + 1}$$

Figure 13. *The formula for the combination of 2 density functions*

The logic behind this formula comes from the idea that the effect of *Attracted Customers with Contingent Shipping* will be proportional to the customers available in the *Pure Customer Behavior.* In other words, if there are a high number of customers in the base condition at the level of the planned threshold, then the customers affected by the contingent free shipping will be high and vice versa. In the base condition, if the planned threshold has more customers compared to ₺100, the weight of the Gamma distribution which represents the attracted customers should be higher than before. On the other hand, if the planned threshold has fewer customers compared to ₺100, the weight of GEV distribution which represents the base condition should be higher than before. This adjustment of weights is successfully done by the parameter "Ratio_T". In this way, all distributions with different contingent shipping thresholds can be calculated. Also, it is important to note that the Gamma distribution should be shifted to the newly selected threshold value to reflect the customer behavior starting at that level.

In the below **Figure 14.,** combined PDFs with various thresholds are presented. The PDFs are calculated using the formula given above.

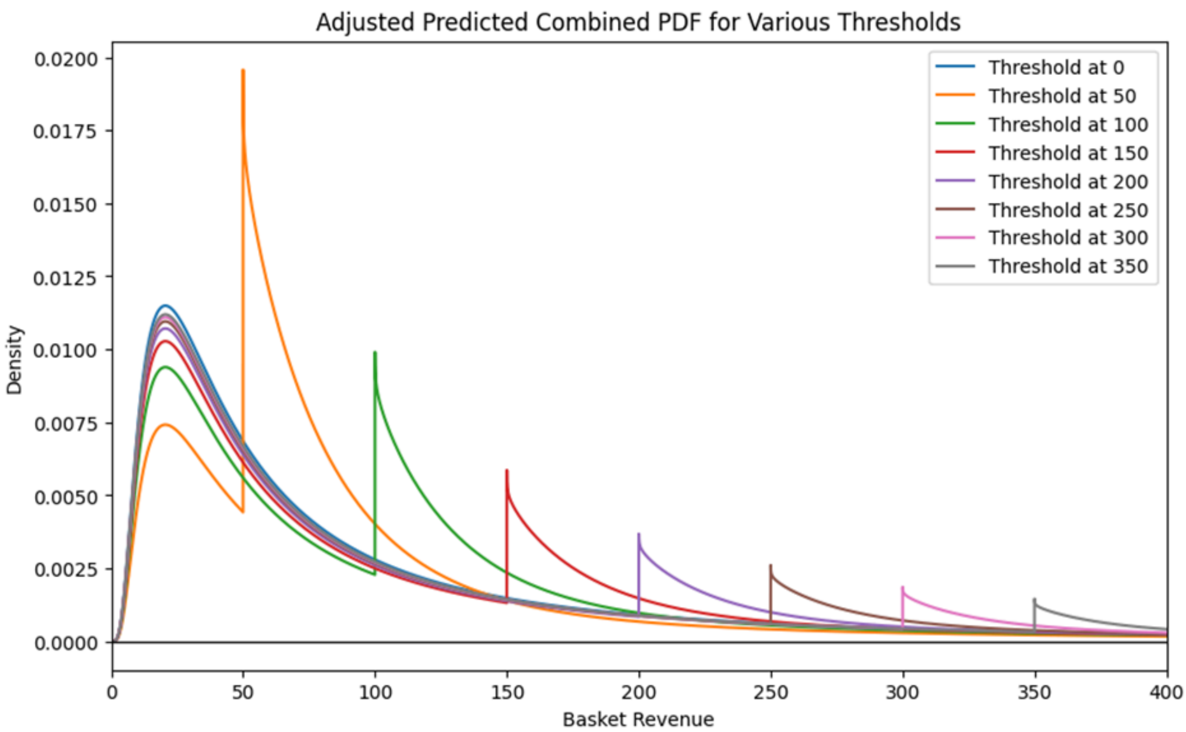


Figure 14. *Combined PDFs for various thresholds*

4.5. Optimal Threshold Selection

Now, it is time to compare different models with respect to their expected revenues. The expected revenues are calculated with the formula given below:

$$\mathrm{mean}_T = \frac{\frac{(a\times \mathrm{Ratio}_T \times \mathrm{meanGEV})+\mathrm{meanGamma}_T}{a\times \mathrm{Ratio}_T+1}}{\int_0^{1271.76} \frac{(a\times \mathrm{Ratio}_T \times \mathrm{GEV}_T)+\mathrm{Gamma}_T}{a\times \mathrm{Ratio}_T+1}\, dt}$$

Figure 15. *The formula for mean parameter calculation*

The mean value of the combined PDF for threshold value T is calculated with the weighted average of theoretical means. Then this value is normalized with the integral of boundaries of the corresponding combined PDF. This normalization is needed because GEV is a heavy-tailed distribution, and this causes a significant deviation if the combined PDF is truncated before infinity. The upper limit of the integral indicates the 90^{th} percentile of the data which will also be the upper limit of the simulations.

Here in **Figure 16.**, the expected basket revenue with respect to different threshold values are given.

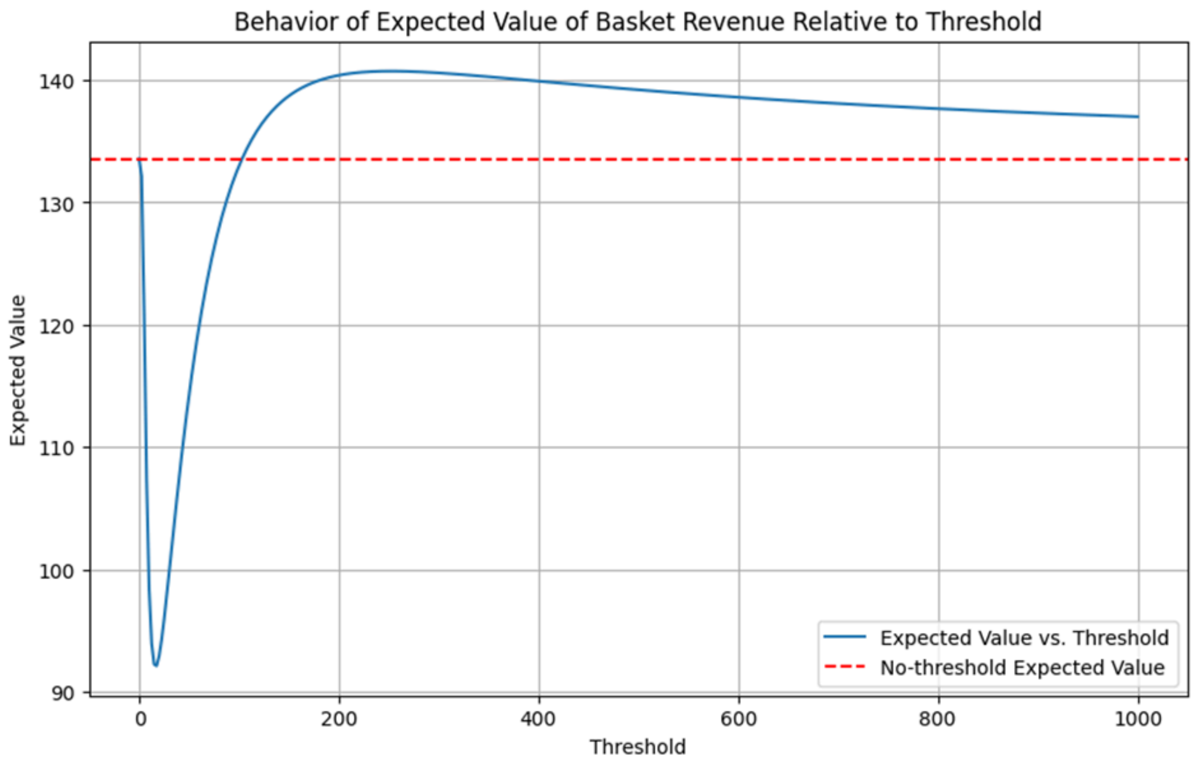


Figure 16. *Expected value of basket revenue with respect to different thresholds*

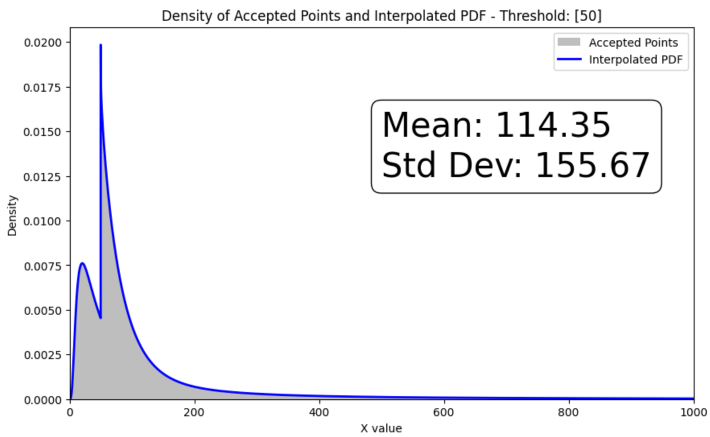
As in **Figure 16.** suggests that the expected value of basket revenue decreases at first due to very low threshold values. In this case, customers tend to separate their orders which leads to low-valued baskets. The expected revenue takes its minimum value at threshold 16.50 with an expected value of 92.02. After that point, the expected basket revenue starts to increase, reaching its maximum at threshold 252.14 with expected value of 140.73. As the threshold goes to infinity, the expected value converges to the red dashed line which indicates the no-threshold expected value. This is quite logical since if the threshold value for contingent free shipping goes too high, customers could lose interest and behave as if there is no campaign at all.

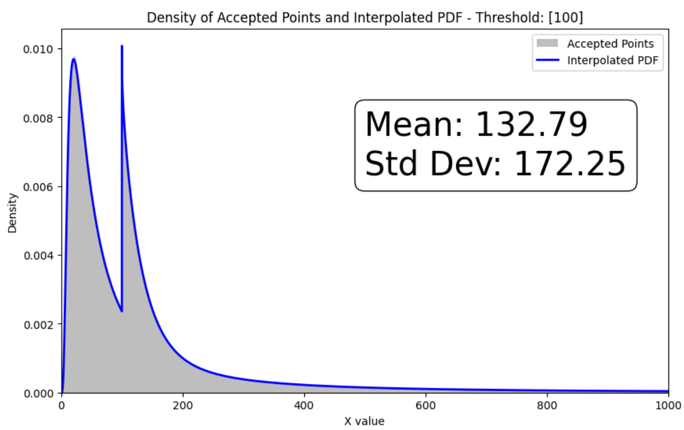
The highest expected basket revenue is reached at threshold level 252.14. According to the model, if the threshold value is set to this point, the expected basket revenue would increase by %5.36. However, further analysis should be conducted to determine how the risk changes with the increasing threshold levels. This will be done in the following chapter with the help of simulations for different scenarios.

5. Comparison of Alternatives and Recommendation

5.1. Numerical Studies and Evaluation Procedure

Simulations with 1 million accepted points have been done to measure the fitness and see the realizations of statistical distributions. The different threshold values are being simulated and their impact on the basket revenue has been assessed. The change and the fluctuations in the mean and the variance of the basket revenues are the key points of the simulation results. In **Figure 17.,** there are the density functions of the simulations with different threshold values.





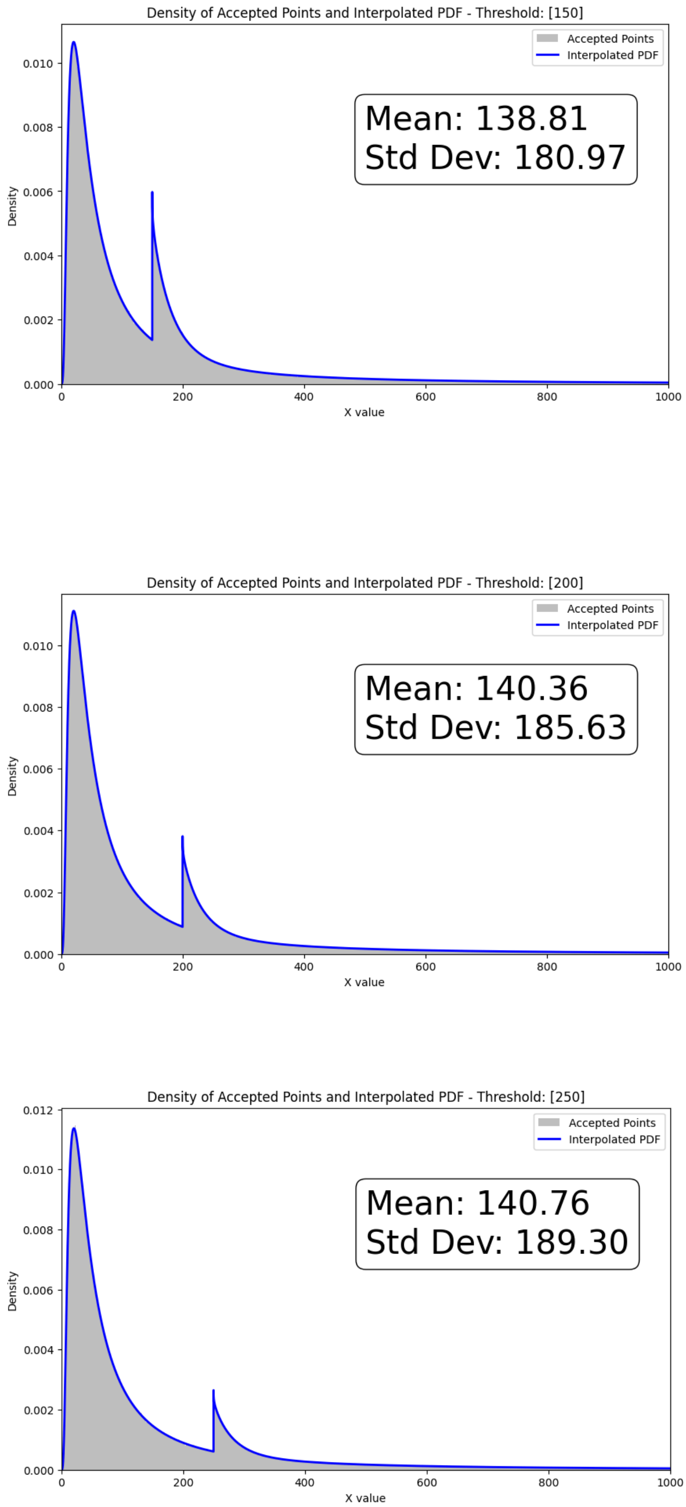


Figure 17. *Simulated data by acceptance-rejection algorithm with different threshold values*

5.2. Proposing a Solution

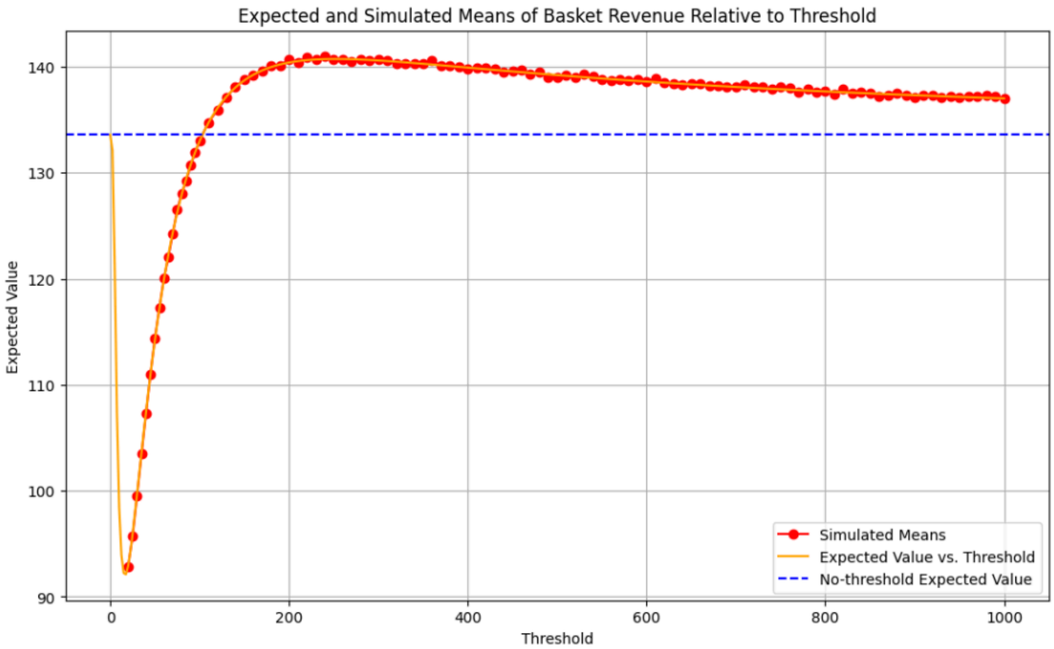


Figure 18. *Expected versus simulated mean parameters with different thresholds*

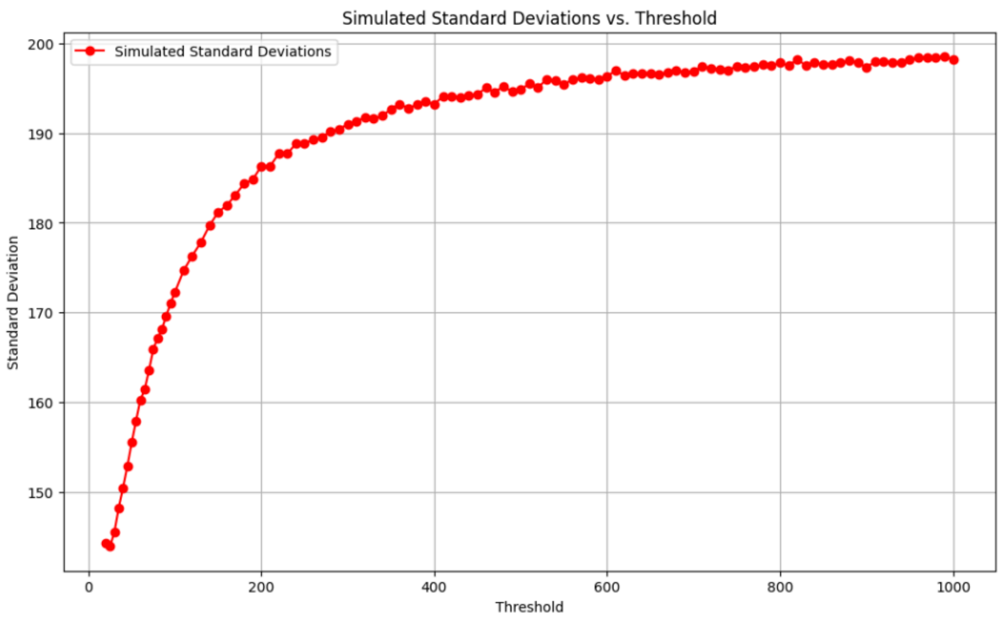


Figure 19. *Standard deviations of the simulated data with different thresholds*

In the above **Figure 18.,** the red dots are the results from the simulations, while the orange line is the fitted statistical distribution. The results from the simulations are the same as the expected values as the graphs above indicate.

- The expected basket revenue peaks at a threshold value of 252.14, after which it begins to decrease. The effect of pure customer behavior starts to dominate the effect of contingent shipping.
- As the threshold value increases, the standard deviation also increases due to less padding around the threshold. Consequently, there is a higher probability of observing both large and small basket revenue values which can be associated with the risk appetite of the company. (can be observed from **Figure 19.**)

Some possible implications are beyond the model that could not be implemented due to the limitations. The simulations can be recurrently done with changing customer bases, the competition dynamics between PO and MP can also be added, and the simulations can be done by filtering the data into different categories.

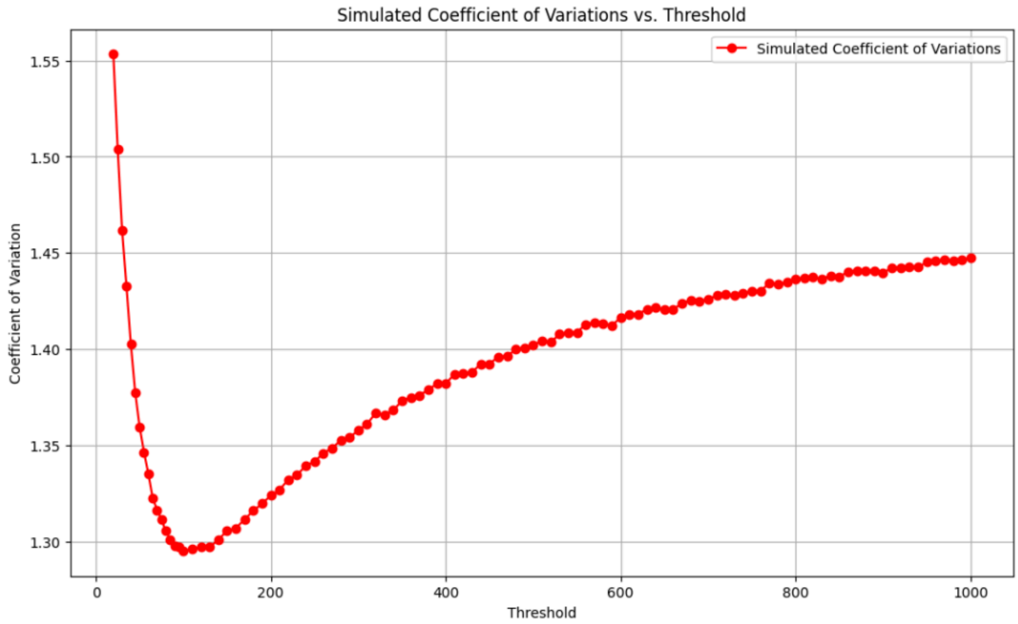


Figure 20. *Coefficients of variations of the simulated data with different thresholds*

As demonstrated in the above **Figure 20.**, the coefficient of variation is the rate of the standard deviation over the mean. It explains the variability around the mean. Low values mean stability while the high values indicate the risk. As it is the lowest around ₺100, the online platform owner may have chosen that value as the threshold for contingent free shipping for risk aversion purposes.

5.3. Assessment of The Recommended Solution

Some requirements should be met to implement the suggested method and the solution efficiently. Firstly, the recommended solution should align with the technical infrastructure of the e-commerce platform. The dynamic threshold adjustments and real-time data processing should be supported by the platform by integration of real-time simulation-based models. The purchasing data should be evaluated in real-time for the statistical model fitting. Secondly, the solution should be suitable for the business strategy and long-term goals. Since the change in the threshold value is a risky operation, the risk appetite of the management should be also aligned with the solution. In addition to those, the logistics providers should be performed with high efficiency to prevent delays in the delivery against the increasing sales. As those requirements are met, the solution can be implemented.

The sustainability of the suggested solution is vital for long-term viability and success. The e-commerce market changes rapidly and the model should also adjust itself fast for sustainability. There should be algorithms in the model that make regular adjustments for continuous improvement with the new data. Not only the maintenance but the cost management will affect the longevity of the project. The cost management of the project should be done carefully since storing data over the long term for the project could be very costly.

In addition to the requirements and sustainability of the project, also the robustness should be assessed as the success metric. The model should be challenged with different scenarios to be adapted to the different market conditions and customer purchase behavior. Also, risk mitigation strategies should be constructed to improve the robustness of the model like system failures, data leakages, and unexpected market shifts. Plans for those risks should be made beforehand to consolidate the model's results.

6. Suggestions for a Successful Implementation

6.1. The Implementation Process of the Proposed Design

Implementing the proposed design is a multi-step process. At the beginning of this process, steps were taken to understand the dimensions of the data obtained from the online sales platform and to decide how the data would be used according to the problem to be solved. Subsequently, as mentioned in the previous chapters of the report, alternative solution approaches were developed, and it was decided that the most significant impact of the different threshold values selected for contingent free shipping on marketplaces and the platform owner would be on their total revenues. To quantify this effect, the available data, as also discussed in the previous chapters of the report, was divided into two meaningful parts (Pure Customer Behavior and Attracted Customers with Contingent Free Shipping). These parts were used to fit the statistical distributions that form the basis of the proposed design and to combine these distributions. The combining process here involves mathematically calculating the relative impact of the statistical distributions obtained from the two meaningful parts on each other, which also can be defined as "weight". In a subsequent stage, it was considered how this combined distribution methodology could be used for different threshold values. As a result of this consideration, using the distribution obtained according to the existing ₺100 threshold value and Pure Customer Behavior in the current dataset, and the value of this distribution at the ₺100 threshold, also known as the reference value, the ratio of the number obtained from this distribution according to a new threshold value to the reference value was calculated, and this ratio was used to update the previously mentioned "weight". Finally, various mathematical models and simulation algorithms were used to conduct simulations with this design, and the results obtained were analyzed in detail to solve the problem.

6.2. Interpretations about the Integration of the Proposed Design into the Overall System

The implementation process described above can generally be used in systems similar to this online sales platform. However, it should be noted that the main focus of this implementation is to determine the threshold value for contingent free shipping and to maximize the total revenue on platforms offering contingent free shipping. Since the main problem to be solved is contingent free shipping-focused, if different online sales platforms have a similar situation, this proposed design is suitable for use. However, different statistical distributions with different parameters may be derived for different datasets, and the process of combining different statistical distributions according to the two meaningful parts in the aforementioned data and the subsequent weight adjustment process may need to be modified.

6.3. Revision of the Proposed Design/Solution

An online sales platform currently using the proposed design might conclude that this solution is insufficient in extreme situations such as changes in customer behavior, increases in operational costs, unexpected movements in the foreign exchange rate, inflation, and other economic reasons. In the event of such circumstances, revising the solution is recommended.

7. Conclusions and Discussion

In order to succeed in this study, the integration of industrial engineering tools, methods, and techniques was crucial. Statistics, statistical modeling, statistical inference and maximum likelihood estimation, systems simulation, data analysis, and data mining were the main industrial engineering tools and techniques used in this study. Furthermore, Python software language and many related software packages were used to process, transform, analyze, and model the data. Also, the KNN algorithm, which is one of the most popular and easy-to-execute algorithms in machine learning, is used for grouping individual points and deducting decision boundaries by using proximity to make classifications.

The most important feature of this proposed design is the ability to analyze customer behavior for different threshold values and to make inferences about how total revenue changes for different threshold values. In addition, by analyzing the simulated data obtained from the calculated combined distribution, it is possible to determine which metrics determine the current policy of the online sales platform and to see the relative impact of changes in the current policy according to these metrics.

The economic impacts of the proposed design appear to be the most important category. Since the main purpose of the solution is to study the impact of contingent free shipping on total revenue, the main aspects related to the solution are closely linked to economic reasons. Given that the online sales platform in the data is a platform in Turkey, economic metrics such as inflation and the unpredictable increase in foreign exchange rates related to imported products are the main aspects of the development process of this solution. Moreover, increased operational costs of the platform owner and shipping costs related to shipping companies are also huge drivers to consider while implementing this design.

Ethically speaking, in cases where online sales platforms also act as retailers (Amazon, Trendyol, etc.), the main benefit of this proposed solution reflects on the platform owner rather than the individual sellers within that platform. The reason for this is the fact that since the platform owner has all the data in the system, such a proposed solution may lead the platform owner to dominate the individual sellers by taking into account its own costs and sales perception.

.

8. References

[1] Huang, W., Shen, G. C., & Liang, C. (2019). The effect of threshold free shipping policies on online shoppers' willingness to pay for shipping. *Journal of Retailing and Consumer Services*, *48*, 105-112. https://doi.org/10.1016/j.jretconser.2019.01.015

[2] Wang, H., Li, D., Jiang, C., & Zhang, Y. (2022). Exploring the Interactive Relationship between Retailers' Free Shipping Decisions and Manufacturers' Product Sales in Digital Retailing. *Sustainability*, *15*(17), 12762. https://doi.org/10.3390/su151712762

[3] Chinonso Ezenwa Etumnu (2023) Free shipping, Applied Economics Letters, 30:16, 2161-2164, DOI: 10.1080/13504851.2022.2094871

[4] Cachon, G. P., Gallino, S., & Xu, J. (2018). Free Shipping Is Not Free: A Data-Driven Model to Design Free-Shipping Threshold Policies. *Available at SSRN 3250971*.