Implementation of Dynamic Pricing Using Machine Learning Predictive Model for Ride-Hailing Service

Sahabaz Khan

Department of Business Administration (MBA), The University of Burdan

MBA4303: Comprehensive Report

Prof. (Dr.) Tanmoy Dasgupta

August 21, 2024

Abstract

In modern era of conducting any business, customer satisfaction has become a must. There is a high competition in every industry either in retail businesses or in e-commerce. No matter what products or services are being offered by a firm there is a high chance that some other firm offering the same. If a company wants to sustain for a longer period it must offer the products or services at right price. That right price will make the customers happy in addition by offering product or services at right pricing the company will achieve a certain competitive advantage over other companies operating in the same industry. This study aims to build a price prediction model utilizing machine learning algorithm on Dynamic Pricing Strategy to have the right price. The model shall be able to predict price in real-time depending on multiple factors responsible for the change in price.

Table of Contents

1 Chapter

1.1 Introduction

Pricing for products or services is tricky for business owners. Businesses want to receive optimal money for their products or services. However, it is not pleasant to price products out of the market. Today, companies have much more information about real-time supply and demand, allowing profit maximization by deploying a strategy called dynamic pricing. If we've ever tried to book online cab during a storm and have watched the price double or even triple, we've seen dynamic pricing in action. While it's not an entirely new concept – companies in the travel and hospitality industries have been using it for years – dynamic pricing is much more relevant in the present age of e-commerce also. (Dublino, 2023).

1.1.1 The concept of dynamic pricing

Dynamic pricing is product pricing depending on various external factors, such as current market demand, the season, supply changes and price bounding. With dynamic pricing, product prices continuously adjust – sometimes in minutes – in response to real-time supply and demand. Companies such as Amazon are one of the largest retailers that use dynamic pricing. Its algorithms continuously adjust and evaluate prices based on real-time external scenario (Dublino, 2023).

Dynamic pricing is applied more to e-commerce than brick-and-mortar businesses. Since brick-and-mortar stores' prices are steadier, they tend to set prices that last for longer periods. In contrast, dynamic pricing relies on real-time trends and supply chain factors. For example, if stock for a particular product drops on an e-commerce site like Amazon, it is likely to see a surge in the price within minutes. Walmart is another example of a major company that uses dynamic pricing to stay competitive in the industry (Dublino, 2023).

1.1.2 The benefits of dynamic pricing

Dynamic pricing gives greater control over pricing strategy. A common argument against dynamic pricing is that it reduces the control over products' prices. In reality, it has the opposite effect. A retailer using dynamic pricing, will have access to real-time price trends across thousands of products in same industry. The business will be able to see competitors' pricing changes and understand a product's supply-and-demand levels. This information will help the business to set the right prices for various products and maximize revenue (Dublino, 2023).

Dynamic pricing allows flexibility without compromising brand value. Many e-commerce retailers shy away from dynamic pricing because of the fear that dynamic pricing will damage the brand value and diminish the customer experience. After all, consumers can easily mistake fluctuating product prices for manipulation or even fraud, right? May be wrong! Brand value can be protected – and even strengthen – by implementing dynamic pricing. Price floor can be set that reflects brand value while gaining the flexibility to stay profitable. Dynamic pricing can also be used as part of seasonal marketing strategies and promotional offers while remaining profitable, which can be challenging with a flat pricing model (Dublino, 2023).

Dynamic pricing saves money in the long run. Dynamic pricing is based on real-time changes in supply and demand and considers market price fluctuations and monitors competitor activity. Using right data and information optimal is set for product prices and stay profitable despite fluctuations. In the long run with dynamic pricing money can be saved. Since web-based software and applications perform all calculations, there's no need to spend time and labor (and therefore money) on manual calculations and related administrative activities. Reduced overhead cost, adding to your ultimate profitability (Dublino, 2023).

Dynamic pricing effectively managed with the right software. Monitoring millions of products and watching real-time supply-and-demand trends is highly complex and challenging. It’s beyond the scope of most businesses. However, the best platforms and dynamic pricing solutions take away the guesswork, automating the process to provide accurate data to set optimal product prices (Dublino, 2023).

Dynamic pricing isn’t error-free, but still in control. Dynamic pricing procedure is based on supply-and-demand changes. As with any technology-based forecast, there is potential for error in dynamic pricing algorithms. However, even if the proposed pricing is inaccurate, it’s still just a proposal. You remain in control and can review the pricing changes your application recommends (Dublino, 2023).

Additionally, the experiences of companies like Amazon, Best Buy and Walmart indicate that potential errors are easily manageable and don’t significantly impact overall profits because the changes are so frequent.

1.1.3 Downsides of dynamic pricing

Dynamic pricing can lead to customer backlash and distrust. While it’s tempting to sit back and leave the pricing to the algorithms, doing so can cause a customer revolt. For example, Bruce Springsteen allowed tickets for his 2023 tour to be handled by Ticketmaster’s dynamic pricing system. The Boss’s blue-collar, advocate-of-the-average-guy brand was seriously tarnished when mid-range seats went on sale for over $4,000 each (Dublino, 2023).

Poor data sources can impact dynamic pricing. Since dynamic pricing is based on real-time data, the data must be as accurate as possible. Otherwise, business will end up with a “garbage in, garbage out” situation that can wreak havoc on profitability and sales volume. Ensure the data going into pricing system is as clean, current and accurate as possible (Dublino, 2023).

Dynamic pricing can alter customer behavior. Once customers realize that prices change in response to specific factors, they may change their behavior accordingly. For example, they may hold off purchasing at peak times and delay until there is less demand. While this can negatively impact profitability, it can be helpful for cash flow (Dublino, 2023).

Dynamic pricing may reduce customer loyalty. If customers are confused or feel taken advantage of by fluctuating prices, they may opt to purchase from a competitor with fixed pricing. To minimize the possibility of diminished customer loyalty, ensure an excellent customer experience. Reduce friction on platform to make purchasing easy, offer free or low shipping prices, and provide easy returns and exchanges with superior customer service (Dublino, 2023).

1.2 Problem Statement

In this research an instance of a ride-sharing company is taken under consideration. The required data has been collected from Kaggle (Dynamic Pricing Dataset, 2024). The aim is to understand the effects of various factors in dynamic pricing strategy to optimize fares based on real-time market conditions and to make an machine learning price predicting algorithm to implement dynamic pricing as well. It shall be noted that, “The company only uses ride duration to decide ride fares currently” (Dynamic Pricing Dataset, 2024). We aim to leverage data-driven techniques to analyze historical data and develop a predictive model that can dynamically adjust prices in response to changing factors.

The historical ride data fromdateset has been provided including features such as the number of riders, number of drivers, location category, customer loyalty status, number of past rides, average ratings, time of booking, vehicle type, expected ride duration, and historical cost of the rides (Dynamic Pricing Dataset, 2024).

Our goal is to build a dynamic pricing model that incorporates the provided features to predict optimal fares for rides in real-time. The model must consider factors such as demand patterns and supply availability (Dynamic Pricing Dataset, 2024).

1.3 Study Background

In present era of conducting business ‘customer’ is the main locus of priority. Businesses do focus on providing value to its customer in order to survive longer in the extremely complex and highly competitive market scenario. Another important part of doing business is to generate revenue as no firm is there to run as a loss maker because there are various internal costs such as cost of procuring raw materials, production cost, costs of human resources, cost of infrastructural resource, cost of procuring financial resources, cost of marketing, and so on. Companies often innovate new products or services and offer it to their customer or they conduct business with existing products in the industry. However it must be reminded that businesses’ or firms’ primary source to generate revenue is selling products or services to customers. In order to have a loyal customer base and to convert potential customers into actual customers businesses must offer their products and services at a price which both serve value to customers and the product quality shall comprise the price of the product.

For every business Pricing of products and services becomes very tricky as there are multiple of factors to price a product properly. Companies shall consider external factors such as the economy of the market in the products will be sold, age and gender of the targeted customers, income level of customers, the wants of products to the customers, geographic location and so on and internal factors such as supply of raw material, manufacturing and production costs, etc. before pricing a specific product.

There are multiple methods of pricing. In brick-and-mortar businesses pricing of products are more static in nature as the cost of storing, and transporting is less prone to change every now and then. But presently in industries such as online service providers and hospital sector the level of information available to both producers and customers is highly increasing in addition to availability of information the factors of pricing is changing more and more frequently. So the requirement of a pricing method which takes these real-time changes into account is increasing.

This research is focused to understand various factors which are responsible for the changing behavior of pricing and in addition there is another aim to provide a price predicting algorithm which can be used in order to predict the price depending on various factors. In this research a data set has been taken from the renowned website Kaggle (Dynamic Pricing Dataset, 2024) to perform certain analysis, testing and training the data.

2 Chapter

2.1 Literature Review

In any study, conceptualization forms an integral part and well defined concept tools and methods and methods of analysis are essential to help researchers in collecting, processing, and analyzing the data and arriving at plausible conclusion.

This chapter attempts to review the relevant literature relating to the field of investigation to form a basis for definition of concepts and to draw meaningful conclusion. For the purpose of clarification and convenience, review has been organized in the following headings-

Different Pricing Strategies,

Dynamic Pricing and Formulating Functions,

Consumer Behavior with Dynamic Pricing,

Dynamic Pricing and Machine Learning,

2.1.1 Different Pricing Strategies

In Pricing Strategy (Sammut-Bonnici & Channon, 2015) it is said that, Pricing strategy is the policy to determine what will be the charge for products and services offered by a firm. Strategic approaches fall broadly into the three categories of cost-based pricing, competition-based common factor among pricing strategies. The strategies differ according to industry and market conditions, the underlying competitive advantage, and in some cases regulatory constraints. Pricing strategy is a key variable in financial modeling to determine the achievable revenues, profits, and the amounts reinvested in the firm’s long-term survival. There are multiple options for pricing strategies are available, which are influenced by key factors. Some of the pricing strategy options are: **(a)** Markup pricing. The strategy involves adding a markup on the product costs. Companies compute the cost of producing a product and add a specific margin. **(b)** Target return on investment pricing. The formula used to calculate the price includes a percentage return on investment that varies with different volumes of production in a given period. **(c)** Perceived value pricing. The price is set to maximize the value that the buyer assigns to the product based on product's utility. The perceived value is a combination of tangible factors and intangible factors. **(d)** Competition-based pricing. In this form prices are decided relevant to those of competitors. The method well apply to medium-share companies competing against high-share competitors. **(e)** Penetration pricing. Also known as promotional pricing, involves temporarily setting prices below the market price or even lower than cost price to ensure maximized rapid market entry into new markets, or the market entry of new products into existing markets to build a strong customer base. Then the customer base is used to generate income from selling the company or its stocks. **(f)** Skimming pricing. This strategy is used to maximize profits by maintaining the highest price possible of new products that face a high demand from specific market segments. The method utilizes a market that is ready to pay a premium for the most recent technologies. The pricing strategy adopted by the firm will depend on the overall corporate strategy, consumer's expectations and behavior, competitor strategies, industry changes, and regulatory boundaries. Other factors that affect pricing strategies are corporate image of the firm, geography in which the firm operates its business, discounts offered by the company depending on demand and supply, Price discrimination (a common practice where demand varies significantly according to circumstances, as in the case of spectator sports and seasonal travel.), etc. Although buyers are price sensitive but the sensitivity depend on situations. Buyers will be fewer prices sensitive when: **(a)** when buyers are unaware of alternatives**. (b)** When buyers are unable to differentiate among product offerings. **(c)** When the purchase use is a low part of discretionary expenditure. **(d)** When the cost is a small proportion of the total cost. **(e)** When costs are shared with another party. **(f)** When costs are related to a cost that has already been incurred. Pricing strategy has been affected by changes in the market structure through retail consolidation, changes in manufacturers' selling policies, advances in technology, and the rapid emergence of internet retailing. Pricing strategies are continuously evolving since the past and with Retail consolidation, Price optimization modeling, Internet pricing disparity the strategies will evolve further in future (Sammut-Bonnici & Channon, 2015).

In Price and Pricing Strategies (Titus, 2013) it is said that, price is one significant factor to achieve success in marketing. In purchasing something, price has great importance to customers. Pricing strategies opted shall be compatible with the rest of the marketing mix. It shall be decided by the management whether to charge the same amount to all similar buyers of identical quantities of a product (a one-price strategy) or to set different charges (a flexible price strategy). Especially retailer organizations use at least some of the following special methods such as price lining, odd pricing and leader pricing. A company must also choose low prices on selected products while having higher prices on all others. 'Price' is the amount of money and other items with utility needed to acquire a product. While utility is an attribute with the potential to satisfy customers' wants. Product's price influences wages, rent, interest and profits besides price are basic regulator of the economic system because it influences the allocation of the factors of production: labor, land, capital and entrepreneurship. The majority of consumers are sensitive to price at certain level but are also concerned with factors, such as brand image, store location, service of the product, quality and value. More attributes that influence consumers to be price sensitive is low income level, small house, large family and member of a minority group. It shall also be kept in minds those perceptions of product quality vary directly with price. Typically, the higher the price, the better the quality is perceived to be. Perceptions of quality may be influenced also by such factors as store reputation and advertising. Price is also a component of value. Value is the ratio of benefits utilized to price and any other incurred costs. Price of products is a major determinant of the market demand. Price has direct effect on firm's competitive position and its market share. So, price has a significant bearing on a company's revenue and profit generation. Depending on firm and the firm's situation a variety of special pricing strategies are there: **(a)** One-Price and Flexible-Price Strategies**.** Under this strategy a seller charges the same price to all similar customers who buy identical quantities of a product. Under a flexible-price strategy, similar customers may pay different prices when buying identical quantities of a product. **(b)** Price lining involves selecting a limited number of prices at which a business will sell related products for the customer to simplify buying decisions**. (c).** Odd pricing. It is a psychological strategy, is commonly used in retailing. Odd pricing sets prices at uneven amounts, rather than at even amounts. Products are priced at $14,995 instead of $15,000. The rationale for odd pricing is that it suggests lower price and, resulting, yields greater sales than even pricing. **(d)** Leader pricing. Temporarily prices are cut on a few items to attract customers. This strategy is called leader pricing. The items on which prices are cut are termed leaders; if the leader is priced below the store's cost, it's a loss leader. Leaders should be well-known, heavily advertised products that are purchased frequently. **(e)** Everyday low pricing and high-low pricing (EDLP) is the hottest retailing pricing trend. Basically, everyday low pricing is done through consistently keeping low prices and few of any temporary price reductions. **(f)**

Resale price maintenance Manufacturers seek to protect the brand's image. Publicly, they state that their control of prices provides middlemen with ample profit margins. In turn, middlemen should be able to give consumers expect sales help and other services when they buy the manufacturers' products from middlemen. **(g)** Reactive and proactive changes. After an initial price is set, a number of situations may influence a firm to change prices for products. If the changes are made to answer the market changes it is reactive change on the other hand if the price changes are made to induce higher competition in the market by the company then it is proactive change. Pricing strategy usually change as a product passes through stages of its life cycle. In pricing innovative products, skimming policy is followed to set high prices initially to skim the maximum amount of revenue from various segments of the market. Or it can use penetration pricing – which sets a low initial price to win a large market share. When the producer is part of a product mix, the firm searches a set of prices to will maximize profits from the total mix (Titus, 2013).

2.1.2 Dynamic Pricing and Formulating Functions

In Dynamic Pricing and Its Forming Factors (Deksnyte, 2012) it is stated that, the determination of proper price still remains a complex task that requires organization's knowledge not only about its operations expenditures but also about its possibilities to foresee products demand and their value with regard to a consumer. Thanks to the advance internet technologies and sales in electronic environment, the information about customers has become more accessible what has determined greater interest in dynamic pricing researchers and their application in different services and industry sectors.

Research literature gives a wide variety of dynamic pricing forming factors. It mentions only assumptions on the ground of which it is possible to distinguish and systematize dynamic pricing forming factors. The factors are – Customer behavior and characteristics; Fair prices: Price fairness is defined as, “consumer’s evaluation and understanding whether the difference between seller’s and other party’s prices is reasonable, acceptable or justifiable” (Maital, 2004; McFadden, 1999); Competition in the market; Demand of the product; The influence of seasonality; The perception of product value (Deksnyte, 2012).

In Dynamic pricing with reference price dependence (Chenavaz, 2016), Regis said a firm that accounts for consumer behavior sets the selling price of a product considering the reference price of consumers. A reference price is usually modeled as depending on past selling prices. That is, past selling prices implicitly constrain the current selling price of a product. Adjusting reference prices effects increase the price elasticity of demand, the demand function becoming flatter. Thus, the reference price effect weakens the market power of the firm. Also, the reference price effect constitutes a main driver of the dynamics of the selling price. But contrary to intuition, selling price dynamics does not systematically imitate reference price dynamics (Chenavaz, 2016).

An article named A 2024 Guide to Dynamic Pricing Strategy (Analytics & Ia, 2024), from Impact Analytics says, dynamic pricing prices adjust based on trends, competition, and availability of stock, etc. Unlike static pricing, which stays same, dynamic pricing enables businesses to win in real time to gain more profit. The core principles of dynamic pricing are: **(a)** Demand-Driven Pricing- at its core this pricing management revolves around the concept of supply and demand. Price changes depending on customer demand for a particular product and services. During peak demand, prices rise, reflecting customer’s will to pay a premium for specific products. Prices go down at low-demand periods to clear inventory. **(b)** Real-Time Adjustments- Dynamic pricing adjusts prices reacting to competitors, trends, and customer behavior. Dynamic price is like having a superpower that is developed to maximize profit in real-time. **(c)** Data-Driven Decision- It thrives on data. Sophisticated math models are used to crunch sales history, competitors prices, and to understand how customers buy. This enables businesses to predict the best prices to make the most money. There are several types of Dynamic Pricing. It is not a one-size-fits-all strategy. Different businesses choose different approaches that fits them the best. Some the types are: **(a)** Segment-Based Pricing- involves setting different prices for different segments of customers. Segmentation depends upon demographics, purchasing behaviors, loyalty status, or geographic location. **(b)** Event-Based Pricing- this involves adjusting prices responding to specific occasions including concerts, sporting events, festivals, or holidays. Prices fluctuate depending on anticipated demand associated with the event. **(c)** Competitor-Based Pricing- involves monitoring and adjusting prices in response to competitors pricing strategies of similar products. Prices set lower than competitors to attract new customers or prices matched to maintain competitiveness. **(d)** Demand-Based Pricing- Involves setting prices based on demand of products or services. **(e)** Location-Based Pricing- involves adjusting prices based on the geographic location of the customer. Prices may vary depending upon factors such as local market conditions, economic factors, or regulatory requirements. **(f)** Discount-Based Pricing- involves offering discounts based on specific conditions, such as bulk purchases, membership status, referral programs, or seasonal sales events. This provides businesses with a range of strategies to optimize pricing decisions in response to changing market dynamics.

Dynamic pricing models are for mastering this pricing strategy. These models lay out the steps, the algorithms, and the decision-making process to adjust prices as quick as possible. **(a)** Algorithmic Pricing Model- algorithms and data analytics are used to analyze factors like demand, competitor pricing, inventory levels, and customer behavior to calculate optimal prices in real time. Algorithms refine pricing strategies using machine learning techniques from feedback and performance metrics. **(b)** Revenue Management Model- models adjust prices based on demand elasticity and capacity constraints. The models are commonly used in industries such as hospitality, airlines, and entertainment, where perishable inventory and limited capacity is available. **(c)** Price Optimization Model- models are profit calculators. Advanced math is used to figure out the best prices that makes most money. This model considers how much people are willing to pay. This keeps balance making money, growing market share, and keeping customers happy. **(d)** Dynamic Discounting Model- deals are offered based on how much someone buys, how often they order, or who they are. This gets them buying more, keeps them happy, and enables to sell more stuff and make more money. Dynamic pricing is rapidly transforming the retail and e-commerce landscape: **(a)** Optimizing the Customer Journey- the strategy goes beyond simply reacting to competitor prices. E-commerce platforms leverage customer behavior data to personalize pricing. **(b)** Maximizing Revenue across Channels: Dynamic pricing isn’t restricted to online stores. Brick-and-mortar retailers utilize it for targeted in-store promotions. **(c)** Enhanced Demand Forecasting: By analyzing historical sales data alongside weather patterns or upcoming events, businesses set prices that anticipate future demand fluctuations. **(d)** Transparency and Building Trust: one must ensure transparency in pricing strategies. This could involve displaying price change history or offering clear explanations for dynamic pricing. Building trust with customers is needed to foster long-term loyalty in this evolving pricing landscape. There are certain benefits of dynamic pricing such as revenue optimization, inventory optimization, competitive advantage, personalized pricing, market adaptability, real-time insights, increased sales, improved customer satisfaction, enhanced price discrimination, effective promotion, and so on. It must be reminded that besides advantage dynamic pricing comes with challenges too, such as- potential consumer backlash, increased complexity, margin erosion, price perception, legal and regulatory risks overdependence on data, customer confusion, lack of transparency, miss brand perception, resistance from traditional customers. Dynamic pricing is a powerful tool for retailers and e-commerce businesses finding a way to maximize revenue, optimize inventory, and stay competitive in today's market. By understanding principles, types, applications, and best practices, full potential of dynamic pricing can be harnessed to drive profitability and success in business (Analytics & Ia, 2024).

In an Article, Price optimization vs. dynamic pricing (Marín et al., 2023), by Triolabs says, even though sometimes these two concepts are understood as same, but they represent different concepts. The main difference is being dynamic pricing is a particular pricing strategy, whereas price optimization can use any kind of pricing strategy to reach its goals. As an example, retailers can dynamically alter the prices of their products in order to match their competitor's price using a dynamic pricing strategy. This strategy would imply changing price frequently but it may not be the best strategy possible. Price optimization techniques focus on finding the price that maximizes a defined cost function (like company's margin). In some highly competitive scenarios where anticipating demand is hard, incorporating dynamic pricing hand crafted rules into a pricing system can allow to take advantage of certain market scenarios (Marín et al., 2023),

2.1.3 Consumer Behavior with Dynamic Pricing

An article, Impact of Dynamic Pricing on Customer Behavior and Loyalty (Strives, 2024) by Upvoty says, prices continuously adapt, influenced by supply, demand, and customer preferences. This is an opportunity to sellers who can utilize dynamic pricing in order to maximize revenue and effectively manage inventory. 55.9% of companies understand the requirement of customer loyalty in managing the challenges of inflation and a potential recession, achieving financial success without detaching the loyal customer base is an essential step for business owners. Visualize the application of personalized pricing strategies, where offers are tailored based on individual customers' browsing and purchasing habits. This approach creates a unique and personalized shopping experience. It's akin to receiving a customized discount, where the store recognizes customer preferences and proposes a special offer on items that align with customer's taste. The significant benefits of dynamic pricing: **(a)** The Strategic Appeal of Dynamic Pricing **-** For the value-conscious customer, dynamic pricing offers a compelling advantage. Personalized discounts and time-sensitive deals are highly attractive, giving more engaging shopping experience. Moving beyond fixed price tags, customers are delighted by the opportunity to secure a bargain. **(b)** The Influence of Pricing on Purchase Timing**-** As prices fluctuate, so does customer decision making. Consumers are led to make purchases quickly to avoid missing short-lived deals. This dynamic ensures products do not remain unsold for longer. **(c)** Tailored Offerings- Data plays a critical role in creating tailored offer structures that align customer preferences. Level of personalization induces a feeling of exclusive attention, boosting the customer's shopping experience. **(d)** The Psychology of Limited-Time Offers- Limited-time offers and changing prices evoke a psychological response, urging customers to act before the opportunity passes. This initiates a sense of urgency and exclusivity, transforming an ordinary purchase into a rewarding experience. Besides creating a positive influence Dynamic pricing have the potential to negatively Impact Customer Loyalty: **(a)** Transparency in Pricing- the reasons behind price changes and the principles guiding your pricing strategy are must to communicate. A lack of transparency undermines trust, which is a critical foundation of customer loyalty. It should be ensured that customers fully comprehend the rationale behind pricing decisions, rather than leaving them in uncertainty. **(b)** Maintain Fairness in Value- during peak demand significant price increases or strategic price changes shall be avoided, as these can negatively impact customer experiences and harm brand's reputation. It is necessary to balance price and value consistently, even at the time of price fluctuation. **(c)** Ethical Use of Data**-** data provided by the customers shall be used responsibly to take informed decisions; businesses shall not exploit individual preferences for profit. Personalized offers shall be tailored in a way that benefits both the customer and the business. Position customers shall be as collaborators in pricing a product, rather than as targets for profit-driven agenda. Dynamic pricing shall be implemented responsibly and precisely, dynamic pricing can significantly transform business landscape. It enhances customer engagement, maximizes profitability, and tailors the shopping experience to individual customer needs. However, the effectiveness of the strategy depends upon maintaining transparency, ensuring fairness, and continuously delivering value. If customers are considered as essential collaborators in pricing strategy, businesses can develop a loyal customer base and secure long-term success (Strives, 2024).

Amirreza Rohani and Mohsen Nazari in their study of, Impact of Dynamic Pricing Strategies on Consumer Behavior (Behera, 2012) said, while some businesses come up with various discount strategies to attract consumers, especially during a recession, both businesses and consumers seem to favor dynamic pricing. The purpose of this study was to investigate how uniform pricing and dynamic pricing influence consumers behavior, in the presence of low involvement and high involvement consumers. The results of study suggested that high involvement consumers responded more positively to dynamic pricing than uniform pricing. Moreover, younger and female consumers are more likely to be involved in obtaining a discount, and high involvement consumers showed more positive feelings, and were more likely to tell others and make repeat purchases from a discount as compared to low involvement consumers. It is a well-known practice that during tough economic time’s businesses drops prices to stimulate demand against competitors and to create the best cash flow possible in the short turn. Among different pricing strategies, however, companies tend to favor dynamic pricing, and consumers seem to accept dynamic pricing. From a company’s perspective, appropriately applied dynamic pricing will increase revenues and profits. The success of dynamic pricing relies on the ability to segment consumers into different groups with different levels of willingness to pay. In particular, the hospitality and airline industries have increasingly employed dynamic pricing since their inventories are perishable, demand can be segmented, the products or services are sold well in advance, and demand fluctuates substantially. From consumers’ perspective, consumers seem to accept the application of dynamic pricing where they are charged different prices for the same service or product since dynamic pricing enables consumers to make a choice over the price. Dynamic pricing has been used as a tool to provide price promotion; for example, consumers receive discounted rates if they accept restrictions, or if they make reservations in advance. In addition, studies have showed that consumers react differently toward price discounts of the same products or services. The concept of consumer involvement plays a significant moderating role. It can be noted that involvement can be used to segment consumers into low, moderate, and high involvement groups which encourages different promotional strategies. Also, the involvement level may influence a consumer's discount receiving behavior, such as high involvement consumers shows more positive feelings from obtaining a discount. There have been limited studies that examine the impact of dynamic pricing on consumer emotions and behaviors. In the study, dynamic and uniform pricing strategies are compared in order to identify which discount strategy consumers prefer. Given today's economic situation, firms are encouraged to use pricing strategies effectively to influence consumers, and online environment enables firms to dynamically manage prices. However, pricing should be made with a careful understanding of their impact on consumers' responses because pricing mistakes can harm firms much more heavily in a downturn than in an upturn. Moreover, it is suggested that individuals with different characteristics perceive the price differently and individual consumers show different reactions to price of the same product in different situations, channels, and occasions of use. The study also examined the role of gender and age in influencing consumers' level of involvement in obtaining a discount. The results indicated that female and younger consumers are more involved in obtaining a discount. The results of the study indicated that consumers highly involved in obtaining a discount respond more positively to dynamic pricing than uniform pricing (Behera, 2012).

2.1.4 Dynamic Pricing and Machine Learning

An article named, Price Optimization using Dynamic Pricing and Machine Learning (Chauhan, 2023), by the Aidream suggests the advancements in data science have enabled the creation of tailor-made products that serves to the unique wants of each individual customer. The initial phase of personalization was introduced through recommender systems, which was used to predict user preferences and suggest products to increase sells and profitability. One of the critical decisions about a product is its price. Nowadays, require sophisticated pricing strategies to perform as a successful business in highly competitive marketplace and to increase customer satisfaction. Businesses face two key challenges: optimizing prices and managing income in a continuously changing environment. So, companies use dynamic pricing to solve the problems simultaneously. The method's efficiency can be significantly increased with artificial intelligence (AI) and machine learning (ML) technologies to open up opportunities and to find new revenue sources, flexibly adapt to market needs, and strengthen customer focus. There are multiple ways to apply machine learning in dynamic pricing. First**,** machine learning algorithms can process real-time data on market demand, competitor pricing, and other factors to adjust prices on-the-fly. Second**,** by analyzing customer data machine learning algorithms can help businesses create personalized pricing strategies to improve customer loyalty and increase sales by offering prices that is tailored to individual customer needs. Third**,** by testing different pricing models, analyzing the results, and identifying the most effective pricing strategies it can help businesses optimize their pricing strategies. Finally, it can predict customer behavior and optimize pricing strategies based on these predictions. By analyzing customer behavior and market trends, businesses can identify patterns and predict future demand for products and services. Typically three types of steps are involved to develop machine learning model. First, understand who the targets are. Next, develop a predictive model that can estimate the likelihood of purchase. Finally, use an optimization algorithm to determine the ideal price and product features as required. The models machine learning includes are: Bayesian model. The user picks a prior value indicating the initial belief about the possible price. Then, whenever a new data point is entered into the algorithm, the initial belief shifts either higher or lower. Here historical pricing data plays the most important feature to decide on the final price. Reinforcement learning model**.** (RL) is a goal-directed model which aims to achieve the highest rewards by learning from environmental data. Decision tree model are classification machine learning models that output a tree-like model of decisions and their possible results. Machine learning techniques enable businesses to implement dynamic pricing on a large scale, considering numerous pricing factors such as price elasticity, and displaying specific prices to customer segments based on their willingness to pay. These results an optimized pricing strategy and maximized profits while remaining competitive in changing market. Overall, machine learning is a tool for deploying dynamic pricing strategies. As machine learning continues to evolve and advance, we can expect to have even more advanced dynamic pricing strategies implemented across the industries (Chauhan, 2023).

Data Pricing in Machine Learning Pipelines (Cong et al.2021) talked about prices that are associated to procurement of data bulding and implementing model and said machine learning can only succeed by collaboration among many factors in multiple steps, such as collecting data, training models by multiple parties and delivering machine learning services to end users. Data is critical in the whole machine learning pipelines. A brief review was conducted on data marketplaces and pricing desiderata. The steps to build a machine learning model are: at first step is collecting training data. A key challenge in pricing raw data sets is to understand the usefulness of a data set. Second is Pricing data labels. In the time of raw data collection, obtaining data labels is critical. Thus, the challenge in pricing data labels is to estimate label accuracy and compensate crowd workers correspondingly it is important to motivate workers to provide accurate data labels. Third is Revenue allocation in collaborative machine learning. Collaborative machine learning collect data sets from different owners who may have different contributions to the learned models. Evenly distributing the revenues is not fair to different data owners, as some contribute more valuable data, while others provide data with less value and quality. Fourth is Pricing machine learning models. Machine learning as a service is a rapidly growing industry. Customers may purchase well-trained machine learning models or build models for their use. The four tasks are connected when machine learning models and data sets are priced in an end-to-end manner. The costs of data procurement and model training also nuance the selling price of machine learning models, as they are part of the manufacturing cost (Cong et al.2021).

In Machine Learning-driven Dynamic Pricing Strategies in E-Commerce (Youbi et al., 2023) said the practice of dynamic pricing is price adjustment in real-time based on various factors. This paper presented a study on usage of machine learning techniques in dynamic pricing to develop an accurate and effective pricing model. Here historical transaction of data from an e-commerce platform is used to identify the most suitable machine learning algorithm. Gradient Boosting Machines (GBM) emerges as the primary model due to the ability to capture complex relationships and provide accurate predictions. The GBM model is trained and tuned using hyper parameter optimization techniques, and its performance is evaluated using Mean Squared Error (MSE) and R-squared (R2) score. An overview of algorithms is discussed as, Gradient Boosting Machines**.** Gradient Boosting Machines combines multiple weak prediction models, such as decision trees, to make a strong predictive model. It works by sequentially adding new models to correct the errors made by the previous models. This process focuses on minimizing the loss function, such as Mean Squared Error (MSE), allowing subsequent models to focus on capturing the remaining errors and enhance overall performance. The final prediction is obtained by aggregating the predictions from all the weak models. Random Forest consists of multiple decision trees. Each tree is trained on a random subset of the original dataset and the final prediction is obtained by aggregating the predictions from all the trees. The randomness helps reduce over fitting and improves the model's generalization ability. The method is known for its robustness and interpretability. Neural Networks is inspired by the functioning of biological neurons, are a class of deep learning algorithms. NN consists of interconnected layers of artificial neurons, also known as nodes or units, organized into input, hidden, and output layers. Each neuron receives inputs, applies an activation function, and produces an output that contributes to the overall prediction. Through the process of forward propagation, the inputs flow through the network, and the model learns the optimal weights for each connection by minimizing a loss function during the back propagation phase. The methodology proposed was a sequence of steps as: Data set acquisition then processing of data after that feature engineering then model selection next is hyperparmeter tuning after that prediction of price and the last of the cycle is model evaluation. The goal was to develop and implement an accurate, effective and efficient pricing model that optimizes revenue and to enhances customer satisfaction (Youbi et al., 2023).

Chaitanya Amballa et al. in “Learning Algorithms for Dynamic Pricing: A Comparative Study” stated about various algorithms for dynamic pricing. Baseline algorithms. The two methods pricing that we consider are Iterated least square (ILS) and Constrained iterated least square (CILS). ILS estimates the revenue curve by applying least squares while CILS achieves asymptotically optimal regret by integrating forced price-dispersion with ILS. Action Space Exploration (ASE) combines ILS with the eta-greedy exploration strategy, which is one of the simplest and most widely used strategies. At each decision instant, action space exploration updates the estimates of the parameters of the revenue curve by using gradient descent on the mean-square error (MSE) loss between the estimated and true revenue curve computed on the price-revenue pairs observed. Parameter Space Exploration (PSE) has been used in reinforcement learning, control and ranking tasks. Thompson Sampling (TS) for learning the unknown parameters. a drawback of TS is that it continues sampling even after the true revenue curve is learnt sufficiently well, potentially leading to unnecessarily large regret. We propose the following two methods for controlling the exploration, and show later on that they lead to smaller regret compared to plain Thompson Sampling. Controlled sampling with stopping criterion is introduced with a stopping criterion to restrict the growth of regret due to prolonged sampling. The stopping criterion compares the optimal price for the latest estimated parameters with the average of optimal prices for estimated parameters from the previous five iterations. Controlled sampling by varying standard deviation (SD): Another alternative for controlling unnecessary sampling is to change SD in the update equations. A smaller value for SD causes the covariance matrix of the posterior distribution to converge to zero faster. This helps in restricting unnecessary exploration, and potentially gives smaller regret.

In an article, IBM (What Is Random Forest? | IBM, n.d.) says Random forest is a machine learning algorithm, trademarked by Leo Breiman and Adele Cutler, which combines the output of multiple decision trees to fetch a single result . Decision trees begin with a fundamental query; these queries comprise the decision nodes inside the tree, serving as a tool for data division. Every inquiry aids in the process of making a final choice. Decision trees seek to find the best split to subset the data, and they are usually trained through the Classification and Regression Tree (CART) algorithm. The split can be assessed using metrics like mean square error (MSE), information gain, and Gini impurity. Bias and overfitting are two issues with decision trees. However, when multiple decision trees are combined random forest algorithm, they predict more accurate results. . The bagging technique is expanded upon by the random forest algorithm, which uses feature randomness in addition to bagging to produce an uncorrelated forest of decision trees. Random forests merely choose a portion of those features, whereas decision trees take into account all potential feature splits. Random forests merely choose a portion of those features, whereas decision trees take into account all potential feature splits. By accounting for all the potential variability in the data, risk of overfitting, bias, and overall variance can be minimized in order to generate more precise predictions. Random forest algorithms have three main hyper parameters which include node size, the number of trees, and the number of features sampled, which are set before training. The random forest algorithm is made up of a collection of decision trees, and each tree in the combination is comprised of a data sample drawn from a training set with replacement, called the bootstrap sample. We shall talk more about the out-of-bag (oob) sample, which is one-third of the training sample that is classified as test data. Then, another randomization is introduced via feature bagging, which decreases decision tree correlation and increases dataset variety. For a regression task, the individual decision trees will be averaged, and for a classification task, the most frequent categorical variable—will yield the predicted class. Lastly, that prediction is confirmed using cross-validation using the oob sample. There is certain benefits random forest algorithm such as Reduced risk of over fitting and provides flexibility by handling both regression and classification tasks with a high degree of accuracy. An additional useful approach for predicting missing values in the random forest classifier is feature bagging. Evaluating a variable's contribution to the model, or its relevance, is made simple by random forest. Typically, the accuracy of the model is measured using the Gini importance and Mean Decrease in Impurity (MDI) when a certain variable is removed. However, mean decrease accuracy (MDA), is another importance measure which identifies the average decrease in accuracy by randomly permutating the feature values in oob samples. However there are some problems in in this algorithm such as It is a time-consuming process, requires more resources, and more complex in nature. The random forest algorithm has been applied across a number of industries such as finance, healthcare, e-commerce, etc., allowing them to make better business decisions. (What Is Random Forest? | IBM, n.d.)

3 Chapter

3.1 Methodology

Methodology means “A science of understanding how a research is performed” (Bouchrika, 2024, Para 5). A methodical explanation of the research problem is achieved through research techniques. It can be taken for granted that it is a science that examines how scientific research is conducted. The methodical, theoretical analysis of the procedures useful in a field of study is called methodology. It comprises the hypothetical analysis of the body of methods and ethics associated with a branch of information. The study of research methodology is the science of doing scientific research. A method that takes sensible measures to clarify the research problem in a methodical manner. Methodology helps to appreciate not only the products of scientific question but the development itself (K, 2022). Now let’s understand briefly the objective, scope, and managerial usefulness of the study we are conducting in section below:

3.1.1 Objective of The Study

This study attempts to understand how pricing is done, depending on various factors. We’ll understand the relationship between pricing of a product and other variables. To have gain knowledge and insight about how pricing can be done. We check if the hypothesis of dynamic pricing using machine learning algorithm is relevant for the said businesses in association to acting variables.

3.1.2 Scope and Managerial Usefulness

As pricing is one of the most crucial and challenging decision that has to be taken in order to grab the attention and satisfaction of customer which in turn make a business profitable.

This study has the scope to help the management to have in-depth and valuable insight from transactional data in order to make a decision from multiple choices about products' pricing. To understand a huge amount of data without any technique must take tremendous time which incurs a large cost to the company. Besides, when data is being studied just by using human calculation and procedures, there is a big chance of miss-calculation and misinterpretation of information which may lead to a destruction of company. Machine learning models give an opportunity to process a massive amount of data that a company has with minimum error in order to help in making informed and error free decision. But it must be remembered any model may or can have a level of error, so to have a product with optimum price ML models be used as a tool only. As the focused customers are human and not machines a human perspective must be kept by the management at time of pricing a specific product. So the model we are going to have at the end of the study will provide an informational aid to the management for choosing a right pricing decision.

3.2 Research Methodology

Now let us focus on research philosophy, approach, strategy, choice, time horizon, and techniques which are taken under consideration to conduct this study one by one.

3.2.1 Philosophy of the study

A study's philosophy is defined as "the collection of ideas, presumptions, and guidelines that guide your methodology" (Jansen, 2024, Para 3). We will talk about the kind of knowledge we hope to acquire in this part. Mainly four types of research philosophies are there as Positivism, Realism, Interpretivism, and Pragmatism. For this study we will focus on Positivism as the data we are using is quantitative in nature and the inferences we will achieve through calculation with these quantitative data. In this study we will try to gain knowledge by taking careful control of the research environment, exploitation of variables and furthermore measurement and analysis of data.

3.2.2 Approach of the Study

Studies' results can be analyzed and explained in a variety of ways depending on whether descriptive writing, numerical measurements, or a combination of the two are used (Taherdoost, 2022). Here we'll continue with more of a deductive approach where we collect, process, analyze, and draw inference from analyzed data to check if and how the strategy of dynamic pricing actually works or in simple words we are going to check the hypothesis of dynamic pricing rather than to observing scenarios and drawing general calculation as done in inductive approaches.

3.2.3 Strategy of the Study

"A research strategy is simply how you aim to achieve your research goal" (MeanThat, 2016, 1:00). Amongst experimental research, survey research, case study, Action research, grounded theory ethnography, archival research we will carry this research forward as experimental research as there are two sets of variables. One set carries price which is dependent variable and will be calculated with respect to the second set of independent variables (factors) and both the sets carry quantitative data (MeanThat, 2018).

3.2.4 Choices of Carrying Out the Research

To conduct a research a choice of method shall be followed based on which the research will be conducted. There are three choices available Mono method, Mixed method, and Multi method (Vizcarguenaga-Aguirre & López-Robles, 2020). For the sake of our research we will focus on mono method where a single kind of method as quantitative research will be followed.

3.2.5 Time Horizon of Research Data

In this research we are going to have historical data of a ride-sharing company on which we will conduct the required analysis to draw required inferences.

3.3 Research Design

A structure that incorporates many research components is called a research design. In order to logically respond to the research questions, it entails applying various data gathering and analysis procedures. Prior to beginning the research process, it would be ideal to decide how best to appropriately address the research questions. This is made possible with the use of the research design. (Bets & ResearchProspect, 2024). To proceed with the study we may now talk about data type, sample size, sampling method, sampling type, and data collection method in the sections below-

3.3.1 Data Type

In this this study we are using quantitative and secondary data for analysis. The data set contains multiple quantitative variables. Which are given in Kaggle with variables such as “Number of riders, number of drivers, customer loyalty status, location category, average ratings, number of past rides, time of booking, expected ride duration, vehicle type, and historical cost of the rides," (Dynamic Pricing Dataset, 2024, Para 2). Gathered data is collected from the said website which is a secondary source of data for us.

3.3.2 Sampling Size of Data

For the sake of this research we have taken seventy five (75%) of data for training our machine learning model and twenty five (25%) of training data is taken as test data.

Train Data. Training data used to determine the performance of the trained model, It acts as power that supplies the model in machine learning. Training data is larger than testing data that to ensure more data is there to train the machine learning model to increase effectiveness of the predictive model. When ML algorithms get data from any record, the algorithms recognize the patterns and create a model for decision making. (GeeksforGeeks, 2023, para. 3).

Test Data. After training the model we need unknown information to test the model. This unknown data is testing data, and it is used as a tool to check the development and efficiency of the algorithms' training also to modify algorithms to advance results further. Testing data must be "unseen" or "unknown" and recent as well. Because training data has already been "learned" by the model. It can be decided if the model is functioning successfully or more training data is required to comply with the standards by observing its performance on fresh and unknown test data. Test data act as the final check to approve the training of the ML model. (GeeksforGeeks, 2023, para. 5).

3.3.3 Samplling Method

In this study random sample of data in a training dataset is selected with replacement. Thus, it is possible to select each of the individual data points more than once. After generating several data samples, models are then trained independently, and depending on the type of task either regression or classification, the average of the predictions confirm more accurate estimate. This method is frequently applied to noisy datasets in order to minimize variance.

3.3.4 Data Collection Method

We have collected quantitative and secondary data from Kaggle with multiple quantitative variables as "Number of riders, number of drivers, customer loyalty status, location category, average ratings, number of past rides, time of booking, expected ride duration, vehicle type, and historical cost of the rides," (Dynamic Pricing Dataset, 2024, Para 2). Gathered data is collected from the said website which is a secondary source of data for us.

3.4 Limitation of The Study

The Study needs more time to be fine-tuned as focus is to make pricing decision for a company. Here predictive modeling has been use with random sampling. But there are multiple sampling methods available and some of the methods shall be compared too for better results. There is a need of extensive training and testing to opt accuracy as well. Although the study works as a foundation to find a way to implement dynamic pricing, but the model which is use to predict the price requires an overall improvement.

4 Chapter

4.1 Analysis and Discussion

In this this study we are using quantitative and secondary data for analysis. The data set contains multiple quantitative variables. Which are given in Kaggle with variables such as "Number of riders, number of drivers, customer loyalty status, location category, average ratings, number of past rides, time of booking, expected ride duration, vehicle type, and historical cost of the rides," (Dynamic Pricing Dataset, 2024, Para 2).

We will discuss every steps of analysis in the section below.

Step 1: Now as we have required dataset we can perform further by Installing needed modules such as pandas, plotly to analyze the data.

```
!pip install pandas
!pip install plotly
```

Step 2 : Importing needed modules such as pandas, plotly to analyse the data

```
import pandas as pan
import plotly.express as pex
import plotly.graph_objects as gr
```

Step 3 : Let us load dataset from file into Data variable to see the data from table in

```
Data = pan.read_csv(r"D:\Dissertation\data set\Dynamic_Pricing_Data.csv")
print(Data.head())
```

Figure 1

Data of the dataset from Step 3

```
   Number_of_Riders  Number_of_Drivers Location_Category  \
0                90                 45             Urban
1                58                 39          Suburban
2                42                 31             Rural
3                89                 28             Rural
4                78                 22             Rural

  Customer_Loyalty_Status  Number_of_Past_Rides  Average_Ratings  \
0                  Silver                    13             4.47
1                  Silver                    72             4.06
2                  Silver                     0             3.99
3                 Regular                    67             4.31
4                 Regular                    74             3.77

  Time_of_Booking Vehicle_Type  Expected_Ride_Duration  \
0           Night      Premium                      90
1         Evening      Economy                      43
2       Afternoon      Premium                      76
3       Afternoon      Premium                     134
4       Afternoon      Economy                     149

   Historical_Cost_of_Ride
0               284.257273
1               173.874753
2               329.795469
3               470.201232
4               579.681422
```

Step 4: Let's see the data type by using 'Data.dtypes()' in Figure 2 in the section below

Figure 2

Multiple Types of Data in the Dataset

```
[5]: Data.dtypes

[5]: Number_of_Riders             int64
     Number_of_Drivers            int64
     Location_Category           object
     Customer_Loyalty_Status     object
     Number_of_Past_Rides         int64
     Average_Ratings            float64
     Time_of_Booking             object
     Vehicle_Type                object
     Expected_Ride_Duration       int64
     Historical_Cost_of_Ride    float64
     dtype: object
```

Step 5: Data Analysis to Explore the Data

Figure 3

Data's descriptive statistics

```
print(Data.describe())
```

	Number_of_Riders	Number_of_Drivers	Number_of_Past_Rides
count	1000.000000	1000.000000	1000.000000
mean	60.372000	27.076000	50.031000
std	23.701506	19.068346	29.313774
min	20.000000	5.000000	0.000000
25%	40.000000	11.000000	25.000000
50%	60.000000	22.000000	51.000000
75%	81.000000	38.000000	75.000000
max	100.000000	89.000000	100.000000

	Average_Ratings	Expected_Ride_Duration	Historical_Cost_of_Ride
count	1000.000000	1000.00000	1000.000000
mean	4.257220	99.58800	372.502623
std	0.435781	49.16545	187.158756
min	3.500000	10.00000	25.993449
25%	3.870000	59.75000	221.365202
50%	4.270000	102.00000	362.019426
75%	4.632500	143.00000	510.497504
max	5.000000	180.00000	836.116419

Step 5: Let's now examine the correlation between factors and the ride's historical cost. To check correlation we use a linear regression approach called 'Ordinary Least Squares (OLS)' is used in statistics and machine learning to determine which straight line best fits a set of data points.

Let's now examine and see the correlation between the anticipated ride time and the ride's historical cost in Figure 4.

Figure 4

The Graph Expected Ride Duration vs. Historical Cost of Ride

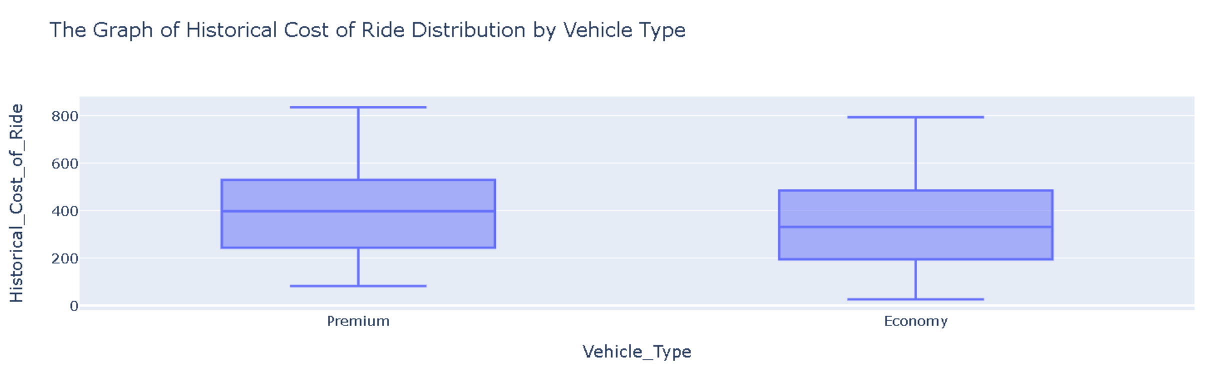
```
[9]: figure = pex.scatter(Data, x='Expected_Ride_Duration',
                  y='Historical_Cost_of_Ride',
                  title='The Graph  Expected Ride Duration vs. Historical Cost of Ride',
                  trendline='ols')
     figure.show()
```

Let's now examine how the historical cost of transportation has been distributed according to the type of vehicle in Figure 5.

Figure 5

The Graph of Historical Cost of Ride Distribution by Vehicle Type

```
figure = px.box(Data, x='Vehicle_Type',
               y='Historical_Cost_of_Ride',
               title=' The Graph of Historical Cost of Ride Distribution by Vehicle Type')
figure.show()
```

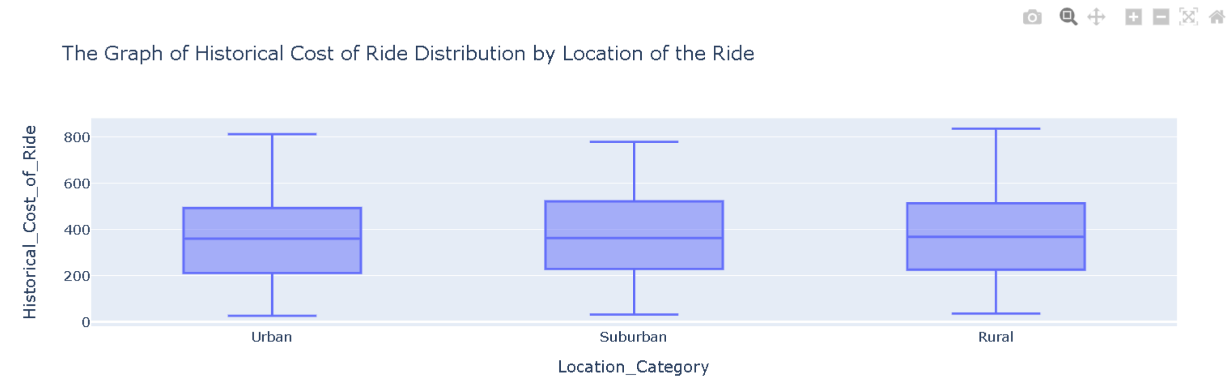


Let's now examine how the historical cost of transportation has been distributed according to the location category in Figure 5.

Figure 5

The Graph of Historical Cost of Ride Distribution by Location of the Ride

```
figure = px.box(Data, x='Location_Category',
             y='Historical_Cost_of_Ride',
             title=' The Graph of Historical Cost of Ride Distribution by Location of the Ride')
figure.show()
```



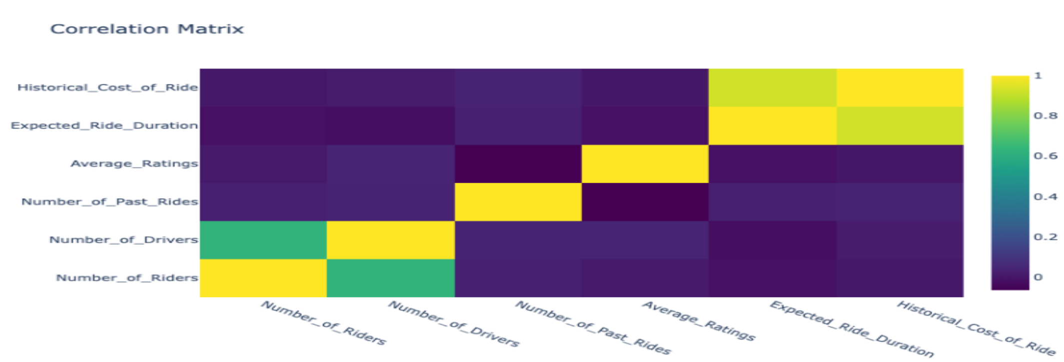
Furthermore the relationship between historical riding cost and other variables can be measured and shown in the above format.

Step 6: Let's now examine the correlation matrix in Figure 6 below

Figure 6

Correlation Matrix

```
corr_matrix = data.corr()

fig = go.Figure(data=go.Heatmap(z=corr_matrix.values,
                                x=corr_matrix.columns,
                                y=corr_matrix.columns,
                                colorscale='Viridis'))
fig.update_layout(title='Correlation Matrix')
fig.show()
```



Step 7: Now well enter into implementing Dynamic Pricing in the sections below

According to the company's statistics, the pricing mechanism it employs determines a ride's price based only on the anticipated time of the voyage. We will now put our dynamic pricing strategy into practice with the goal of dynamically adjusting the ride charges in accordance with the data's observed levels of supply and demand. It will identify times of strong demand and low supply, which will result in price increases, and times of low demand and high supply, which will result in price decreases. Let’s implement dynamic pricing strategy using Python shall using the code shown as in Figure 7.

Figure 7

Implementation of dynamic pricing strategy using Python

```
import numpy as np

# Demand_multiplier computation for high and low demand depending on percentile
high_demand_percentile = 75
low_demand_percentile = 25

data['demand_multiplier'] = np.where(Data['Number_of_Riders'] > np.percentile(Data['Number_of_Riders'], high_demand_percentile),
                                     Data['Number_of_Riders'] / np.percentile(Data['Number_of_Riders'], high_demand_percentile),
                                     Data['Number_of_Riders'] / np.percentile(Data['Number_of_Riders'], low_demand_percentile))

# Calculate supply_multiplier based on percentile for high and low supply
high_supply_percentile = 75
low_supply_percentile = 25

Data['supply_multiplier'] = np.where(Data['Number_of_Drivers'] > np.percentile(Data['Number_of_Drivers'], low_supply_percentile),
                                     np.percentile(Data['Number_of_Drivers'], high_supply_percentile) / Data['Number_of_Drivers'],
                                     np.percentile(Data['Number_of_Drivers'], low_supply_percentile) / Data['Number_of_Drivers'])

# Define price adjustment factors for high and low demand/supply
demand_threshold_high = 1.2  # Higher demand threshold
demand_threshold_low = 0.8  # Lower demand threshold
supply_threshold_high = 0.8  # Higher supply threshold
supply_threshold_low = 1.2  # Lower supply threshold

# Calculate adjusted_ride_cost for dynamic pricing
Data['adjusted_ride_cost'] = Data['Historical_Cost_of_Ride'] * (
    np.maximum(Data['demand_multiplier'], demand_threshold_low) *
    np.maximum(Data['supply_multiplier'], supply_threshold_high)
)
```

In the Figure 7 first the demand multiplier in the code above was first determined by comparing the number of riders to percentiles that denoted high and low demand levels. The demand multiplier is determined by dividing the total number of riders by the high-demand percentile in the event that the number of riders above the percentile for high demand. In the event where the number of riders is less than the low-demand percentile, the demand multiplier is calculated by dividing the total number of riders by the low-demand percentile.

The supply multiplier was then determined by comparing the number of drivers to the high and low supply percentiles. The supply multiplier is determined by dividing the high-supply percentile by the total number of drivers in the event that the number of drivers surpasses the low-supply percentile. In contrast, the supply multiplier is determined by dividing the low-supply percentile by the total number of drivers in the event that the number of drivers falls below the low-supply percentile.

In the end, we determined the dynamic pricing adjusted ride cost. The historical cost of the ride is multiplied by both the maximum supply multiplier and the upper threshold (supply_threshold_high) as well as the maximum demand multiplier and a lower threshold (demand_threshold_low). The thresholds function as ceilings or floors to regulate the price adjustments, and this multiplication guarantees that the modified ride cost reflects the combined effect of supply and demand multipliers.

Step 8: Let's now determine the profit margin we achieved by applying this dynamic pricing method in figure 8.

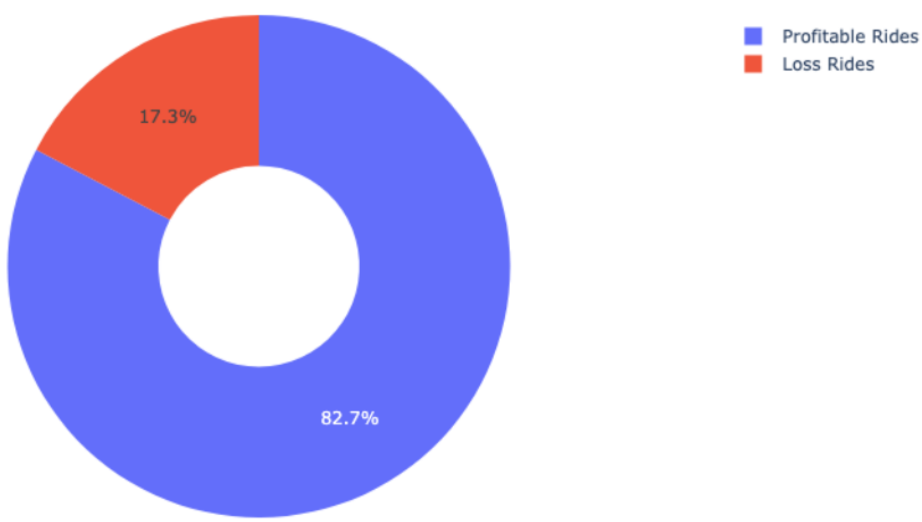
Figure 8

Code for profit margin following the use of this dynamic pricing technique

```
# Determine each ride's profit percentage.
Data['profit_percentage'] = ((Data['adjusted_ride_cost'] - Data['Historical_Cost_of_Ride']) / Data['Historical_Cost_of_Ride'])

# Determine which rides are profitable and have a positive profit percentage.
profitable_rides = Data[Data['profit_percentage'] > 0]

# Find the loss rides where the profit margin is lower than zero.
loss_rides = Data[Data['profit_percentage'] < 0]


import plotly.graph_objects as plg

# Calculate the count of profitable and loss rides
profitable_count = len(profitable_rides)
loss_count = len(loss_rides)

# Create a donut chart to show the distribution of profitable and loss rides
labels = ['Profitable Rides', 'Loss Rides']
values = [profitable_count, loss_count]

figure = go.Figure(data=[go.Pie(labels=labels, values=values, hole=0.4)])
figure.update_layout(title='Profitability of Rides (Dynamic Pricing Strategy vs. Historical Ride Pricing)')
figure.show()
```

Figure 9

Profitability of Rides (Dynamic Pricing Strategy vs. Historical Ride Pricing)



Step 9: Let's now examine the relationship based on the dynamic pricing method between the estimated ride duration and the ride cost as shown in Figure 10

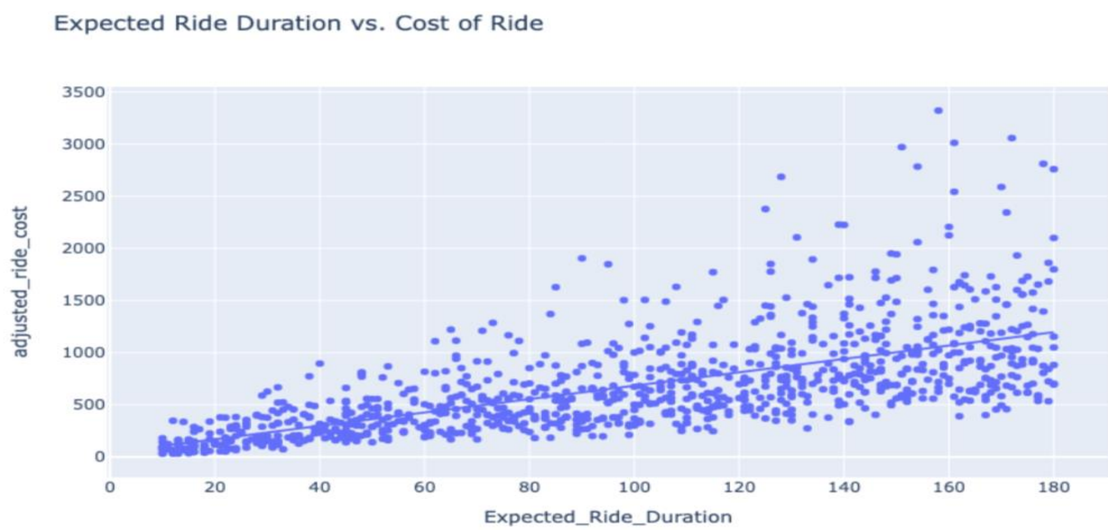
Figure 10

Code to examine the relationship based on the dynamic pricing method between the estimated ride duration and the ride cost

```
figure = plg.scatter(Data,
                     x='Expected_Ride_Duration',
                     y='adjusted_ride_cost',
                     title='Expected Ride Duration vs. Cost of Ride',
                     trendline='ols')
figure.show()
```

Figure 11

Expected Ride Duration vs. Cost of Ride Graph



Step 10: Training a Predictive Model

Now that a dynamic pricing approach has been put into place, let's train a machine learning model using the code below sections. Let's preprocess the data first before training the model as per Figure 12.

Figure 12

Data Preprocessing

```
import pandas as pan
import numpy as np
from sklearn.preprocessing import StandardScaler

def data_preprocessing_pipeline(Data):
    #Determine the categories and numerical features.
    numeric_features = Data.select_dtypes(include=['float', 'int']).columns
    categorical_features = Data.select_dtypes(include=['object']).columns

    #Address missing values in numerical characteristics
    Data[numeric_features] = Data[numeric_features].fillna(Data[numeric_features].mean())

    #Detect and handle outliers in numeric features using IQR
    for feature in numeric_features:
        Q1 = Data[feature].quantile(0.25)
        Q3 = Data[feature].quantile(0.75)
        IQR = Q3 - Q1
        lower_bound = Q1 - (1.5 * IQR)
        upper_bound = Q3 + (1.5 * IQR)
        Data[feature] = np.where((Data[feature] < lower_bound) | (Data[feature] > upper_bound),
                                 Data[feature].mean(), Data[feature])

    #Handle missing values in categorical features
    Data[categorical_features] = Data[categorical_features].fillna(Data[categorical_features].mode().iloc[0])

    return Data
```

We have preprocessed the data by implementing a pipeline in the code above. More details about it are available here. Since the type of vehicle is an important factor, let's first turn it into a numerical feature before continuing as per Figure 13.

Figure 13

Converting Values into Numbers

```
[ ]: Data["Vehicle_Type"] = Data["Vehicle_Type"].map({"Premium": 1,
                                                   "Economy": 0})
```

To anticipate the cost of a ride, let's now split the data and train a machine learning model as per Figure 14.

Figure 14

Train-Test-Split

```
#splitting data into train and test data sets
from sklearn.model_selection import train_test_split
x = np.array(Data[["Number_of_Riders", "Number_of_Drivers", "Vehicle_Type", "Expected_Ride_Duration"]])
y = np.array(Data[["adjusted_ride_cost"]])

x_train, x_test, y_train, y_test = train_test_split(x,
                                                    y,
                                                    test_size=0.2,
                                                    random_state=42)

# Reshape y to one dimensional array
y_train = y_train.ravel()
y_test = y_test.ravel()

# Training a random forest regression model
from sklearn.ensemble import RandomForestRegressor
model = RandomForestRegressor()
model.fit(x_train, y_train)
```

Step 11: Test The Model.

Now let's use some input values to test this machine learning model as per coding below in Figure 15.

Figure 15

Testing The Model

```
def get_vehicle_type_numeric(vehicle_type):
    vehicle_type_mapping = {
        "Premium": 1,
        "Economy": 0
    }
    vehicle_type_numeric = vehicle_type_mapping.get(vehicle_type)
    return vehicle_type_numeric

# Predicting using user input values
def predict_price(number_of_riders, number_of_drivers, vehicle_type, Expected_Ride_Duration):
    vehicle_type_numeric = get_vehicle_type_numeric(vehicle_type)
    if vehicle_type_numeric is None:
        raise ValueError("Invalid vehicle type")

    input_data = np.array([[number_of_riders, number_of_drivers, vehicle_type_numeric, Expected_Ride_Duration]])
    predicted_price = model.predict(input_data)
    return predicted_price

# Example prediction using user input values
user_number_of_riders = 50
user_number_of_drivers = 25
user_vehicle_type = "Economy"
Expected_Ride_Duration = 30
predicted_price = predict_price(user_number_of_riders, user_number_of_drivers, user_vehicle_type, Expected_Ride_Duration)
print("Predicted price:", predicted_price)
```

The model above gives value of Predicted price: [244.44059707]

Step 12: A comparison between the actual and anticipated outcomes. Let's see Figure 16 below and Figure 17 below.

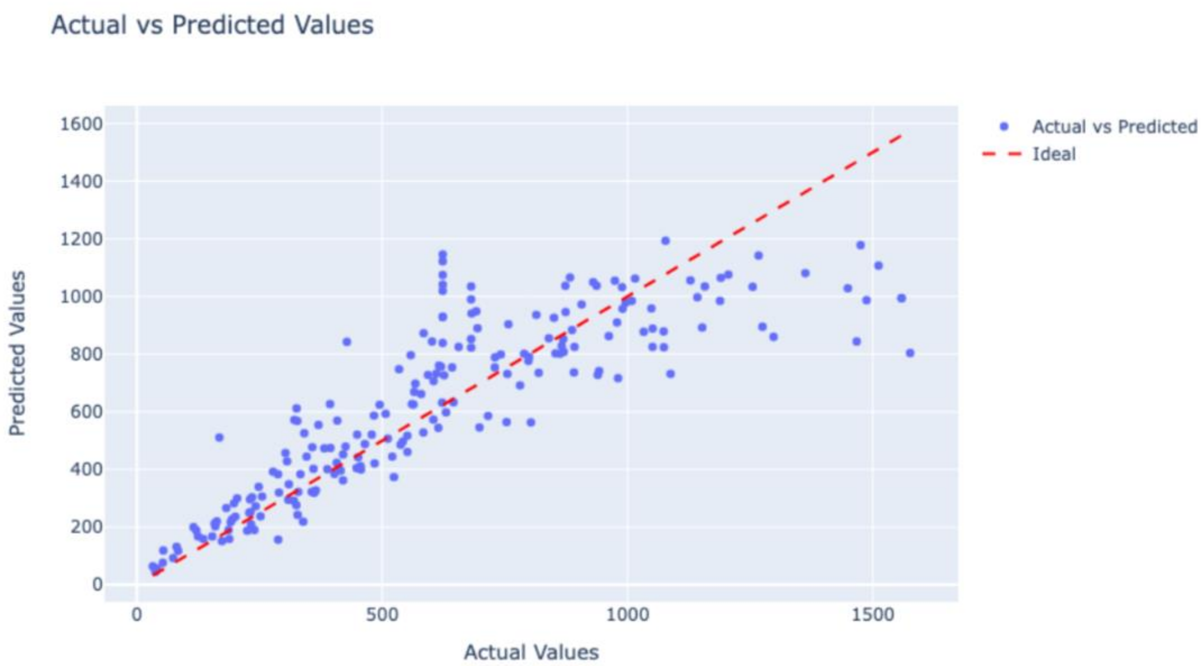
Figure 16

Comparison Between the Actual and Anticipated Outcomes

```
import plotly.graph_objects as plg

# Predict on test set
y_pred = model.predict(x_test)

# Create a scatter plot with actual vs predicted values
figure = plg.Figure()

figure.add_trace(plg.Scatter(
    x=y_test.flatten(),
    y=y_pred,
    mode='markers',
    name='Actual vs Predicted'
))

# Add a line representing the ideal case
figure.add_trace(plg.Scatter(
    x=[min(y_test.flatten()), max(y_test.flatten())],
    y=[min(y_test.flatten()), max(y_test.flatten())],
    mode='lines',
    name='Ideal',
    line=dict(color='red', dash='dash')
))

figure.update_layout(
    title='Actual vs Predicted Values',
    xaxis_title='Actual Values',
    yaxis_title='Predicted Values',
    showlegend=True,
)

figure.show()
```

Figure 17

Actual vs Predicted Values



This is the process by which Python and machine learning can be utilized to construct a data-driven dynamic pricing strategy. (Kharwal, 2023).

Conclusion

By setting prices for products at a level that strikes a balance between supply and demand, dynamic pricing strategies seek to maximize revenue and profitability. Businesses can use it to dynamically modify prices according to many criteria such as the time of day, day of the week, consumer segments, inventory levels, seasonal variations, pricing strategies of competitors, and market conditions. This Python post about dynamic pricing strategy was hopefully enjoyable for you. Please feel free to post insightful queries in the section below the comments.

References

Amballa, Chaitanya & Gugulothu, Narendhar & Kumar, Manu & Bhat, Sanjay. (2020). Learning Algorithms for Dynamic Pricing: A Comparative Study.

Analytics, I., & Ia, T. (2024, May 10). *Dynamic Pricing Strategy: A 2024 Guide*. Impact Analytics. https://www.impactanalytics.co/blog/what-is-dynamic-pricing

Behera, R. (2012b). Impact of Dynamic Pricing Strategies on Consumer Behavior. *Journal of Management Research*, *4*(4). https://doi.org/10.5296/jmr.v4i4.2009

Bets, A., & ResearchProspect. (2024, June 24). *How to Write a Research Design – Guide with Examples*. ResearchProspect. https://www.researchprospect.com/how-to-write-a-research-design

Burney, s.M.Aqil & Saleem, Hussain. (2008). *Inductive and Deductive Research Approach.* 10.13140/RG.2.2.20028.90249.

Business.com : Dublino, J. (2023, November 6). *What Is Dynamic Pricing, and How Does It Affect E-commerce?* business.com. https://www.business.com/articles/what-is-dynamic-pricing-and-how-does-it-affect-ecommerce

Chauhan, N. S. (2023, May 3). *Price Optimisation using Dynamic Pricing and Machine Learning*. The AI Dream. https://www.theaidream.com/post/price-optimisation-using-dynamic-pricing-and-machine-learning

Chenavaz, R. (2016). Dynamic Pricing with Reference Price Dependence. *Economics*, *10*(1). https://doi.org/10.5018/economics-ejournal.ja.2016-22

Cong, Z., Luo, X., Pei, J., Zhu, F., & Zhang, Y. (2022). Data pricing in machine learning pipelines. *Knowledge and Information Systems*, *64*(6), 1-5. https://doi.org/10.1007/s10115-022-01679-4

Deksnyte, Indre & Lydeka, Prof. (2012). Dynamic Pricing and Its Forming Factors. *International Journal of Business and Social Science*, Vol. 3, No. 23, pp. 213

Dr. Swarooprani. K : K, N. D. S. (2022). An Study of Research Methodology. *International Journal of Scientific Research in Science Engineering and Technology*, 537–543. https://doi.org/10.32628/ijsrset2293175

Dublino, J. (2023, November 6). *What Is Dynamic Pricing, and How Does It Affect E-commerce?* business.com. https://www.business.com/articles/what-is-dynamic-pricing-and-how-does-it-affect-ecommerce

Dynamic Pricing Dataset. (2024, January 28). Kaggle. https://www.kaggle.com/datasets/arashnic/dynamic-pricing-dataset/data

GeeksforGeeks. (2023, November 29). *Training data vs Testing data*. GeeksforGeeks. https://www.geeksforgeeks.org/training-data-vs-testing-data

Imed Bouchrika (2024, May 16). How to Write Research Methodology in 2024: Overview, Tips, and Techniques.research.com. https://research.com/research/how-to-write-research-methodology

Jansen, D. (2024, August 12). *Research Philosophy & Research Paradigms*. Grad Coach. https://gradcoach.com/research-philosophy/

Kharwal, A. (2023, December 28). *Dynamic Pricing Strategy using Python*. Thecleverprogrammer. https://thecleverprogrammer.com/2023/06/26/dynamic-pricing-strategy-using-python/

Marín, G., Toscano, M., Brandes, M., Alfaro, P., & Ríos, B. (2023, August 17). *How Machine Learning is reshaping Price Optimization*. Tryolabs. https://tryolabs.com/blog/price-optimization-machine-learning

MeanThat. (2016, March 17). 3.4 How To Choose A Research Strategy [Video]. YouTube. https://www.youtube.com/watch?v=kqkVStFLzL4

MeanThat. (2018, August 5). *3.5 Research Strategy: Experiment* [Video]. YouTube. https://www.youtube.com/watch?v=piuiBvUTXLM

Strives, M. (2024, February 9). *Impact of Dynamic Pricing on Customer Behavior and Loyalty — Upvoty*. Upvoty. https://www.upvoty.com/blog/impact-of-dynamic-pricing-on-customer-behavior-and-loyalty

Taherdoost, H. (2022). What are Different Research Approaches? Comprehensive Review of Qualitative, Quantitative, and Mixed Method Research, Their Applications, Types, and Limitations. *Journal of Management Science & Engineering Research*, *5*(1), 53–63. https://doi.org/10.30564/jmser.v5i1.4538

Tanya Sammut-Bonnici and Derek F. Channon : Sammut-Bonnici, T., & Channon, D. F. (2015). Pricing Strategy. *Wiley Encyclopedia of Management*, 1–3. https://doi.org/10.1002/9781118785317.weom120162

Titus, S. (2013). Price And Pricing Strategies. *ideas.repec.org*. https://ideas.repec.org/a/blg/reveco/v65y2013i4p194-200.html

Vizcarguenaga-Aguirre, I., & López-Robles, J. R. (2020). Mono, mixed or multiple strategy approach: a descriptive study of the latest published articles in the International Journal of Project Management. In *3rd International Conference on Research and Education in Project Management – REPM 2020*.

http://eprints.rclis.org/40506/1/2020_Vizcarguenaga-Aguirre_et_al_Researches_REPM2020.pdf

What Is Random Forest? | IBM. (n.d.). https://www.ibm.com/topics/random-forest#:~:text=The%20random%20forest%20algorithm%20is%20made%20up%20of%20a%20collection,replacement%2C%20called%20the%20bootstrap%20sample.

Youbi, R. E., Messaoudi, F., & Loukili, M. (2023). *Machine Learning-driven Dynamic Pricing Strategies in E-Commerce*. https://doi.org/10.1109/icics60529.2023.10330541