Ca' Foscari
University
of Venice

Single Cycle Degree Programme
in Global Development and Entrepreneurship

Final Thesis

Factors Influencing Student Preferences and Willingness to Pay for Online SAT Preparation Services in Vietnam

Supervisor
Ch. Prof. Lisa CROSATO

Graduand
Hong Hanh LE

Matriculation Number
897250

Academic Year
2023 / 2024

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

1. Rationale of the Research

In the past 2 years, the importance of the SAT in Vietnam's university admissions process has been increasing, which has created a significant demand for SAT preparation resources. This emphasis on standardized testing stems from several key factors:

Firstly, there is a growing competition for university places as Vietnam boasts a rapidly expanding higher education system, attracting a large number of students vying for limited university slots. As a result, universities increasingly utilize standardized tests like the SAT as a crucial factor in their admissions decisions.

Secondly, Vietnamese students are increasingly seeking admission to universities abroad, particularly in the United States and other Western countries. These universities often require standardized tests like the SAT as part of their application process.

In addition, many Vietnamese universities are gradually adopting a more holistic approach to admissions, incorporating factors beyond solely academic performance. However, the SAT remains a significant component of the evaluation process, providing a standardized measure of academic ability.

This growing demand for SAT preparation coincides with a recent significant change in the test format. The SAT transitioned to a fully digital format in 2023, requiring students to adapt to a new testing environment characterized by computer-based testing (CBT). This shift presents unique challenges for Vietnamese students who may not be as familiar with CBT compared to their counterparts in other countries with established digital testing infrastructure.

Furthermore, the COVID-19 pandemic has significantly impacted educational practices globally, accelerating the adoption of online learning methodologies. This trend has fostered a growing acceptance of virtual education platforms, including those offering SAT preparation courses. In Vietnam, the pandemic further highlighted the need for flexible and accessible learning solutions, making online SAT prep courses a potentially attractive option for students seeking to navigate the new digital format of the test.

Therefore, investigating the willingness to pay for online SAT prep courses in Vietnam presents a timely and relevant research opportunity. By understanding student preferences and their willingness to pay within this evolving landscape, we can gain valuable insights into the potential success of online SAT prep offerings in the Vietnamese market. This research will contribute to the development of effective pricing strategies and course features that cater to the specific needs of Vietnamese students preparing for the digital SAT.

2. Aims and Objectives

The purpose of this thesis is to investigate the willingness of Vietnamese students to pay for online SAT prep courses and from that, provide insights into the pricing strategies most effective for online SAT prep courses in the Vietnamese market, considering student preferences and their willingness to pay for features that cater to the digital testing format.

In order to achieve the above goal, this thesis will need to carry out the following

specific tasks:

First, establish a comprehensive theoretical framework on consumer behavior and willingness to pay that incorporate relevant theories and existing research on the topic.

Second, select a suitable analysis model to guide the data analysis based on the theoretical framework. The model will identify the key factors hypothesized to influence student willingness to pay for online SAT prep courses in the Vietnamese market.

Third, design a comprehensive questionnaire to collect primary data on student preferences and willingness to pay. It will target potential customers in Vietnam and gather data on their awareness of the digital SAT format, preferred learning methods, and maximum willingness to pay for online courses with specific features and benefits.

Fourth, statistically analyze the factors influencing student willingness to pay by developing a research model that utilizes appropriate statistical techniques to identify significant relationships between variables and provide insights into student decision-making patterns.

Finally, propose practical recommendations for optimizing online SAT prep courses in the Vietnamese market based on the findings of the data analysis.

3. Scope of the Research

This study focuses on the Vietnamese market for online SAT prep courses. It specifically examines the willingness to pay of potential customers within the context of the recent digitalization of the SAT and the growing popularity of online learning post-pandemic.

4. Research Questions

The study aims to provide answers to these questions

- How do Vietnamese students perceive the value of online learning platforms in preparing for the digital SAT format?

- What factors influence the willingness of Vietnamese students to pay for online SAT prep courses?

- What pricing strategies are most effective for online SAT prep courses in the Vietnamese market, considering student preferences and their willingness to pay for features that cater to the digital format?

- Can we identify different customer segments in the market as well as their characteristics to develop a tailored marketing strategy for each one?

5. Research Methods

This research utilizes a survey questionnaire to collect primary data on student preferences and willingness to pay for online SAT prep courses. The questionnaire targets potential customers in Vietnam and gathers data on their awareness of the digital SAT format, preferred learning methods, and maximum willingness to pay for online courses with specific features and benefits.

Afterward, data will be collected and cleaned from the survey run in April 2024 and then used to build a regression model analyzing these factors.

6. Structure of the Thesis

In addition to the introduction and conclusion, the structure of the thesis includes the following chapters:

Chapter 1: Theoretical Framework

Chapter 2: Market Overview

Chapter 3: Research Methodology

Chapter 4: Data Description

Chapter 5: Research Result and Regression Analysis

Chapter 6: Discussion and Recommendations

CHAPTER I: THEORETICAL FRAMEWORK

1. Customer behavior and willingness to pay (WTP)

1.1. Definitions

Customer behavior is a multifaceted field with various definitions emphasizing different aspects of how consumers interact with products, services, and brands. Listed below are some key definitions from prominent sources, which highlight different aspects of customer behavior:

American Marketing Association (AMA) defines customer behavior as "the study of how individuals, groups, and organizations select, purchase, consume, and dispose of goods, services, ideas, or experiences to satisfy their needs and desires." Understanding what motivates customers is crucial in customer behavior analysis. This area delves into the underlying factors that drive consumers to seek out specific products or services. Needs can be basic, like hunger or thirst, or more complex, such as the need for social connection or self-improvement. Desires, on the other hand, are often associated with wants and aspirations. Marketers aim to identify both needs and desires within their target audience, tailoring their offerings and messaging to address these motivations and create a compelling value proposition.

Francis J. Aguilar, a marketing scholar, defines customer behavior as "the observable actions of people towards products and services." It implies that customer behavior can also be studied through the lens of observable actions. This approach focuses on quantifiable data points that reveal how customers interact with products, services, and brands in the real world. Examples include tracking website visits, analyzing online reviews, measuring purchase frequency, or observing customer behavior in physical stores. By collecting and analyzing this data, businesses gain valuable insights

into customer preferences, buying habits, and pain points, enabling them to optimize their offerings and improve customer experiences.

Valarie Zeithaml and Mary Jo Bitner, service marketing experts, define it as "the study of the processes involved when individuals or groups select, purchase, consume, and dispose of products, services, ideas, or experiences to satisfy their needs and desires." This focuses on the cognitive processes customers go through when considering a purchase. It examines the various stages involved, from identifying a need or desire to evaluating options, making the purchase decision, consuming the product or service, and finally, post-purchase evaluation and potential disposal. Marketers leverage this knowledge to develop targeted messaging that resonates with customers at each stage of the decision-making journey, influencing their choices and guiding them towards the desired outcome – a purchase.

Theodore Levitt, a marketing pioneer, offered a broader perspective: "Customer behavior is the human behavior that reflects a person's relationship with products and services." Customer behavior goes beyond just transactions. This area examines the emotional and psychological connections customers form with brands and products. Factors like brand loyalty, brand trust, and emotional attachment are explored here. Understanding these deeper relationships allows businesses to create emotional connections with their customers, fostering brand advocacy and repeat business.

Willingness to pay (WTP) is a crucial concept within customer behavior, closely linked to the decision-making processes and value perception aspects. It refers to the maximum price a customer is willing to pay for a product, service, or experience to satisfy their needs or desires. Understanding WTP is essential for businesses to develop effective pricing strategies, optimize resource allocation, and ultimately maximize profits.

WTP is intricately linked to the decision-making processes customers go through when considering a purchase. During the evaluation stage, customers assess the product's features, benefits, and potential drawbacks, comparing these to their needs and desires. They then consider the price point and weigh it against the perceived value they expect to receive. Ultimately, a customer's WTP aligns with the maximum price they are comfortable paying to fulfill their needs or achieve the desired outcome.

Marketing efforts can strategically impact WTP through various strategies that influence customer perceptions and value judgments. One such strategy is positioning. By crafting a clear and differentiated brand message that emphasizes a product's unique benefits and competitive advantages, marketers can create a perception of increased value, justifying a higher price point in the customer's mind. Research by Aminu & Ahmad (2018) demonstrates that effective positioning strategies can lead to premium pricing, as customers are willing to pay more for products perceived as superior or offering exclusive features.
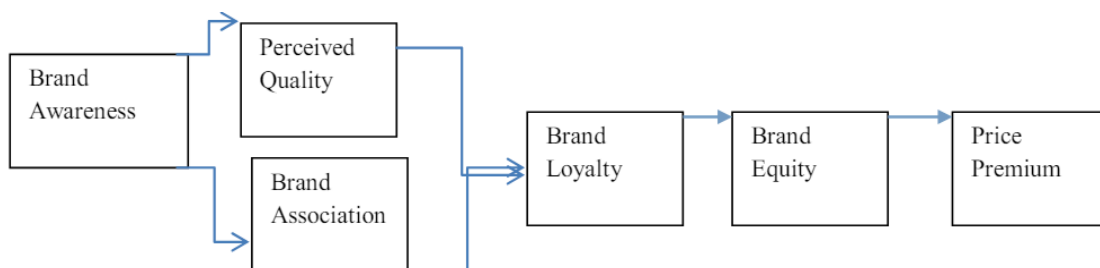
Figure 1: Effective positioning strategy

(Source: Aminu & Ahmad (2018))

Branding also plays a significant role in influencing WTP. Building strong brand trust and positive brand associations fosters customer loyalty and a willingness to pay a premium for products associated with a trusted and reputable brand. This phenomenon is well-documented in consumer behavior research, as studies by Park et

al. (2010) highlight how brand loyalty can lead to a higher tolerance for price increases, as customers perceive additional value from the brand experience itself.

However, marketing strategies that solely rely on price-based tactics, such as promotional strategies involving discounts and offers, can have a double-edged effect on WTP. While initial discounts might increase purchase appeal and influence WTP, over-reliance on such tactics can lead to price sensitivity in the long run. Customers may become accustomed to lower prices and develop a resistance to paying the full price, ultimately hindering the brand's ability to command a premium price point. Research by Blattberg & Neslin (2000) supports this notion, suggesting that frequent promotions can erode brand image and lead to customer expectations of lower prices, making it difficult to raise prices in the future.

Understanding WTP and its relationship with customer behavior allows marketers to develop strategic approaches that go beyond simply manipulating price points. By leveraging effective positioning strategies, building strong brand loyalty, and employing promotional tactics judiciously, businesses can create a compelling value proposition that influences customer perceptions, justifies premium pricing, and ultimately maximizes customer lifetime value.

1.2. Factors influencing WTP

Consumer purchasing behavior encompasses the actions of individuals in selecting, acquiring, and utilizing goods and services to fulfill their needs and desires. This process involves various stages and is influenced by a multitude of factors, including personal traits, decision-making patterns, shopping preferences, brand loyalty, and retailer preferences. Each purchase decision is shaped by these factors, beginning with identifying desired commodities, then assessing their utility, estimating available budget, considering market prices, and finally making a consumption choice. Additionally, external influences such as social, cultural, economic, personal, and

psychological factors also play significant roles in consumer purchases (Blackwell, Miniard, and Engel, 2006).

Consumers, whether individuals or households, engage in these processes to acquire products for personal use. There are two main types of consumers: personal consumers and organizational consumers (Krishna, 2010). The study of their actions in obtaining, using, and disposing of products and services is known as consumer behavior. This field investigates the patterns of what, when, how, and why people make purchases (Kotler, 2004). The consumer decision-making journey typically involves stages of recognition, information search, evaluation, purchase, and feedback, ultimately leading to the selection of a product or brand from various available options. However, the factors influencing consumer behavior can vary due to a range of environmental and individual factors. Overall, consumer behavior is a complex and dynamic process that is influenced by a multitude of internal and external factors, making it a crucial area of study for marketers and businesses (Blackwell, Miniard, and Engel, 2006).

Consumer buying behavior is influenced by two major factors. These factors are individual and environmental. The major categories of individual factors affecting consumer behavior are demographics, consumer Knowledge, perception, learning, motivation, personality, beliefs, attitudes and life styles. The second category of factors is environmental factors. Environmental factors represent those items outside of the individual that affect individual consumers decision making process. These factors include culture, social class, reference group, family and household. The above mentioned factors are the major determinants behind the decision of consumers to opt a given good or service (Blackwell, Miniard, and Engel, 2006).

Marketing begins with identifying customer needs and culminates in ensuring their satisfaction. When the focus is on the customer, understanding consumer behavior

becomes imperative (Khan, 2007). In contemporary marketing, customers have abundant choices, making marketer influence crucial in their purchasing decisions. To effectively influence consumers, marketers must delve into the study of consumer behavior, which encompasses cultural, social, personal, and psychological factors that impact buying decisions. This study specifically explores psychological factors.

Consumers are pivotal in the economic ecosystem of a nation, as they drive demand for goods and services. Without their effective demand, the economy could collapse. A customer can be an individual, a company, or any other entity that purchases goods and services from another entity. Consumer behavior involves various mental, emotional, and physical activities undertaken when selecting, purchasing, using, and disposing of products and services to fulfill needs and desires. Understanding these behaviors is essential for marketers and businesses to thrive (Khan, 2007).

1.3. Measuring WTP

Understanding customer behavior and their willingness to pay is a cornerstone of successful marketing strategies. While directly observing a customer's WTP in a real-world purchase scenario provides valuable data, marketers often require more nuanced and controlled methods to assess this crucial concept. Several research techniques can be employed to effectively measure WTP and gain valuable insights into customer preferences.

One approach involves direct methods. Here, researchers directly ask respondents to state their maximum WTP for a specific product or service description. This method offers clear information but may suffer from hypothetical bias, as customers might not always accurately predict their actual purchase behavior. Studies by Van Campenhout & De Wulf (2005) acknowledge this limitation and suggest combining direct methods with other approaches for a more comprehensive picture.

Indirect methods, on the other hand, employ hypothetical scenarios to assess WTP without directly asking for a price point. Techniques such as conjoint analysis present respondents with different product options varying in features and prices. By analyzing their choices and preferences, researchers can estimate the relative value customers place on specific product attributes and ultimately infer their WTP for different product configurations. Similarly, Van Westendorp's Price Sensitivity Meter uses a series of price points to gauge customer reactions, identifying the price range perceived as too expensive, acceptable, and potentially too cheap. This information helps businesses understand the price sensitivity of customers and identify the optimal price point that maximizes revenue.

Beyond these established techniques, the rise of big data and advanced analytics offers new possibilities for measuring WTP. By analyzing past customer behavior data, including purchase history, price sensitivity to past promotions, and online browsing patterns, businesses can develop predictive models to estimate WTP for new products or services. However, this approach requires careful data collection and analysis practices to ensure ethical data usage and avoid potential bias within the models.

Various research techniques, both traditional and innovative, are available to marketers for measuring WTP. Direct methods offer clear data points, while indirect methods provide valuable insights into customer preferences for different product attributes. By strategically utilizing these tools, businesses can gain a deeper understanding of customer behavior and their WTP, enabling them to develop effective pricing strategies, optimize resource allocation, and ultimately achieve long-term success in today's competitive landscape.

1.4. WTP for educational services, especially online courses

Specifically in the context of online educational services (e-learning), WTP is influenced by a combination of several factors on different levels

We first need to look at the whole picture of the environment that the individual is living in, specifically, the role of economic and socio-cultural factors.

Regarding economic factors, macroeconomic stability, characterized by low inflation and steady economic growth, enhances consumer confidence, leading to higher WTP for education. When the economy is stable, individuals are more likely to invest in long-term benefits such as education (Ang, 2001). Individuals with higher disposable income are also generally more willing to pay for educational services. According to Kahneman and Deaton (2010), income positively correlates with the ability to spend on non-essential items, including education. Higher-income groups perceive less financial strain and are more likely to invest in personal development. In addition, education is often considered a normal good, meaning demand increases as income increases (Varian, 1992). Thus, higher income levels lead to higher WTP for educational services. This is why the availability of financial aid, including scholarships and grants, can alleviate the financial burden on students, making them more willing to pay the remaining costs. Research by Dynarski (2002) shows that financial aid positively impacts enrollment rates by reducing the effective price of education. Besides, the presence and cost of substitute educational options, such as traditional in-person courses or other online platforms, significantly influence WTP. When substitutes are perceived as more expensive or less valuable, individuals may be willing to pay more for the preferred option (Tirole, 1988). This principle is evident in online education, where competitive pricing and perceived value can drive consumer choices.

Individuals' willingness to pay for an educational service is also highly influenced by social and cultural factors. In cultures that place a high value on education and continuous learning, individuals are more likely to invest in educational services. Hofstede's (1980) cultural dimensions theory suggests that societies with a high score on long-term orientation emphasize perseverance and thrift, leading to greater investment in education. The norms and expectations of the community regarding education also influence individual decisions. In communities where higher education is highly valued, individuals are more likely to invest in educational services (Coleman, 1988). An individual may also get influenced by peers and family. Research by Granovetter (1973) on the strength of weak ties highlights how social networks and peer recommendations can affect educational choices and WTP. Positive endorsements from peers can enhance perceived value and increase WTP for online courses. Meanwhile, familial encouragement and financial support can lead to higher WTP for educational services (Hossler, Schmit, & Vesper, 1999).

On the individual level, we need to consider demographic factors, their past experiences with the educational services they are considering to purchase, and their perceived value of what is important to them based on these experiences.

Demographic factors, in the context of educational services, also result in differences in WTP. Different age groups exhibit varying levels of WTP for educational services. Younger individuals may be more willing to invest in education for career advancement, while older individuals might seek education for personal fulfillment or skill enhancement. Research by Chao and Schor (1998) indicates that life stage significantly impacts consumer spending patterns. Higher levels of prior education often correlate with a greater willingness to pay for further education. Educated individuals recognize the benefits of continuous learning and are more inclined to invest in it (Pascarella & Terenzini, 2005). In addition, the location of the individual

can indirectly influence WTP, as we need to take into account the economic condition of the geographical regions.

Past experiences also matter, as positive previous interactions with a product or brand enhance future WTP (Kumar et al., 2010). From their own experience, each individual would develop their own preferences, as well as their own perception of what is important to them, which can directly influence WTP (Lancaster, 1966). Then, they can assess if a course is considered high quality by comparing it to their wants and needs. The perceived quality of a product can raise WTP (Garvin, 1987).

2. Customer Segmentation

Since there are many potential customers with different willingness to pay and perceived value in the market, one solution is customer segmentation.

2.1. Definition

Customer segmentation is a strategic marketing process that involves dividing a broad consumer or business market, typically consisting of existing and potential customers, into sub-groups of consumers based on some type of shared characteristics. These characteristics can be demographic, psychographic, geographic, behavioral, or technographic in nature. The goal of segmentation is to identify groups of customers who have similar needs, preferences, and behaviors, allowing companies to tailor their products, services, and marketing efforts to meet the specific demands of each segment more effectively. According to Wedel and Kamakura (2000), segmentation enables businesses to achieve a more precise understanding of their market, leading to more effective targeting and positioning strategies.

2.2. History of development

The concept of customer segmentation has evolved significantly over time. The early foundations of market segmentation can be traced back to the work of Wendell R. Smith in 1956, who introduced the idea of product differentiation and market segmentation as alternative marketing strategies. Smith's seminal paper argued that segmentation could help businesses better meet the needs of different consumer groups, thereby enhancing their competitive advantage (Smith, 1956). This marked the beginning of a more scientific approach to understanding and serving diverse customer bases.

In the 1960s and 1970s, the development of more sophisticated data collection and analysis techniques enabled marketers to implement segmentation strategies more effectively. The advent of computer technology allowed for the processing of large amounts of consumer data, which facilitated the identification of distinct market segments. Research by Wind and Cardozo (1974) emphasized the importance of understanding consumer behavior patterns and using segmentation to create targeted marketing campaigns. Their work highlighted the role of segmentation in improving marketing efficiency and effectiveness.

The 1980s and 1990s saw further advancements in segmentation methodologies, driven by the increasing availability of consumer data and the development of advanced statistical techniques. The use of cluster analysis, factor analysis, and other multivariate techniques became more prevalent, allowing marketers to identify more nuanced and sophisticated segments. According to Dibb and Simkin (1991), these advancements enabled a more detailed and accurate understanding of consumer needs and preferences, leading to more personalized and effective marketing strategies.

In recent years, the rise of digital technology and big data has revolutionized customer segmentation. The ability to collect and analyze vast amounts of data from various sources, including online behavior, social media interactions, and transactional data, has enabled marketers to create highly detailed and dynamic segments. This has led to the emergence of micro-segmentation, where marketers can target extremely specific groups of consumers with tailored messages and offers. Research by Wedel and Kannan (2016) highlights the impact of big data and predictive analytics on the evolution of segmentation, noting that these technologies allow for real-time insights and more agile marketing strategies.

Overall, the development of customer segmentation has been marked by continuous advancements in data collection, analysis techniques, and technological capabilities. From its early conceptualization by Wendell Smith to the current era of big data and predictive analytics, segmentation has become an essential tool for marketers seeking to understand and meet the diverse needs of their customers. By enabling more targeted and personalized marketing efforts, segmentation continues to play a crucial role in driving business success.

2.3. The importance of customer segmentation

Research by the American Marketing Association (AMA) emphasizes the importance of segmentation for successful marketing. They highlight that segmentation allows businesses to "identify and profile distinct groups of customers who are likely to exhibit similar needs and wants." (AMA, n.d.). This targeted approach contrasts with a one-size-fits-all strategy, which often fails to connect with specific customer groups and leads to wasted marketing resources.

Academic research delves deeper into the benefits of customer segmentation. Studies by scholars like Solomon et al. (2016) demonstrate how segmentation enhances

marketing effectiveness. By understanding the factors influencing customer decision-making within each segment, businesses can craft compelling messaging and product offerings that address the specific needs and motivations of those groups. For instance, a segmentation strategy targeting young adults might prioritize marketing messages that emphasize social proof and brand experiences, while a segment of working professionals might be more receptive to messaging focused on efficiency and time-saving benefits.

Segmentation also drives improved product development. By analyzing the characteristics of each segment, businesses can identify unmet needs and develop products or services that cater specifically to those segments. Research by Green and Wind (1973) highlights how segmentation can inform product innovation, allowing businesses to "tailor product offerings to the specific needs of different market segments" and ultimately gain a competitive advantage.

Furthermore, customer segmentation fosters a superior customer experience. Targeted communication creates a sense of personalization, building stronger customer relationships and loyalty. Studies by Reichheld (2006) emphasize the importance of customer loyalty for business success, demonstrating that "loyal customers are more likely to repurchase, recommend the company to others, and be less price-sensitive." Segmentation allows businesses to tailor their interactions with each segment, fostering a positive experience that builds loyalty and increases customer lifetime value.

Finally, customer segmentation facilitates optimized resource allocation. By focusing marketing efforts on the most profitable customer segments, businesses can maximize their return on investment (ROI). Academic research by Lilien and Rangaswamy (1998) explores the role of segmentation in marketing resource allocation and

highlights how "targeting marketing resources towards high-potential customer segments" can significantly improve marketing effectiveness.

Customer segmentation is a powerful tool for businesses to understand their target market, develop effective marketing strategies, and ultimately achieve long-term success. By leveraging research findings that emphasize its benefits, businesses can create targeted messaging, refine product offerings, personalize customer experiences, and optimize resource allocation, ensuring they connect with the right customers in the most impactful way.

2.4. Segmentation in the market for online educational services

Market segmentation in the market for online educational courses involves dividing the diverse pool of potential learners into distinct groups based on various characteristics such as demographics, psychographics, behaviors, and preferences. According to research by Dibb and Simkin (2001), segmentation is essential in understanding the different needs and preferences of learners and tailoring educational offerings to meet these specific requirements. Demographic segmentation, including factors such as age, gender, education level, and occupation, allows online educational providers to target different groups with courses that align with their interests and career goals (Coleman, 2010). For example, courses targeting working professionals may focus on career advancement or skill development relevant to their industry, while courses for retirees may emphasize personal enrichment or hobbies (Bourne, 2003).

Psychographic segmentation, which considers learners' values, attitudes, and lifestyles, further refines the segmentation process. Research by Smith (1956) suggests that understanding learners' motivations for pursuing online education, whether for career advancement, personal growth, or hobbyist interests, enables providers to develop courses that resonate with each segment's unique aspirations and preferences.

18

Behavioral segmentation, based on learners' past interactions and engagement with online courses, helps providers identify different learner personas and tailor their marketing and communication strategies accordingly (Kotler & Keller, 2012). For instance, learners who prefer self-paced learning may respond better to targeted messages promoting flexible course structures, while those who seek interaction and collaboration may be attracted to courses with live online sessions and discussion forums (Pine & Gilmore, 1999).

Geographic segmentation also plays a role in online educational course offerings, particularly in addressing language preferences, regulatory requirements, and cultural differences across different regions (Hofstede, 1980). Providers may offer courses in multiple languages or localize content to appeal to learners in different countries or regions. Technographic segmentation considers learners' technological preferences and proficiency levels, ensuring that courses are accessible and user-friendly across various devices and platforms (Nijssen & Frambach, 2000). For example, courses designed with mobile optimization may cater to learners who prefer to access content on smartphones or tablets, while those with interactive features may appeal to tech-savvy learners seeking a more engaging learning experience (McDonald & Dunbar, 2012).

In summary, market segmentation in the market for online educational courses enables providers to understand and cater to the diverse needs and preferences of learners. By segmenting the market based on demographics, psychographics, behaviors, geography, and technology usage, providers can develop targeted course offerings, marketing strategies, and user experiences that resonate with different learner segments. Academic research underscores the importance of segmentation in driving the success and effectiveness of online educational programs, ultimately leading to improved learner satisfaction, engagement, and outcomes.

CHAPTER II: MARKET OVERVIEW

This chapter introduces context for the study by examining the current market for SAT prep online courses in Vietnam. SAT, a standardized test that is gaining importance in college admissions both globally and in Vietnam, is introduced by discussing its definition and key milestones relevant to this study to develop an understanding of its situation in the Vietnamese market. Understanding the test, especially in the specific context of the nation, situates how potential customers behave when considering purchasing a preparation course, particularly how its value would be perceived and what are the main influencing factors on customers' willingness to pay. The chapter also includes a market research using the PESTEL framework and highlights a previous research on portraying the characteristics of potential customers of the market.

1. Current market for SAT prep online courses in Vietnam

1.1. What is SAT?

The SAT (Scholastic Assessment Test) is a standardized test widely used for college admissions in the United States. It assesses students' readiness for college and is typically taken by high school juniors and seniors. The SAT measures skills in reading, writing, and mathematics, and is scored on a scale of 400 to 1600, combining scores from the two sections: Evidence-Based Reading and Writing (EBRW) and Math. It's an important factor in college admissions, often alongside other criteria such as GPA, extracurricular activities, and recommendation letters.

The College Board, which officially administers the SAT, defines it as follows:

"The SAT is a standardized test that measures a student's readiness for college. It assesses the knowledge and skills that students need to succeed in college and beyond."

This standardized test is gathering its importance in the Vietnamese education market due to the shifts in its format, as well as in the overall market in recent years.

1.2. Growth Potential of the market (PESTEL analysis)

Vietnam's educational landscape is witnessing a growing emphasis on the SAT for university admissions: more and more students are intending to get their higher degrees in universities abroad; at the same time, more and more Vietnamese universities are adopting a new admission criteria requiring the certificate of SAT score. This increased importance has fueled a corresponding rise in demand for SAT preparation resources. However, a recent shift to a fully digital SAT format presents a new challenge for Vietnamese students, requiring them to adapt to a different learning environment. This coincides with a post-pandemic rise in the adoption of online learning methodologies. Below is a detailed PESTLE analysis to provide an overview of the Vietnamese market for online SAT prep courses.

Political Factors

The Vietnamese government's emphasis on standardized testing for university admissions directly influences the demand for SAT prep services. Continued focus on the SAT as a key selection criterion will likely boost market growth.

In addition, government initiatives promoting international student mobility and collaborations with foreign universities can further increase demand for SAT preparation, especially for students seeking admission abroad.

Economic Factors

As Vietnam's economy continues to grow, disposable income levels are expected to rise. This can lead to increased spending on educational resources like SAT prep

courses. At the same time, the expanding middle class creates a larger pool of potential customers who can afford SAT prep services, contributing to market expansion.

Social Factors

The social landscape in Vietnam presents several factors influencing the potential growth of the SAT prep market. Firstly, a growing emphasis is being placed on the importance of higher education, particularly internationally recognized degrees. This shift in societal priorities fuels the demand for SAT preparation, as it serves as a gateway to prestigious universities abroad and a potential path towards better career prospects.

Vietnamese culture places a strong emphasis on meritocracy, where success and social mobility are heavily linked to educational attainment. Scoring well on standardized tests like the SAT is seen as a crucial gateway to prestigious universities, both domestically and internationally. Vietnamese society places high value on academic achievement, creating a sense of pressure and competition among students. Scoring well on the SAT becomes a matter of not only personal ambition but also social standing. This competitive atmosphere combined with cultural emphasis on education fuels the demand for effective SAT prep resources, as students are driven to outperform their peers and gain admission to the most coveted universities, making students and families willing to invest heavily in anything that can potentially improve their test scores and secure admission to top universities.

Furthermore, the increasing importance of English language proficiency in the globalized job market acts as another social driver. Students are increasingly motivated to pursue SAT preparation to improve their English skills, recognizing its value beyond simply achieving a high score on the test. This trend is likely to continue as English fluency becomes a more sought-after skill in the Vietnamese workforce. The growing awareness of the globalized job market and the perceived benefits of an

international education further fuel the demand for SAT prep services. Parents and students recognize the value of a foreign degree in enhancing career prospects and social standing. The SAT serves as a critical hurdle for students seeking admission to prestigious international universities, leading families to prioritize SAT preparation to ensure their children achieve a competitive score.

Technological Factors

On February 25, 2022, the College Board officially announced that the SAT will transition to a fully computer-based format (also known as the Digital SAT), starting in March 2023 for international students, including those in Vietnam.

In addition to the testing format, the test date, question structure, and difficulty will also change. A notable change is the "adaptive" nature of the test, which means that the difficulty of the subsequent section will depend on the student's performance in the previous section.

A summary of the key changes:

- The SAT will be administered entirely on computers, eliminating the traditional paper-and-pencil format.
- The SAT will be offered more frequently throughout the year, with more flexible testing dates and locations.
- The test will continue to assess reading, writing, and language skills but with a revised question structure and new question types.
- The test will adapt to the student's performance, adjusting the difficulty level of questions based on their responses in previous sections.
- The scoring system will remain the same, with a maximum score of 1600.

The College Board has stated that these changes aim to modernize the SAT to align with the evolving learning and testing landscape, enhance the accessibility and

flexibility of the SAT for students worldwide, and provide a more accurate and personalized assessment of student skills and potential.

This recent shift of the SAT to a fully digital format presents both a challenge and an opportunity for the Vietnamese SAT prep market. On the one hand, students require adaptation to this new testing environment, potentially increasing the demand for online SAT prep courses that offer digital learning tools and practice tests specifically designed for the computer-based format. This digitalization aligns perfectly with the growing internet penetration in Vietnam, making online platforms more accessible and convenient for students nationwide. This convergence of the digital SAT with the expanding internet infrastructure creates a fertile ground for online SAT prep providers to reach a wider audience and cater to the evolving needs of students preparing for the test.

Environmental Factors

In areas with limited access to physical SAT prep centers, online platforms can be a crucial alternative, potentially expanding the market reach and making SAT preparation more accessible to students in geographically disadvantaged locations. This is particularly relevant in rural areas or regions with limited educational infrastructure.

Legal Factors

Government regulations governing online education providers and data privacy can impact the operating environment for online SAT prep companies. Clear regulations regarding intellectual property rights for course materials and question banks are essential for protecting the interests of both providers and students.

While Vietnam possesses a legal framework for protecting intellectual property rights (IPR), its enforcement can be less strict compared to some developed countries. This

relative laxity presents both potential advantages and disadvantages for online SAT prep providers.

On the one hand, the less stringent enforcement might allow online platforms to utilize unlicensed content in their question banks or course materials, potentially reducing operational costs. However, this practice poses significant ethical and legal risks.

On the other hand, the lack of strict enforcement can create challenges for online SAT prep providers who invest heavily in developing original question banks and course content. The risk of unauthorized copying and distribution of their intellectual property can be a significant concern, potentially hindering innovation and investment in high-quality course materials.

Therefore, while the less stringent enforcement might offer some short-term cost advantages, it also creates a less secure environment for online SAT prep providers who prioritize ethical practices and original content creation. Long-term success in this market likely hinges on building a reputation for quality and innovation, which ultimately requires robust intellectual property protection measures and stricter enforcement mechanisms.

Potentials and Challenges

Based on this analysis, the SAT prep market in Vietnam exhibits significant potential for growth driven by several key factors. The continued emphasis on standardized testing in university admissions creates a strong demand for preparation resources. As Vietnam's economy continues to flourish and disposable income levels rise, students and families are increasingly able to invest in educational resources like SAT prep courses. Additionally, the expanding middle class creates a larger pool of potential customers who can afford these services.

Furthermore, the growing awareness of the importance of higher education, particularly internationally recognized degrees, fuels the demand for SAT preparation. Students recognize the SAT as a gateway to prestigious universities and a potential path towards better career prospects. This trend is further amplified by the growing importance of English language proficiency in the globalized job market, motivating students to improve their English skills through SAT preparation.

However, despite the promising outlook, the market also faces certain challenges. There is a potential lack of awareness about the availability and benefits of online SAT prep options, particularly in areas with limited exposure to digital learning platforms. Additionally, competition from established traditional in-person prep centers needs to be considered. Ensuring the quality and effectiveness of online courses is crucial to gain student trust and market share. Finally, navigating the complexities of legal and regulatory frameworks governing online education and data privacy is essential for online SAT prep providers operating in Vietnam.

1.3. Competitors

The current Vietnamese SAT prep market features established in-person centers, individual online tutors, and self-study resources, each offering varying levels of accessibility, cost, and personalized attention.

In-person extra-class centers

In-person extra-class centers have long been the dominant players in the Vietnamese SAT prep market. They offer structured classroom instruction, often led by experienced teachers or tutors familiar with the SAT format. These centers provide students with a physical learning space, access to study materials, and the opportunity for in-person interaction with instructors and peers. This traditional approach can be

beneficial for students who thrive in a structured classroom environment and value personalized guidance.

However, in-person centers also present limitations that online platforms have the potential to address. Firstly, their physical setup restricts accessibility for students in remote areas or those with limited transportation options. Additionally, the cost of attending these centers can be significant, especially for comprehensive or long-term programs. This financial barrier can exclude students from lower-income families or those who need to prioritize other expenses. For this reason, in recent years some centers are starting to develop their online presence and set up online classes.

Furthermore, a concerning practice within some in-person centers involves the use of unlicensed materials. This can include textbooks, practice questions, or even entire question banks created by other authors or test prep companies without proper licensing agreements. While this practice might reduce operational costs for the centers, it raises ethical and legal concerns regarding intellectual property rights. The lack of strict enforcement within Vietnam's legal framework might create an environment where such practices go unchecked, potentially hindering innovation and investment in original content creation within the SAT prep industry.

The reliance on unlicensed materials can also raise questions about the quality and effectiveness of the preparation offered by these centers. Students might be exposed to outdated content or question formats that don't align with the most recent SAT revisions or the digital testing environment. This can disadvantage them on the actual test and undermine the value proposition of these in-person programs.

While in-person centers have established themselves as a familiar and accessible option for many Vietnamese students, their limitations in terms of reach, cost, and potential use of unlicensed materials highlight the potential advantages of online SAT prep platforms. These platforms can offer comprehensive and up-to-date content

specifically designed for the digital SAT format, potentially providing a more standardized and reliable learning experience for students across the country, regardless of their location or financial constraints.

Online tutors

Online tutors represent another segment within the Vietnamese SAT prep market. These are often students who have successfully completed the SAT themselves and leverage their experience to offer individual online tutoring sessions. This approach provides a high degree of personalized attention and caters to specific student needs. Tutors can adapt their teaching methods to individual learning styles and address student weaknesses in a targeted manner. Additionally, the online format eliminates geographical limitations, allowing students to connect with tutors from anywhere in the country or even internationally.

However, online tutors also face limitations that online platforms can potentially overcome. Firstly, the cost of individual tutoring can be quite high, especially for students seeking long-term or comprehensive preparation. This financial barrier might exclude students from lower-income families or those who need to manage their resources carefully.

Furthermore, similar to in-person centers, a significant concern regarding online tutors involves the use of unlicensed materials. This can include textbooks, practice questions, or even entire question banks created by other authors or test prep companies without proper licensing agreements. While this practice might reduce costs for the tutors, it raises ethical and legal concerns regarding intellectual property rights. The relative laxity in enforcement within Vietnam's legal framework might create an environment where such practices are not always addressed, potentially hindering the development of original content and innovative teaching methods within the online tutoring space.

It's important to note that while online tutors often rely on third-party materials, their unique advantage lies in their recent experience with the SAT. Having recently taken the test themselves, their knowledge and understanding of the current format, question types, and testing strategies are likely to be more up-to-date compared to materials found in older textbooks or question banks. This can be particularly beneficial for students preparing for the digital SAT, as online tutors can provide insights and guidance specifically tailored to the computer-based testing environment.

However, the lack of standardized quality control within the online tutoring market remains a concern. Students might encounter tutors with varying levels of expertise or teaching styles, potentially leading to inconsistent learning experiences. Additionally, the scheduling flexibility offered by online tutors can sometimes be limited by their individual availability, potentially creating logistical challenges for students seeking consistent support.

While online tutors offer personalized attention and potentially up-to-date insights, the limitations in cost, potential use of unlicensed materials, and lack of standardized quality control highlight the potential advantages of online SAT prep platforms. These platforms can provide standardized, high-quality content specifically designed for the digital SAT format, potentially offering a more reliable and consistent learning experience for students, regardless of their budget or the availability of experienced online tutors in their region.

Self-study materials

Some online resources offer a free and readily available option for self-directed learning.

First, Khan Academy and the College Board have collaborated to offer a comprehensive digital SAT preparation program, providing students with accessible

resources to excel on the exam. Its video tutorials cover all tested areas, including reading, writing, and math. Students can access practice questions, personalized learning plans, and performance tracking tools. Khan Academy's greatest strengths lie in its accessibility and affordability, making it a valuable resource for students seeking a self-directed learning approach. However, the platform might lack the in-depth explanations and personalized feedback often found in paid prep materials or individual tutoring sessions.

The second good and free source is Bluebook, the College Board's official SAT desktop application, which provides a compilation of SAT questions written by official, qualified test-makers. This resource offers students the opportunity to practice with authentic test materials and familiarize themselves with the question types and testing format. While the Blue Book is a valuable tool for developing test-taking skills, it lacks comprehensive explanations or strategic guidance, which might leave students struggling to understand the reasoning behind correct answers or develop effective test-taking strategies.

Another popular source is books. Numerous SAT prep books are available, offering in-depth explanations of concepts, practice questions with detailed answer keys, and test-taking strategies. Popular options like "The Complete Guide to SAT Reading" by Erica Meltzer or "Bringing Home the Score" by Mike Barrett provide structured learning paths and valuable insights from experienced educators. While these books can be highly effective for self-motivated students, they require a significant financial investment, which is why their digital versions are usually sold illegally in Vietnam at extremely low price compared to their retail price, concerning, again, regulations on intellectual property rights. They also might not cater to individual learning styles as effectively as personalized instruction from tutors or online platforms.

2. Target customers

In a previous user research conducted by the author by interviewing current and potential customers to identify customer characteristics, two main customer personas were portrayed: the parent and the student.

However, for the nature of the business, DSAT16, the conductor of the survey, finds it more difficult to connect with the parents. As a result, for feasibility reasons, the research will focus mainly on the behavior and willingness to pay of the students.

In a previous research in July, 2023, DSAT16 has come up with a portrait of its target market. This market segment encompasses ambitious 15-16 year old high school students who aspire to gain admission to prestigious universities, both in Vietnam and abroad. Driven by the desire to excel academically and secure scholarships, these students prioritize standardized testing, particularly the SAT, as a crucial stepping stone towards their educational goals.

These students are likely highly motivated and well-rounded individuals who excel not only in academics but also actively participate in extracurricular activities. They are likely self-directed learners who value a logical and strategic approach to studying, seeking resources and methodologies that optimize their time management and maximize their score improvement potential.

Within this segment, there exists a diversity of learning styles and preferences. Some students might excel in traditional classroom settings, thriving on lectures and guided exercises. Others might be more independent learners, gravitating towards online resources and self-paced study strategies. However, a common thread unites this market segment: a shared ambition to achieve a high score on the SAT, particularly to achieve maximum or near maximum score in the Math Section and to score high in the Reading and Writing Section. This focus stems from the understanding that Asian

students with good mathematical ground are usually expected to find the Math section way less challenging than those from other countries, and from the expectation that the Language Section would be more challenging for them as English is not their first language.

These students are having some frustrations with their preparation for the SAT. First, as they strive to both achieve high academic results at school and maintain a balanced life by participating in social activities, their schedules are usually tight and hence require some time flexibility in their course, which is also the reason why online courses are becoming more popular. However, they are also concerned about the quality of the course, as well as the fact that they might be easily distracted when studying online, which might not lead to their desired results.

Second, while taking the SAT (or when taking practice tests or doing exercises to prepare for the SAT), they often fall into the traps of the test makers. As a result, even though they studied hard, they still make mistakes that prevent them from improving their score, making them feel discouraged of being stuck at the bottleneck. Therefore, these students understand that they need a clear approach, as well as tips and tricks, to understand the logic behind each question to provide the correct answer. They will be motivated if their improvement in SAT is visible through their higher score every time they take another test or practice test.

Finally, the students often report having trouble with time pressure. The SAT consists of 98 questions with a time allowance of 2 hours and 14 minutes (excluding break time). This extended duration can be mentally draining, leading many students to struggle with finishing the test on time. This time constraint often triggers panic and test anxiety, creating a vicious cycle that further hinders their performance. Additionally, even students who perform well under timed conditions during practice may experience a significant decrease in their actual test scores due to the heightened

anxiety associated with the real exam setting. As a result, they look for a prep course that provides them with tips and tricks to help them answer the questions more quickly but also more accurately, as well as some mental preparation to get used to the test environment, avoiding anxiety on the official test day.

To illustrate the portrait of the targeted customer segment, DSAT16 built a customer persona using a fictional student Nguyen Van An:
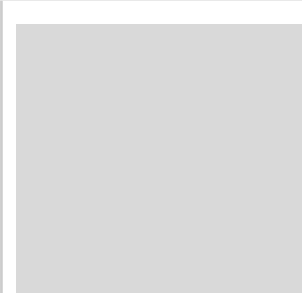
NGUYEN Van An

15-16 years old

Student of Hanoi-Amsterdam high school for the gifted

An is currently an 11th grade student in Hanoi-Amsterdam highschool. He has a dream of studying abroad in the US; hence he wants to start studying for the SAT this summer so that he can get the certificate this year. His target score exceeds 1500, facilitating his eligibility for scholarships. Having already undertaken the SAT once, achieving a score of 1400, he identified the Reading and Writing (R&W) section as an area necessitating improvement. Besides studying, he is also an active student who participates in many extracurricular activities such as scientific research and organizing club events to build a good profile.

Interests	Influences	Goals
➢ High academic achievement ➢ Life balance	➢ Teacher profile ➢ Tips and tricks for the test ➢ Interesting lecture	➢ Increase SAT score ➢ Burst through bottleneck zone ➢ High SAT score to study abroad in the US
Needs & Expectations	**Motivations**	**Pain Points**
➢ Logical approach to the test ➢ Preparation strategy ➢ Lectures and exercises	➢ Visible score improvement ➢ Clear of what to do	➢ Tight schedule ➢ Easily falls into test-makers' traps ➢ Time pressure

Table 1: Customer Persona

(Source: Author)

This customer persona serves as the basis upon which the questionnaire and hypotheses about potential customers are built. These will be discussed further in Chapter III.

This chapter explored the Vietnamese market for online SAT prep courses, laying the foundation for understanding customer behavior and willingness to pay. We began by introducing the SAT itself, outlining its key features and its growing importance in Vietnamese college admissions. By examining the test within the Vietnamese context, we gained insights into how potential customers might perceive the value of SAT prep courses and the factors influencing their purchasing decisions.

Furthermore, a PESTEL analysis provided a comprehensive view of the social, political, economic, technological, environmental, and legal factors shaping the Vietnamese SAT prep market. Additionally, we reviewed previous research that helps us understand the characteristics of potential customers in this market. By analyzing the test, market dynamics, and potential customer behavior, this chapter has established a strong foundation for further research into the factors influencing Vietnamese students' willingness to pay for online SAT prep courses.

The next chapter builds on the discussion of the described context by reviewing existing literature on the influences of different factors on customers' behavior and willingness to pay, especially for online educational services in a post-pandemic era.

CHAPTER III: RESEARCH METHODOLOGY

In this chapter, the methodological research employed for this dissertation will be presented and explained in detail.

1. Designing the Questionnaire

This study adopts quantitative research methodology in analyzing the impact of personal, psychological and product-related factors on the behavior of Vietnamese high school students in purchasing online educational services to prepare for the SAT exam. It is based on a survey organized by the author in April, 2024

At the time working for DSAT16, an online SAT preparation course provider in Vietnam, the author had the chance to conduct several interviews with both existing and potential customers to build the customer personas. In the process, important ideas unfold about the customers' interests, needs and expectations, influences, goals, motivations and frustrations.

Then, based on the content of the interviews as well as the aforementioned theoretical framework, the survey is designed with 4 main parts:

Part 1 collects data about the respondents' socio-demographic characteristics such as age (by asking about which grade they are in), gender and location. Questions regarding the respondents' income levels or financial situations are left out as the main subject of the survey are high school students, who are receiving financial support from their parents. The survey conductor is concerned that asking about their parents' income may lead to some privacy concerns, or they might underreport their situations because they are not fully aware of their parents' financial status. In addition, income alone does not reflect the full picture, and it would be complicated to analyze other factors as well in this dissertation.

Part 2 assesses students' awareness of the SAT test, their plans for taking the test, and their target score. It also gathers information on their previous experience with the SAT or official practice tests. This data helps establish the baseline knowledge and aspirations of the target population.

Part 3 explores students' comfort level with online learning and their past experiences with different SAT preparation methods. This information provides insights into student preferences and potential barriers to online learning adoption.

Part 4 focuses on the factors that students consider important when choosing an online SAT prep course. It utilizes a Likert scale to gauge the relative importance of various features, including reviews, question banks, instructor qualifications, and cost. This data reveals the key drivers of student value perception and helps identify the most relevant selling points for online courses.

Furthermore, this part includes a question about the maximum amount that students are willing to pay for a comprehensive online SAT prep course with specific features. This data point directly addresses the research objective of understanding student WTP. Additionally, it explores the potential interest in a subscription model as an alternative pricing strategy.

Finally, this part allows students to suggest a new feature they believe would enhance their online learning experience. This question can provide valuable insights into student needs and potential areas for course improvement.

The questionnaire was then sent to the thesis supervisor as well as discussed within DSAT16 to get their feedback for adjustment. Afterward, due to the length of the questionnaire, it was decided that the socio-demographic section should be moved to the end to ensure to gather the responses of the other three sections, which are more central for the primary purpose of the study.

The preview of the questionnaire is included in the appendix of this thesis.

2. Conducting the Survey

The final survey questionnaire is then distributed through the social media channel of the online SAT course provider (mainly Facebook). It is then also distributed in various SAT self-study groups in Vietnam to reach more audiences and to ensure a more diversified and representative sample of the population.

The tool to collect responses for the survey is Google Form. In addition, to create an incentive to fill and send the survey, DSAT16 decided to create several self-study documents as well as a free trial of the software to attach to the thank-you note after the students submit their responses.

3. Ordinal Regression

To establish some notation and review the concepts involved in ordinal regression, let Y be an ordinal outcome with J categories. Then $P(Y \leq j)$ is the cumulative probability of Y less than or equal to a specific category $j = 1, \ldots, J - 1$. Note that $P(Y \leq J) = 1$. The odds of being less than or equal to a particular category can be defined as $\frac{P(Y \leq j)}{P(Y > j)}$ for $j = 1, \ldots, J - 1$ since $P(Y > J) = 0$ and dividing by zero is undefined. The log odds is also known as the logit:

$$log\frac{P(Y \leq j)}{P(Y > j)} = logit(P(Y \leq j))$$

3.1. Ordinal Logistic Regression Model

The ordinal regression model can be defined as

$$logit(P(Y \leq j)) = \beta_{j0} + \beta_{j1}x_1 + \ldots + \beta_{jp}x_p,$$

where β_{j0}, β_{j1}, . . . , β_{jp} are model coefficient parameters (i.e., intercepts and slopes) with p predictors for $j = 1, . . . , J - 1$. Due to the parallel lines assumption, the intercepts are different for each category but the slopes are constant across categories, which simplifies the equation above to

$$logit(P(Y \leq j)) = \beta_{j0} + \beta_1 x_1 + . . . + \beta_p x_p.$$

3.2. Model's goodness of fit

To assess the model's goodness of fit, we use the Hosmer–Lemeshow test, which is a statistical test for goodness of fit and calibration for logistic regression models. The test assesses whether or not the observed event rates match expected event rates in subgroups of the model population. Models for which expected and observed rates in subgroups are similar are called well-calibrated. The test was named after its developers, statisticians David Hosmer and Stanley Lemeshow, and it was popularized by their textbook on logistic regression.

The test utilizes the Pearson chi-squared method to test if the observed and expected proportions differ significantly. The null hypothesis is that the observed and expected proportions are the same across all subgroups. The alternative hypothesis is that the observed and expected proportions are not the same. Therefore, if the obtained p-value is greater than 0.05, we can conclude that there is no evidence to reject the null hypothesis, which suggests that the ordinal regression model has an adequate fit to the data. Otherwise, if p-value is smaller than or equal to 0.05, we can conclude that there is evidence to reject the null hypothesis, which means that the model does not fit the data well.

3.3. Proportional odds assumption

The key assumption in ordinal regression is that the effects of any explanatory variables are consistent or proportional across the different thresholds, hence this is usually termed the assumption of proportional odds (or the parallel lines assumptions, since it assumes the coefficients are equal). This assumes that the explanatory variables have the same effect on the odds regardless of the threshold. If a set of separate binary logistic regressions were fitted to the dât, a common odds ratio for an explanatory variable would be observed across all the regressions. In ordinal regression there will be separate intercepts terms at each threshold, but a single odds ratio for the effect of each explanatory variable.

4. Cluster Analysis

Cluster analysis is a versatile and exploratory data analysis technique used to identify natural groupings or clusters within a dataset. It is also known as segmentation analysis or taxonomy analysis and is particularly useful when the groupings within data are not previously known. This technique is exploratory in nature, focusing solely on discovering and describing structures and patterns in the data without distinguishing between dependent and independent variables.

In this case, we are running a cluster analysis to help with market segmentation by analyzing the respondent data to create informed targeted marketing strategies by identifying customer segments with similar characteristics — current SAT score, targeted SAT score, and willingness to pay for the course. The main methodology would be agglomerative hierarchical clustering.

4.1. Computing the Distances

First step in a cluster analysis is to calculate the distance matrix among the data points. Euclidean distance is typically the most familiar and straightforward method for

calculating the distance between two points in any dimensional space. It represents the length of the direct line segment that connects the points, computed as the square root of the sum of the squared differences between their respective coordinates. Mathematically expressed:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \ldots + (p_n - q_n)^2}$$

Euclidean distance is straightforward to compute and understand but has limitations when applied to multivariate data. One major drawback is that it doesn't consider the correlation between highly correlated variables. Consequently, Euclidean distance gives equal weight to these variables, which essentially measure the same characteristic, leading to an undue emphasis on that characteristic.

An alternative method is to adjust the contribution of individual variables to the distance measure based on their variability, a concept utilized by the Mahalanobis distance. Developed by Indian statistician PC Mahalanobis, this statistical measure is widely used in multivariate statistics. Unlike Euclidean distance, Mahalanobis distance considers correlations between variables (Ghorbani, 2019). It is computed as the square root of the product of the difference vector, the inverse covariance matrix, and the transpose of the difference vector.

Given a probability distribution Q, with $\mu = (\mu_1, \mu_2, \ldots, \mu_N)^T$ being the vector of mean values of independent variables (where T indicates that the vector should be transposed), and covariance matrix Σ, the Mahalanobis of the data point with vector $x = (x_1, x_2, \ldots, x_N)^T$ from Q is (Maesschalck, Jouan-Rimbaud & Massart, 2000):

$$D = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}$$

The Mahalanobis distance between two points x and y with respect to Q is:

$$D(x, y) = \sqrt{(x - y)^T \Sigma^{-1} (x - y)}$$

Mahalanobis distance accounts for data distribution and covariance, making it more suitable for real-life scenarios, such as analyzing customer behavior, where it can better represent dissimilarities by considering correlations between purchase patterns. This capability to capture meaningful distinctions in complex, correlated multivariate datasets showcases the advantage of Mahalanobis distance over traditional distance metrics.

4.2. Agglomerating the Clusters

Once the distance matrix is established, the second step of hierarchical clustering is to start grouping clusters, which requires calculating not only the distance between data points but also the distance between clusters (or between a point and a cluster). There are four main methods to achieve this: single-linkage clustering (computing the minimum of object distances), complete-linkage clustering (computing the maximum of object distances), average-linkage clustering (computing the average of object distances), and Ward's linkage clustering (computing the increase in the "error sum of squares" after fusing two clusters into a single cluster, hence minimizing the total within-cluster variance). While conducting the cluster analysis, we will set up a dendrogram for each of these methods to identify which method works best and use it for our model.

4.3. Evaluation and Assessment

After setting up the dendrograms, we need to determine the number of clusters for our analysis. There are various indices for evaluating and assessing the clustering results, and we can use Charrad, Ghazzali, Boiteau and Niknafs's NbClust Package, which provides 30 indices for determining the optimal number of clusters in a data set

and proposes the best clustering scheme from different results obtained, as a basis to our evaluation and assessment.

The table below summarizes indices implemented in NbClust and the criteria used to select the optimal number of clusters.

	NbClust Index	Reference	Criteria
1	"kl"	Krzanowski and Lai, 1988	Maximum value of the index
2	"ch"	Calinski and Harabasz, 1974	Maximum value of the index
3	"hartigan"	Hartigan, 1975	Maximum difference between hierarchy levels of the index
4	"ccc"	Sarle, 1983	Maximum value of the index
5	"scott"	Scott and Symons, 1971	Maximum difference between hierarchy levels of the index
6	"marriot"	Marriot, 1971	Maximum value of second differences between levels of the index
7	"trcovw"	Milligan and Cooper, 1985	Maximum difference between hierarchy levels of the index
8	"tracew"	Milligan and Cooper, 1985	Maximum value of absolute second differences between levels of the index
9	"friedman"	Friedman and Rubin, 1967	Maximum difference between hierarchy levels of the index
10	"rubin"	Friedman and Rubin, 1967	Minimum value of second differences between levels of the index
11	"cindex"	Hubert and Levin, 1976	Minimum value of the index
12	"db"	Davies and Bouldin, 1979	Minimum value of the index

13	"silhouette"	Rousseeuw, 1987	Maximum value of the index
14	"duda"	Duda and Hart, 1973	Smallest n_c such that index > criticalValue
15	"pseudot2"	Duda and Hart, 1973	Smallest n_c such that index < criticalValue
16	"beale"	Beale, 1969	n_c such that critical value of the index >= alpha
17	"ratkowsky"	Ratkowsky and Lance, 1978	Maximum value of the index
18	"ball"	Ball and Hall, 1965	Maximum difference between hierarchy levels of the index
19	"ptbiserial"	Milligan, 1980-1981	Maximum value of the index
20	"gap"	Tibshirani et al., 2001	Smallest n_c such that criticalValue >= 0
21	"frey"	Frey and Van Groenewoud, 1972	The cluster level before that index value < 1.00
22	"mcclain"	McClain and Rao, 1975	Minimum value of the index
23	"gamma"	Baker and Hubert, 1975	Maximum value of the index
24	"gplus"	Rohlf, 1974 Milligan, 1981	Minimum value of the index
25	"tau"	Rohlf, 1974 Milligan, 1981	Maximum value of the index
26	"dunn"	Dunn, 1974	Maximum value of the index
27	"hubert"	Hubert and Arabie, 1985	Graphical method
28	"sdindex"	Halkidi et al., 2000	Minimum value of the index
29	"dindex"	Lebart et al., 2000	Graphical method
30	"sdbw"	Halkidi and Vazirgiannis, 2001	Minimum value of the index

4.4. Segmenting and Profiling

After deciding the optimal number of clusters, the next step is to identify and describe each segment based on the dimensions previously used to compute the distance matrix.

Finally, this thesis will base on the collected data to create a thorough customer profile of each segment, including their socio-demographic profile, psychographic profile, and their needs and wants.

CHAPTER IV: DATA DESCRIPTION

1. Translating and cleaning data

Since the original survey is conducted entirely in Vietnamese, the first step of the data analysis process is to clean the data and translate it into English (the main language of this thesis).

For simplicity purposes, only quantitative data and some of the qualitative data (those that can be transformed into one or multiple binary variables) are kept in the statistical model. All data received under the form of long-text response will be analyzed later to provide more insights for the suggestions and recommendations on pricing and marketing strategies.

The data is translated and cleaned using Python. The process is mostly straightforward for categorical variables. However, there is one variable that requires some more attention to detail: the methods that the students have experienced with for the purpose of preparing for the SAT. It is a check-box question that allows the students to tick all the methods they have tried before, with an option of "Others", where they can add in any method that was not mentioned in the list.

By its nature, the variable value for each response returns a list of methods of SAT prep that the student has tried before. To make the analysis process more simple, it was decided that this variable should be turned into 4 dummy ones, each representing a different method mentioned in the question answer list, receiving "Yes" if the student has tried the method, and "No" if he or she has not.

So what about those who chose "Others". Of the 110 received responses, only 7 ticked this option. Their open-ended responses were all regarding the fact that they have never had any experience with preparing for the SAT, which is consistent with the fact that they did not tick any of the other boxes, or that the values for the new dummy

variables were all "No". At this point, it was safe to keep the dummy variables and discard the original one.

In addition, for the variable giving information about the student's location, the survey sample collected responses from 22 provinces in Vietnam (where there are 63 provinces in total). In which, 64.6% of the respondents are from Hanoi, 12.4% are from Ho Chi Minh City (HCM), and the remaining of the respondents are from the remaining 20 cities, each city accounting for less than 3% of the respondents. To simplify the model, this variable was turned into a categorical one that received 3 values, "Hanoi", "HCM", and "Others".

In the end, it was noticed that although the survey stated clearly that it is made for students, there were still 2 two respondents claiming they are parents. Since these are minority and insignificant (less than 1% of the total respondents), they are dropped out from the analysis.

2. Data Description

For simplicity purposes, the variables in the dataset, as well as their values, are coded as described in the table below:

	Variable name	Description	Coded values
V1	SAT_aware	Whether the respondent is aware of the SAT test and its importance in college admissions	1 — Yes 0 — No
V2	SAT_in_1_year	Whether the respondent plans to take the SAT within one year	1 — Yes 0 — No
V3	SAT_plan	Whether the	1 — Apply for universities

		respondent plans to use your SAT score to apply for universities in Vietnam or to study abroad	in Vietnam 2 — Apply for universities abroad 3 — Both
V4	test_before	Whether the respondent has ever taken the SAT or completed an official full-length practice test before	1 — He or she has already taken the official test at least once. 2 — He or she has already taken one of the official practice tests but not the official one 3 — He or she has taken neither of the tests mentioned
V5	current_score	The interval into which the respondent's current SAT score (or estimated) falls. This is left blank for those who have not taken the test or the practice test.	Minimum score is 400. Maximum score is 1600.
V6	targeted_score	The interval into which the respondent's aimed SAT score (or estimated) falls.	1 — Under 1000 2 — From 1000 to 1090 3 — From 1100 to 1190 4 — From 1200 to 1290 5 — From 1300 to 1390 6 — From 1400 to 1490 7 — From 1500 to 1600
V7	comf_study_online	How comfortable the respondent is with learning online	1 — Very uncomfortable 2 — Somewhat uncomfortable 3 — Neutral 4 — Somewhat comfortable 5 — Very comfortable
V8	met_online_mat	Whether the respondent has utilized	1 — Yes 0 — No

		any online materials to prepare for the SAT before	
V9	met_books	Whether the respondent has used any books to prepare for the SAT before	1 — Yes 0 — No
V10	met_tutor	Whether the respondent has tried any tutoring services to prepare for the SAT before	1 — Yes 0 — No
V11	met_center	Whether the respondent has been to any extra-class center to prepare for the SAT before	1 — Yes 0— No
V12	imp_stud_reviews	The respondent's perception of the significance of old students' reviews when considering online SAT preparation courses	1 — Not at all important 2 — Not very important 3 — Neutral 4 — Somewhat important 5 — Very important
V13	imp_test_bank	The respondent's perception of the significance of the volume of the course's question bank bank when considering online SAT preparation courses	1 — Not at all important 2 — Not very important 3 — Neutral 4 — Somewhat important 5 — Very important
V14	imp_teacher_result	The respondent's perception of the significance of the teacher's SAT score when considering	1 — Not at all important 2 — Not very important 3 — Neutral 4 — Somewhat important 5 — Very important

		online SAT preparation courses	
V15	imp_teacher_exp	The respondent's perception of the significance of the teacher's teaching experience when considering online SAT preparation courses	1 — Not at all important 2 — Not very important 3 — Neutral 4 — Somewhat important 5 — Very important
V16	imp_fee	The respondent's perception of the significance of the course's fee when considering online SAT preparation courses	1 — Not at all important 2 — Not very important 3 — Neutral 4 — Somewhat important 5 — Very important
V17	imp_ipr	The respondent's perception of the significance of the course's respect to intellectual property rights when considering online SAT preparation courses	1 — Not at all important 2 — Not very important 3 — Neutral 4 — Somewhat important 5 — Very important
V18	perceived_value	The maximum the respondent would be willing to pay for a comprehensive online SAT prep course that offers some listed features	1 — Under 2.000.000 VND 2 — From 2.000.000 VND to under 4.000.000 VND 3 — From 4.000.000 VND to under 6.000.000 VND 4 — From 6.000.000 VND to under 8.000.000 VND 5 — From 8.000.000 VND to under 10.000.000 VND 6 — More than 10.000.000 VND
V19	payment_method	The respondent's preferred method of payment	1 — One-time payment 2 — Monthly payment

V20	request_feature	The additional course's feature that the respondent thinks would create the most value for his or her preparation for the SAT	1 — Full-length practice tests + Result reviews 2 — Flashcards for vocabulary 3 — Daily notification 4 — Larger question bank 5 — Others
V21	gender	The respondent's gender	0 — Female 1 — Male
V22	grade	Which grade the respondent is in if he or she is a highschool student, or whether the student is a university student	0 — Grade 10 1 — Grade 11 2 — Grade 12 3 — University student
V23	location	The city that the respondent is currently living in	1 — Hanoi 2 — HCM 3 — Other cities

Table 3: Data description

(Source: Author)

CHAPTER V: ANALYSIS RESULT

1. Summary of Variables

The bar plots below summarize the data collected from the survey by variables.

100% of the survey respondents were aware of the importance of the SAT in college admission in and out of Vietnam. 92% of them also plan to take the SAT within one year since the day of the survey. 85% of the respondents plan to use the SAT score to apply for universities in Vietnam, and 69% plan to use it to apply for universities abroad.
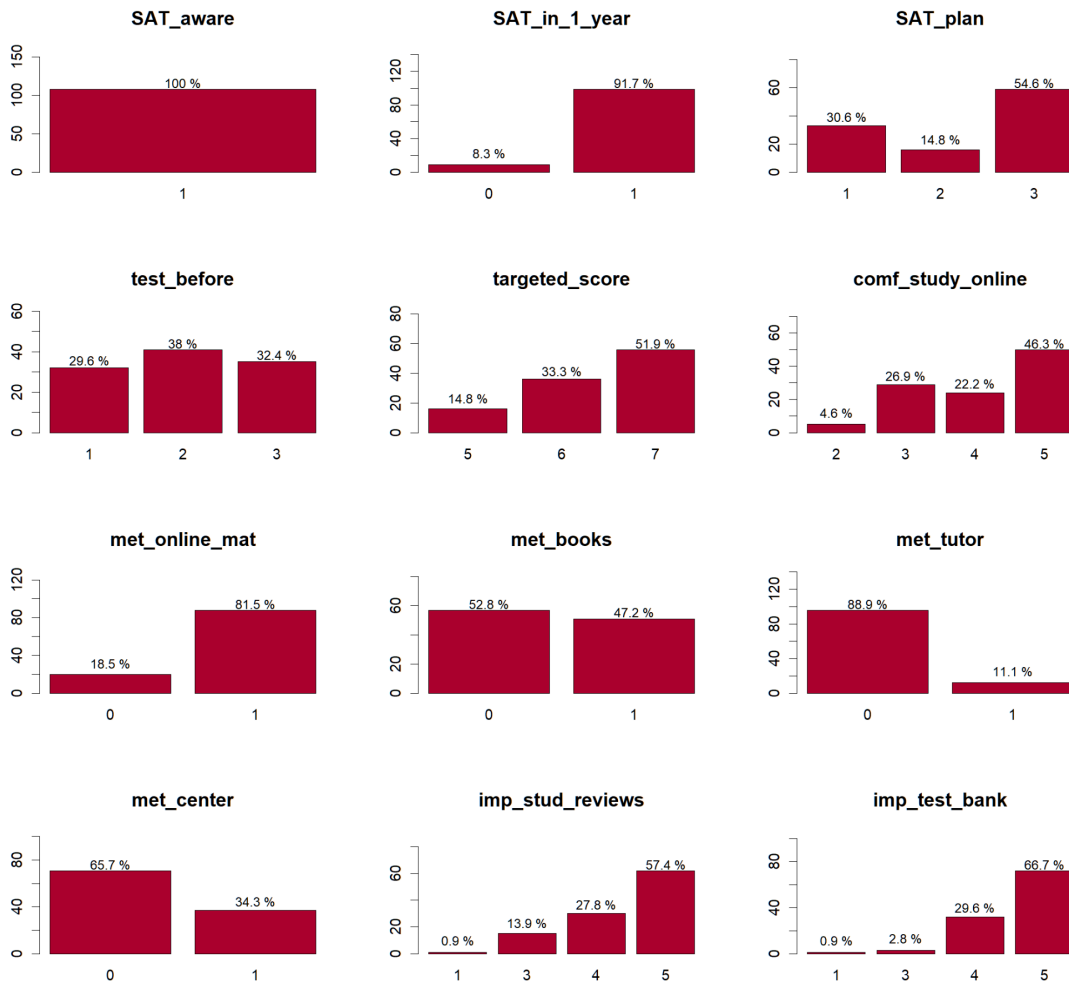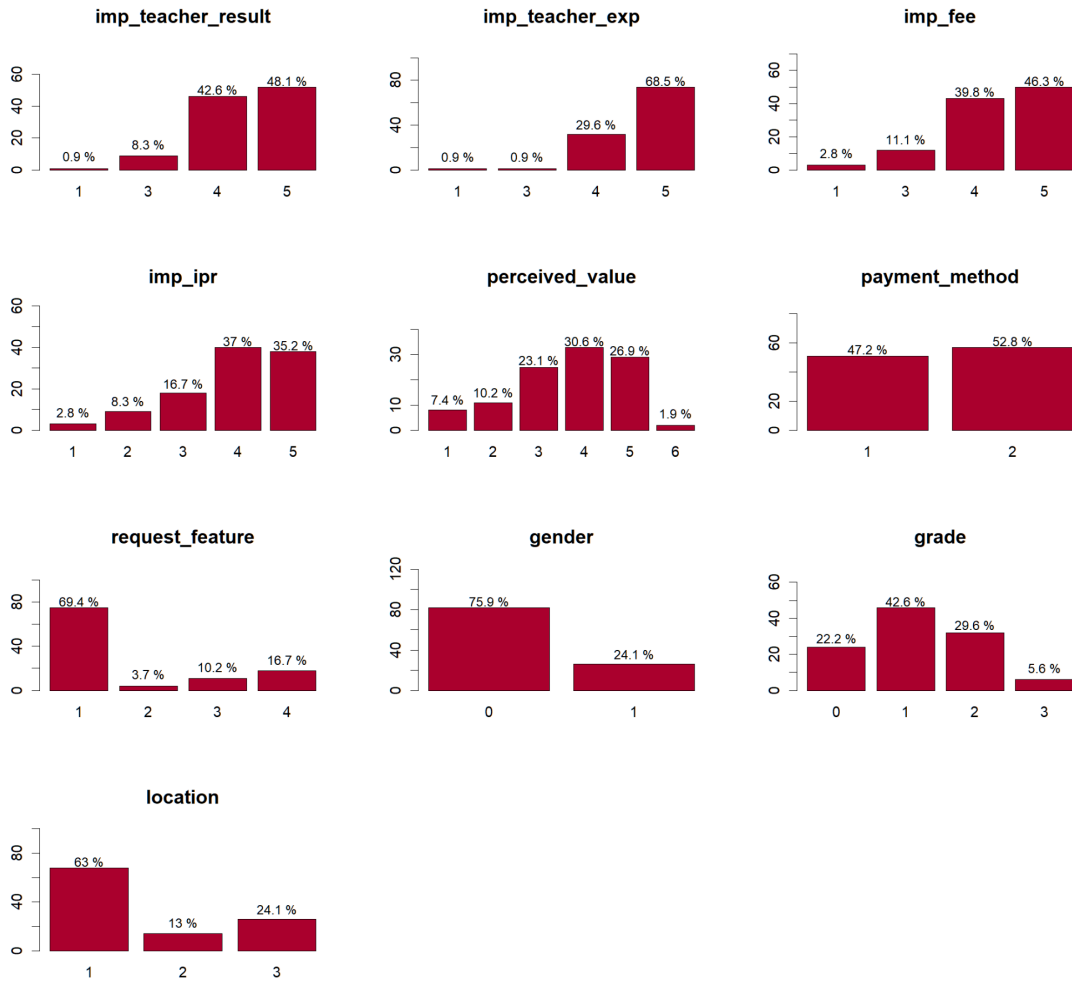
Figure 2: Summary of Variables

(Source: Author)

Of all the respondents, 30% have already taken the official test and would like to retake it to improve their score, while 38% only have experience with the official practice tests by College Board, the test administrator, and 32% have never taken any test or practice test before. For those who have taken the official SAT or one of the official practice tests, their current scores range from 980 to 1530, with a median score of 1270.

Current SAT scores
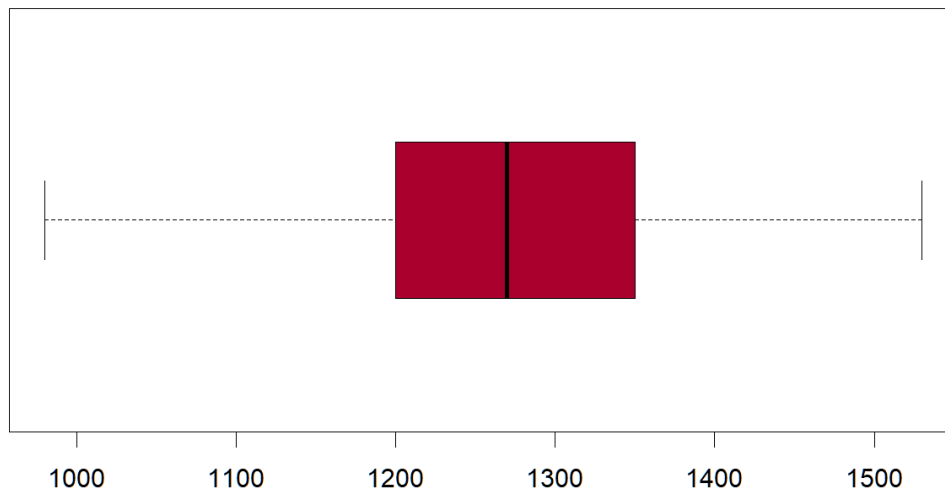
Figure 3: Current SAT score boxplot

(Source: Author)

The culture of high expectations in academic achievement is also reflected in the data sample. 100% of the students participating in the survey reported that they aim to get at least 1300 for the SAT; 85% reported to aim for a score of higher than 1400, and more than half (52%) of the students reported to aim for a score higher than 1500, compared to the maximum score one can get in the SAT being 1600. Notice that according to College Board, "The average SAT score is around 1050. Any score above that would be above average. A score of 1350 would put you in the top 10% of test takers and help make your application competitive at more selective schools." Meanwhile, a 1500 SAT score puts the student at the 98th percentile, meaning he or she scored higher than 98% of all other test takers. To make the dreams of 50% of the students to get in the 98th percentile is simply impossible.

These post-Covid students also reported to be pretty unworried with studying online: only 5% reported to be somewhat uncomfortable with studying online, while 27%

reported to be neutral and more than two thirds (69%) reported to be somewhat comfortable or very comfortable with studying online. This is also reflected in the fact that most students (82%) have some experience with online materials to prepare for the SAT, compared to other more traditional methods like studying from books (47%), using tutoring services (11%) and taking extra classes at SAT prep centers (34%).

Regarding the respondents' perception of value, the most important features are the size of test banks and the teacher's teaching experience: more than 97% of the students rated these factors as somewhat important or very important. The following factors are reviews from old students, the teacher's SAT score and the tuition fee, at around 85% of the students rating these factors as somewhat or very important. Not surprising at all, intellectual property rights are still not as valued as other factors. Nonetheless, it is a good sign that 73% of the respondents rated this factor to be somewhat or very important.

Regarding their willingness to pay for the prescribed online course in the questionnaire, the majority of respondents (80%) agreed that they are willing to pay from 4.000.000 VND to 10.000.000 VND for the course (or approximately from 150 to360 EUR, in May 2024 exchange rate). About half of them prefer monthly payment; however, this is not the most feasible solution as the number of domestic credit card holders in Vietnam is still relatively low (Trinh, Tran & Vuong, 2020). The Global Financial Index Report (World Bank, 2020) shows that the credit card penetration rate in Vietnam in 2020 was 4.7% (World Bank, 2020), which means about 2.65 million workers own credit cards — a relatively modest number compared to 9 million urban households (General Statistics Office of Vietnam, 2020), 46 million bank users (General Statistics Office of Vietnam, 2020) or 15.8 million people in need of consumer loans with annual consumer credit growth of up to 20%.

The sample witnessed a skew in gender as more than three quarters of the survey respondents are female and nearly two thirds of them live in Hanoi. Most of the respondents (95%) are high school students who are preparing for their bachelor application, and the remaining 5% are those who already graduated from high school but intend to change their major or to apply for admission in a better university, either in Vietnam or abroad.

Finally, when given the chance to request one feature to add to the mentioned course to bring the most value for their preparation, nearly 70% of the students requested more full-length practice tests, implying a huge potential to respond to this demand as there are currently only 6 official practice tests on the desktop application Bluebook by College Board. In addition, 17% requested daily notification, which expresses their concern of distraction when studying online.

2. Ordinal Regression Analysis

In our ordinal regression model, the response variable is the willingness to pay for the described online course of the respondents, which is an ordinal variable that receives values "Low", "Medium", and "High", respective to the range that the customers are willing to pay for the course. We use 19 regressors to fit the model. They are: SAT_in_1_year, SAT_plan, test_before, current_score, targeted_score, comf_study_online, met_online_mat, met_books, met_tutor, met_center, imp_stud_reviews, imp_test_bank, imp_teacher_result, imp_teacher_exp, imp_fee, imp_ipr, grade, gender, and location.[1]

In the variables reflecting the importances that the respondents assign to each factor when considering purchasing an online course, due to the issue of having too few observations in the extreme levels (in this case, respondents rarely rated the factors as "not at all important" or "not very important"), the categories are aggregated to form

---

[1] Refer to the description of the variables in Chapter IV.

new clusters for these variables to make the model more manageable. Therefore, the new values of the variables imp_stud_reviews, imp_fee, and imp_ipr would be "Low", "Medium", and "High", reflecting a low, medium, or high sense of importance that they assign to each factor, respectively. On the other hand, since the distribution of the variables imp_test_bank, imp_teacher_result, and imp_teachre_exp are even more extreme (very few people would say they are neutral when considering these factors), their are grouped into "Low" importance and "High" importance.

Of the 19 regressors, only current_score, targeted_score, and comf_study_online are treated as continuous quantitative variables. The remaining 16 dependent variables are treated as categorical variables when fitting the model.

After calculating the p-values, we obtain the coefficients as below:

	Value	SD	t-value	p-value	
SAT_in_1_year | 1	0.936918	0.8873831	1.05582	0.29104990	
SAT_plan | 2	1.442670	0.7944641	1.81590	0.06938510	*
SAT_plan | 3	-0.206967	0.5252579	-0.39403	0.69355933	
test_before | 2	1.011040	0.5443902	1.85720	0.06328313	*
test_before | 3	4.699158	0.2502210	18.78003	0.00000000	***
current_score	0.002809	0.0005434	5.16911	0.00000024	***
targeted_score	-0.083055	0.2361943	-0.35164	0.72510941	
comf_study_online	0.168776	0.2323991	0.72623	0.46769674	
met_online_mat | 1	1.737000	0.7581895	2.29098	0.02196432	**
met_books | 1	-0.486524	0.5133414	-0.94776	0.34325250	
met_tutor | 1	0.957548	0.6424016	1.49058	0.13607310	
Met_center | 1	1.090680	0.5177558	2.10655	0.03515631	**
imp_stud_reviews | Low	0.006593	0.6706607	0.00983	0.99215672	

imp_stud_reviews | Medium	-0.175171	0.5534239	-0.31652	0.75160680	
imp_test_bank | Low	-0.066124	0.5337953	-0.12388	0.90141418	
imp_teacher_exp | Low	-0.379335	0.6532604	-0.58068	0.56145604	
imp_teacher_result | Low	0.519921	0.7157423	0.72641	0.46758872	
imp_fee | Low	0.427449	0.6809104	0.62776	0.53016082	
imp_fee | Medium	0.378829	0.4950304	0.76526	0.44411400	
imp_ipr | Low	-0.671041	0.7387414	-0.90836	0.36368934	
imp_ipr | Medium	-0.465566	0.6958007	-0.66911	0.50342610	
grade | 1	-0.337630	0.6264584	-0.53895	0.58992115	
grade | 2	0.690313	0.6510895	1.06024	0.28903423	
grade | 3	-1.040787	1.1400663	-0.91292	0.36128577	
gender | 1	-0.494793	0.5245661	-0.94324	0.34555685	
location | 2	0.472987	0.7333414	0.64498	0.51894295	
location | 3	1.021647	0.5407262	1.88940	0.05883849	*

Intercept					
Low | Medium	6.7773	0.2134	31.7585	0.00000000	***
Medium | High	8.3818	0.3105	26.9963	0.00000000	***

Table 4: Ordinal logistic regression coefficients table

(Source: Author)

Here we can see that the coefficients for the variables met_online_mat and met_center are statistically significant at 99% confidence level; test_before3 and current_score at 95% confidence level; and SAT_plan2, test_before2, and location3 at 90% confidence level.

2.1. Result interpretation

To make the interpretation more intuitive, we exponentiate the coefficients.

	Value	Exponentiated value
SAT_in_1_year | 1	0.936918	2.552
SAT_plan | 2	1.442670	4.232
SAT_plan | 3	-0.206967	0.813
test_before | 2	1.011040	2.748
test_before | 3	4.699158	109.855
current_score	0.002809	1.003
targeted_score	-0.083055	0.920
comf_study_online	0.168776	1.184
met_online_mat | 1	1.737000	5.680
met_books | 1	-0.486524	0.615
met_tutor | 1	0.957548	2.605
Met_center | 1	1.090680	2.976
imp_stud_reviews | Low	0.006593	1.007
imp_stud_reviews | Medium	-0.175171	0.839
imp_test_bank | Low	-0.066124	0.936
imp_teacher_exp | Low	-0.379335	0.684

imp_teacher_result | Low	0.519921	1.682
imp_fee | Low	0.427449	1.533
imp_fee | Medium	0.378829	1.461
imp_ipr | Low	-0.671041	0.511
imp_ipr | Medium	-0.465566	0.628
grade | 1	-0.337630	0.713
grade | 2	0.690313	1.994
grade | 3	-1.040787	0.353
gender | 1	-0.494793	0.610
location | 2	0.472987	1.605
location | 3	1.021647	2.778

Table 5: Exponentiated values of coefficients

(Source: Author)

According to the table, we can interpret the coefficients that are statistically significant as followed:

For students who plan to use the SAT result to apply for both Vietnamese and international universities, the odds of being willing to pay more for the described online course is 4.2 times that of students who only plan to study abroad with the SAT score, holding constant all other variables.

For students who have never taken the official SAT test but have taken at least one full-length practice test, the odds of being willing to pay more for the described online course is 2.8 times that of students who have already taken the SAT before, holding constant all other variables.

For students who have taken neither the official SAT test nor the official full-length practice test, the odds of being willing to pay more for the described online course is 110 times that of students who have already taken the SAT before, holding constant all other variables.

For every one unit increase in a student's current SAT score, the odds of being willing to pay more for the described course increases by 0.03%, holding constant all other variables. Since the scores vary in intervals of 10 (1390-1400-1410, there is no such score as 1401 or 1402), we can reinterpret: for every 10 unit increase in student's current SAT score, the odds of being willing to pay more for the described course is 3% higher (($1.003^{10} - 1$)*100%), holding constant all variables.

For students who have used online materials to practice with the SAT before, the odds of being willing to pay more for the described online course is 5.7 times that of those who never used online materials for SAT preparation, holding constant all other variables.

For students who have taken extra classes in SAT prep centers, the odds of being willing to pay more for the described online course is 3 times that of those who have never been to any SAT prep center, holding constant all other variables.

For students based in small cities in Vietnam, the odds of being willing to pay more for the described online course is 2.8 times that of those living in the capital, holding constant all other variables.

2.2. Goodness of fit testing

To test the goodness of fit of the model, we run the Hosmer-Lemeshow test.

```
Hosmer-Lemeshow Test:

                  Chi-sq  df  pr(>chi)
ordinal(Hosmerlem)  17.037  17    0.4518

HO: No lack of fit dictated

rho: 93.33%
```

Figure 4: Hosmer-Lemeshow test result

(Source: Author)

We obtain a high p-value (p > 0.05), which indicates that there is no significant difference between the observed and expected frequencies, suggesting a good fit of the model to the data.

2.3. Proportional odds likelihood test

One key assumption of ordinal logistic regression is that the relationship between each pair of outcome groups remains consistent. This means that the coefficients describing the relationship between the lowest category and all higher categories of the response variable are the same as those for the next lowest category and all higher categories, and so on. This is known as the proportional odds assumption or the parallel regression assumption. Because the relationship is consistent across all pairs of groups, there is only one set of coefficients.

If this assumption were not true, different sets of coefficients would be required to describe the relationship between each pair of outcome groups. Therefore, it is important to evaluate whether the proportional odds assumption is valid. Some software packages offer statistical tests for this purpose. However, these tests are criticized for often rejecting the null hypothesis (that the coefficients are the same), indicating that the parallel slopes assumption does not hold, even when it does (see Harrell 2001, p. 335). We could not find a function in R to perform these common

tests. However, Harrell suggests a graphical method to assess the parallel slopes assumption. The graph displays (linear) predictions from a logit model, used to model the probability that y is greater than or equal to a specific value (for each level of y), using one predictor (x) variable at a time. To create this graph, the Hmisc library is needed.

Essentially, we run 3 logistic regression models for each of the response variables (Y >= "Low"), (Y >= "Medium"), and (Y >= "High"), and compare their coefficients to see if the lines are parallel (i.e., if the proportional odds assumption holds).

The first model (with the predicted variable being (Y >= "Low")) will obtain undefined values for all of its coefficients, since the odds of (Y >= "Low") is equal to 1 in any case. Therefore, we only need to observe the difference between the coefficients of the later two models. To help demonstrate this, we normalized all the first set of coefficients to be zero so there is a common reference point.

		N	Low	Medium	High
SAT_in_1_year	0	9	Inf	0	-0.4700036
	1	99	Inf	0	-1.3612583
SAT_plan	1	33	Inf	0	-0.8754687
	2	16	Inf	0	-0.7621401
	3	59	Inf	0	-1.7117168
test_before	1	32	Inf	0	-1.5242801
	2	41	Inf	0	-1.2135423
	3	35	Inf	0	-1.1856237
current_score	0	35	Inf	0	-1.1856237
	[980, 1210)	20	Inf	0	-1.7917595

	[1210, 1340)	28	Inf	0	-1.3350011
	[1340, 1530]	25	Inf	0	-1.0245043
targeted_score	5	16	Inf	0	-1.9771627
	6	36	Inf	0	-1.4759065
	7	56	Inf	0	-1.0454597
comf_study_online	2	5	Inf	0	-0.9808293
	3	29	Inf	0	-1.7869862
	4	24	Inf	0	-1.3350011
	5	50	Inf	0	-1.0649124
met_online_mat	0	20	Inf	0	-1.7346011
	1	88	Inf	0	-1.2247636
met_books	0	57	Inf	0	-1.0916436
	1	51	Inf	0	-1.5108917
met_tutor	0	96	Inf	0	-1.1895841
	1	12	Inf	0	-2.3025851
met_center	0	71	Inf	0	-1.1478833
	1	37	Inf	0	-1.5941704
imp_stud_reviews	High	62	Inf	0	-1.5712861
	Low	16	Inf	0	-0.7884574
	Medium	30	Inf	0	-0.9808293
imp_test_bank	High	72	Inf	0	-1.3623252
	Low	36	Inf	0	-1.1451323
imp_teacher_exp	High	74	Inf	0	-1.5388739
	Low	34	Inf	0	-0.7576857

imp_teacher_result	High	52	Inf	0	-1.6345176
	Low	56	Inf	0	-0.9734491
imp_fee	High	50	Inf	0	-1.0245043
	Low	15	Inf	0	-1.0986123
	Medium	43	Inf	0	-1.6919949
imp_ipr	High	38	Inf	0	-1.9275610
	Low	30	Inf	0	-1.0560527
	Medium	40	Inf	0	-0.9315582
grade	0	24	Inf	0	-1.5020552
	1	46	Inf	0	-1.0010320
	2	32	Inf	0	-1.7267270
	3	6	Inf	0	-0.6931472
gender	0	82	Inf	0	-1.2135423
	1	26	Inf	0	-1.5892352
location	1	68	Inf	0	-1.4677098
	2	14	Inf	0	-0.2876821
	3	26	Inf	0	-1.6739764
Overall		108	Inf	0	-1.2845117

Table 6: Proportional odds assumption testing

(Source: Author)

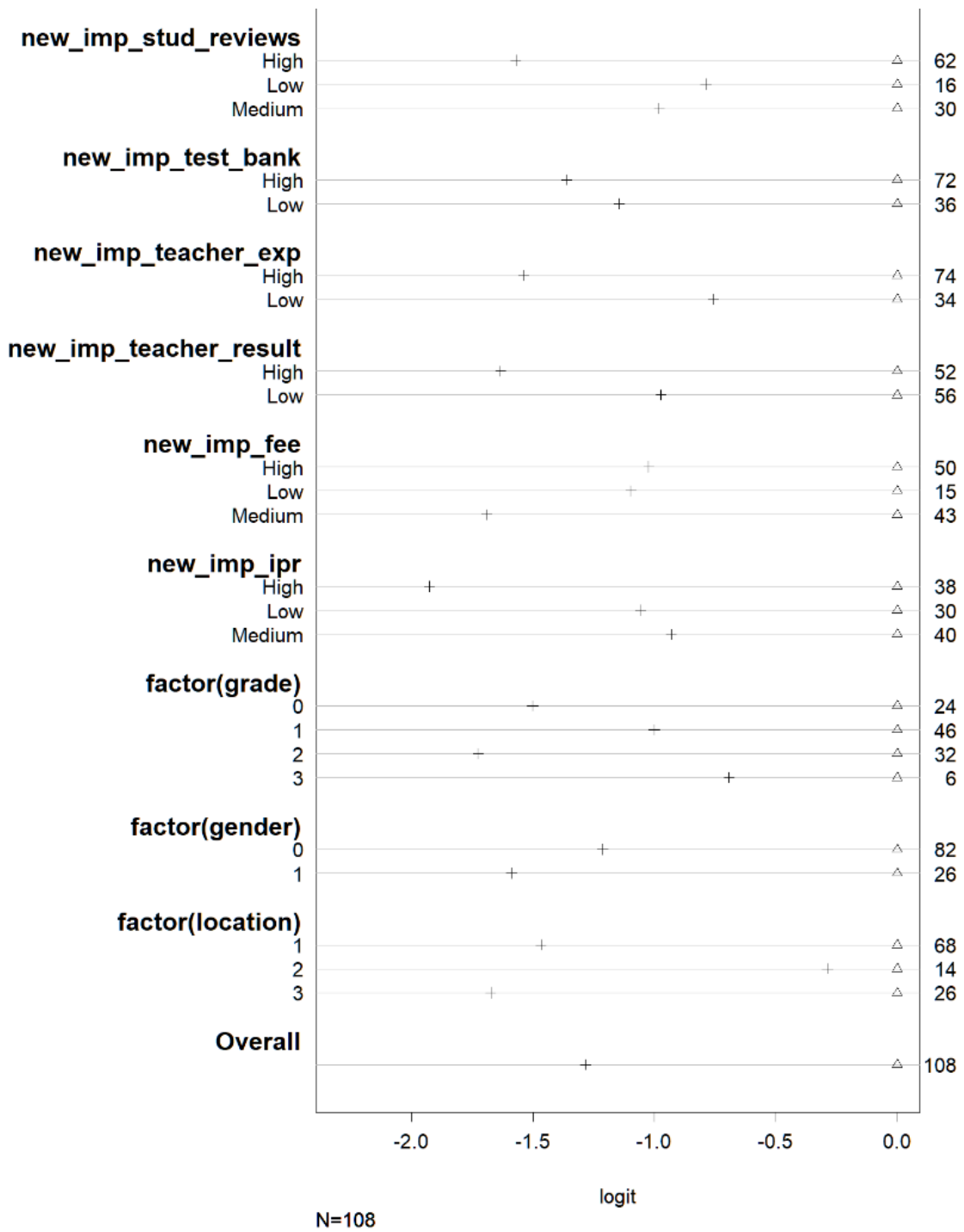From here we can create a plot for easier visualization.
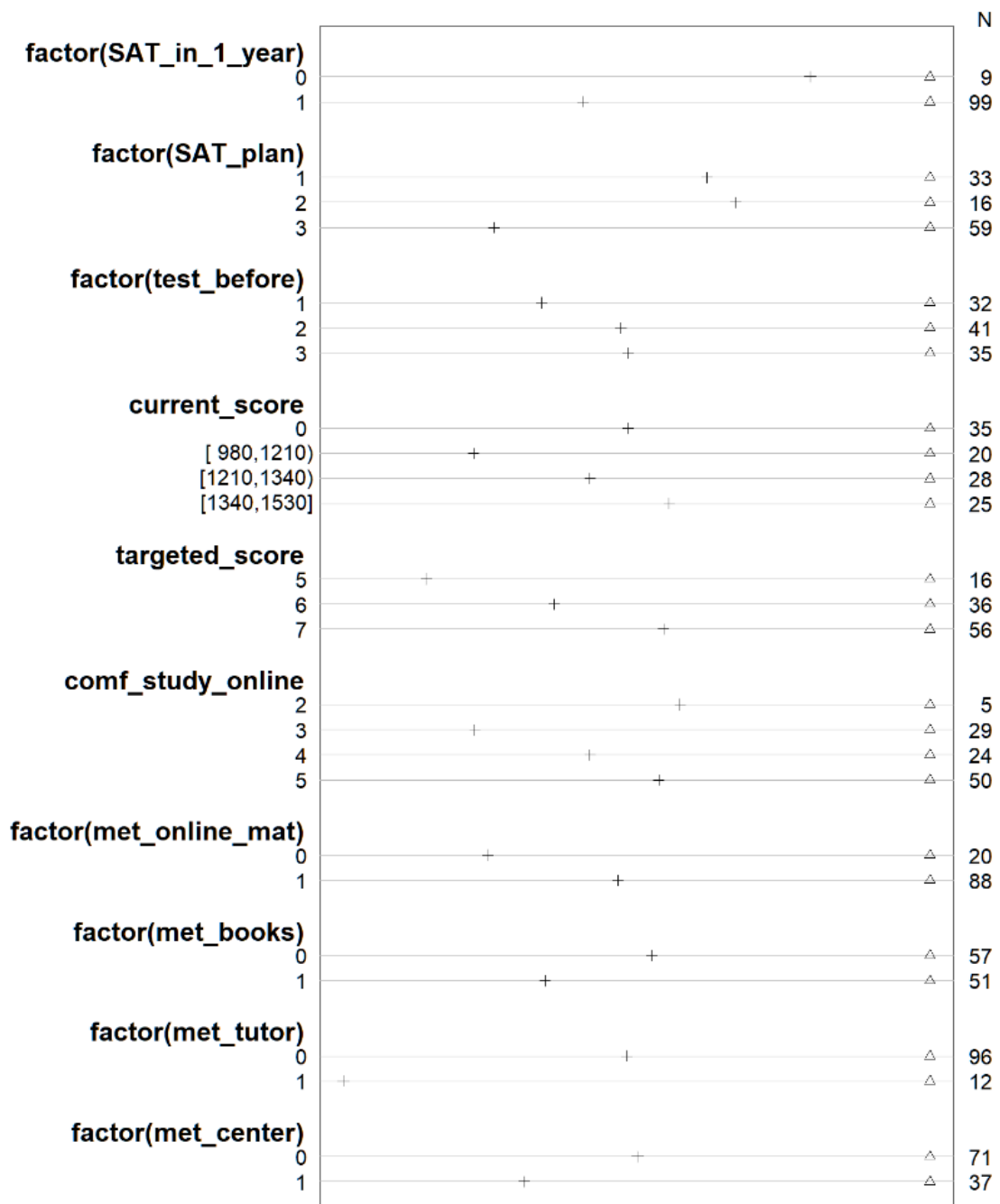
Figure 5: Parallel slopes assumption visualization

(Source: Author)

Looking at the coefficients for the variables test_before, met_online_mat, met_books, met_center, imp_test_bank and gender, we see that the distance between the sets of coefficients is similar.

In contrast, the distances between the estimates for the remaining variables are different (i.e., the markers are placed much further from each other in each of the lines), suggesting that the proportional odds assumption may not hold.

However, if we look more closely, in some variables that have 3 or more levels, there are 2 levels that the markers are close to each other. For example: in the variable imp_stud_reviews, the markers on the "Low" and "Medium" levels are closer together, and further from the marker on the "High" level. The similar phenomenon can also be observed in the variables SAT_plan, current_score, comf_study_online, imp_fee, imp_ipr, grade, and location.

3. Cluster Analysis

We apply the cluster analysis algorithm using the agglomerative hierarchical clustering method to identify the clusters of the data points based on the three dimensions: the student's current score, his or her targeted score and his or her willingness to pay. Then based on the analysis result, we can characterize the clusters according to the variables useful in segmenting the market.

After calculating the Mahalanobis distance matrix, we try to agglomerate the clusters with each of the four methods: single-linkage, complete-linkage, average-linkage, and Ward's linkage.

Single-linkage
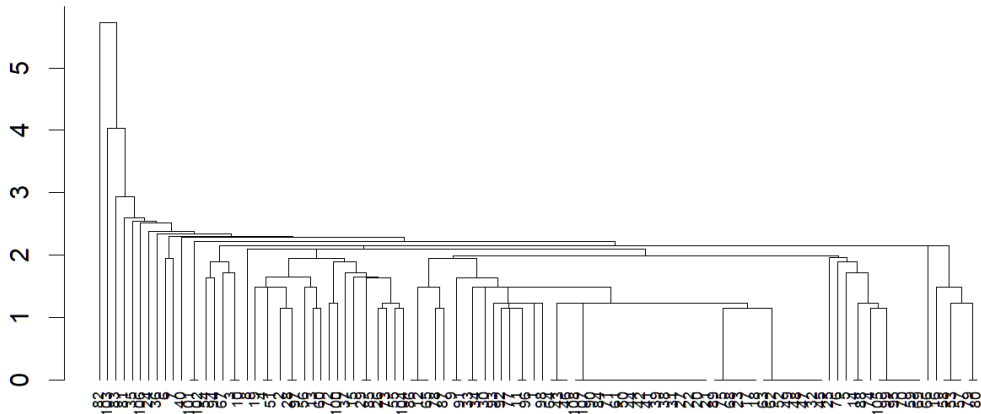
hclust (*, "single")

Figure 6: Single-linkage clustering

(Source: Author)

This is not a good clustering option as most of the units collapse into one cluster with some units from large distances, implying that these are outliers. The distances at the beginning steps of the agglomeration process are high, which means that we are joining dissimilar units; meanwhile, afterward, the distances are small, suggesting the lack of clear separation. We discard this method and move on to the next one.

Complete-linkage
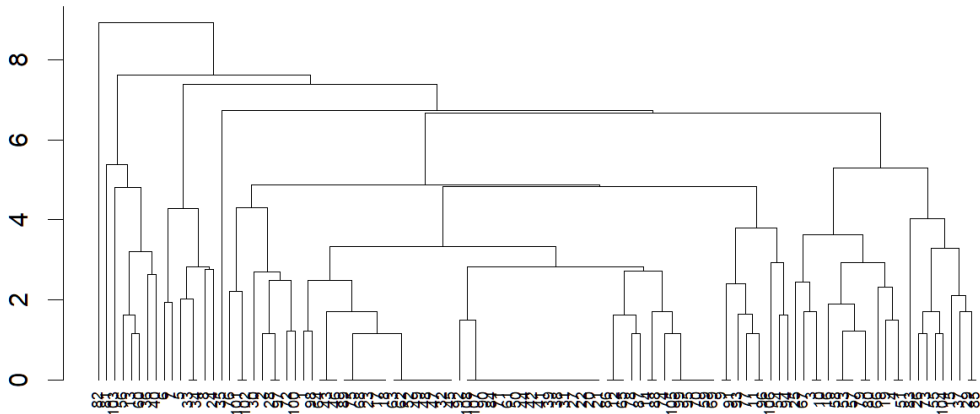
hclust (*, "complete")

Figure 7: Complete-linkage clustering

(Source: Author)

With the complete-linkage method, we still face the problem of outlier. The joined units are still dissimilar, as the distances at the beginning are relatively large. The hierarchy in the dendrogram still does not appear to be clear.

Average-linkage
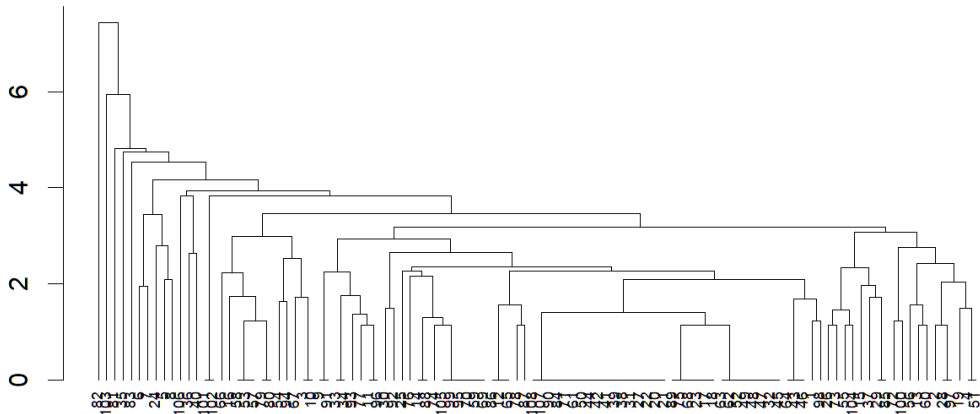
hclust (*, "average")

Figure 8: Average-linkage clustering

(Source: Author)

The average-linkage method gives results similar to the single-linkage clustering: there are still outliers, and the distances are also not large enough to make separations clear at any step.
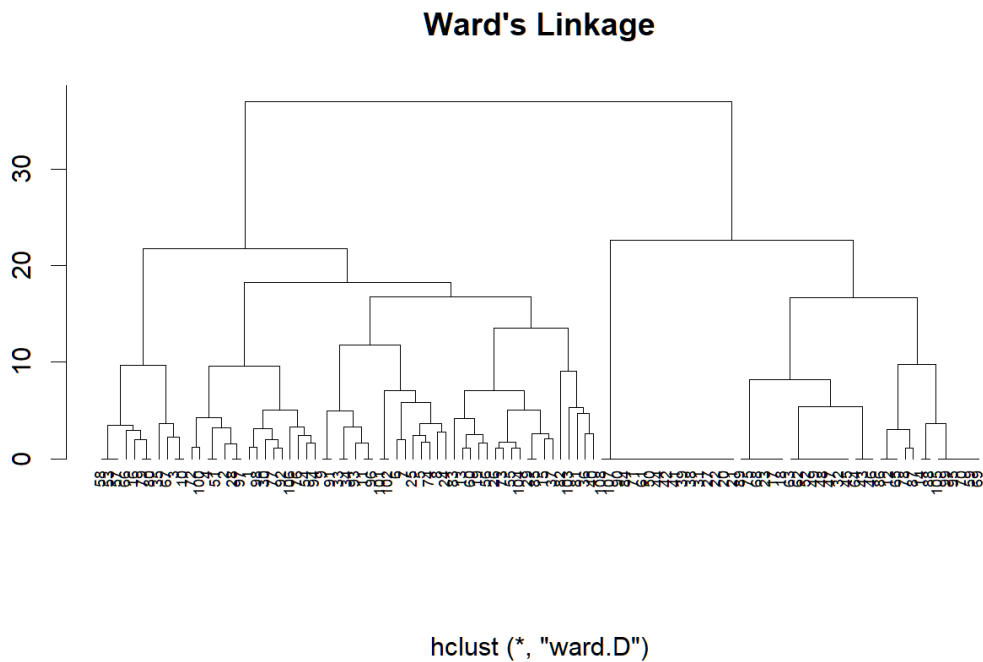
Figure 9: Ward's linkage clustering

(Source: Author)

With the Ward method, we obtain clusters which are much more separated, the distances of the agglomeration are far larger than observed according to other methods of aggregation. On the right part of the plot, these distances are very short at the beginning of the clustering process and then increase. However, in the left part of the plot, we have large distances at the beginning.

Therefore, we keep the clustering solution provided by the Ward method.

3.1. Choosing the Optimal Number of Clusters

To decide on the number of clusters, we can apply the thirty indices provided by the NbClust package. Most indices are based on comparing all possible solutions and finding the one providing clusters which are more compact (within) but also well separated (between).

75

Setting the minimum and maximum acceptable values for the number of clusters as 3 and 7, respectively, we obtain the result that among all indices:

- 9 proposed 3 as the best number of clusters
- 6 proposed 4 as the best number of clusters
- 3 proposed 5 as the best number of clusters
- 2 proposed 6 as the best number of clusters
- 3 proposed 7 as the best number of clusters

Looking at the numerical results, we can see that the best number of clusters is 3 with 9 proposals, and 4 is the second best with 6 proposals. We will proceed with the second-best 4-cluster solution because it is likely to provide a more variegated segmentation..

This is how the separation into four clusters looks like:

Ward's Linkage

hclust (*, "ward.D")
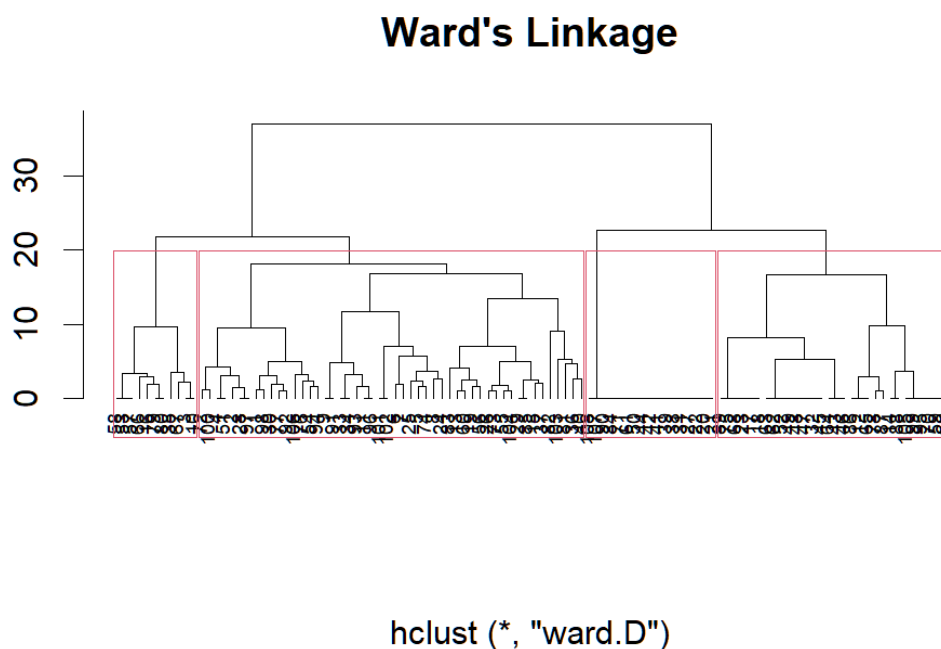
Figure 10: Ward's linkage clustering with 4 clusters
(Source: Author)

3.2. Market Segmentation

Let's proceed onto cluster identification. To illustrate these characteristics better, let's look at the data:

	student _reviews	test_bank	teacher_ result	teacher_ exp	fee	ipr	size
clu1	3.960	4.440	4.160	4.640	4.240	3.260	50
clu2	4.818	4.909	3.364	3.909	4.455	3.818	11
clu3	4.667	4.567	4.733	4.733	3.833	4.500	30
clu4	5.000	5.000	5.000	5.000	5.000	5.000	17

Table 7: Cluster identification

(Source: Author)

According to the matrix, we can see that the first cluster places a high value on most aspects of the online course, with the exception of intellectual property rights responsibility (3.260). Among all factors, they seem most interested in the teacher's teaching experience and the volume of the course's question bank, implying that these students focus more on the practice part of the course. Since they make up approximately 50% of the sample, we can call these the "standard customers"

The second cluster seems to value other students' reviews and the course's test bank the most (4.818 and 4.909, respectively), and place less emphasis on the teacher's actual result and teaching experience (3.364 and 3.909, respectively). This suggests that these students are more oriented toward a practice-based approach (utilizing the test bank rather than instructions) to prepare for the test. We can call these "practice-focused students".

For the third cluster, while they value all aspects to a certain degree, their highest ratings are for factors related to the instructor's credibility (on average, they evaluate

the importance of the teacher's result and teacher's credibility at 4.733 on the Likert scale). This suggests a strong focus on the quality of instruction over other considerations, especially the course fee (3.818). We can call these "instruction-focused students".

The fourth and final cluster represents those who consider all factors to be extremely important (average rating of 5.0 on a 5-point scale) when choosing an online course. In other words, student reviews, test banks, teacher results, teacher experience, course fee, and intellectual property rights are all very important considerations for this group. For this reason, we can call these students the "perfectionists".

Now let's set up the profile for each segment.

3.3. Segment Profiling

	Female	Male
Cluster 1	76.00	24.00
Cluster 2	54.55	45.45
Cluster 3	83.33	16.67
Cluster 4	76.47	23.53
Sample	75.93	24.07

Table 8: Clusters' gender composition (%)

(Source: Author)

	Grade 10	Grade 11	Grade 12	University students
Cluster 1	16.00	42.00	40.00	2.00
Cluster 2	36.36	36.36	9.09	18.18
Cluster 3	23.33	50.00	23.33	3.33

Cluster 4	29.41	35.29	23.53	11.76
Sample	22.22	42.59	29.63	5.56

Table 9: Clusters' age composition (%)

(Source: Author)

	Hanoi	**HCMC**	**Others**
Cluster 1	60.00	10.00	30.00
Cluster 2	36.36	27.27	36.36
Cluster 3	76.67	10.00	13.33
Cluster 4	64.71	17.65	17.65
Sample	62.96	12.96	24.07

Table 10: Clusters' geographical location composition (%)

(Source: Author)

	No	**Yes**
Cluster 1	4.00	96.00
Cluster 2	18.18	81.82
Cluster 3	13.33	86.67
Cluster 4	5.88	94.12
Sample	8.33	91.67

Table 11: Cluster's response to if they plan to take the SAT in 1 year (%)

(Source: Author)

	Abroad	**Both**	**In Vietnam**
Cluster 1	14.00	50.00	36.00
Cluster 2	18.18	63.64	18.18
Cluster 3	6.67	66.67	26.67
Cluster 4	29.41	41.18	29.41
Sample	14.81	54.63	30.56

Table 12: Clusters' plan to use SAT score (%)

(Source: Author)

	Neither	**Official test**	**Practice test**
Cluster 1	26.00	38.00	36.00
Cluster 2	54.55	9.09	36.36
Cluster 3	36.67	20.00	43.33
Cluster 4	29.41	35.29	35.29
Sample	32.41	29.63	37.96

Table 13: Clusters' past experience with the test or practice test (%)

(Source: Author)

	1300-1390	**1400-1490**	**1500-1600**
Cluster 1	10.00	34.00	56.00
Cluster 2	0.00	45.45	54.55
Cluster 3	23.33	26.67	50.00
Cluster 4	23.53	35.29	41.18
Sample	14.81	33.33	51.85

Table 14: Clusters' targeted SAT score (%)

(Source: Author)

	Somewhat Uncomfortable	**Neutral**	**Somewhat Comfortable**	**Very Comfortable**
Cluster 1	6.00	32.00	22.00	40.00
Cluster 2	9.09	36.26	27.27	27.27
Cluster 3	0.00	20.00	30.00	50.00
Cluster 4	5.88	17.65	5.88	70.59
Sample	4.63	26.85	22.22	46.30

Table 15: Clusters' level of comfort with online learning (%)

(Source: Author)

	Online materials	**Books**	**Tutors**	**Centers**
Cluster 1	74.00	48.00	10.00	30.00
Cluster 2	90.91	45.45	9.09	18.18
Cluster 3	86.67	43.33	13.33	53.33
Cluster 4	88.24	52.94	11.76	23.53
Sample	81.48	47.22	11.11	34.26

Table 16: Clusters' past experience with other learning methods (%)

(Source: Author)

	Monthly payment	One-time payment
Cluster 1	50.00	50.00
Cluster 2	72.73	27.27
Cluster 3	50.00	50.00
Cluster 4	52.94	47.06
Sample	52.78	47.22

Table 17: Clusters' preference of payment method (%)

(Source: Author)

	Daily notifications	Vocabulary Flashcards	Full-length practice tests	Larger question bank
Cluster 1	8.00	6.00	58.00	24.00
Cluster 2	9.09	0.00	81.82	9.09
Cluster 3	3.33	3.33	83.33	6.67
Cluster 4	11.76	0.00	70.59	17.65
Sample	7.41	3.70	69.44	16.67

Table 18: Clusters' requested features to add value to the online course (%)

(Source: Author)

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Sample mean	Sample median
3.52	3.82	3.70	3.82	3.65	4

Table 19: Level of willingness to pay of sample and clusters

(Source: Author)

We notice that the characteristics of cluster 1, the standard customers, are not very different from that of the sample. Hence we will mainly focus on highlighting the distinctiveness of the remaining three clusters.

Psychographic profile

We can notice that the perfectionists are in much more of a rush than the practice-based and instruction-based students: 15% of clusters 2 and 3 do not plan to take the SAT in the near future, while only 5% of the perfectionists report so. They also have a clearer idea of their future, which is why most of them identify clearly where they would go to for university (in Vietnam or abroad). Meanwhile, cluster 2 and 3 still are open for more opportunities (65% of them stated that they want to use the SAT to apply for both Vietnamese and international universities). However, while the practice-focused students favor studying abroad and studying in Vietnam equally, instruction-focused students are clearly more inclined toward studying in Vietnam (27% want to study in Vietnam, compared to 7% who want to study abroad).

Surprisingly, the practice-focused students actually haven't had that much practice yet. More than half of them have never taken any test or practice test for the SAT, and only 9% have taken the official test before. Despite this, they set very ambitious goals: 100% aim to get a score of 1400 or more. On the other hand, it's the instruction-focused students who have the most experience with practice tests, with 43% having done at least one of these, compared to 36% of the practice-focused students. They also tend to set very ambitious goals — 50% want to get a score of 1500 or more. The perfectionists are the ones that have the most experience, with more than 70% having taken at least a test or a practice test before. Probably thanks to their experience that they tend to set more realistic goals than other clusters: 60% of them would be already content if they get a score between 1300 and 1500. They are also the segment that are most comfortable with online learning, with 71% reporting a high level of comfort.

Instruction-focused students have a notably good rate with 50% reporting a high level of comfort. Practice-focused students are the least comfortable with online learning, with only 27% reporting that they are very comfortable with the concept (along with a common explanation that their discomfort is caused by easy distraction and difficulty concentrating in an online environment). Regardless, they are most likely to have tinkered with online materials for SAT preparation: 91% have utilized some SAT prep online materials before, compared to only 74% of the standard customers have done this. However, they are less apt to try other methods, especially going to prep centers (only 18% have been to one of those, compared to 53% of instruction-based students).

Demographic profile

Overall, the demographic profile of the perfectionist segment is pretty close to that of the sample. However, cluster 2 and 3 are quite distinct demographically.

Regarding gender, the practice-focused segment has a much higher proportion of male while the instruction-focused cluster is the most female-dominant. The majority of cluster 2 is either Grade 10 highschool students or university freshmen, which is consistent with the aforementioned observation that they are quite new to the test (university freshmen here are usually those who are either unhappy with their choice and would like to change majors, or looking for an opportunity to study abroad). Many of those in cluster 3 are Grade 11 students (50%), which means that they have had some time to familiarize themselves with the test, but still have quite some time to prepare.

In addition, while instruction-focused students are predominantly based in Hanoi (77%), most of the practice-focused students live in Ho Chi Minh City or smaller cities. This suggests a preference for different learning styles related to geographical regions.

Willingness to pay

Generally, the median customers from the sample and from each cluster are all willing to spend from 6.000.000 VND to 8.000.000 VND for the described course. However, cluster 2, 3, and 4 are all willing to pay more for the course than the standard customers. This finding suggests an incentive to get to know the customers, their requirements, as well as their learning styles.

Of all the clusters, only the practice-focused customers express a clear preference for monthly payment over one-time payment. Their most demanded feature requests are for full-length practice tests, which is not surprising at all given their practice-oriented approach, and daily notification, which can be explained by their excuse for their discomfort studying online - distraction.

Full-length practice tests are also highly requested among all other segments, especially by instruction-focused learners (83% express that they would add the most value to their studying). Another feature that deserves some attention is the question bank, which is particularly requested by the perfectionists and the standard customers.

CHAPTER VI: DISCUSSIONS AND RECOMMENDATIONS

1. Summary of the study

1.1. Ordinal Logistic Regression

In this thesis, an ordinal logistic regression analysis was carried out to investigate the main factors influencing students' willingness to pay for online SAT preparation courses. The dependent variable, students' willingness to pay was categorized based on the maximum amount they are willing to pay for a described online SAT course, ranging from low willingness to pay to high willingness to pay. The ordinal logistic regression was significant, as the test of the full model against a model with only the intercept was statistically significant. This showed that at least one of the regression coefficients in the model is not equal to zero and that the predictors as a set reliably distinguished between customers of various willingness to pay. In addition, we have also tested and verified that the proportional odds assumption is satisfied, and that this model represents an adequately good fit to the data. This shows that the result of the present study can be generalized.

Several factors were initially considered as potential determinants of customers' willingness to pay for the course. However, as the result of the study points out, the students' future plan with the test score, their past experience with the test, their current score, their past experience with other learning methods including self-learning using online materials and taking extra classes at study centers, and their geographical location were the main factors influencing their willingness to pay for the described online SAT course.

The odd ratio for the student's future plan with university application using SAT score is greater than 1 for students who plan to use the SAT result to apply for both Vietnamese and international universities, compared to those who only use the SAT

score to apply to study abroad. For these students who apply for both, their perceived value of getting a high SAT score is higher, hence they are willing to pay more for the online course.

The odd ratios for the students' past experience with the exam are significant. Specifically, compared to those who have already taken the SAT before, the odds of being willing to pay more for the described online course is 2.8 times for those who have never taken the official SAT test but have taken at least one full-length practice test, and 110 times for those who have taken neither the official SAT test nor the official full-length practice test. This observation can be explained by the level of investment that the students have already put into their study: if they have already taken the test before, it means that they have already invested in their SAT preparation and hence, would not be willing to spend much more on it; conversely, if they have never taken any test or practice test, it means that they are new to the journey and still willing to invest in their education.

The odd ratios for the students' past experience with other learning methods including self-learning using online materials and taking extra classes at study centers are greater than 1, which is consistent with Kumar's (2010) research findings that having some prior knowledge or experience with a similar product or service can enhance future WTP.

For every 10 unit increase in a student's current SAT score, the odds of being willing to pay more for the described course is 3% higher, holding all variables constant. This result is unexpected: conventionally, students with a lower starting point should be willing to spend more money to improve their score, holding all others constant. Hence, further research may be required to understand the dynamics of this observation.

The odd ratio of the students' geographic location is significant and greater than 1. This means that those living in smaller cities have higher odds of being willing to pay more for online courses compared to those living in the capital, holding all other variables constant. This result can be explained by the lack of alternative options in smaller cities, compared to the capital, making the cost of substitutions higher, which in turn enhances the students' WTP.

1.2. Cluster Analysis

This thesis also conducted a cluster analysis that can provide insights for customer segmentation. By employing the Mahalanobis distance technique along with Ward's linkage clustering method, we were able to identify 4 unique segments, each characterized by their perceived value of various factors when considering purchasing an online course for SAT preparation.

The analysis revealed that the segment of the standard customers, which account for nearly 50% of the sample, has characteristics that are similar to the sample. Meanwhile, the other 3 smaller segments exhibit unique traits, both psychographically and demographically. We have also found that these three segments, on average, have a higher willingness to pay than the standard customers, which is a main reason why we should dive deeper into their distinctive features.

One of the smaller segments consists of individuals who demand the highest quality of the course across all factors. These are the students who are the most comfortable with online learning. They have a lot of experience with the test before, intend to take the test (again) in the near future, and they already have a clear idea of their plan to use the test score (whether to apply for university in Vietnam or abroad). They are also among the clusters who exhibit the highest willingness to pay for the course.

The remaining two segments are distinguished by their preference of learning style , and accordingly, are divided into practice-based and instruction-based students. Different from the perfectionists, both of these segments have less experience with the test, more time to prepare, and they tend to set more ambitious goals. They also seem to have not made up their mind regarding their future plan with the SAT score, which is why they tend to leave a lot of opportunities open.

However, there are a lot of differences between these segments. Demographically, the practice-based segment has a higher male ratio while the instruction-based one seems to be more female-dominant; their age and geographic location composition are also different.

Most of the instruction-based students are located in the capital, Hanoi, which may help explain why they are more comfortable with studying online and have a lot of experience with extra-class centers. By contrast, the practice-based students are predominantly based in smaller cities, which could be the reason explaining their low level of comfort with studying online. Despite this, they still tend to tinker a lot with online material for SAT preparation.

2. Recommendations

2.1. Based on the ordinal regression model

This model suggests several significant factors that may influence the student's willingness to pay for the course. Among all of them, there is no particular factor that a course provider can have direct control over. However, in case the course provider decides to increase their prices, they can achieve it by simply selecting its customers by asking questions regarding the students' past experiences with the SAT, with other learning methods, as well as their future plans with the exam result. The course provider can also manage their marketing strategy to target students in smaller cities

rather than those living in the capital, as this is also a significant factor that positively influences the students' maximum amount that they are willing to pay for the course.

2.2. Based on the cluster analysis

The conducted cluster analysis suggests that a course provider should tailor their marketing strategy based on their targeted customers. For example, if the company is trying to target the mass market (i.e., the standard customers), they should focus more on delivering the message about the teacher's teaching experience and the course's test bank, rather than trying to explain the course's responsibility regarding intellectual property rights to the customers. In this case, they will also need to expect that these customers are not willing to pay as much as those from other segments.

Conversely, the course provider can also decide which segment to target by analyzing their own strengths and weaknesses. For example, if the teacher has a high score in SAT and many years of teaching experience, they can use these strengths as the big idea in their marketing campaign to target the instruction-based students.

Another way to make use of this result is that when consulting to a potential customer, the course provider can pose questions to get to know more about the customer's expectations, wants, and needs, to identify if they can serve this customer, and if so, which feature they should try to promote to them.

CONCLUSION

The object of the study was to investigate the willingness of Vietnamese students to pay for online SAT prep courses and from that, provide insights into the pricing strategies most effective for online SAT prep courses in the Vietnamese market, considering student preferences and their willingness to pay for features that cater to the digital testing format. A survey was conducted to collect data from 110 Vietnamese students who are looking for an online course for SAT preparation.

Due to the ordinal nature of the dependent variable, an ordinal logistic regression was used in the analysis. The test result used to check if the present model with explanatory variables included is an improvement over the intercept only model was statistically significant at 5% level of significance. The ordinal logistic regression model used in the analysis satisfied all the necessary assumptions. Based on the literature, several factors were initially considered as potential determinants of the students' willingness to pay for the online SAT course described in the survey. However, the result of the empirical analysis revealed that, the students' future plan with the test score, their past experience with the test, their current score, their past experience with other learning methods including self-learning using online materials and taking extra classes at study centers, and their geographical location were the main factors influencing their willingness to pay for the described online SAT course.

Afterward, to determine if the sample base can be divided into different clusters of observation that have similar traits, a cluster analysis was carried out. By calculating various indices, it was concluded that a four-cluster model is optimal for this analysis. After identifying the distinctive characteristics of each cluster and naming them accordingly, we proceeded to profile each segment based on their demographic and psychographic data, as well as their willingness to pay for the online course.

REFERENCES

1. American Marketing Association(AMA) Dictionary, https://www.ama.org/topics/marketing-definition/

2. Aguilar, F. J. (1968). *Consumer Behavior: A Framework for Analysis*. New York: Random House.

3. Zeithaml, V. A., & Bitner, M. J. (2003). *Services Marketing: Integrating Customer Focus Across the Firm* (3rd ed.). New York, NY: McGraw-Hill.

4. Levitt, T. (1986). *The Marketing Imagination*. New York, NY: Free Press.

5. Stobierski T. (2020), *Willingness to pay: What it is & how to calculate*, available https://online.hbs.edu/blog/post/willingness-to-pay

6. Van Campenhout, M., & De Wulf, G. (2005), *Modeling willingness to pay for a quality differentiated product: A comparison of direct versus indirect questioning modes*, Journal of Economic Behavior & Organization, 57(2), 281-301.

7. Garba Aminu, Suleiman & Ahmad, Norzieiriani. (2018), *An Evaluation of the Effects of Brand Equity on Consumer Willingness to Pay Price Premium*, International Journal of Marketing Research Innovation. 2. 64-72. 10.46281/ijmri.v2i1.106.

8. Park, Cheol & Macinnis, Deborah & Priester, Joseph & Eisingerich, Andreas & Iacobucci, Dawn. (2010), *Brand Attachment and Brand Attitude Strength: Conceptual and Empirical Differentiation of Two Critical Brand Equity Drivers*, Journal of Marketing. 74. 10.2307/25764280.

9. Blattberg, Robert & Malthouse, Edward & Neslin, Scott. (2009), *Customer Lifetime Value: Empirical Generalizations and Some Conceptual Questions*, Journal of Interactive Marketing. 23. 157-168. 10.1016/j.intmar.2009.02.005.

10. Blackwell, Roger & Miniard, Paul & Engel, James & Rahman, Zillur. (2017), *Consumer Behavior*.

11. Alderson, Wroe (1965), *Dynamic Marketing Behavior: A Functionalist Theory of Marketing*, Homewood, IL: Irwin

12. Assael, Henry R., and A. Marvin Roscoe (1976), *Approaches to Market Segmentation Analysis*, Journal of Marketing 40 (October): 67–76.

13. Beales, Howard, Michael B. Mazis, Steven C. Salop, and Richard Staelin (1981), *Consumer Search and Public Policy*, Journal of Consumer Research 8 June, p11–22.

14. Cox, D.R. (1972), *Regression Models and Life Tables*, Journal of the Royal Statistical Society B34: 187–202.

15. Dewsnap, B. and Jobber, D. (2002), *A social psychological model of relations between marketing and sales*, European Journal of Marketing, Vol. 36 No. 7/8, pp. 874-894., available https://doi.org/10.1108/03090560210430854

16. Dominguez, Louis V., and Albert L. Page (1981) *Stratification in Consumer Behavior Research: A Re-Evaluatio*n, Journal of the Academy of Marketing Science 9 (Fall): 250–273.

17. Hawkins, D. I., & Mothersbaugh, D. L. (2010), *Consumer behavior: Building marketing strategy*, Boston: McGraw-Hill Irwin.

18. Imam, F. (2013), *Gender difference in impulsive buying behavior and post purchasing dissonance under incentive conditions*, Journal of business strategies, Volume 7,No 1, pp 23-29.

19. Kotler, P. and Armstrong, G. (2004) *Principles of Marketing*, 10th Edition, Pearson-Prentice Hall, New Jersey.

20. Khan, M. A. (2007), *Consumer Behaviour and Advertising Management* Delhi: New Age International.

21. SAT Nationally Representative and User Percentiles, https://research.collegeboard.org/reports/sat-suite/understanding-scores/sat

22. What Is a Good SAT Score?, https://satsuite.collegeboard.org/sat/scores/understanding-scores/what-is-a-good-score

23. Trinh N. H., Tran H. H. & Vuong Q. D. H. (2021), *Perceived Risk and Intention to Use Credit Cards*, available https://koreascience.kr/article/JAKO202109554061585.pdf

24. World Bank. (2020), *Global Financial Index Report*, retrieved December 12, 2020, from https://globalfindex.worldbank.org/

25. General Statistics Office of Vietnam (2020), *Data of Population and Employment*, retrieved December 12, 2020, from https://www.gso.gov.vn/en

26. Vietnam Investment Review (2022), *Financial titans seeking to prompt credit card adoption*, https://vir.com.vn/financial-titans-seeking-to-prompt-credit-card-adoption-96518.html

27. P. C. Mahalanobis (1936), *On the generalized distance in statistics*, proceedings of the National Institute of Sciences (Calcutta), 2, pp. 49–55.

28. G. M. Mimmack, S. Mason and J. Galpin (2001), *Choice of distance matrices in cluster analysis: defining regions*, Journal of Climate 14, 2790–2797

29. Charrad M., Ghazzali N., Boiteau V. & Niknafs A. (2015), *Package 'NbClust'*, https://cran.r-project.org/web/packages/NbClust/NbClust.pdf

30. Ordinal Logistic Regression | R Data Analysis Examples, UCLA: Statistical Consulting Group, available https://stats.oarc.ucla.edu/r/dae/ordinal-logistic-regression/

31. Agresti, A. (1996), *An Introduction to Categorical Data Analysis*, New York: John Wiley & Sons, Inc

32. Agresti, A. (2002), *Categorical Data Analysis*, Second Edition, Hoboken, New Jersey: John Wiley & Sons, Inc.

33. Harrell, F. E, (2001), *Regression Modeling Strategies*, New York: Springer-Verlag.

34. Liao, T. F. (1994), *Interpreting Probability Models: Logit, Probit, and Other Generalized Linear Models*, Thousand Oaks, CA: Sage Publications, Inc.

35. Powers, D. and Xie, Yu, *Statistical Methods for Categorical Data Analysis,* Bingley, UK: Emerald Group Publishing Limited.

36. Hosmer, David W., Lemesbow, Stanley (1980), *Goodness of fit tests for the multiple logistic regression model*, Communications in Statistics - Theory and Methods. 9 (10): 1043–1069. doi:10.1080/03610928008827941. ISSN 0361-0926.

37. Lemeshow, Stanley, Hosmer, David W. (1982). *A review of goodness of fit statistics for use in the development of logistic regression models*, American Journal of Epidemiology. 115 (1): 92–106. doi:10.1093/oxfordjournals.aje.a113284. ISSN 1476-6256

38. Hosmer, David W., Lemeshow, Stanley, Sturdivant, Rodney X. (2013). *Applied logistic regression. Wiley series in probability and statistics* (3 ed.). Hoboken, NJ: Wiley. ISBN 978-0-470-58247-3

39. Using Statistical Regression Methods in Education Research, https://www.restore.ac.uk/srme/www/fac/soc/wie/research-new/srme/modules/mod5/3/index.html

40. Smith, W. R. (1956). *Product differentiation and market segmentation as alternative marketing strategies.* Journal of Marketing, 21(1), 3-8.

41. Wedel, M., & Kamakura, W. A. (2000). *Market Segmentation: Conceptual and Methodological Foundations (2nd ed.).* Boston, MA: Kluwer Academic Publishers.

42. Kumar, V., Anand, A., & Song, H. (2010). *Future of retailer profitability: An organizing framework.* Journal of Retailing, 86(1), 6-18.

43. Lancaster, K. J. (1966). *A new approach to consumer theory.* Journal of Political Economy, 74(2), 132-157.

44. Garvin, D. A. (1987). *Competing on the eight dimensions of quality.* Harvard Business Review, 65(6), 101-109

45. Pascarella, E. T., & Terenzini, P. T. (2005). *How College Affects Students: A Third Decade of Research (Vol. 2).* San Francisco, CA: Jossey-Bass.

46. Chao, A., & Schor, J. B. (1998). *Empirical tests of status consumption: Evidence from women's cosmetics.* Journal of Economic Psychology, 19(1), 107-131.

47. Hossler, D., Schmit, J., & Vesper, N. (1999). *Going to College: How Social, Economic, and Educational Factors Influence the Decisions Students Make.* Baltimore, MD: Johns Hopkins University Press.

48. Granovetter, M. S. (1973). *The strength of weak ties.* American Journal of Sociology, 78(6), 1360-1380.

49. Coleman, J. S. (1988). *Social capital in the creation of human capital.* American Journal of Sociology, 94(S1), S95-S120.

50. Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values.* Beverly Hills, CA: Sage Publications.

51. Dynarski, S. (2002). *The consequences of merit aid.* National Bureau of Economic Research (No. w9400).

52. Varian, H. R. (1992). *Microeconomic Analysis* (3rd ed.). New York, NY: W.W. Norton & Company.

53. Kahneman, D., & Deaton, A. (2010). *High income improves evaluation of life but not emotional well-being.* Proceedings of the National Academy of Sciences, 107(38), 16489-16493.

54. Ang, S. H. (2001). *Crisis marketing: A comparison across economic scenarios.* International Business Review, 10(3), 263-284.

55. American Marketing Association (AMA). (n.d.). Customer segmentation, https://www.ama.org/marketing-news/how-segmentation-provides-the-roadmap-to-success/

56. Green, P. E., & Wind, Y. (1973). *Multiattribute decisions in marketing: A model and method.* Marketing Science, 2(3), 367-390.

57. Lilien, G. L., & Rangaswamy, A. (1998). *Marketing segmentation and marketing models.* Journal of Marketing, 62(1), 1-17.

58. Reichheld, F. F. (2006). *The one number you need to grow.* Harvard Business Review, 81(12), 46-55.

59. Solomon, M. R., Bamossy, G., & Askegaard, S. (2016). *Consumer behavior: Buying, having, and being* (12th ed.). Pearson Education Limited.

60. Dibb, S., & Simkin, L. (2001). *Marketing Briefs: A Revision and Study Guide* (2nd ed.). Oxford: Butterworth-Heinemann.

61. Coleman, J. (2010). *The New Science of Retailing: How Analytics are Transforming the Supply Chain and Improving Performance.* Hoboken, NJ: John Wiley & Sons.

62. Bourne, L. (2003). *Project Management for Flat Organizations: Cost Effective Steps to Achieving Successful Results.* Boca Raton, FL: CRC Press.

63. Kotler, P., & Keller, K. L. (2012). *Marketing Management* (14th ed.). Upper Saddle River, NJ: Prentice Hall.

64. Pine, B. J., & Gilmore, J. H. (1999). *The Experience Economy: Work is Theatre & Every Business a Stage.* Boston, MA: Harvard Business School Press.

65. Nijssen, E. J., & Frambach, R. T. (2000). *Creating Customer Value Through Strategic Marketing Planning: A Management Approach.* Boston, MA: Kluwer Academic Publishers.

66. McDonald, M., & Dunbar, I. (2012). *Market Segmentation: How to Do It, How to Profit from It* (4th ed.). Chichester, UK: John Wiley & Sons.

APPENDIX

Survey Questionnaire

I. Awareness & Interests

1. Are you aware of the SAT test and its importance in college admissions?

   ☐ Yes   ☐ No

2. Are you planning to take the SAT within 1 year from now?

   ☐ Yes   ☐ No

3. Do you plan to use your SAT score to apply for universities in Vietnam or to study abroad?

   ☐ Universities in Vietnam

   ☐ Study abroad

   ☐ Both

   ☐ Other

4. Have you ever taken the SAT or completed an official full-length practice test before?

   ☐ I have already taken the SAT and I want to retake it

   ☐ I have not taken the SAT but I have completed one of the official full-length practice tests

   ☐ I have not taken the SAT and I have not taken an official full-length practice test

5. If you have already taken either the SAT or an official full-length practice test, what is your current score? _____

6. What is your target score? (1600 is maximum)

☐ 1500-1600

☐ 1400-1490

☐ 1300-1390

☐ 1200-1290

☐ Under 1200

II. Preferred Learning Method

1. How comfortable are you with learning online?

☐ Very comfortable

☐ Somewhat comfortable

☐ Neutral

☐ Somewhat uncomfortable

☐ Very uncomfortable

2. Can you provide some explanation for your answer to the above question?

3. Which of the following methods to study for the SAT have you tried before? (Choose all that apply)

☐ Online materials (Khan Academy, Bluebook, ...)

☐ Books

☐ Tutor

☐ In-person extra classes

☐ Other

III. Value Perception

1. When considering online SAT prep courses, how important are these factors to you?

	Very Important	Somewhat Important	Neutral	Not Very Important	Not At All Important
Reviews from old students					
Question bank of the course					
Teacher's SAT score					
Teacher's teaching experience					
Tuition fee					
Intellectual Property Right (teacher creating his own question bank instead of using unlicensed contents)					

2. What is the maximum you would be willing to pay for a comprehensive online SAT prep course that offers:

    -   Video lecture

    -   Lecture notes

- Question bank created completely by the teacher, fully licensed

- Official test interface

- Additional reading materials

☐ More than 10.000.000 VND

☐ From 8.000.000 VND to under 10.000.000 VND

☐ From 6.000.000 VND to under 8.000.000 VND

☐ From 4.000.000 VND to under 6.000.000 VND

☐ Under 4.000.000 VND

3. Would you be interested in a subscription model for the course, with monthly payments instead of a one-time fee?

☐ Yes

☐ No

4. If yes, what is your preferred monthly price range?

5. Do you want to request a new feature for an online SAT prep course to help you study more effectively? (Choose the one you think would bring the most value to your study)

☐ Flashcards for vocabulary

☐ Full-length practice tests + Result reviews

☐ More questions in the question bank

☐ Daily notification to remind you to study

☐ Others

IV. Socio-demographic information

1. Gender: ☐ Male ☐ Female ☐ Other

2. You are in: ☐ Grade 10 ☐ Grade 11 ☐ Grade 12 ☐ Other

3. Which city are you living in?

   ☐ Hanoi ☐ Ho Chi Minh City ☐ Other

4. Do you have any other suggestions or feedback for us?