Profiling Consumers’ Online Shopping and Following Social Media Influencers Behaviors

Ming-Yi Wu¹
ORCID: https://orcid.org/0000-0002-4510-8651
E-mail: m.wu@northeastern.edu

¹Northeastern University, Boston, MA, USA

Abstract
In order to develop effective digital marketing strategies, this study explores how demographic factors may affect consumers' online shopping and following social media influencers behaviors. Market segmentation theory and consumer demographics theory provide the theoretical foundation for this study. By analyzing survey data collected from 6,034 U.S. adults with decision tree analysis, there are a number of significant findings. First, educational level, income, and age category are the important predictors for consumers’ buying things online using a desktop or laptop computer behavior. Second, age category, geographic location (urban, suburban, rural) are the predictors for consumers’ buying things online using a smartphone behavior. Third, educational level, age, and geographic location are the predictors for consumers’ preferring online shopping over in-store shopping behavior. Fourth, age category and race-ethnicity are the predictors for following social media influencers behavior. Fifth, gender and age category are the predictors for purchasing something after seeing an influencer’s posts behavior. Finally, age category and gender are the predictors for purchase decisions getting impacted by influencers. The results provide valuable insights about consumer behaviors online, market segmentation, and influencer marketing strategies.

Keywords: Consumer Profiling; Consumer Demographics; Market Segmentation; Social Media Influencer; Decision Tree.

This work is licensed under a Creative Commons Attribution 4.0 International License: https://creativecommons.org/licenses/by/4.0/
INTRODUCTION

In the last few years, e-commerce has become an integral part of global retail. There was a surge in e-commerce amid the COVID-19 pandemic. Some of the pandemic-inspired consumption trends, such as online shopping, online grocery shopping, and work from home will continue as the new normal as we enter the post-pandemic era (Wu, 2022c). Thus, the market of e-commerce will continuously grow. As Laudon and Traver (2024) noted, “in 2022, almost 215 million U.S. consumers are expected to spend about $1.3 trillion, and businesses about $8.5 trillion, purchasing products and services via a desktop/laptop computer, mobile device, or smart speaker. A similar story has occurred throughout the world” (p. 6). Therefore, organizations which sell products and services online want to know who are the e-commerce consumers and how they behave in the online environment. In order to develop effective marketing strategies (e.g., market segmentation strategies, influencer marketing strategies), e-commerce companies can profile their online consumers’ behaviors.

Today’s e-commerce environment has become a social commerce environment. Therefore, it’s important to profile not only consumers’ online shopping behaviors, but also following social media influencers behaviors. As Vilkaite-Vaitone (2024) noted, contemporary brands are shifting away from traditional advertising methods and turning towards social media influencers as a means of promoting their products to their audiences because of their higher levels of authenticity and credibility. For example, social media influencers can affect consumers’ green/sustainable product purchase intentions because of their perceived importance and credibility among their followers. Brands can partner with the right social media influencers to encourage support for green initiatives and promote green products (Zhao et al., 2024). Because influencer marketing has recently attracted scholarly attention (e.g., Vilkaite-Vaitone, 2024; Zhao et al., 2024), how consumers’ purchase decisions and buying behaviors are affected by social media influencers can be further examined.

Although a number of previous studies (e.g., Gatziolis et al., 2022; Ketipov et al., 2023; Tawang et al., 2008; Vilkaite-Vaitone, 2024; Wu, 2022c; Zhao et al., 2024) were conducted to explore consumer behaviors in the e-commerce/social commerce environment, consumer behaviors are still evolving in the post-pandemic era. To contribute to scholarly literature with updated findings and develop effective digital marketing strategies, this study aims to (1) profile consumer’s online shopping behaviors, (2) explore the factors that affect consumers’ following social media influencers’ behaviors, and (3) examine how consumers’ purchase decisions and buying behaviors are affected by influencers in the social commerce environment. Next section of this article reviews literature about consumer profiling in e-commerce, social media influencers’ impact on consumer behaviors, market segmentation theory, and consumer demographics theory.

LITERATURE REVIEW

Consumer profiling in e-commerce

Consumer/customer profiling is a technique which is frequently used by market researchers. In general, organizations can profile the consumers in the market. In specific, organizations can profile their customers who purchase their products or use their services. As Wu (2018) noted, consumer relations and customer relations are closely related concepts in the e-commerce environment, because of the nature of online shopping. Different from in-store shopping, on-line consumers are facing higher level of uncertainties. “Either online consumers in general or customers in specific cannot physically see the products when making purchases online. In addition, online shoppers need to pay first and wait for the products being delivered to them. Therefore, it is very important for online vendors to reduce uncertainties and build trust with consumers and customers” (p. 34). Previous e-commerce studies (e.g., Beatty et al., 2011; Wu, 2018) use the terms, consumer and customer, interchangeably. Therefore, both terms, consumer profiling and customer profiling, are used interchangeably in this article.

According to Maya (2021), “customer profiling is the process of defining who your customers are in order to understand their needs or wants better. This information can then be used to make customer-specific marketing decisions to increase customer satisfaction and loyalty. A customer profile
is basically a description of your ideal customer or customer segment” (Para 1). The customer profile usually includes information about customer demographics, customer activities (e.g., purchase and service histories), product preferences, and customer retention history. Consumer/customer profiling can be done by using various research methods, such as demographic analysis, psychographic analysis, customer surveys, and focus group interviews. All in all, the purposes for consumer/customer profiling are helping organizations, including e-commerce companies, to understand who are their customers, what are their customers’ needs and wants, and how to reach their customers, how to do market segmentation effectively, and how to customize or personalize their products and services.

Consumer/customer profiling is extremely important for e-commerce companies. As Laudon and Travor (2024) noted, “the first principle of marketing sales is 'know the customers who is online, who shops online and why, and what do they buy” (p. 319). Because it is critical for e-commerce companies to know how demographic and psychological factors may affect consumer behaviors online, a number of studies about consumer/customer profiling in e-commerce were conducted previously (e.g., Gatziolis et al., 2022; Ketipov et al., 2023; Mohamed et al., 2022; Ndawanga et al., 2008; Wu, 2022b; 2023; Yadav et al., 2012). For example, Wu (2022c) analyzed consumers’ spending more time shopping online behaviors and online grocery shopping behaviors amid the COVID-19 pandemic by analyzing survey data collected from 857 US adults in 2020.

The results of Wu’s study suggested that perceived threat/concern is the most influential variable for predicting consumers’ overall online shopping behaviors, followed by annual family income and age. Overall, consumers who have higher level of perceived threat/concern, higher family income, and younger are more likely to shop online. The results also suggested that education is the most influential variable for predicting consumers’ online grocery shopping behaviors, followed by perceived threat/concern and race. Overall, consumers with higher educational level, higher concern/perceived threat, and White and Asian Americans are more likely to order grocery online. Wu (2023) analyzed U.S. consumers’ (N = 1,033) choice of consumption channels during the business re-opening period amid the pandemic and found that perceived risk for food delivery and take out is the most important factor that predicts consumers’ ordering food delivery and takeout behaviors. Ketipov et al. (2023) analyzed how personality traits may predict consumers’ user behavior in e-commerce. The results of decision tree and random forest analyses suggested that there is a significant relationship between key personality traits (e.g., extraversion, agreeableness, emotional stability) and specific online shopping activities. For example, more extroverted individuals would react positively if they gain additional articles or accessories appropriate to the already selected product...For these individuals, the ability to access and contribute to user comments is considered essential in their purchase decision-making process” (p. 97).

In summary, various factors (e.g., demographics, psychological factors, social influence) can affect online consumers’ decision-making process and purchasing behaviors. Recent studies (e.g., Needle, 2023; Ruvio and Iacobucci, 2023) suggested that social media influencers have significant impacts on consumer behaviors in today’s e-commerce/social commerce environment. Next section of this paper reviews literature in this area.

Social media influencers’ impacts on consumer behaviors

In today’s e-commerce/social commerce environment, consumers would rather to trust other consumers’ opinions, reviews, and recommendations online than companies’ marketing communications. Previous studies (e.g., Vilkaitė-Vaitone, 2024; Zhao et al., 2024; Wu, 2022b) suggested that social media influencers have become important mediators in consumers’ decision-making process, particularly for younger consumers (e.g, Needle, 2023; Reinkikainen et al., 2020; Saima & Khan, 2021). As Wu (2022b) noted, organizations can partner with social media influencers and have them serve as their brand ambassadors and foster positive electronic word of mouth (eWOM) in the e-commerce environment. According to Ruvio and Iacobucci (2023), influencers are “consumers who have the ability to influence the opinions and behaviors of other consumers” (p. 288). They are the opinion leaders or experts in a specific consumption domain. There are different types of influencers, such as meg influencer, macro-influencer, mid-influencers, and micro-influencers, depending on number of
followers on social media platforms, such as Instagram, YouTube, and TikTok. In green/sustainable consumption literature, influencers can be categorized into two categories, informers and entertainers. Zhao et al.'s (2024) research findings suggested that informers have a powerful green endorsement effect than entertainers.

Today's consumers are more likely to trust peer recommendations than brand ads, and social media influencers can be very powerful in this respect. As Baker (2020) noted, social media influencers can impact consumers' purchase decisions and play a pivotal role in marketing for a variety of brands. For example, organizations can partner with social media influencers to build brand awareness, promote environmentally friendly actions and sustainable lifestyles, endorse their products and services, strengthen customer relationships, and improve buying decisions with unbiased opinions. Thus, organizations need to work with the “right” influencers on the right social media platforms based on business/marketing goals, influencer types, and consumer demographics. Although multiple factors, such as cultural (e.g., Zhou et al., 2021), psychological (see Vilkaite-Vaitone, 2024; Zhao et al., 2024), and demographic factors (Shieber, 2020; Wu, 2022b) can affect the effectiveness of influencer marketing, many organizations develop their strategies based on consumer demographics because young consumers’ (e.g., Generation Y, Z) purchase decisions are affected by social media influencers (Wu, 2022b).

Previous studies have suggested that there are generational differences in information and communication technologies (ICTs) use, social media use (e.g., Fietkiewicz et al., 2016; Wu, 2022a) and following social media influencers behaviors (e.g., Hazari & Sethna, 2023; Needle, 2023). For example, Fietkiewicz et al. (2016) compared how different generations (e.g., Generation X vs. Generation Y vs. Generation Z) use different social media platforms by surveying 373 participants and found that these three generations have different preferences for social media platforms. Specifically, Generation Z are more likely to use Instagram, an online photo- and video-sharing platform, than older generations. Wu (2022a) compared how five generations (e.g., Generation X, Y, Z, Boomer, Silent Generation) use ICTs and social media in the U.S. by using the digital divide theory (e.g., Friemel, 2016; Pearce & Rice, 2013; 2014; Rice & Pearce, 2015) which describes the socioeconomic gap between those with and without access and usage of ICTs (e.g., computers, Internet, email, smartphone, social media) as a guiding theory. The author found that the oldest generation, the silent generation, is left behind other generations (e.g., Generation X, Y, Z, Boomer) in terms of ICT ownership and use and social media use (e.g., ever use social media). Thus, the digital divide/grey divide still exists. In addition, different generations have different preferences for social media platforms. For example, more Generation Z uses Snapchat, Twitter, and Instagram.

Shieber (2020) noted that 70% of Generation Y (born between 1981 and 1996) and Z (born between 1997 and 2012) consumers learn about products they are interested in buying on social media, while 56% have purchased a product after seeing a post from someone they follow. As Needle (2023) noted, influencer marketing has high return on investment (ROI) because Generation Zers trust influencers. Needle’s research findings suggested that 33% of Gen Zers have bought a product based on an influencer’s recommendation in the last three months. In terms of social media platforms, Instagram and TikTok are the most preferred platforms. According to Santora (2023), 72% of marketers use Instagram for influencer marketing. The popularity of TikTok is also growing, as 61% marketers are using it. As Hazari and Sethna (2023) noted, Instagram is the most popular platform for brands to utilize the influencer marketing strategies, such as lifestyle marketing and brand influencer advertising to engage with Generation Zers.

In summary, with appropriate consumer profiling, e-commerce companies can develop effective market segmentation and influencer marketing strategies. Market segmentation theory and consumer demographics theory provide the theoretical foundation for this study.

Market segmentation theory

Market segmentation theory (e.g., Kotler, 1989; 1999) is frequently cited in market research literature (e.g., Mothersbaugh et al., 2020; Ruvio & Iacobucci, 2023; Smith & Albaum, 2005) and used by researchers. According to Ruvio and Iacobucci (2023), marketing strategies involve a 3-step process:
segmentation, targeting, and positioning (STP). Segmentation, the first step, is defined as “the division of the broader consumer market into subgroups of consumers based on shared characteristics or needs” (p. 23). As Kotler (1989) noted, market segmentation can be used as an analytical act that precedes the development of marketing strategies. Instead of viewing the market as a mass market, marketers can create differentiation strategies for the segmented markets, the micromarkets, and the individual markets of just one customer, because the market is not homogeneous. Kotler (1999) further argued that many markets can be broken down into different segments. In addition, any market can be segmented in several different ways, such as demographic segmentation, benefit segmentation, occasion segmentation, usage level segmentation, and lifestyle segmentation.

According to Kotler, “demographic segmentation means grouping people who share a common demographic makeup: affluent senior citizens, young low-income minorities, and so on” (p. 26). Similarly, Ruviu and Iacobucci proposed four market segmentation bases, including geographic (e.g., region, country, climate, density, postal code), demographics (e.g., age, gender, income, education), psychographics (e.g., lifestyle, activities, interest, personality, values, attitudes), and behavioral (e.g., benefit thoughts, usage, occasion, buyer stage, involvement, loyalty level) segmentation. Among these methods, many marketers use demographic segmentation because it provides the basic information about the consumers. In addition, demographics are easy to obtain and are predictable over time (see Smith & Albaum, 2005; Ruviu & Iacobucci, 2023).

**Consumer demographics theory**

Consumer demographic theory (e.g., Sheth, 1977; Martins & Brooks, 2010; Martins et al., 2012) provides another theoretical foundation for this study. Martins and Brooks (2010) identified several perspectives that may explain consumer behaviors, such as the economic, psychological, sociological, and demographic perspectives. They argued that the demographic perspective is very important because “demographic events are often triggers for the consumption of goods and services during the life cycle” (p. 87). The basic assumption of consumer demographic perspective/theory is that demographic variables, such as gender, age/generation, income, and education, affect consumers’ perceptions, motivations, decision making process, lifestyles, behaviors, and the nature of consumer markets (Martins et al., 2012). Some researchers (e.g., Mothersbaugh et al., 2020) also include a geographic variable (e.g., urban, suburban, second city, and rural) for geo-demographic analysis.

Although some studies (Gatziolis et al., 2022; Ketipov et al., 2023; Wu, 2022c; 2023) have been done to profile consumer behaviors online, consumer behaviors are still evolving in the post-pandemic era. For example, consumers can choose to shop online or shop in retail stores as we enter into the new normal. Thus, updated insights about consumer behaviors online should be provided. To build on previous studies (Ketipov et al., 2023; Mothersbaugh et al., 2020; Wu, 2022c; 2023) and help e-commerce companies to understand online consumers, this study aims to profile consumers’ online shopping and following social media influencers behaviors by following Mothersbaugh et al.’s (2020) geo-demographic analysis approach with predictive analytics. By doing so, we can better understand what are the most important factors that may affect consumers’ online shopping behaviors and following social media influencers behaviors.

**Research questions**

Based on the literature review, six research questions guide this study. What are the demographic factors that predict U.S. consumers’...
METHOD

Procedure

The results of this study are based on Pew Research Center's American Trends Panel (ATP) Survey data (Pew Research Center, 2022). The data of this study was collected from web-based surveys. This survey was conducted in July (July 5 – July 17), 2022 in the United States. The data collection point was chosen for analysis because it's noteworthy to track consumers’ online shopping behaviors in the new normal. With the availability of the COVID-19 vaccines, all of the retail stores have been reopened over two years since the outbreak of the COVID-19 pandemic (as of July, 2022). Thus, consumers have the choice to either shop online or shop in the physical stores. It would be interesting to analyze consumers’ online shopping behaviors with data collected at this point of time.

Samples

Participants were 6,034 U.S. adults, including 2,800 (47.0%) male and 3,157 (53.0%) female. Participants’ age categories include 1,071 (17.8%) 18-29, 2,077 (34.5%) 30-49, 1,546 (25.6%) 50-64, and 1,322 (22.0%) 65+ years old. Participants’ ethnicities include 3,874 (65.1%) White, 716 (12.0%) Black, 860 (14.4%) Hispanic, 329 (5.5%) Asian, and 171 (2.9%) other races. Respondents reported diverse educational levels: 2,117 (35.2%) have high school degree or less, 1,859 (30.9%) have some college education, 2,037 (33.9%) have college degree or above. Respondents' self-reported residential areas include: 1,326 (22.1%) urban, 3,086 (51.3%) suburban, and 1,599 (26.6%) rural. The data was weighted to reflect the demographic distribution in the United States.

Data analysis

To answer the research questions, decision tree analysis with the CRT method was conducted in SPSS, version 28. To identify the top predictors, the maximum tree depth is set at 3. As Strickland (2014) noted, the goal of decision tree “is to create a model that predicts the value of a target variable based on several input variables” (p. 87). Comparing with traditional classification methods, such as logistic regression, the advantage of decision tree analysis is that the results are easier to interpret visually. In addition, the machine learning algorithm can clearly identify the most important predictor for the outcome variable in the model (IBM, 2010). With these unique strengths, decision tree analysis can be used as a powerful data mining tool for market research (see Topal, 2019; Wu, 2022; 2023).

The independent variables/predictors for the decision tree analysis include gender, age category, race, education, income tier (e.g., low income, middle income, upper income), and geographic area/USR (e.g., urban, suburban, rural). The dependent variables for decision analyses are categorical variables which measure consumers’ online shopping and social media use behaviors.

RESULTS AND DISCUSSION

RQ1: Online Shopping Using a Desktop or Laptop Computer Behavior

The result of decision tree analysis suggests that educational level is the most influential variable for predicting consumers’ online shopping using a desktop or laptop computer behavior, followed by annual family income tier and age category. Among the participants (N = 5,266), 71.5% (n = 3,763) ever purchased things online using a desktop or laptop computer, whereas 28.5% (n = 1,503) didn’t. The first split suggests that participants are categorized into two Nodes, with Node 1 (college graduate+) and Node 2 (some College, high school graduate or less) based on educational level. Among the participants with some College or high school degree or less (Node 2), those with higher family income (middle income, higher income tier) are more likely to shop online than those with lower family income. Among the lower income group (Node 3), age category is the splitting variable suggesting that the youngest age group (Node 5, 18-29) and the oldest age group (Node 6, 65+) are more likely to shop online by using a desktop or laptop computer. The prediction accuracy rate is 72.7%%. Figure 1 summarizes the decision tree analysis results.
In general, the results are similar to Wu (2022c) because both studies suggest that consumers with higher educational level and income are more likely to shop online. The result suggesting that income is a factor which affects consumers’ ever shopping online using a computer behavior is also consistent with previous research findings (e.g., Favrio & Anderson, 2022; Wu, 2022c). As Favrio and Anderson (2022) noted, online shopping behaviors vary by household income. U.S. adults with higher family income are more likely to use a computer to shop online than those with lower income. This result can be explained by device divide (Rice & Katz, 2003; Rice et al., 2023). As Rice and Katz (2003) found, the distinctions among PC-based Internet usage and mobile phone usage are primarily influenced by income and education. As Rice et al. (2023) noted, “in the early years of mobile Internet, PCs were expensive and required considerable infrastructure, while ‘dumb’ mobile phones were fairly affordable, portable, and did not require an Internet connection” (p.2). Now, the device divide is gradually withering. However, it does not completely go away. It explains why consumers with higher income and educational levels are more likely to own a computer and have broadband Internet connection at home and are more likely to buy things online using a desktop or laptop computer.

RQ2: Online Shopping Using a Smartphone Behavior

The result of decision tree analysis suggests that age category is the most influential variable for predicting consumers’ online shopping using a smartphone behavior. Other predictors are geographic area/USR and race-ethnicity. Among the participants (N = 5,292), 79.3% (n = 4,194) ever purchased things online using a smartphone, whereas 20.7% (n = 1,098) didn’t. The first split suggests that...
participants are categorized into two Nodes, with Node 1 (18-29; 30-49) and Node 2 (50-64; 65+) based on age category. Among the participants who are younger (Node 1), those who live in suburban and rural area (Node 4) are more likely to shop online using a smartphone than those who live in the urban area (Node 3).

Among Node 2 (50-64; 65+), participants aged 50-64 (Node 5) are more likely to shop online by using a smartphone than participants who are 65+ (Node 6). Among Node 4, race-ethnicity is the splitting variable. The percentage for White and Black adults (Node 9) to shop online using a smartphone is slightly higher than Hispanic and Asian adults (Node 10). Among participants aged 50-64 (Node 5), those who live in urban areas (Node 11) are more likely to shop online using a smartphone than those who live in rural or suburban areas (Node 12). Among participants aged 65+ (Node 6), race-ethnicity is the splitting variable. Black and Asian Americans (Node 14) are more likely than White and Hispanic Americans to use a smartphone to shop. The prediction accuracy rate is 79.4%. Figure 2 summarizes the decision tree analysis results.

Figure 2. Decision Tree Predicting Consumers' Online Shopping Using a Smartphone Behavior

The result that participants' geographic area/USR is a predictor for their using smartphone to shop behavior varies by age category is interesting. For those who are younger (18-29; 30-49), those who live in the suburban and rural area are more likely to shop online using a smartphone than those who live in urban areas. This result may be explained by young consumers’ convenience motives for shopping online. Wu’s (2018) focus group interview results suggested that some young consumers, college students, chose to shop online because there is no store around where they live. Thus, it’s more convenient for consumers to shop online using a smartphone than driving long distance to purchase what they want. There are less physical stores for some brands in rural and suburban areas than urban areas. As Wu (2022a) found, the percentages for young consumers, Generation Yers and Zers, to own a smartphone are higher than the older generations. If there is no store around, young consumers can use their smartphones to order products online. However, the result related to geographic area is the opposite for those who are older (50-64).

In the 50-64 age group, those who live in the urban area are more likely to be mobile shoppers than those who live in the suburban and rural area. The result related to the older age group may be explained by digital divide, which is a combination of age/generational divide and urban vs. rural divide. As Wu (2022a) noted, generational digital divide, grey divide, still exists in the U.S. Older generations
(e.g., Boomers and Silent Generation, are less likely to own a smartphone than the younger generations (e.g., Generation Y and Z). Vogels (2021) noted the rural vs. urban vs. suburban digital divide persists in the United States. For example, rural (80%) and suburban (84%) adults are less likely than urban (89%) adults to own a smartphone. The combination of generational divide/grey divide and rural vs. urban divide may explain why older consumers (50 - 64) who live in the rural or suburban areas are less likely to shop online by using a smartphone.

Race-ethnicity is a splitting variable for (1) young consumers who live in the rural and suburban area (Node 4) and (2) those who are 65+ (Node 6). However, how different ethnic groups are categorized are different in Node 4 and Node 6. Previous studies (e.g., Atske & Perrin, 2021; Bartikowski et al., 2018) suggested that there is a device divide/mobile only divide among different ethnic groups. Atske and Perrin (2021) found that home broadband adoption, computer ownership varies by race and ethnicity in the U.S. The percentage for White Americans to own a computer and have home broadband connection is higher than Black and Hispanic adults. For example, 80% White adults owned a desktop or laptop computer, compared with 69% of Black adults and 67% of Hispanic adults. Similarly, 80% White adults have home broadband connection, compared with 71% of Black and 65% of Hispanic adults. Nevertheless, there is "no" statistical difference in smartphone ownership in smartphone ownership among White (85%), Black (83%), and Hispanic (85%) adults.

As Bartikowski et al. (2018) noted, there is a mobile-only Internet access divide between ethnic minority and majority consumers, because the percentage for minority ethnic groups (e.g., Black and Hispanic American) to have broadband Internet connection and computers are lower than the majority (e.g., White). Therefore, ethnic minority may only have high speed Internet connection via smartphones, instead of broadband connect at home. Thus, ethnic minority consumers tend to use smartphone apps. (e.g., email, social networking, or listening to music) more frequently than majority consumers, and more likely use smartphones for purchasing online. The device divide/mobile only divide may explain why more Black Americans in the 65+ group (Node 14) shop online using smart phone than White (Node 13), because the result is consistent with Bartikowski et al. However, the mobile-only digital divide cannot explain why Asian and Black consumers are categorized together in Node 14 and White and Hispanic Americans are categorized together in Node 13. Previous studies (e.g., Atske & Perrin, 2021; Bartikowski et al., 2018) didn't indicate Asian Americans have less access to computers or home broadband than White Americans.

RQ3: Online Shopping vs. In-Store Shopping

The result of decision tree analysis suggests that educational level is the most influential variable for predicting consumers’ preferring online shopping over in-store shopping behavior, followed by age category and geographic area/USR. Among the participants (N = 5,296), 39.7% (n = 2,105) prefer to buy online, whereas 60.3% (n = 3,191) prefer to buy from a physical store. The first split suggests that participants are categorized in to two Nodes, with Node 1 (college graduate+, some college) and Node 2 (high school graduate or less) based on educational level. Among the participants with higher educational level (Node 1), those in age group 30-49 (Node 3) are more likely to prefer to shop online than those in other age groups (Node 4). Among the 30-49 age group (Node 3) who prefer to shop online, those who live in the urban and suburban areas (Node 5) are more likely to prefer to shop online than those who live in the rural areas (Node 6). The prediction accuracy rate is 61.5%. Figure 3 summarizes the decision tree analysis results.

The result that about 4 out of 10 (39.7%) participants would prefer online shopping over in-store shopping is similar to Tighe (2023) which suggested that 43% of U.S. said they would prefer to shop mostly online rather than in-store, making it the country with highest online shopping preference as of early 2023. The percentage of preferring online shopping is higher in the U.S. than in other countries, such as Australia, Finland, and New Zealand. Generally speaking, consumers with higher educational levels and in the mid-age group (30 - 49) would prefer online shopping over in-store shopping. The results that more mid-aged consumers who live in the urban and suburban areas would prefer online shopping than those who live in the rural area may be explained by rural vs. urban digital divide. As Vogels (2021) noted, rural Americans have consistently lower levels of technology ownership
(e.g., home broadband, smartphone, tablet, desktop/laptop computer) than urbanities and suburbanites.

In addition, rural residents go online less frequently than urban residents. This finding may also be explained by Skrovan’s (2017) research finding which suggested that more rural shoppers would prefer in-store shopping over online shopping than urban residents because they would like to see, touch, feel and try out items in stores. As the author noted, rural consumers often need to drive long distances to shop and may consider in-store shopping an event and a time investment, so they want to see and try on the products to make sure to get the items right. How geographic location may affect consumers’ preference for shopping channels (online vs. in-store) may vary by age groups, device type (computer vs. mobile), and consumers’ motivations (e.g., convenience motive, economic motive, safety motive). Geographic location alone only has modest influence on consumer behaviors online.

**RQ4: Following Social Media Influencers Behavior**

The result of decision tree analysis suggests that age category is the most influential variable for predicting consumers’ following social media influencers behavior, followed by race-ethnicty. Among the participants (N = 3,694), 43.6% (n = 1,610) follow influencers or content creators on social media, whereas 56.4% (n = 2,084) do not. The first split suggests that participants are categorized in to two Nodes, with Node 1 (18-29; 30-49) and Node 2 (50-64; 65+) based on age category. Among the participants in the younger age categories (Node 1), those in 18-29 (Node 3) are more likely to follow influencers than those in 30-49 age group (Node 4). Among the 30-49 age group (Node 4), race-ethnicity

**Figure 3. Decision Tree Predicting Consumers’ Preference for Online vs. In-Store Shopping Behavior**

The result of decision tree analysis suggests that age category is the most influential variable for predicting consumers’ following social media influencers behavior, followed by race-ethnicty. Among the participants (N = 3,694), 43.6% (n = 1,610) follow influencers or content creators on social media, whereas 56.4% (n = 2,084) do not. The first split suggests that participants are categorized in to two Nodes, with Node 1 (18-29; 30-49) and Node 2 (50-64; 65+) based on age category. Among the participants in the younger age categories (Node 1), those in 18-29 (Node 3) are more likely to follow influencers than those in 30-49 age group (Node 4). Among the 30-49 age group (Node 4), race-ethnicity
is the splitting variable. Black, Hispanic, and Asian (Node 6) are more likely to follow influencers or content creators than White (Node 5). The prediction accuracy rate is 70.6%. Figure 4 summarizes the results.

Figure 4. Decision Tree Predicting Consumers’ Following Influencers Behavior

The result that younger consumers are more likely to follow social media influencers is consistent with previous studies (e.g., Dopson, 2023). As Dopson (2023) noted, the number of people active on social media differs by generation. Therefore, certain demographics are more attuned to influencer recommendations than others. For example, Generation Zers are most influenced by social media influencers. The percentage for older generations to follow influencers is much lower. For example, only 9% of Boomers follow social media influencers, compared to 47% of Gen Z. Thus, organizations’ influencer marketing strategies may work better for younger generations. However, the result regarding gender difference is different from Dopson’s argument that more men (95%) than women (93%) follow social media influencers. However, gender is not a significant predictor. Instead, race-ethnicity is a significant predictor for following influencers. Generally speaking, the percentage for ethnicity minorities (e.g., Asian, Black, Hispanic Americans) to follow influencers are higher than White Americans. This result is consistent with Faverio and Anderson’s (2022) research finding suggesting that there are race and ethnicity differences in following influencers. About 6 in 10 (59%) Hispanic social media users follow influencers, compared with 44% of Black users and a third of White users.
RQ5: Purchase Decisions Getting Impacted by Influencers

The result suggests that age category is the most influential variable for predicting consumers' purchase decisions getting impacted by influencers a lot, a little, or not at all, followed by gender. Among the participants (N = 4,006), 2.9% (n = 116) participants’ purchasing decisions are impacted by influencers a lot; 35.8% (1,436) are impacted by influencers a little; 61.3% (2,454) are not impacted at all. It implies that about 4 out of 10 (38.7%) of participants’ purchasing decisions are impacted by social media influencers to certain extent (a lot and a little combined). Among all participants, age category is the most important predictor. More younger participants', 18-29 and 30-49 (Node 1), purchase decisions are impacted by influencers to some extent (a lot or a little) than those in the older age group, 50-64 and 65+ (Node 2). Among Node 1, gender is the splitting variable. More women’s (Node 3) purchasing decisions are impacted by influencers than men’s (Node 4). Among women's group (Node 3), more youngest participants', 18-29 (Node 5), purchase decisions are impacted by influencers than those in the 30-49 age group. Prediction accuracy rate for this model is 63.3%. Figure 5 summarizes the results.

Figure 5. Decision Tree Predicting Consumers' Purchase Decisions Getting Impacted by Social Media Influencers
The result suggests that majority (60.9%) of young females’ (18-29) purchasing decisions are affected by influencers to certain extent (a lot and a little combined). The high percentage (60.9%) implies that social media influencers’ posts do have significant impacts on consumers’ purchase decisions in the young females’ consumer segment. This result is consistent with previous studies (e.g., Faverio & Anderson, 2022; Santora, 2023). As Faverio and Aderson (2023) noted, younger social media users are more likely to say that influencers affect their purchasing habits a lot or a little. This is particular common for young women (18 – 29) to say that influencers or content creators affect what they purchase at least a little.

RQ6: Ever Purchased Something After Seeing Influencers Post About It

The result suggests that gender is the most influential variable for predicting consumers’ actual purchase behavior, followed by age category. Among the participants (N = 3,728), 32.2% (n = 1,202) participants ever purchased something after seeing a social media influencer or content creator post about it; 67.8% (n = 2,526) didn’t. The first split suggests that participants are categorized in to two Nodes, with Node 1 (man) and Node 2 (women) based on gender. More women (38.8%) than men (23.3%) have done this. Among women (Node 1), more younger participants, 18-29 and 30-49 (Node 3), than older participants, 50-64 and 65+ (Node 4), have done so. Among Node 3, age category serves as the splitting variable again. More participants in the 18-29 age group (Node 5) ever purchased something than those in the 30-49 age group (Node 6). The prediction accuracy rate is 68.7%. Figure 6 summarizes the results.

Figure 6. Decision Tree Predicting Consumers’ Ever Purchasing Something After Seeing an Influencer Post About It on Social Media Behavior
The results that younger participants are more likely to purchase products which are posted by social media influencers are consistent with previous research findings (e.g., Needle, 2023; Santora, 2023). In this study, the youngest age group (18-29) includes Generation Zers and younger Generation Yers. Interesting, majority (53.9%) of the female participants in the 18-29 age group have done so. As Santora (2023) noted, Generation Z is most likely to be influenced by influencer recommendations. Generation Zers prefer to discover new products via influencers. About 32% Generation Z social media users made a purchase as a result of influencers’ recommendations. Therefore, organizations which sell products to female participants in younger generations, especially Generation Zers, may utilize influencer marketing strategies.

CONCLUSIONS
Significances and Contributions

By analyzing survey data collected from 6,034 U.S. adults with decision tree analysis, this study provides empirical findings about how demographic factors may affect consumers’ online shopping and following social media influencers behaviors. The findings have significant theoretical, methodological, and practical contributions.

First, this study has theoretical contributions because the results support consumer market segmentation theory (e.g., Kotler, 1989; 1999) and consumer demographic theory (e.g., Sheth, 1977; Martins & Brooks, 2010). Both theories have been frequently cited in consumer behaviors and marketing literature (e.g., Mothersbaugh et al., 2020; Ruvio & Iacobucci, 2023) and applied in academic research (e.g., Wu, 2022c; 2023). However, consumer behaviors were significantly disrupted amid the COVID-19 pandemic and are still evolving in the new normal. Therefore, it would be important to evaluate the applicability of existing marketing theories in the post-pandemic era. The results of this study suggest that various demographic factors, such as gender, age, educational level, family income, and geographic areas/USR, are significant predictors for consumers’ online shopping and following social media influencers behaviors. Therefore, this study’s empirical findings suggest that consumer demographic theory and demographic approach of market segmentation are still valid and can be applied to explain consumer behaviors online in today’s e-commerce/social commerce environment.

Second, this study has methodological contributions by using decision tree for data analysis. Decision tree analysis, a predictive analytics method, is considered as a powerful data mining method. In this study, decision tree analysis was performed to identify the most influential demographic predictors for consumer behaviors online. By doing so, we can profile consumers’ online shopping and following social media influencers behaviors with predictive analytics.

Finally, this study has practical contributions. Some important questions which e-commerce companies’ executives and marketing practitioners frequently ask are: Who are the online consumers? What are the factors that affect consumers’ online shopping behaviors? How do they shop online (e.g., using a computer or using a smartphone)? Do consumers prefer online shopping or in-store shopping if they have the choice? What are the factors that affect consumers’ following social media influencers behaviors? To what extent that consumers’ purchase decisions and buying behaviors are affected by social media influencers? The results of this study have answered all these questions.

Based on the results of this study, strategic recommendations for market segmentation, e-commerce site design, and influencer marketing are provided. First, organizations can develop their market segmentation and influencer marketing strategies by using demographic segmentation. As the statistical results suggested, several demographic variables (e.g., gender, age category, educational level, race-ethnicity, and geographic area/USR) are significant predictors for consumers’ online shopping behaviors, following influencer behaviors, and having purchase decisions affected by influencers. Thus, consumer demographics do matter in the online environment. Specifically, age category is the most important predictor for consumers’ following social media influencers and purchasing something after seeing influencers’ posts behaviors. Therefore, organizations can utilize influencer marketing strategies to target at the younger consumer segments (e.g., Generation Y and Z).
In addition, organizations may partner with influencers in different ethnic groups because race-ethnicity is an important predictor for consumers’ following social media influencers behavior. As Wu (2022b) noted, diversity is an emerging trend in influence marketing. Second, organizations need to make their e-commerce sites to be mobile friendly. As the results suggested, about 8 out of 10 (79.3%) of the participants ever purchased something by using a smartphone, compared with that about 7 out of 10 (71.5%) participants saying that they ever purchased using a desktop or laptop computer. In addition, there might be *device divide/mobile only divide* among different race-ethnicity groups and in different geographic areas (urban, suburban, rural). By doing so, consumers who only own a smartphone, instead of a computer, can shop online with ease. Finally, organizations can partner with influencers on Instagram and TikTok to engage with young females, foster positive eWOM, and drive business growth, because many younger females are following social media influencers and actually purchase something after seeing influencers’ posts.

**Limitations and Suggestions for Future Studies**

Although there are significant contributions, this study has limitations. First, this study is purely a quantitative study. The decision tree analysis method can identify the predictors for consumers’ online shopping and following social media influencers behaviors. However, the quantitative results can't fully explain “why” minority ethnic groups are categorized differently in some nodes when shopping online using a smartphone is examined. Specifically, previous studies (e.g., Atske & Perrin, 2021; Bartikowski et al., 2018) which examined *race-ethnicity digital divide* only included White, Black, and Hispanic, instead of Asian American in the samples. This study has Asian in the sample. However, the quantitative results can't explain why older Asian consumers (65+) are more likely to shop online by using a smartphone than White and Hispanic consumers. Thus, future studies may use qualitative research methods, such as focus group interviews or in-depth interviews, to get additional insights about how consumers in different ethnic groups shop online by using different devices (e.g., computers vs. smartphone).

Second, the age group category is pre-coded in the secondary dataset. Therefore, a specific age category may include two generations, instead of a specific generation. For example, the 18 -29 age category includes Generation Zers (18 – 25 in 2022) and younger Generation Yers (26 – 29 in 2022). In this dataset, each respondent's actual age (e.g., 18, 30 years old) is not available in the age variable. In order to dive deeper into generational differences in consumer behaviors online, future studies may categorize/code consumers’ age based on generation (e.g., Generation X, Y, Z, Boomer, Silent Generation).

Finally, this study mainly focus on how demographic factors may affect consumer behaviors online. Future studies may further explore how cultural and psychological factors, such as personality, lifestyle, values, and attitudes, may be used as predictors for consumer behaviors online.

**Research ethic statement**

This manuscript is original, and not submitted to any other journals.

**Author contribution statement**

The sole author of the article was responsible for the entire elaboration of the research.

**Funding**

The author declares that no financial support was received for the research, authorship, and/or publication of this article.

**Disclosure statement**

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Acknowledgements
The author would like to acknowledge Dr. Carl Zangerl, Dr. Edward Powers, Dr. Mikhail Oet and the anonymous reviewer for their suggestions for an earlier version and Pew Research Center for proving data for academic use.

References


IBM (2010). *Predictive modeling with IBM SPSS Modeler student guide.* Course Code: 0A032. IBM.


Profiling consumers' online shopping and following social media influencers behaviors


Vilkaite-Vaitone, N. (2024). From likes to sustainability: How social media influencers are changing the way we consume. Sustainability, 16(4), 1393. https://doi.org/10.3390/su16041393


