

Predicting the Impact of Short-Form Video Content on Gen Z's Brand Awareness, Attitudes, Behavior and Purchase Decisions Using Machine Learning

School of business, RV University, Bangalore, Karnataka

Sivamugunthan Ashok

Sivamugunthan342@gmail.com

Abstract

The study investigates and predicts the impact of short form video content on brand awareness, consumer attitude, interactive behavior, pre purchase behavior and purchase decision among generation Z using machine learning approach. The primary objectives are to know the brand awareness, customers attitudes, interactive behavior, pre-purchase behavior and purchase decision in short form video, to know the impact of social demographic variables on brand awareness, customers attitudes, interactive behavior, pre-purchase behavior and purchase decision in short form video, to know the impact of brand awareness, customers attitudes, interactive behavior, pre-purchase behavior on purchase decisions using machine learning predictive modeling, to know the impact of one variable on another in a sequential manner, leading from Brand awareness to decision-making. A quantitative research method was conducted, by employing structured surveys and personal interviews to collect data from generation Z respondents. The findings reveal that brand awareness, consumer attitudes, and pre-purchase behavior significantly influence purchase decisions. Although interactive behavior does not show a strong correlation with purchase decision making. The obtained result highlights the effectiveness of (SFV) in framing the viewer's psychological aspects. Marketers should focus on enhancing transparency, informativeness, clarity, trust on short form video rather than only relying on interactive features.

Keywords: Short-form video content, Brand awareness, Consumer attitude, Interactive behavior, Pre-purchase behavior, Purchase decision, Generation Z, Social demographic variables, Psychological influence, Transparency, Informativeness, Trust, Digital marketing, Machine learning, Predictive modeling.

1. Introduction

This research has analyzed and predicted the impact of (SHV) content on brand awareness, consumer attitudes, pre-purchase behavior, interactional behavior and purchase decisions among Gen Z. Through this analysis, the research provides insights on understanding how marketers can strategically manipulate this insight to build strong and effective brand communications. 61 percent of Generation Z are consuming video content under 60 seconds, on social media platforms [4]. Research shows that Gen Z people's attention span roughly lasts about 8 seconds, which is quite shorter compared to other generations [1]. Therefore, the study narrowed down to Generation Z of people born between the mid 1990s to early 2010's.

The brisk progression of social media has emerged the trend to watch (SFV) content, enabling the brand to engage with consumers. Specifically, Generation Z are the major audience to consume short form videos to have quick bursts of entertainment. Prior research spotlights the effect of (SFV) in escalating brand recognition, fostering authenticity impulse purchases, and improving engagement through interactive features [4].

Although a gap still exists in recognizing how particular elements of (SFV) content such as trending memes, branding consistency, brand recall, user-generated content, influencer endorsements, and shoppable links directly impact

consumer purchase decisions. Campaigns using short form video led to increase brand awareness by 67% in under three weeks [19]. Moreover, shoppable links have significantly shortened the customer decision making process by promoting impulse purchase [9].

The emerging short form content consumption behavior of Generation Z has significantly influenced the marketing strategies across fast moving consumer goods (FMCG) industries. In contrast to traditional advertisements, short form videos provide great opportunities to brands to go viral with visually appealing content. Advertisements in short form video are 2.5x times more engaging than long form content for generation Z. A global survey conducted by Accenture resulted in 89% of generation Z having discovered products via short form videos to make purchases.

2. Research problem

The proliferation of (SFV) content has significantly transformed digital marketing strategies, enabling the brand to transform itself as innovative and dynamic to engage with consumers. The industrial data shows that short form video can drive immediate consumer action, so academic research has yet to provide a comprehensive analysis on how different elements like influencer endorsements, emotional aspect and interactivity affect generation Z via short form video. [16]. There are various features have influenced to do this research, Despite the extensive adoption of short form video marketing, there is some limited understanding on how particular short form video content's characteristics like memes, consistent branding, shoppable links, limited time offers, transparency, influencer endorsements, ectara effects the consumer psychology such as brand awareness, consumer attitudes, behavior aspects and purchase decisions among Generation Z. Meanwhile the poorly implemented short-form video strategies can lead to losing potential customers and detract from the intended marketing message [2].

3. Scope

This study explores and predicts the effect of (SFV) content on brand awareness, consumer attitudes, behavior, and purchase decisions among Gen Z. The respondents of this study are generation Z people who responded to previous research that has demonstrated the efficiency of (SFV) in enhancing brand recall, regulating the impulse purchases, fostering authenticity, and increasing engagement through interactive features [4]. Thus, this study explores these factors while ultimately connecting them from the initial stage of brand awareness followed by consumer attitudes, behavior to the final stage of purchase decision.

4. Objectives

This study aims to examine the pathways through which brand awareness, consumer attitudes, interactive and pre-purchase behaviors in short-form video environments influence final purchase decisions. The specific aim of this research is to:

1. To know the brand awareness, customers attitudes, interactive behavior, pre-purchase behavior and purchase decision in (SFV).
2. To know the impact of social demographic variables on brand awareness, customers attitudes, interactive behavior, pre-purchase behavior and purchase decision in (SFV).
3. To know the impact of brand awareness, customers attitudes, interactive behavior, pre-purchase behavior on purchase decisions using machine learning predictive modeling .
4. To know the impact of one variable on another in a sequential manner, leading from Brand awareness to decision-making.

5. Relevance of the study

This study is highly relevant in today's digitalized world where (SFV) content is dominating the social media platform. Gen Z people are the present digitised generation spending their significant portion of daily time in consuming content from (SFV) on social media platforms. Moreover, brands are highly investing in digital marketing, their limited amount of research conducted to evaluate the factors in consumer journey amidst gen Z. So, this research aims to fill this gap by analyzing several factors such as user generated content, interactive features, brand recall, trust, trending memes, shoppable links. Marketers will understand deeply about the short form video by utilizing the techniques and strategies shown in this research. Additionally, this study evaluates the function of social demographic variables like age, gender and area towards the (SFV) to make personalized marketing strategies. Overall, this study is particularly relevant to the brand marketers, who market the brand via short form video to younger consumers to increase brand reachability.

6. Literature review

This part gives an outline of previous research to get an overview of consumers purchase decisions in short form videos over the aspects like brand awareness, consumer attitudes and behavior. Specifically, the first section focuses on the Short-Format Videos on Social Media Platforms. Then, the second section is about who are Generation Z, followed by short form videos in brand awareness, consumer Attitudes, Interactional behavior, Pre purchase Behavior and purchase Decisions.

6.1 Generation Z

As per previous research, Gen Z are the people who were born between 1997 to 2012, they are also called as digital natives who are proficient in using technology and social media [5]. The focus of this study is narrowed to the age group between 1996 and 2007 because of access to use and consume social media content with low restrictions by their parents. The consumption of (SMV) through social media by gen Z daily makes the platform more dominant by influencing the habit, attitudes, and purchase behaviors. Generation Z pays about 3.5 hours daily for watching online videos, in which 60% of the time is spent watching (SFV) content on present social media platforms [6].

6.2 Short-format videos on social media platforms

According to [3], in 2024 current social media platforms will have three billion monthly active users worldwide. In this digitised period, the escalation of (SFV) content has impacted digital marketing, particularly in the zone of grabbing the attention of Gen Z people. (SFV) is an engagement content that has a length range between a few seconds to two minutes [4]. The consumption of (SFV) has gradually increased because of its vertical orientation and it hugely contains user generated contents. Across social media platforms, entry barriers for brand and individual's creators are very weak. Furthermore, algorithms play an important role in recommending and suggesting relatable video by using artificial intelligence to personalize video feeds. This strategy retains the users and increases content exposure [1]. On the marketing aspect short form video offers a high return on investment (ROI). Research shows that 44 percent of the marketers had found that (SFV) provides the most effective content format that leads to 57 percent increase in consumer awareness. This statistical data shows that (SFV) is dominating engagement building and awareness quickly. The most effective aspects of short videos are grabbing attention in the first 3 seconds, making creative elements like cyclic videos and soundless storytelling.

6.3 SFV in brand awareness

Shorter videos have arrived on social media platforms that are optimized for mobile-first consumption to address the shorter attention spans. Shorter video content has come into force with a huge dominance in shaping consumer behavior towards brand awareness and quick initial engagement, while long form videos contribute more to brand loyalty [23]. Brand awareness is one of the crucial steps in the consumer decision making journey. Short form video provides content in the way by grabbing and building brand image in the minds of consumers quickly and effectively. Through the frequent exposure and virality, brand elements like brand logos, slogans are repeatedly shown in advertisements, which increase brand familiarity and recall. User generated content and influencer generated content are more similar due to their perceived relatability [4].

6.4 Consumer attitudes on SFV

Consumer attitude is the factor which defines how marketing makes a connection with consumer's feeling and perception. In the short form video platform attitudes are often shaped with characteristics, psychology, culture and perception. Shorter videos have a potential in altering the perception of consumers due to their quick attention grabbing, entertainment value, and emotional connect. Most attitudes in social media are often formed through consumer's awareness towards the brand and it leads to pre purchase behavior and helps in making purchase decisions. Authenticity is one of the important factors in shaping the favorable consumer attitude, especially Generation Z due to preferences such as genuine feelings, unscripted and humanized content. As per the research by [9], influencers who express vulnerability and personal narratives in their (SFV) video content are more likely to gain a stronger and more reliable audience.

6.5 User-generated content

According to (Lee Choi,2021) 68 percent of the Generation Z consumers are trusting user generated content over brand created content due to its Genuity. Glosser's short form video campaign was successful by finding User generated content has a good relationship between brand authenticity and loyalty [24]. So, this study had focused on finding the type of relationship that the user generated content and authenticity of the brand have. User generated content is playing a crucial role in shaping brand awareness through short form videos. When content creators and consumers involve a brand intentionally or organically, it helps the brand to expand their reach without spending much on marketing and advertising. Most of the brands use influencers in their brand advertisements. This strategy helps the brand to connect with the specific audience of the influencer. Influencers act as brand ambassadors in promoting the products. If the influencer's authenticity and relatability are good, it enhances brand recall and viewers perception.

6.6 Interactive behavior

Interactive behavior refers to the way consumers engage with the short form video beyond primary consumption. Factors such as likes, shares, comments, saves, polls are the major aspects evaluating the interactive behavior of consumers. According to the research by [21] these interactive factors alone can't make a connection between brand and consumers, building emotional and symbolic connection with consumers is very crucial. Furthermore, there are several factors that drive interactive behavior with short form video such as emotional resonance, trend participation, relatability, authenticity, call to action (CTA). Interactive behavior often associated with brand advocacy; this includes sharing or remixing the brand in short form video that endorses the content and helps in reaching more audience. Moreover, brands that reply to comments are more likely to increase in customer trust and loyalty [20]. Influencers play an important role in shaping the interactive norms in (SFV). When the influencers engage in creating challenges to their followers it tends to mimic or respond that creates a ripple effect. This mimicry associated with social relationship and aspirational behavior increases the interaction and pushes the campaign to become viral [4].

6.7 Pre purchase behavior

Pre purchase behavior encloses the aspects such as cognitive, emotional, behavioral process that consumers practice before making a purchase decision. In the current generation short form video is more likely to have behavior shaped by rapid exposure to obtain information, user friendly and effective content that stimulate the curiosity and intention to make a purchase. Instagram reels act as a stimulating curiosity due to its frictionless exposure which often initiates the pre purchase behavior and stimulates curiosity. Consumers are more likely to trust a product when the other people are using and endorsing it in shorter videos. (UGC) and influencer endorsement are showcased as social proof which positively affects the pre-purchase confidence [21]. Emotional engagement in short form video content is influencing the viewer's perceptions towards the product value and brand trustworthiness. Fear of missing out (FOMO) is a strong factor that motivates the buyer's and shapes their pre-purchase behavior through limited time offers, viral challenges and showcasing exclusive experiences. The (SFV) viewers are coercively to explore the product when they perceive it due to a social trend [4].

6.8 SFV and purchase decisions

Short-form video marketing influences purchase intention through Applicable content, Entertainment content, Emotional content and purchase motivation through Perceived value [8]. Basically, a purchase decision is forming when curiosity and evaluation is changing into action. According to the research done by Meta, direct response factors such as swipe up to buy, tap to shop, product tags, shoppable links significantly increase the click through and conversion rate. When consumers perceive the brand as authentic, they are more likely to purchase the product. When user generated content and influencer collaborations in short form video are unscripted, still there will be higher sense of trust and social validation towards consumers. Influencers are significantly impacting the purchase decision, particularly in short form video environments. According to research by Meta branding short form video with influencers are achieved 35% higher conversion rates compared to brand content.

7. Rationale of the study

Social media channels are dominating Generation Z's social media consumption. (SFV) are broadly adopted by Generation Z, but the main challenge is difficulty in conveying the complex information [9]. Generation Z grew up in social media era like YouTube, Instagram, Facebook, Twitter/x, where they consume a lot of short form video contents in raw and unfiltered formats. The reason behind considering the factors like Trust, entertainment and information in this study is because it plays significant roles in shaping Generation Z's purchase decision in short form videos [22]. There is Generational behavior gap between Generation Z and the older generation in perceiving authenticity. Generation Z prefers storytelling and vulnerability, while older generations favor straight forward information [20]. So this study touches upon storytelling and considers it as a factor in affecting gen Z's attitude. Shorter videos perform better with brand awareness and initial engagement while long form video contributes more to brand loyalty [23]. And the length of a video impacts engagement but does not necessarily influence purchase intention. Thus, the study analyzes the effect of SFV content on gen Z purchase journey.

8. Review summary

The previous literature review indicates SFV as an effective promotional instrument in the digitized world, especially in shaping the consumer behavior and preferences of the generation Z. The high digital engagement and low attention span characterizing this demography which demonstrates the visually stimulating content such as reels, shorts. Existing studies have shown that short form videos significantly contribute to brand awareness, particularly when they include trending elements, frequent branding, and influencer collaboration. Prior studies also suggest that consumer attitudes are positively influenced by factors such as entertainment, information, consistent branding, transparent pros and cons. Similarly, user generated content in short form video is more likely increasing the brand credibility. In generation Z perception user generated content is often viewed as more trustworthy and relatable compared to traditional advertising. Moreover, interactive behaviors such as liking, sharing, commenting, participating in polls and challenges have strong relationships with consumer engagement and are assumed to have deeper brand connection for the consumer during the pre-purchase stage.

Only a few studies have done empirical research on finding the sequential path from brand exposure to consumer purchase action. Thus, the study decided to focus on finding the impact of Brand awareness, Consumer attitude, Intentional behavior, Pre purchase behavior and purchase decision amidst gen Z people which deeply examine the sequential path from awareness to action.

9. Methodology

9.1 Design of the research

This study had an approach of quantitative research which focused on studying the impact of short form video content on brand awareness, consumer attitudes, interactive behavior and purchase decision on gen Zs. A primary data collection technique was conducted to gather reliable data from generation Z respondents via structured survey and

personal interview. This quantitative method enables the research paper to have statistical analysis by providing insights on the relationship between the brand awareness, consumer attitudes, and interactive behavior on purchase decisions

9.2 Sampling method

This study contains a non-probability sampling technique to collect responses. Target respondents for this research were selected based on their age, specifically targeting individuals aged between 15 to 25 years who fall under the generation Z group. This survey has taken place in social media such Instagram, WhatsApp, LinkedIn. Also reached respondents in Real Time face to face interaction for collecting responses. This method made the data collection part very effective within a short period of time by reaching a larger audience and made the data reliable for doing diverse analysis.

9.3 Method of data collection

Data has been collected via online surveys by google forms. By ensuring there is no major bias from respondents, the survey had informal interviews. The study's primary data is entirely related to the quantitative survey data. The advantage over here is using an online platform for survey questionnaire distribution and data collection enabling the researcher to ensure the quality and manage the survey response effectively for further analysis.

9.4 Research instrument

Instrument used for this research is a structured google form questionnaire. The data had been divided into five main domains named Brand awareness, consumer attitude, Interactional behavior, pre-purchase behavior and purchase decision. Most of the survey questions are close ended with 5 points Likert scale, which range between 1 (strongly disagree) to 5 (strongly agree). The responses were pretested to check the reliability with 30 individuals to ensure and avoid bias from respondents.

9.5 Variables and operational definitions

This study contains five domains in which four are explanatory variables and one outcome variable. The explanatory variables are Brand awareness, which measures the aspects like sustainability, trending meme, trust towards the SHV; consumer attitude, which evaluates the feelings of the consumers towards (SFV); interactional behavior, evaluates the interaction of consumers such as polling, commenting, liking, sharing on short form video; pre-purchase behavior which records the information seeking process during the purchase mindset. The dependent variable is purchasing decision which is considered as the final consumer action to measure how likely they buy the product that is shown via short form video.

9.6 Data analysis technique

To analyze the collected data, SPSS, Python, and Excel were used. The following analyses were conducted:

1. Descriptive Analysis:
Conducted to understand the primary pattern and demographic characteristics of the respondents.
2. Reliability Test:
Cronbach's Alpha was calculated to check the internal consistency and reliability of the constructs.
3. Sampling Adequacy and Factorability:
KMO Test and Bartlett's Test of Sphericity were conducted to assess the suitability of the data for factor analysis.

4. Dimensionality reduction:
PCA test was performed to reduce the dimensionality of the dataset and to create meaningful components.
5. Correlation analysis:
Used to identify the relationships between the independent variables and the dependent variable.
6. Non-parametric tests:
MWU Test and Kruskal-Wallis Test were conducted to check the mean rank differences and Z-scores across different groups.
7. One-Way ANOVA test:
Performed to observe the differences and patterns across demographic groups.
8. Multicollinearity Check:
VIF scores were obtained to check the multicollinearity among independent variables.
9. Regression Analysis:
Both simple linear regression and multiple linear regression were used to test the sequential and direct influences between the variables such as brand awareness, consumer attitudes, interactive behavior, pre-purchase behavior, and the purchase decision variable.
10. Machine Learning Diagnostic Checks:
Under a machine learning approach, additional model evaluation techniques were conducted using Python and SPSS, including:
 - 1) R^2 to measure the model variance explanation.
 - 2) Obtained RMSE value to check the average prediction error.
 - 3) Scatter plot of Actual vs Predicted Values to visualize prediction accuracy.
 - 4) Residual plot to check for homoscedasticity of residuals.
 - 5) Residual distribution plot to observe the normality of residuals.
 - 6) Q-Q Plot to further assess the normality of residuals.

10. Statistical hypotheses

The following null hypothesis are tested for possible rejections

1. **H01** There is no influence of social demographic variable (Gender, Area) on total score on Brand awareness, consumer attitude, interactive Behavior, Pre purchase Behavior and Purchase Decision.
2. **H02** Brand awareness gained through SFV does not affect the consumer attitude.
3. **H03** Consumer attitude toward SFVs does not affect interactive behavior.
4. **H04** Consumer attitude toward SFVs does not affect Pre-Purchase Behavior.
5. **H05** Pre-Purchase Behavior in SFV does not affect the purchase decision.
6. **H06** Interactive behavior in SFV does not affect the purchase decision.

7. **H07** Brand awareness gained through SFV does not affect the purchase decision.
8. **H08** Consumer attitude towards SFV does not affect the purchase decision.
9. **H09** There is no effect of brand awareness, consumer attitude, interactional behavior, pre purchase behavior on purchase decision.

11. Empirical results

This part highlights the result from the 100 respondents collected via online survey forms and personal interview. The data is collected from the age group of 18 to 22. The collected data is used for analysis to find the impact of SFV content on brand awareness, consumer attitude, interactional behavior, pre purchase behavior and purchase decision.

11.1 Data pre processing

The first step was to clean the data in Microsoft Excel. During this step, all the duplicate values have been deleted. Outliers were identified and removed based on response patterns. Categorical variables were coded with numerical values and null values are replaced with mean values. The variables are renamed according to serial numbers for better convenience. Moreover, this process gives a clean and consistent dataset to further proceed with valid statistical analysis.

11.2 Principal component analysis

11.2.1 KMO and Bartlett's test

To know the stability of the dataset for factor analysis, the KMO and Bartlett's Test of Sphericity were conducted.

Table 1: KMO and Bartlett's Test

KMO Measure of Sampling Adequacy.		.733
Bartlett's Test of Sphericity	Approx. Chi-Square	968.472
	df	136
	Sig.	<.001

The obtained KMO value was 0.733, which is higher than the minimum threshold score of 0.6. A KMO value closer to 1 indicates that the partial correlation between variables is small, thus the sample test data is good enough for conducting factor analysis which more likely gives reliable results. Bartlett's Test of Sphericity shows that approx. chi-Square value is 968.472 which indicates that the correlation among variables is good enough to each other to perform factor analysis. The degrees of freedom is 136 which indicates the number of relationships tested between the variables. The P- value in table 1 is <0.01 which means the overall result is significant.

11.2.2 Factor analysis

Rotated component matrix helps us to understand whether the variables are grouped into the correct construct. The Rotated Component Matrix and Cronbach's Alpha values shown in Table 2 were used to determine the construct validity and internal consistency reliability of the variables used in this study. The rotated factor Loadings value specifies that all the variables are loaded significantly. All the variables included in the analysis showed factor Loadings greater than 0.50, which means the variables have a good level of correlation between its respective constructs. Therefore, the variable is correctly grouped under the constructs. Specifically, for the construct Consumer Attitude, variables have a Loading value range from 0.511 to 0.815, which indicates a strong relationship with the construct. Similarly, variables under Brand Awareness have Loadings ranging from 0.597 to 0.765, followed by Interactive behavior variables ranging between 0.629 to 0.798. The factor Loadings value closer to 1 indicates the variable has good fit towards their respective constructs.

To know the reliability of each construct, Cronbach's Alpha test was conducted. This test is used to find the internal consistency of the constructs. The construct can be used for further analysis if it has Cronbach's Alpha value of 0.7 or higher. In table 2 Component 1 has Cronbach's Alpha value of 0.812 which indicates good internal consistency. Component 2 has Cronbach's Alpha value of 0.870 specifies that the construct is excellent in internal consistency. Similarly, component 4 with 0.790 Cronbach's Alpha value is good in internal consistency. Component 3 with 0.662 Cronbach's Alpha value is slightly below the ideal threshold but still the construct can be used for further analysis.

Table 2: Rotated component matrix: Reliability test

	Component				Cronbach's Alpha
	1	2	3	4	
CA 3	.815				0.812
CA 5	.800				
CA 6	.742				
CA 1	.715				
CA 4	.643				
CA 2	.511				
BA 5		.765			0.870
BA 3		.727			
BA 1		.692			
BA 4		.659			
BA 2		.597			
IB 3			.798		0.662
IB 1			.757		
IB 2			.629		
PPB 2				.847	0.790
PPB 1				.818	
PPB 3				.537	

11.3 Normality of measuring instruments

After performing the data cleansing process, Normality test was conducted to identify the skewness and kurtosis of the variables as shown in Table 3. It gives insights about the distribution aspect of the dataset. This test helps to identify whether the data is following a normal distribution, which is very crucial for doing many statistical analyses like Multiple linear regression, Structural equation model. Skewness shows the symmetry of the distribution. Symmetry indicates positive and negative deviations from the mean.

Negative skewness values lead to left skewness, and it is said to be tail on the left. Positive values indicate the data is right skewed and it is called the tail on the right. Skewness is found from the third central moment, which refers to cubing the deviations from the mean and all mirrored values are deleted. The data should not be skewed because it leads to incorrect P value and poor confidence levels, thus many statistical tests like t-test, regression assume the data to be normally distributed.

Table 3: Skewness and Kurtosis

Variables	Skewness	Kurtosis
Brand Awareness	-.865	.352
Consumer Attitude	-1.083	.746
Interactional Behavior	.205	-.083
Pre purchase Behavior	-.791	.357
Purchase Decision	-.359	-1.040

In table 3 Brand awareness has moderately negative skewness value of -0.865 which indicates more respondents leaned towards higher agreement. Similarly, pre purchase behavior has respondents who responded as higher agreement. Consumer attitude has a substantial negative skewness value of -1.083 which specifies most respondents held strong positive attitudes. Interactive behavior obtained an almost symmetric value of 0.205 where responses were normally distributed. Purchase decision has mild negative skewed value of -0.359 where more respondents agreed to the purchasing aspects via short form video with reasonable balanced distribution.

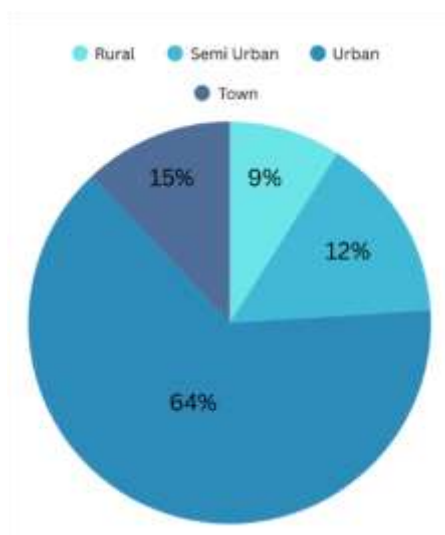
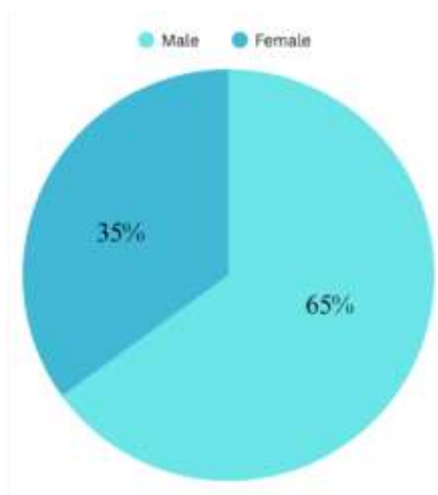
According to (Xhignesse, 2021) the values of this test must be between -2 and 2 to be considered as the data is not skewed in huge levels and it follows a normal distribution. As shown in table 3 the values are within the range, thus it is following normal distribution, and the data is suitable for doing both parametric and non-parametric tests. Kurtosis value in table 3 measures the tailed Ness and peakedNess of the distribution. A kurtosis value close to 0 specifies normal distribution, positive value indicates the data have peak distribution while negative value indicates the data is flatter distribution. In this study the kurtosis value is within ± 2 , thus the data is almost normally distributed.

11.4 Summary of social demographic variables

Table 4 shows the summary of the demographic characteristics of the 100 respondents who have participated in this study.

Table 4: Summary of social demographic variable: counts and percent

Demographic	Group	Counts	Percentage
Age	18	7	7.0
	19	44	44.0
	20	33	33.0
	21	9	9.0
	22	7	7.0
Gender	Male	65	65.0
	Female	35	35.0
Profession	Student	100	100.0
Area	Rural	9	9.0
	Semi urban	15	15.0
	Urban	64	64.0
	Town	12	12.0
Education attainment	School	2	2.0
	Undergraduate	94	94.0
	Postgraduate	1	1.0
	Other	3	3.0
SFV Watch frequency	Never	4	4.0
	Rarely	7	7.0
	Sometimes	21	21.0
	Often	26	26.0
	Always	42	42.0
SFV Shopping frequency	Rarely	16	16.0
	Sometimes	28	28.0
	Often	29	29.0
	Always	27	27.0



Most of the respondents are 19 (44%) and 20 (33%) -year-old. A smaller percentage of participants are aged 18 (7%), 21 (9%), and 22 (7%). This age distribution shows that the data mostly consist of young adults. This makes the study relevant by evaluating the impact of short form video content on Brand Awareness, Consumer Attitude, Interactive Behavior, Pre Purchase Behavior and purchase decision among generation Z. The data consist of 65% male and 35% female respondents, indicating male as the major audience.

All the respondents of this study are students, that exactly justifying the study is focusing on youth adults who are generation Z. About 64% of the respondents are from urban areas, followed by 15% from semi-urban, 12% from towns, and 9% from rural areas. Here respondents from urban areas are dominating. Almost 94% are undergraduates, while 2% have schooling level education, 1% are postgraduates and 3% belong to the 'other' category. This indicates the data consist of educated young audiences.

About 42% of the respondents watch short form video frequently, 26% of the respondents watch short form video often and 21% watch sometimes. Only 11% of the respondents do not or rarely watch short form videos. This indicated generation Z are more likely to consume content via short form video. Shopping behavior of the respondents via short form video is quite active. About 29% of the respondents shop often, 27% of the respondents always shop, followed by 28% of the respondents shop sometimes. Only 16% of the respondents shop rarely. This indicates almost all generation Z people make purchases after watching short form videos.

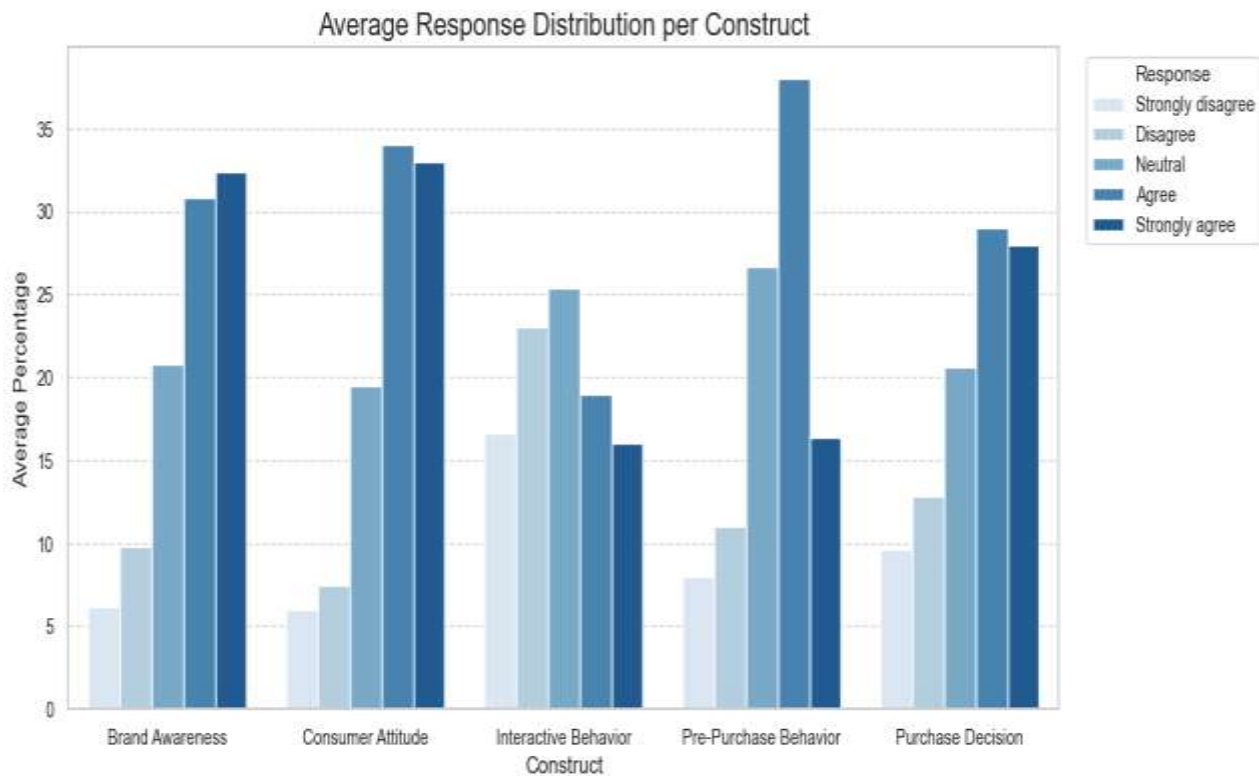
11.5 Summary of constructs

Table 5 shows the percentage distribution of the respondents across all the variables related to the constructs such as Brand Awareness, Consumer Attitude, Interactive Behavior, Pre Purchase Behavior, and Purchase Decision. This response has the Likert scale with 5 points ranging from strongly disagree to strongly agree.

Table 5: Summary of construct variables: *percentage*

Construct	Variables	Strongly disagree	disagree	Neutral	Agree	Strongly agree
Brand Awareness	BA 1	3.0	6.0	23.0	42.0	26.0
	BA 2	6.0	9.0	19.0	31.0	35.0
	BA 3	4.0	5.0	22.0	35.0	34.0
	BA 4	10.0	19.0	19.0	22.0	30.0
	BA 5	8.0	10.0	21.0	24.0	37.0
Consumer Attitude	CA 1	6.0	10.0	21.0	33.0	30.0
	CA 2	4.0	5.0	31.0	29.0	31.0
	CA 3	8.0	5.0	19.0	31.0	37.0
	CA 4	5.0	7.0	26.0	26.0	36.0
	CA 5	8.0	4.0	10.0	44.0	34.0
	CA 6	5.0	14.0	10.0	41.0	30.0
Interactive Behavior	IB 1	22.0	20.0	34.0	10.0	14.0
	IB 2	13.0	36.0	27.0	13.0	11.0
	IB 3	15.0	13.0	15.0	34.0	23.0
Pre-Purchase Behavior	PPB 1	8.0	10.0	16.0	42.0	24.0
	PPB 2	7.0	10.0	39.0	31.0	13.0

	PPB 3	9.0	13.0	25.0	41.0	12.0
Purchase Decision	PD 1	6.0	16.0	27.0	33.0	18.0
	PD 2	11.0	15.0	29.0	28.0	17.0
	PD 3	13.0	17.0	13.0	41.0	16.0
	PD 4	16.0	9.0	7.0	13.0	55.0
	PD 5	2.0	7.0	27.0	30.0	34.0



11.5.1 Brand awareness

Overall, participants showed positive perception towards brand awareness via short form video content. On an average about 60% of participants responded agree and strongly agree across all five variables under Brand awareness. In table 4 we can see 35% of the respondents strongly agreed and 31% of the respondents agreed for (BA 2), which indicates that they remember brands better when they collaborate with influencers. 37% of the respondents strongly agreed and 24% of the respondents agreed for (BA 5), which shows that they are more likely to have trust on the brands when they have repeated exposure. Only 52% of the respondents are either strongly agreed or agreed for (BA 4), which specifies that seeing a brand frequently in short-form videos makes it more recognizable. About 70% of the respondents either strongly agreed or agreed for the variable (BA 2) means, Trending challenges/memes make brands more memorable to them. Overall short form video content is an effective medium for increasing brand recall and recognition, particularly when the content has influencer collaboration, repeatedly endorsed and engaging factors like trending memes.

11.5.2 Consumer attitude

The Consumer attitude construct shows the feelings of respondents towards credibility, informativeness, and positivity of short form videos with brand and product promotion. Across the six variables under Consumer attitudes majority of the respondents indicated that they had a positive attitude towards short form video content. A large position (60%) of the respondents either agreed or strongly agreed for the positive feelings towards influencer recommendation,

brand information, product Comparison, testimonials, transparent pros and cons, and consistent branding via short form video content. Therefore, Generation Z consumers appreciate authentic, informative and visually appealing short form video content. Influencers and user driven content significantly create brand credibility which gives a positive attitude towards the brand.

11.5.3 Interactional behavior

This construct shows how likely respondents were engaging towards short form video through the features such as polls, links, and social factors (likes, shares, comments). The results obtained from the respondents for this construct have a lower level of engagement compared to other constructs. For instance, the variable (IB1) 20% of the respondents disagreed and 22% of the respondents strongly disagreed, only 14% strongly agreed which indicates polls or challenges shown in short form video do not create a huge impact on consumer's interactive behavior. Similarly, 50% of the respondents either strongly disagreed or disagreed for clicking the shoppable links shown in the short form video. The variable (IB 3) has slightly better results, with 34% of the respondents agreed and 23% of the respondents strongly agreeing. Thus, we can conclude that social validation cues such as Likes, comments, shares have a positive impact on shaping consumers' interactional behavior.

11.5.4 Pre purchase behavior

The results obtained from the respondents for the construct Pre purchase behavior showed a moderate to high acceptance towards short form video content. The variable (PPB 1) shows that 64% of the respondents either strongly agreed or agreed, which indicates respondents discover new brands or products through short form video. (PPB 2) showed a high neutral response, so there is a huge uncertainty in defining whether consumers find user generated content in short form video as authentic and trustworthy. Moreover, the variable (PPB 3) obtained a result of 53% of respondents who either strongly agree or agree on making impulse purchases due to limited time offers. In conclusion short form video content plays a crucial role in making pre purchase decisions to purchase the product.

11.5.5 Purchase decision

This construct (PD 4) shows 68% of the respondents either strongly agreed or agreed, which indicates respondents purchase a product from short form video if it looks useful and necessary for them. While the constructs (PD2) and (PD 3) show similar kind of acceptance towards purchasing the product via short form video if they have digitised advertisement and good mindset. Overall, we can conclude that short form video significantly influences the purchase decision.

11.6 Normality check

Table 6: Shapiro-Wilk Test; *Statistic value, P value*

Metrics	Values
Test statistic	0.9816
p-value	< 0.01

The Shapiro-Wilk test was conducted to check normality of the data. As shown in table 6 the test obtained a statistic value of 0.9816 and p value of < 0.01. The p value is lesser than 0.05, so null hypothesis is rejected which has the assumption that data is normally distributed. Therefore, the data is **not normally distributed** and **non-parametric test** was decided to make further analysis.

11.7 Descriptive statistics for gender

In this study Mann- Whitney U Test, a non-parametric test was conducted to explore the gender-based variations across all constructs. The reason to use nonparametric tests is due to the small sample size and slight deviation in the

normality, which have a chance of affecting the parametric test like T test. The Mann- Whitney U Test does not expect the data to be normally distributed and it is reliable for comparing medians between two independent variables.

In this test, the mean rank specifies the average rank scored by each group based on the collected responses. The Z score mentioned in table 7 shows how different the male and female are based on the data and used to calculate the P value which shows the significance. Moreover, it shows the expectation under the assumption that there is no difference between both groups. A higher Z score indicates larger difference between the male and female; Z score close to 0 means the groups are very similar.

Table 7: Impact of brand awareness, consumer attitude, Interactive behavior, Pre purchase behavior, Purchase decision on gender : Mann- Whitney U Test for Gender; Mean rank, Z score, p value

Construct	Gender	N	Mean	Std dev	Median	Mean Rank	Z	P Value
Brand Awareness							2.348	.019
	Male	65	17.83	4.837	19.00	45.52		
	Female	35	20.23	3.191	21.00	59.74		
Consumer Attitude							1.507	.132
	Male	65	23.25	5.543	25.00	53.69		
	Female	35	22.06	5.116	24.00	44.57		
Interactive Behavior							1.838	.006
	Male	65	8.38	2.827	9.00	46.65		
	Female	35	9.69	3.104	9.00	57.66		
Pre-purchase Behavior							.685	.494
	Male	65	10.08	3.164	11.00	49.05		
	Female	35	10.74	2.049	11.00	53.19		
Purchase Decision							.693	.488
	Male	65	17.45	4.555	18.00	49.03		
	Female	35	18.03	3.769	18.00	53.23		

Among all the variables, brand awareness and interactive behavior is statistically significant. For brand awareness, the mean rank for female group is 59.74 and male group is 45.52, with Z score value of 2.348 and p value is 0.019. This shows that female gender has higher brand awareness via short form video compared with male gender. Similarly, females had a higher mean rank with 57.66 than male with 46.65, with Z score value of 1.838 and a p value of 0.006 for Interactional behavior. This indicates that female gender is more likely to engage in interactions in short form video contents. In contrast there was no significant difference found in consumer attitude ($P=0.132$), Pre purchase behavior ($P=0.494$), and Purchase Decision ($P=0.488$) due to the P value greater than 0.05.

To analyze the difference between geographical areas such as Rural, semi urban, Urban, and Town, the Kruskal - Wallis Test Area was conducted for each variable. As shown in table 8 the result obtained was Interactive behavior is not

statistically significant towards geographical area which shows all groups engage similarly. In brand awareness urban and semi urban areas have the highest mean rank value of 59.63 and 48.33 compared to other areas such as rural (11.33), town (17.50) which indicates people in urban and semi urban are most aware of brands that are promoted via short form video content. Similarly, Consumer attitudes also showed significant P value (0.001) in which urban and semi urban people dominated with mean rank values of 60.95 and 45.53, which indicates people in urban and semi urban areas have a positive attitude towards short form video content. Meanwhile, Pre purchase behavior is significant in showcasing the highest mean rank scored by both urban and semi urban areas with values of 55.74 and 53.90 which indicates people in urban and semi urban areas are influenced in pre purchase behavior to make final decisions, as compared to town and rural groups. Finally, purchase behavior also showed a significant difference with the P value of 0.001, in which urban and semi urban people have the highest mean rank value of 59.28 and 46.73 which indicates that urban and semi urban people are more likely to make decisions through short form video content to purchase the product. People in Rural and town areas are not influenced by short form video to make purchase decisions.

Table 8: Impact of brand awareness, consumer attitude, Interactive behavior, Pre purchase behavior, Purchase decision on area: Kruskal - Wallis Test; Mean rank, p value

Construct	Area	N	Mean	Std dev	Median	Mean rank	P Value
Brand Awareness							.001
	Rural	9	12.22	1.986	12.00	11.33	
	Semi urban	15	19.07	2.251	19.00	48.33	
	Urban	64	20.03	3.846	21.00	59.63	
	Town	12	15.75	5.578	17.50	33.88	
Consumer Attitude							.001
	Rural	9	16.22	4.042	15.00	16.67	
	Semi urban	15	23.20	2.783	24.00	45.53	
	Urban	64	24.56	4.594	26.00	60.95	
	Town	12	18.08	6.487	20.00	26.38	
Interactive Behavior							.615
	Rural	9	8.56	2.877	9.00	50.56	
	Semi urban	15	9.73	3.634	9.00	55.00	
	Urban	64	8.97	2.731	9.00	51.26	
	Town	12	7.25	3.194	9.00	40.79	
Pre-purchase Behavior							.007
	Rural	9	8.44	0.882	8.00	24.50	
	Semi urban	15	10.60	2.444	11.00	53.90	
	Urban	64	10.88	2.616	11.00	55.74	
	Town	12	8.33	3.962	9.50	37.79	
Purchase Behavior							.001
	Rural	9	13.44	2.068	13.00	21.11	

Semi urban	15	17.33	3.812	18.00	46.73
Urban	64	18.91	5.108	19.00	59.28
Town	12	14.50	4.286	14.00	30.42

Through Mann-Whitney U and Kruskal-Wallis tests, the null hypothesis (H01) is rejected.

H01 There is no influence of social demographic variable (Gender, Area) on total score on Brand awareness, consumer attitude, interactive Behavior, Pre purchase Behavior and Purchase Decision.

11.7 Correlation analysis

Table 9: Correlation between variables: *Pearson correlation matrix*

Variable	Brand Awareness	Consumer Attitude	Interactional Behavior	Pre-Purchase Behavior	Purchase Decision
Brand Awareness	1				
Consumer Attitude	.568**	1			
Interactional Behavior	.236*	.331**	1		
Pre-Purchase Behavior	.564**	.538**	.331**	1	
Purchase Decision	.777**	.758**	.225*	.643**	1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

To study the relationship between the variables such as Brand Awareness, Consumer Attitudes, Interactional Behavior, Pre purchase behavior and Purchase Decisions, Pearson correlation analysis was conducted. The table 9 shows the result obtained from this test was, all the variables have statistically significant relationships between them.

Brand awareness showed a strong positive correlation with purchase decision ($r=0.568$, $P < 0.01$), indicating that people with higher brand awareness are significantly more likely to make a purchase. Additionally, Brand awareness also has positive correlation with consumer attitude ($r=0.568$, $P < 0.01$) and Pre purchase behavior ($r=0.564$, $P < 0.01$) shows that greater awareness increases likelihood of exploring the product before purchase. There is a weaker relationship observed between brand awareness and interactive behavior but still it is significant. Similarly, Interactive behavior has a weaker relationship between purchase decisions, but it is statistically significant at 0.05 level. Overall, all the variables have an excellent relationship between each other, and all the variables have sequential relationships except Interactional behavior on pre-purchase behavior which supports the null hypothesis 2,3,4,5,6 to be rejected.

11.8 Multicollinearity testing

Table 10: Multicollinearity test; VIF score

Variables	VIF	Interpretation
Brand Awareness	1.730773	No multicollinearity
Consumer Attitude	1.699205	No multicollinearity
Interactional Behavior	2.903278	Low multicollinearity
Pre purchase behavior	3.187599	Moderate multicollinearity

Before conducting linear regression, it is very crucial to check the presence of multicollinearity among the explanatory variables. Multicollinearity refers to a relationship in which two or more explanatory variables are highly correlated with each other which can inflate the standard errors and statistical significance of the predictor in the regression model. To evaluate the multicollinearity, the VIF score is calculated for each independent variable, as shown in table 10.

11.9 Simple linear regression

Simple linear regression was conducted to understand the relationship between a single independent variable which is called as predictor and a dependent variable which is also called as outcome variable. The regression equation is $y = \beta_0 + \beta_1x + \varepsilon$ in which y is dependent variable, x is independent variable, β_0 is intercept means value of dependent variable when x is 0, ε refers to error which means difference between predicted value and actual value. Through this model the null hypothesis 2,3,4,5,6,7,8 was rejected.

Table 11: Model summary: Brand awareness towards consumer attitude

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.322	.315	.660	<.001

The model obtained value in Table 11 has an R square value of 0.332, which indicates that 32.2% of the variation in purchase decision is explained by brand awareness. This specifies a moderate level of explanatory power which is acceptable in social science and behavioral research

Table 12: Coefficient: Brand awareness towards consumer attitude

Variables	Coefficients Beta	t	Sig.
Constant	10.011	5.186	<.001
Brand Awareness	0.687	6.826	<.001

a. Dependent Variable: Consumer Attitude

In the Table 12, the equation derived from this model is $Y = 10.011 + 0.687 * X$, where Y is the predicted Consumer attitude score, and X is the brand awareness score. 0.687 is the regression coefficient (slope), this means for every 1 unit increase in brand awareness, the consumer attitude score is increased by 0.687 units. Thus, brand awareness is shaping the consumer attitude.

Table 13: Model summary: Consumer attitude towards Interactive Behavior

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.110	.101	1.289	<.001

As mentioned in the Table 13 Consumer attitude is an independent variable and Interactive behavior is a dependent variable. The model obtained an R square value of 0.110, which indicates 11.0% of the variation in interactive behavior is explained by consumer attitude. While this R square value is lower, it is quite common in behavioral research.

Table 14: Coefficient: Consumer attitude towards Interactive behavior

Variables	Coefficients Beta	t	Sig.
Constant	4.672	3.792	<.001
Consumer Attitude	0.183	3.475	<.001

a. Dependent Variable: Interactive Behavior

The Table 14 shows that derived regression equation of interactive behavior is $4.672 + 0.183 * \text{Consumer attitude}$. The regression coefficient is 0.183 and P value is <0.001 which indicates this result is statistically significant. For every 1 unit increase in Consumer attitude, the Interactive behavior score is increased by 0.183 units, which indicates consumer attitude is shaping the interactive behavior.

Table 15: Model summary: Consumer attitude towards Pre purchase behavior

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.289	.282	1.401	<.001

As shown in Table 15, Consumer attitude is an explanatory variable, and pre purchase behavior is an outcome variable. The result extracted R square value of 0.289, which indicates 28.9% of the variance in pre purchase behavior is explained by consumer attitude. This specifies a moderate level of explanatory power, which is good in behavioral studies. The P value of this model is <0.001 which is statistically significant.

Table 16: Coefficient: Consumer attitude towards Pre purchase behavior

Variables	Coefficients Beta	t	Sig.
Constant	3.873	3.700	<.001
Consumer attitude	0.282	6.317	<.001

a. Dependent Variable: Pre purchase behavior

The derived regression equation is $\text{pre purchase behavior} = 3.873 + 0.282 * \text{Consumer attitude}$. As shown in Table 16 the coefficient for Consumer attitude is 0.282 with P value of <0.001 which is statistically significant. For 1 unit increase in Consumer attitude, the Pre purchase behavior is increased by 0.282 units, which indicates consumer attitude tends to demonstrate higher levels of pre purchase behavior.

Table 17: Model summary: Brand awareness towards Purchase decision

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.604	.600	.974	<.001

The table 17 shows that R square obtained from this model is 0.604 which means 60.4 percent of the variance in purchase decision is explained by brand awareness which has a very good explanatory power. The P value of the model is <0.001 which is statistically significant.

Table 18: Coefficient: *Brand awareness towards Purchase decision*

Variables	Coefficients Beta	t	Sig.
Constant	3.724	3.180	.002
Brand awareness	0.746	12.222	<.001

a. Dependent Variable: Purchase decision

The derived regression equation is Purchase decision = 3.724 + 0.746 * Brand awareness. As per the table 18 the coefficient for brand awareness is 0.746 with P value of <0.001 which is statistically significant. For 1 unit increase in Brand awareness, Purchase decision is increased by 0.746 units, which indicates Brand awareness is positively associated with purchase decision.

Table 19: Model summary: *Consumer attitude towards Purchase decision*

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.575	.571	1.222	<.001

The table 19 shows that R square obtained from this model is 0.575 percent which indicates 57.5 percent of the variance in consumer attitude is explained by purchase decision which has a good explanatory power. The P value of this model is <0.001 which is statistically significant.

Table 20: Coefficient: *Consumer attitude towards Purchase decision*

Variables	Coefficients Beta	t	Sig.
Constant	3.911	3.191	.002
Consumer attitude	0.602	11.517	<.001

a. Dependent Variable: Purchase decision

The derived regression equation is Purchase decision = 3.911 + 0.602 * Consumer attitude. As shown in table 20 the regression coefficient for brand is 0.602 with P value of <0.001 which is statistically significant. For 1 unit increase in Consumer attitude, Purchase decision is increased by 0.602 units, which shows that Consumer attitude can influence the purchase decision.

Table 21: Model summary: *Interactive behavior towards Purchase decision*

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.051	.041	1.001	<.024

As mentioned in Table 21, the model obtained an R square value of 0.051, which indicates only 5.1.0% of the variation in interactive behavior is explained by consumer attitude. The R square value is lower in this study, and it is not statistically significant with a P value of < 0.24. Thus, in this case the Interactive behavior cannot be used to study the Purchase decision.

Table 22: Coefficient: *Interactive behavior towards Purchase decision*

Variables	Coefficients Beta	t	Sig.
-----------	-------------------	---	------

Constant	14.785	11.191	<.001
Interactive behavior	0.324	2.287	0.24

a. Dependent Variable: Purchase Decision

The Table 22 shows that derived regression equation of interactive behavior is $14.785 + 0.183 * \text{Interactive behavior}$. The regression coefficient is 0.324 and P value is <0.024 which indicates that the result is statistically significant. For every 1 unit increase in Interactive behavior, the Purchase decision score is increased by 0.324 units, which indicates Interactive behavior is also one of the factors in influencing the Purchase decision.

Table 23: Model summary: *Pre purchase behavior towards Purchase decision*

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.414	.408	1.359	<.001

The regression model obtained in Table 23 has an R square value of 0.414, which indicates that 41.4% of the variation in purchase decision is explained by Pre purchase behavior. This specifies a moderate level of explanatory power which is acceptable in the aspect of behavioral research.

Table 24: Coefficient: *Pre purchase behavior towards Purchase decision*

Variables	Coefficients Beta	t	Sig.
Constant	7.607	6.079	<.001
Pre purchase behavior	0.974	8.320	<.001

a. Dependent Variable: Purchase decision

In the Table 24, the regression equation derived from this analysis is $Y = 9.607 + 0.687 * X$, where Y is the predicted Pre purchase behavior score, and X is the purchase decision score. 0.974 is the regression coefficient (slope), this means for every 1 unit increase in Pre purchase behavior, the Purchase decision score is increased by 0.687 units. Thus, pre purchase behavior is influencing the purchase decision.

11.10 Multiple linear regression

11.10.1 Ordinary least square model

Multi linear regression was conducted to test the null hypothesis 9 which states that there is no effect of Brand awareness, consumer Attitude, Interactional Behavior, Pre purchase behavior on Purchase Decision. In this model Purchase decision is dependent variable rest of them are independent variables.

Table 25: Model summary: *Brand awareness, Consumer attitude, Interactive behavior, Pre purchase behavior towards Purchase decision*

Model	R Square	Adj R Square	Durbin Watson	Sig.
1	.776	.767	1.249	<.001

As shown in the Table 25; The model has obtained an R square value of 0.776 which indicates that 77.6 percent of the variation in Purchase decision is explained by Brand awareness, Consumer attitude, Interactive behavior, Pre purchase behavior, which indicates a very good explanatory power.

Table 26: ANOVA: Brand awareness, Consumer attitude, Interactive behavior, Pre purchase behavior towards Purchase decision

Model		Sum of sq	Df	Mean sq	F	Sig
1	Regression	1411.510	4	352.877	82.318	<.001 ^b
	Residual	407.240	95	4.287		
	Total	1818.750	99			

The table 26 ANOVA test indicates that the overall model is statistically significant. The result shows that F value is 82.318 and P value is <0.001, which indicates that the variation explained by the regression model is statistically significant and proves not due to random aspects. The sum of square for regression is 1411.510 which is higher than residual sum of square (407.240) which indicates that the independent variables such as Brand awareness, Consumer attitude, Interactive behavior, Pre purchase behavior has a major proportion of the total variance in purchase decision.

Table 27: Coefficient: Brand awareness, Consumer attitude, Interactive behavior, Pre purchase behavior towards Purchase decision

Variables	Coefficients Beta	t	Sig.
Constant	-.020	-0.19	.985
Brand Awareness	.428	7.020	<.001
Consumer Attitude	.343	6.832	<.001
Interactive behavior	-.123	-1.631	.106
Pre purchase behavior	.284	2.974	<.001

a. Dependent Variable: Purchase decision

As shown in table 27 Brand awareness, Consumer behavior, Pre purchase behavior was statistically significant, and Interactive behavior is not statistically significant. The regression equation derived for the brand awareness towards purchase decision is purchase decision = -0.20 + 0.428 * brand awareness, which indicates every 1-unit increase in brand awareness, purchase decision is increased by 0.428 units. Similarly, every 1-unit increase in consumer attitude and Pre purchase behavior, purchase decision is increased by 0.343 and 0.284 units. Even though Interactive behavior is not significant individually, still the overall model is statistically significant. Therefore, through this multi linear regression test null hypothesis H09 is rejected.

H09 There is no effect of Brand awareness, consumer Attitude, Interactional Behavior, Pre purchase behavior on Purchase Decision.

11.10.1 Machine learning approach

This study used a machine learning approach to make a model and predict the outcome using Multiple linear regression (MLR) technique. The approach was initiated with importing the dataset using the pandas library. The explanatory variables such as (Brand awareness, Consumer attitude, Interactive behavior, Pre purchase behavior) were used to predict purchase decision. A constant value was added to the predictors which denoted the intercept of the model using the add_constant() function from the Statsmodel python library.

To ensure the unbiased model performance, the dataset was split into training and testing sets in a ratio of 30:70 using the train_test_split() function from the python library scikit-learn. This step helped to proceed with the further process.

Through the ordinary least square summary the independent variable ‘Interactive behavior’ was found to be statistically insignificant. So, it was removed to improve the model’s performance and reliability.

The trained model’s performance was measured using metrics such as R-square, Root Mean square error (RMSE) for both testing and training sets. These metrics helped in checking the model’s variability and accuracy. Visual evaluation was also conducted to

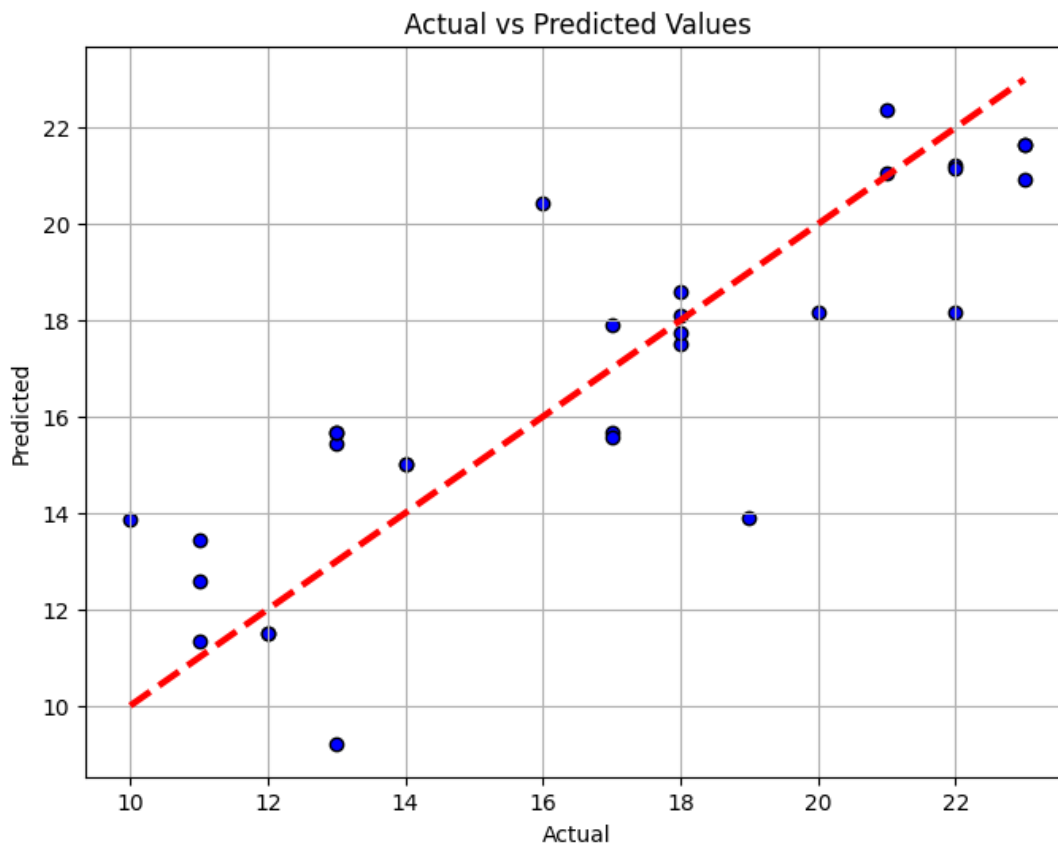
observe the actual versus predicted accuracy, residual errors, residual normality. Additionally, diagnostic check checks were conducted to identify the presence of multicollinearity using the variance inflation factor (VIF), Durbin-Watson statistic was conducted to check the autocorrelation in residuals. Overall, this process helps in building a regression model by ensuring the assumptions of multi linear regression were satisfied.

After conducting the analysis using test data, as shown in table 28 the study found that 73 percent of the variability in the Purchase decision was explained by the independent variables which is indicated by the Test R squared value. The Mean squared error and Root mean squared error indicating that the model has a reasonably good fit and predicting accuracy.

Table 28: Diagnostic check on test data; *R square, MSE, RMSE*

Metrics	Values
R-squared (Accuracy)	0.73
Mean Squared Error	4.68
Root Mean Squared Error (RMSE)	2.16

Diagram 3: Scatter plot of Actual vs Predicted value; *Test data*



The diagram 3 shows the relationship between the actual and predicted values of the purchase decision on the test dataset. The scattered point in the graph is not much deviated from the linear red line (representing the ideal position when actual value is equal to predicted value), which indicates there is a reasonable alignment between actual and predicted purchase decision. By considering these results, the data collection was started for proceeding further analysis with the train dataset.

Table 29: Actual versus predicted value comparison; *Train data*

Actual values	Predicted values
8	5.076547
18	15.525455
17	16.347071
18	18.708469
21	20.780439
17	15.944616
19	17.994350
23	21.783870
19	17.925678
17	17.925678

The table 29 shows the first ten rows of the actual and predicted values from the training dataset. The predictions were gained using the trained model by the predict () function (i.e., df_model_TR1.predict(X_train)). The table 29 shows that the predicted values are almost closer to the actual values, which indicates that the model has an effective prediction. There is a minor deviation between actual versus predicted values, but it is in acceptable position.

Table 30: Diagnostic check on train data; *R square, MSE, RMSE*

Diagnostic Metrics	Values
R-squared (Accuracy)	0.79
Mean Squared Error	3.86
Root Mean Squared Error (RMSE)	1.96

Table 30 shows the diagnostic metrics for the train data, where the model has explained 79 percent of the variance in the purchase decision by the brand awareness, Consumer attitude, and Pre purchase behavior. The obtained Mean square error and root mean square error were low, which indicates that the model has good predictive performance on the train dataset. Overall, due to high R squared value and low RMSE value the model is well fitted, which indicates that the model is not overfitting or underfitting. Therefore, the model is reliable in predicting the purchase decision based on the predictors.

Diagram 4: Scatter plot of Actual vs Predicted value; *Train data*

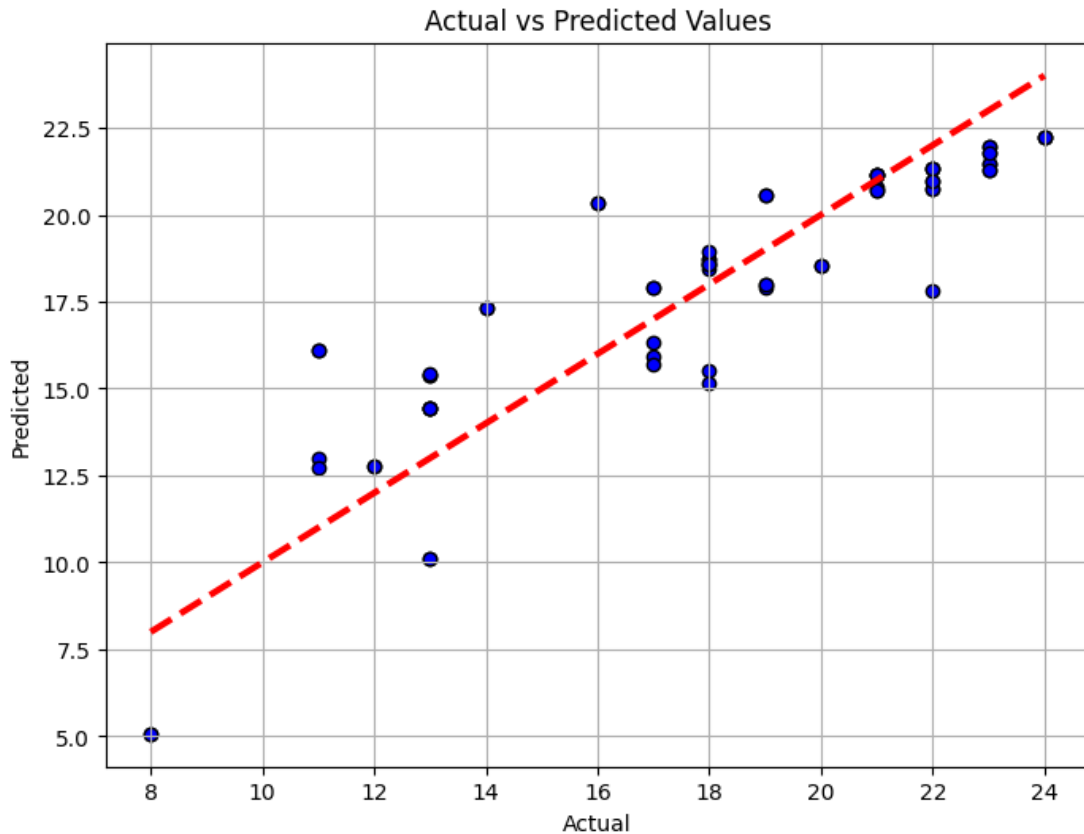


Diagram 4 shows the comparison between the actual and predicted values extracted from the training dataset. The red linear line indicates the ideal position of the actual values when the predicted values are perfectly matching. When the data points are distributed near the red line, the predictive accuracy of the model is excellent. In the Diagram 4, most points lie close to the line, which indicates that the model can predict outcomes with low error from the training data. Although, some outliers are observed, especially at the lower actual values range between 8 and 14, where the points are scattered widely. This indicates the model performs slightly lower compared to higher actual values range around 18, where the predicted values are positioned near to the ideal line.

Diagram 5: Residual plot

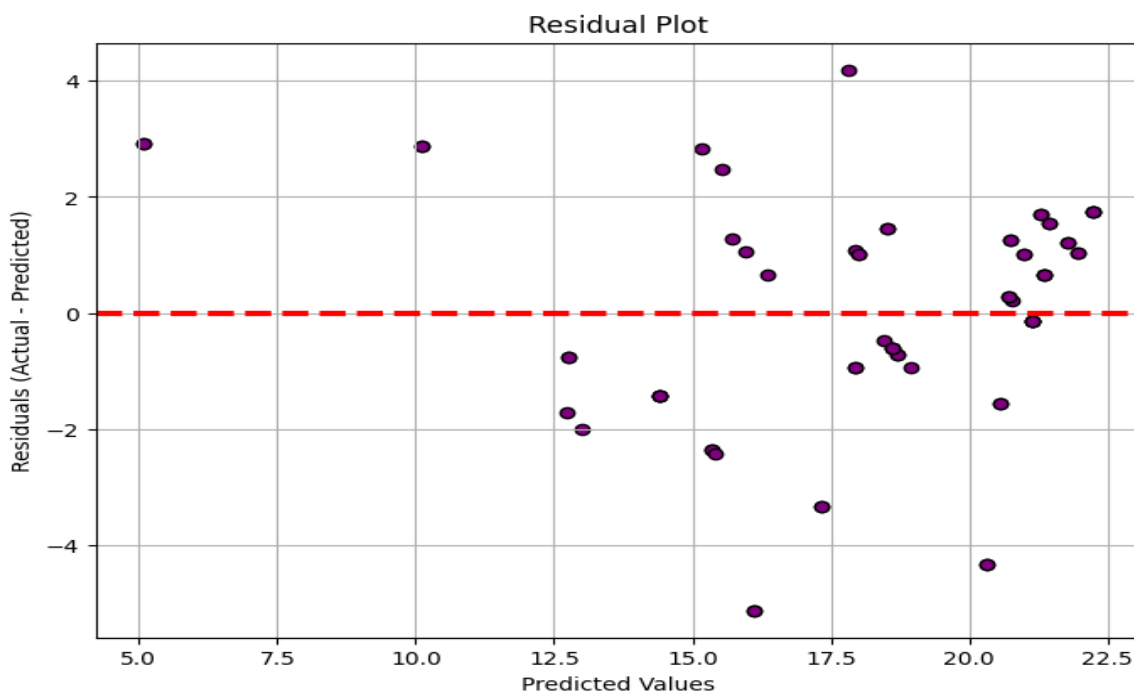


Diagram 5 shows the distribution of residuals across the Stright ideal line which illustrates the deviation of the residuals from zero. When the residuals are scattered around the zero line, the model does not suffer from the significant systematic bias. In this case, lower predicted residual values range between 5 and 15 has higher deviation which indicates the presence of slight variability. Meanwhile, higher predicted residual values range above 17.5 are more likely to be positioned around zero which indicates that the predicting accuracy is good in that range. Additionally, there is no clear pattern or trend such as funnel shape or curvature is formed by the residuals, indicating that the model has satisfied the homoscedasticity assumption.

Diagram 6: Residual distribution

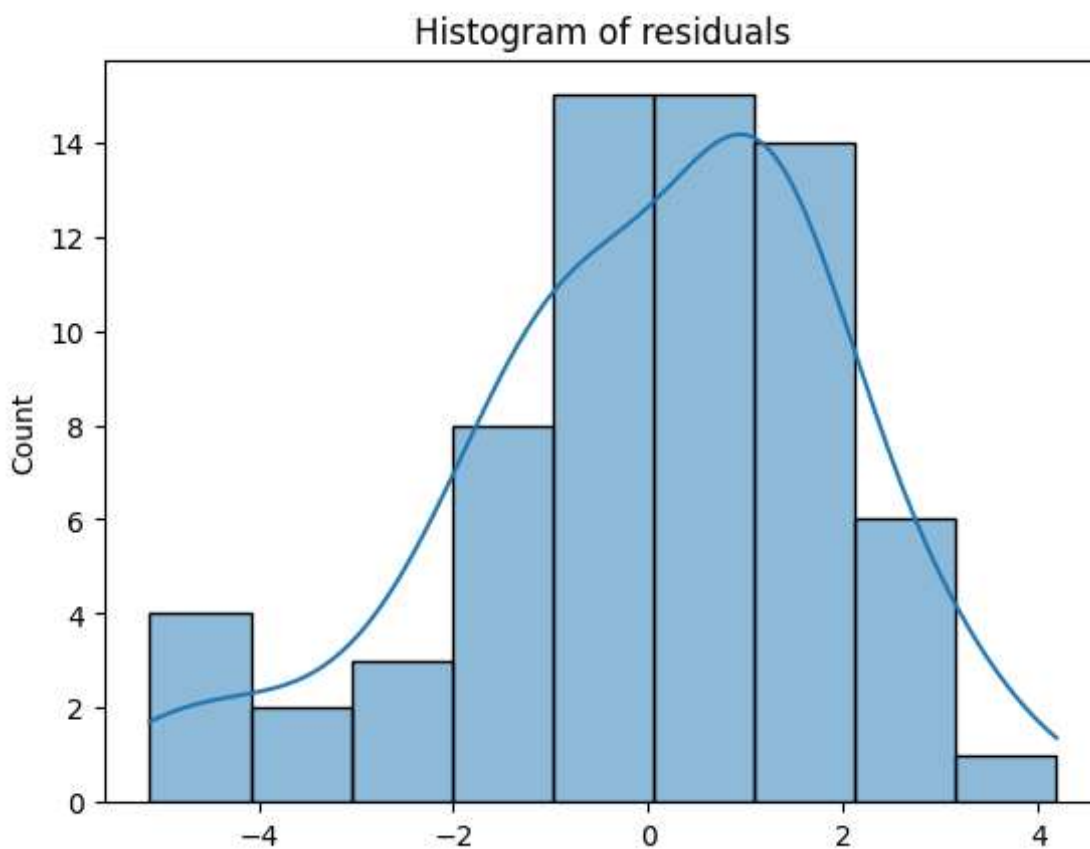


Diagram 6 shows the histogram with distribution of the residuals. Most of the residuals are centered around zero. The histograms are creating a bell-shaped curve which is critical in indicating that the residuals are normally distributed. Although, there is a slight deviation on the left side (negative residuals), indicating that the model has a minor skewness as shown in table 3. However, the overall pattern follows the assumption of normality, which is very crucial in evaluating the model performance.

Diagram 7: Quantile-Quantile plot

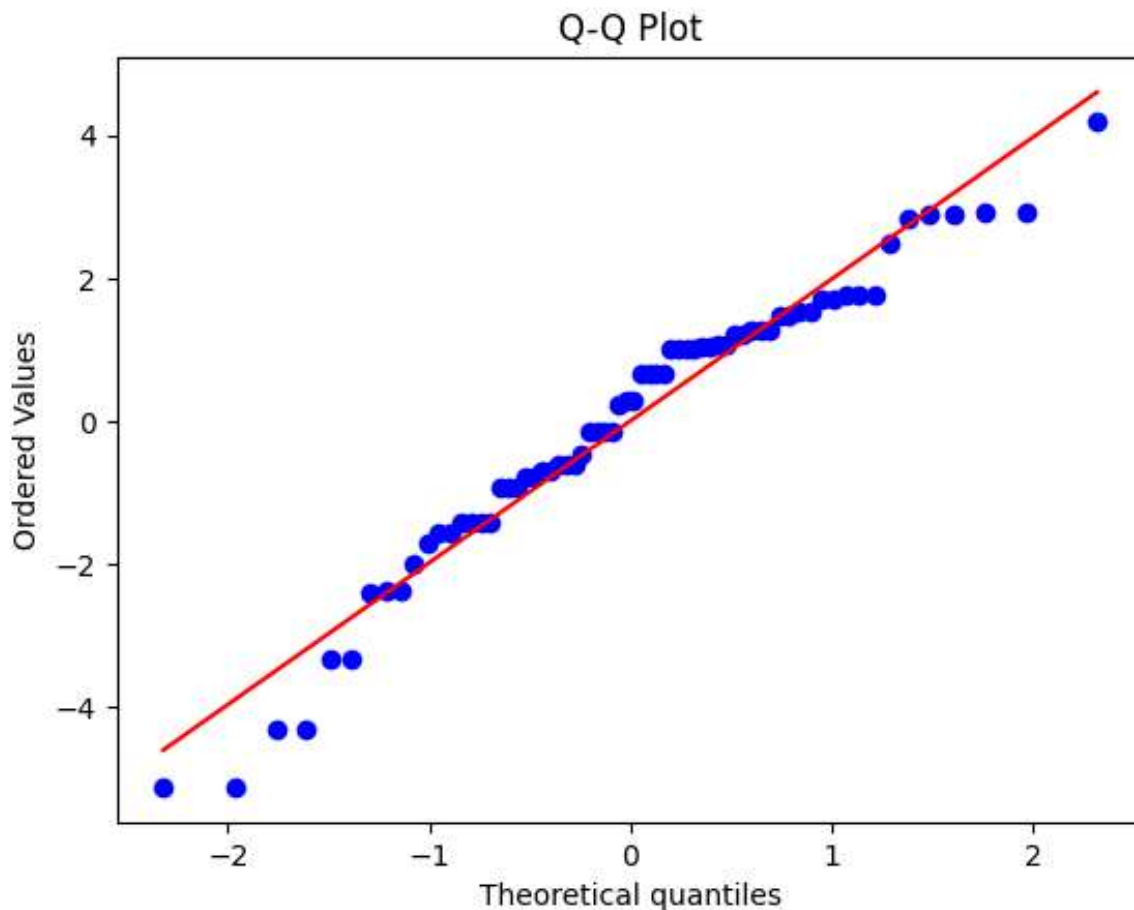
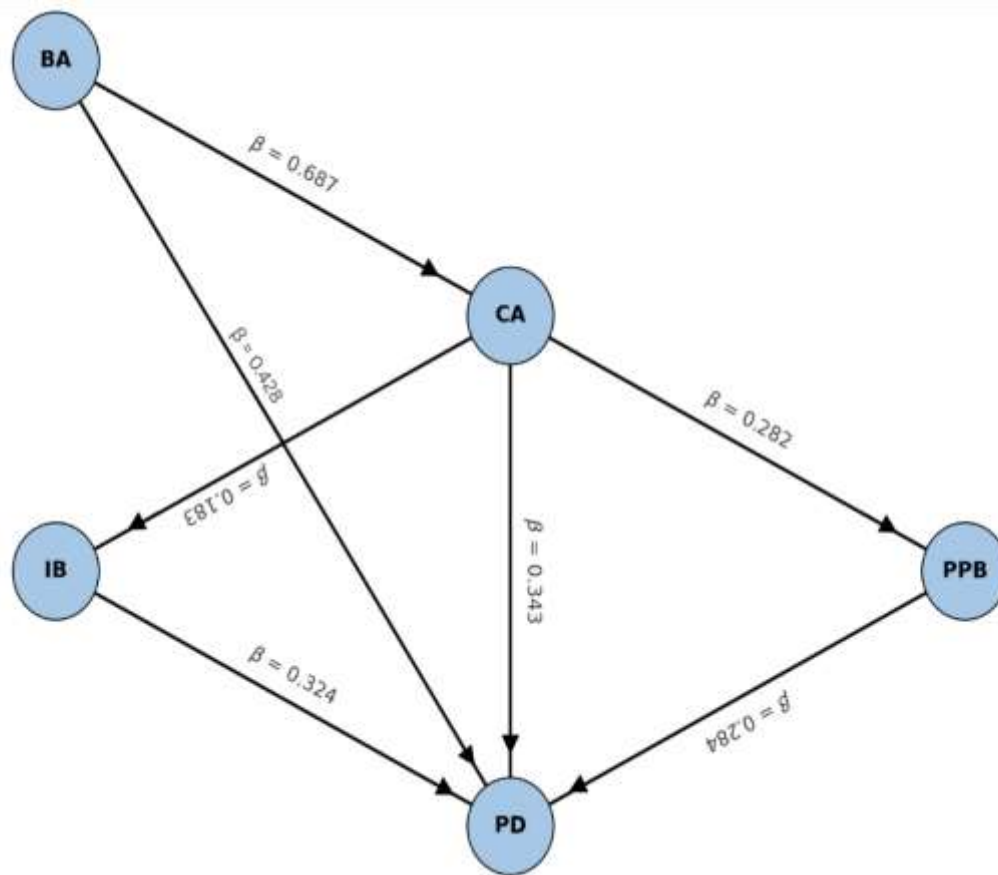


Diagram 7 illustrates the Quantile-Quantile (Q-Q) plot of the residuals, which measures the normality of the error distribution. In this plot, the ordered values (the sorted residuals from the train data) are plotted against the theoretical quantiles (the expected values if the data follows perfect normal distribution). The red linear line indicates perfect normality. The residual points are closely following the red linear line, especially in the central position of the distribution. Some minor deviations are noted at the lower and upper ends, which indicates slight variations in the normality. However, these deviations are normal and do not create an impact on the overall assumption of the normality.

11.11 Path analysis

Diagram 8: Path Diagram



As mentioned in table 27, the result obtained from this hypothesis testing shows several relationships between constructs. Brand awareness shows a significant relationship and influence on consumer attitude, with a higher path coefficient value of 0.687 and has a higher effect size of 0.474, which indicates brand awareness plays a crucial role in shaping the consumer attitude through short form video content. Similarly, consumer attitude affects interactive behavior and pre purchase behavior with a path coefficient of 0.183 and 0.282, which indicates medium to large effect on consumer behavior. Additionally, While Interactive behavior has a direct influence on purchase decision, it is not statistically significant (P value = 0.24), which indicates the engagement factor alone cannot lead to making a purchase decision.

Table 31: Hypotheses test and Result

Path	Path coefficients	P Val	R sq	Effect size	Effect interpretation	Description	Decision
BA --- CA	0.687	<.001	.322	.474	Large	Significant	Reject
CA --- IB	0.183	<.001	.110	.123	Medium	Significant	Reject
CA --- PPB	0.282	<.001	.282	.392	Large	Significant	Reject
IB --- PD	0.324	0.24	.051	.053	Small	Not Sig	Not Reject

PPB --PD	0.284	.004	.414	.706	Large	Significant	Reject
BA ---PD	0.428	<.001	.604	1.525	Very Large	Significant	Reject
CA ---PD	0.343	<.001	.575	1.352	Very Large	Significant	Reject

12. Findings

The findings of the study produce a crucial understanding of the impact of short form video content on Generation Z's brand awareness, consumer attitudes, behavior, and purchase decisions. The insights gained through this research are found to be strongly beneficial in understanding how SFV content affects brand recall, consumer behavior, interactive behavior, pre-purchase behavior, and purchase decision of Generation Z. From the findings of the hypothesis testing, the study has confirmed that the null hypothesis H01 was rejected since it was demonstrated that social demographic variables of gender and area have a marked effect on brand awareness, consumer attitude, interactional behavior, pre-purchasing behavior, and purchase decision. This result is aligning with previous research that has shown demographic factors have a moderating influence on online participation behavior [22].

H02 was rejected, confirming that brand awareness significantly affects consumer attitude, as evidenced in a path coefficient of 0.687 and a significant effect size of 0.474. These results are aligning with previous research which identified the importance of positive consumer attitudes which relatively increased the levels of involvement and information search behavior [14].

Interestingly, the interactive behavior was positively influencing the purchase decision (H06), which is not significant ($p = 0.24$), and therefore, H06 was accepted. This indicates that primary interaction elements such as liking, sharing, or commenting is not creating an impact on consumer's purchase behavior, which aligns with the previous research by McInnes (2024), who had highlighted that interactive behavior alone is not sufficient for converting the consumers to customers. Additionally, pre-purchase behavior (H05) has significantly influenced purchasing decision ($p = .004$), which rejects the null hypothesis, showcasing the process of pre-purchase behavior in consumer buying decisions. H07 and H08 were rejected, confirming that brand awareness and consumer attitude have strong influences on purchases with very high effect sizes (BA→PD is 1.525 & CA→PD is 1.352). The obtained results are showcasing the importance of building a strong significant brand awareness and favorable consumer emotions via short-form video platforms [19].

Furthermore, H09 was rejected by the analysis, which indicates that brand awareness, attitude, interactive behavior, and pre-purchase behavior are having a significant role in influencing the purchase decision. Pearson correlation matrix highlighted the crucial correlations between these variables in which there was a highest correlation between purchase decision and brand awareness ($r = 0.777$, $p < 0.01$). Finally, diagnostic testing confirmed that multicollinearity was not a problem ($VIF < 3.2$), and residual analysis showed that normality and homoscedasticity assumptions were satisfied. Machine learning diagnoses, i.e., R-squared values (Train: 0.79, Test: 0.73), confirmed that the models possess robust predictive capabilities. Overall, the study shows the power of short-form video content as a strong marketing channel among Gen Z by utilizing the significant influence of primary factors in the consumer purchase journey from awareness to end-purchase decision. The findings highlight the cruciality of prioritizing emotional engagement, transparent communication from the brand, and focusing on information-rich content rather than interactive attributes.

13. Practical implementations

These findings of the study provides several crucial practical strategies to the businesses and marketers who can enhance their digital marketing performance via short form video content, especially when targeting the generation Z consumers.

13.1 Focus on building strong brand awareness

While brand awareness created through short form videos was found to be statistically significant in influencing both consumer attitude and purchase decision, marketers should concentrate on consistent and recognizable branding in short form video content [4]. This can be achieved by adding memorable logos, catchy taglines, and storytelling which emotionally connects with the viewers.

13.2 Design videos that influence consumer attitude

This study found that consumer attitude is positively affecting the interactive and pre purchase behavior. So, marketers should focus on creating short form videos which are relatable, provide clear and sufficient information, product comparisons, real customer testimonials or success stories, and be transparent. Videos should show authenticity, creativity and transparency to encourage consumer sentiments [14].

13.3 Promote Pre purchase engagement

As pre-purchase behavior has a strong relationship with purchase decisions, brands should add elements in their short form videos that engage viewers to seek more information, visiting websites, and reading comments. Adding call to action (CTA), interactive links, and easy navigation options for the consumers while watching short form videos can guide them in further purchase decisions.

13.4 Move beyond simple engagement metrics

Meanwhile, interactive behavior like likes, shares and comments are crucial in defining success of the video, this research found that they alone do not guarantee the sales. Thus, brands should not solely focus on interactive aspects. Additionally, brands can focus on optimizing the deeper behavior such as trust, authenticity, limited time offers, website clicks, and information.

13.5 Personalize content based on consumer demographics

Since demographic factors such as gender and area showed positive influence on brand awareness and consumer behavior. Thus, the marketers should focus on personalizing the video content based on the target segments. When there is a small variation on message, visuals in short form video can make the content more relatable and impactful.

14. Conclusion

This research aimed to study how short-form video (SFV) content affects brand awareness, consumer attitude, interactive behavior, pre-purchase behavior, and purchase decision among Generation Z consumers. By analysing the survey data and using diverse statistical methods, the findings show that short form videos play a vital role in creating a strong and positive influence at different stages of the consumer purchase journey. Primarily, the study found that brand awareness created via short form videos are significantly impacting the consumer attitude. This result was aligned with the previous research by [4]. which states that when consumers are repeatedly exposed to brands shown in short form videos, they are more likely to develop stronger positive perceptions about the brand. A strong consumer attitude was found to be positively affecting the consumers in terms of trust, feelings, brand recall and audacity. These findings are consistent with previous research which underscore generation Z's need for authenticity and emotional connection in advertisements.

Meanwhile, interactive behavior such as liking, sharing, commenting and link clicking on short form videos was corresponding with pre purchase behavior and not with the final purchase decision. This is a crucial insight because it shows that only engaging the consumers in short form video, marketers and influencers cannot convert it into sales. This result was supported by the previous study by [9]. which states that interaction must be combined with positive brand experiences and informative content to influence actual buying actions. This study also found that pre-purchase behavior plays a major role when consumers spend time on knowing the product details and reading reviews after watching a short form video, they are more likely to consider the product if they have positive results. Thus, creating short videos which promote product exploration and provide clear and genuine information can be a victorious strategy for brands [21]. Moreover, demographic factors such as gender and area have a positive impact on the generation Z purchase journey, thus the study suggests that short-form video can be personalized according to gender and area while marketing the brand.

Overall, this study provides strong evidence that short form video is a powerful marketing tool to promote the product by influencing generation Z consumers. Stages in consumer purchase journey from brand awareness to the final purchase decision should be focused carefully while creating the short form video. In this competitive era, it is very essential for a brand to consider aspects such as authenticity, reliability, informativeness, loyalty, trust and emotional engagement in their short-form videos to achieve and retain consumers. This study offers valuable insights to the developing digital marketing field and ensures that short form videos are not just trend makers, it can connect and influence the generations. Future marketers and researchers should explore how to enhance the effectiveness of short form video by personalizing the content.

15. Reference

- [1] Shatto B, Erwin K. Moving on From Millennials: Preparing for Generation Z. J Contin Educ Nurs. 2016 Jun 1;47(6):253-4. [doi: 10.3928/00220124-20160518-05](https://doi.org/10.3928/00220124-20160518-05).
- [2] Wang, W. (2024). Exploring the impact of social media short-form videos on adolescents' psychological and psychosocial well-being. SHS Web of Conferences, 199, 02016. <https://doi.org/10.1051/shsconf/202419902016>
- [3] Statista. (2025, March 26). Most used social networks 2025, by number of users. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number>
- [4] Stepanova, D. (2024). EFFECTIVENESS OF SHORT-FORM VIDEO MARKETING IN SOCIAL MEDIA STRATEGY. In Patteri Entrepreneurship Society, Patteri Entrepreneurship Society (p. 81 pages) [Thesis]. https://www.theseus.fi/bitstream/handle/10024/859812/stepanova_daria.pdf?sequence
- [5] Jayatissa, K. a. D. U. (2023). Generation Z – A New Lifeline: A Systematic Literature review. Sri Lanka Journal of Social Sciences and Humanities, 3(2), 179–186. <https://doi.org/10.4038/sljssh.v3i2.110>
- [6] How Video Consumption Is Changing in 2024 <https://blog.hubspot.com/marketing/how-video-consumption-is-changing>
- [7] De Liège - ULiège > Hec Liège > Hec Liège Wouters, S. U. (2024, September 1). A comparative study of long versus short video advertisements on social media and their impact on consumer purchase intention. MatheO - Master Thesis Online. <https://matheo.uliege.be/handle/2268.2/21576>
- [8] Zhao, J. (2023). THE EFFECTS OF SHORT VIDEO MARKETING ON CONSUMER'S PURCHASE INTENTION – a CASE OF FLORASIS CORPORATION, LTD. In Siam University, Siam University [Thesis; PDF]. <https://e-research.siam.edu/wp-content/uploads/2023/11/33.-Zhao-Jingnan-MBA-2023-IS-The-Effects-of-Short-Video-Marketing.pdf>
- [9] McInnes, A. (2024). Is short-form video content the answer to enhanced consumer engagement for solopreneurs? [Thesis]. University of Northumbria. <https://copyfolio.s3.us-east-1.amazonaws.com/clvy4t20f0000kmrh0krg4itp/clz0j2uijp000j0ck05xsw12tx.pdf>
- [10] Nguyen, T. T., Veer, E., & University of Canterbury. (2024). Why people watch user-generated videos? A systematic review and meta-analysis. In International Journal of Human-Computer Studies (Vol. 181, p. 103144) [Journal-article]. <https://doi.org/10.1016/j.ijhcs.2023.103144>
- [11] NOLEGEIN-Journal of Global Marketing. (2024). Research Gate. <https://doi.org/10.37591/njgm>
- [12] Islam, M. M., & Kabir, M. A. (2025). The Impact Of Short Video Content And Social Media Influencers On Digital Marketing Success: A Systematic Literature Review Of Smartphone Usage. AIM, 1(02), 01–17. <https://doi.org/10.70937/faet.v1i02.46>

- [13] Macwan, C., Chandni J. Vidani, & Jignesh Vidani. (2024). Exploring the influence of reels and short videos on the reading and listening habits of Generation Z: A comprehensive study. *International Journal of Business and Management Practices (IJBMP)*, 3–3, 309–330. <https://journal.multitechpublisher.com/index.php/ijbmp>
- [14] Araujo, C. J. G., 1, Perater, K. A. S., 2, Quicho, A. M. V., 3, & Antonio E. Etrata, Jr. (2022). Influence of TikTok video advertisements on Generation Z's behavior and purchase intention. *INTERNATIONAL JOURNAL OF SOCIAL AND MANAGEMENT STUDIES (IJOSMAS)*, 140–141. <https://www.researchgate.net/publication/358467845>
- [15] Morreale, C. (2022). The influence of social media video marketing on consumers purchase intentions: the role of video length. In L. Dessart (Ed.), *Master Thesis [Thesis]*. HEC-Ecole de gestion de l'Université de Liège. <http://hdl.handle.net/2268.2/14267>
- [16] Maenhout, W. (2022). LENGTH DOES MATTER: THE EFFECTIVENESS OF LONG-FORM VERSUS SHORT-FORM VIDEO MARKETING ACCORDING TO ADVERTISING CONTEXT AND VIEWER AGE. In L. Mulier & Laura De Kerpel, *Master's Dissertation [Thesis]*. https://libstore.ugent.be/fulltxt/RUG01/003/117/738/RUG01-003117738_2023_0001_AC.pdf
- [17] Hasbullah, N. N., Kiflee, A. K. R., Ramachandran, K., Anwar, S., & Sulaiman, Z. (2023). Advertising value of online advertisement on sustainable brand purchase intention: The moderating role of user-generated video among Gen Z in Malaysia. *International Journal of Electronic Commerce Studies*, 14(3), 49–68. <https://doi.org/10.7903/ijecs.2260>
- [18] Van Tuan, P., Chan, L. A., Huy, D. T. N., Anh, N. Q., Trang, N. Q., Giang, P. C., & Uyen, N. P. (2023, January 1). Effects of Advertising Value and Attitude on the Purchase Intention Through Short-Form Video Advertising on Social Media: Empirical Research on Gen Z in Vietnam [Slide show]. <https://doi.org/10.46431/mejast.2023.61>
- [19] Anagnostou, E. (2024). Viral marketing in the cosmetics industry: Strategies for creating viral short-form video ads on social media (By Haaga-Helia University of Applied Sciences) [Thesis]. https://www.theseus.fi/bitstream/handle/10024/855430/Anagnostou_Evangelia.pdf?sequence=2&isAllowed=y
- [20] McInnes, A. (2024b). Is short-form video content the answer to enhanced consumer engagement for solopreneurs? [Thesis]. University of Northumbria. <https://copyfolio.s3.us-east-1.amazonaws.com/clvv4t20f0000kmrh0krg4itp/clz0j2ujp000j0ck05xsw12tx.pdf>
- [21] Hammar, I.-N., & Johansson, S. (2024). Exploring the dynamics of Short-Format videos on social media platforms on brand perception and brand loyalty [Thesis].
- [22] Plötz, S. X. (2022). A GEN Z STUDY: THE INFLUENCE OF SUSTAINABLE TIKTOK VIDEOS ON CONSUMER'S ATTITUDE AND PURCHASE INTENTION IN GERMANY (By L. F. Martinez & L. M. Martinez).
- [23] La Rocca, N. (2024). A comparative study of long versus short video advertisements on social media and their impact on consumer purchase intention. In N. Steils, *Master Thesis [Thesis]*. HEC-Ecole de gestion de l'Université de Liège. <http://hdl.handle.net/2268.2/21576>
- [24] Buzzhubsocial. (2023, July 10). How Glossier leveraged User-Generated Content for Social Media success. *Medium*. <https://medium.com/@buzzhub/how-glossier-leveraged-user-generated-content-for-social-media-success-4b3bc74a4bcb>

16. Appendices

16.1 Survey questions

Note: SHV refers to short form video.

Table 32: Consumer Awareness: Questionnaire

Q no	Brand Awareness	Strongly disagree	disagree	Neutral	agree	Strongly agree

BA_1	I remember brands that promote sustainability in SFVs as more responsible.	1	2	3	4	5
BA_2	Trending challenges/memes make brands more memorable.	1	2	3	4	5
BA_3	I remember brands better when they collaborate with influencers.	1	2	3	4	5
BA_4	Seeing a brand frequently in short-form videos makes it more recognizable to me.	1	2	3	4	5
BA_5	I am more likely to trust a brand if I have seen it multiple times in different short-form videos.	1	2	3	4	5

Table 33: Consumer Attitude: Questionnaire

Q no	Consumer Awareness	Strongly disagree	disagree	Neutral	agree	Strongly agree
CA_1	I feel positive about SFVs when influencers recommend a product.	1	2	3	4	5
CA_2	I trust brands more when their short-form videos provide clear and sufficient information about the product.	1	2	3	4	5
CA_3	I feel positive about SFVs that include direct product comparisons (e.g., brand A vs. brand B)	1	2	3	4	5
CA_4	I feel positive about SFVs when they show real customer testimonials or success stories.	1	2	3	4	5
CA_5	Transparent pros/cons videos reduce perceived risk in short form video.	1	2	3	4	5
CA_6	I feel consistent branding (colors, logos) improves brand recall.	1	2	3	4	5

Table 34: Interactive behaviour: Questionnaire

Q no	Interactive Behavior	Strongly disagree	disagree	Neutral	agree	Strongly agree
IB_1	I am more likely to interact with SFVs that include polls or challenges.	1	2	3	4	5
IB_2	I interact more when SFVs have shoppable links (e.g., 'Tap to Shop').	1	2	3	4	5
IB_3	I am more likely to interact with a short-form video if it has positive views, likes, or comments.	1	2	3	4	5

Table 35: Pre purchase behaviour: Questionnaire

Q no	Pre purchase Behavior	Strongly disagree	disagree	Neutral	agree	Strongly agree
PPB_1	I discover new brands/products through short-form videos.	1	2	3	4	5
PPB_2	I find user-generated content (UGC) in short-form videos (SFVs) to be authentic and trustworthy.	1	2	3	4	5
PPB_3	Limited time offers in short form videos push me to buy immediately.	1	2	3	4	5

Table 36: Purchase Decision: Questionnaire

Q no	Purchase Decision	Strongly disagree	disagree	Neutral	agree	Strongly agree
PD_1	How likely are you to purchase a product after seeing it in a short-form video?	1	2	3	4	5
PD_2	I prefer purchasing products featured in short-form videos over those you see in traditional advertisements?	1	2	3	4	5
PD_3	When I have a good mindset with a short-form video, I am more likely to purchase the product	1	2	3	4	5
PD_4	I purchased a product because a short-form video made it look useful or necessary?	1	2	3	4	5
PD_5	The branding of a product significantly influences my decision to purchase it.	1	2	3	4	5

16.2 Python Libraries used in this study

Table 37: Python Libraries

S no	Libraries	Python code
1	Numerical python	import numpy as np
2	Panel data analysis	import pandas as pd
3	Mathematics	import math
4	Matplot Library	import matplotlib.pyplot as plt
5	Statistical model	import statsmodels.api as sm
6	Scientific python	from scipy import stats
7	Scientific python tool kit learn	from sklearn.model_selection import train_test_split from sklearn.metrics import mean_squared_error, r2_score

16.3 Abbreviations

Table 38: Abbreviations

S.no	Abbreviations	Expansions
1	BA	Brand awareness
2	CA	Consumer attitude
3	IB	Interactive behavior
4	PPB	Pre purchase behavior
5	PD	Purchase decision
6	SFV	Short form video
7	UGC	User generated content
8	KMO	Kaiser-Meyer-Olkin
9	PCA	Principal Component Analysis
10	VIF	Variance Inflation Factor
11	R ²	R-squared
12	RMSE	Root Mean Squared Error
13	Q-Q	Quantile-Quantile
14	Sig.	Significance
15	df	Data freedom
16	SLR	Simple linear regression
17	MLR	Multiple linear regression
18	Adj R ²	Adjusted R square
19	MWU	Mann-Whitney U Test
20	KWT	Kruskal-Wallis Test

16.4 Python code used for this study

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
import statsmodels.api as sm
from scipy import stats
import matplotlib.pyplot as plt
import math
import seaborn as sns
```

```
sfv= pd.read_excel(
'/Users/sivamugunthanashok/Desktop/Capstone project/RP SHV LIN REG DATASET TS (1).xlsx'
)
sfv.head()
```

```
x_features=sfv.columns
x_features=['D1_TS', 'D2_TS', 'D3_TS', 'D4_TS']
```

```
x_sfv=sfv[x_features]
x_sfv
```

```
#Adding constant to the model to build MLR
from sklearn.model_selection import train_test_split

x=sm.add_constant(x_sfv)
y=sfv["D5_TS"]
```

```
df_model_T1=sm.OLS(y,x).fit()
df_model_T1.summary2()
```

```
x_features=sfv.columns
x_features=['D1_TS', 'D2_TS', 'D4_TS']
x_sfv=sfv[x_features]
x_sfv
```

```
x=sm.add_constant(x_sfv)
y=sfv["D5_TS"]
```

```
# 80% training and 20% testing
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_state=42)

print("Shape of X_train:", X_train.shape)
print("Shape of X_test:", X_test.shape)
```

```
df_model_T2=sm.OLS(y_test,X_test).fit()
df_model_T2.summary2()
```

```
# Predict using the trained model on train data
y_test_pred = df_model_T2.predict(X_test)
```

```
# Create a table comparing actual and predicted for training data
test_comparison_table = pd.DataFrame({
    'Actual (Test)': y_test.values,
    'Predicted (Test)': y_test_pred
})
```

```
# Display first 10 rows
print(test_comparison_table.head(10))
```

```
# R-squared and Mean Squared Error for train data
test_r2 = r2_score(y_test, y_test_pred)
test_mse = mean_squared_error(y_test, y_test_pred)

print(f'Test R-squared (Accuracy): {test_r2:.2f}')

print(f'Test Mean Squared Error: {test_mse:.2f}')
print(f'Root Mean Squared Error (RMSE): {math.sqrt(test_mse):.2f}')

# Scatter plot of Actual vs Predicted
plt.figure(figsize=(8,6))
plt.scatter(y_test, y_test_pred, color='blue', edgecolor='k')
plt.plot([y_test.min(), y_test.max()], [y_test.min(), y_test.max()], 'r--', lw=3)
plt.xlabel('Actual')
plt.ylabel('Predicted')
plt.title('Actual vs Predicted Values')
plt.grid(True)
plt.show()

# Residuals = Actual - Predicted
residuals = y_test - y_test_pred

plt.figure(figsize=(8,6))
plt.scatter(y_test_pred, residuals, color='purple', edgecolor='k')
plt.axhline(0, color='red', linestyle='--', lw=3)
plt.xlabel('Predicted Values')
plt.ylabel('Residuals (Actual - Predicted)')
plt.title('Residual Plot')
plt.grid(True)
plt.show()

import pandas as pd
from sklearn.decomposition import PCA
from factor_analyzer import FactorAnalyzer, calculate_bartlett_sphericity, calculate_kmo
import numpy as np

# Load the Excel file
df = pd.read_excel("RP SFV SURVEY DATASET.xlsx", sheet_name="Sheet1")

# Filter columns starting with D1_, D2_, D3_
cols = [col for col in df.columns if col.startswith(('D1_', 'D2_', 'D3_'))]
data = df[cols].dropna()

# Bartlett's Test & KMO
bartlett_test = calculate_bartlett_sphericity(data)
kmo_all, kmo_model = calculate_kmo(data)
print("Bartlett's Test:\n", bartlett_test)
print("KMO Test:\n", kmo_model)

# Step 1: Fit PCA with enough components (use eigenvalue > 1 later to select)
fa = FactorAnalyzer(n_factors=len(cols), rotation=None)
fa.fit(data)
ev, v = fa.get_eigenvalues()

# Determine number of components with eigenvalue > 1
n_components = sum(ev > 1)

# Fit Factor Analyzer with Varimax rotation and selected components
fa = FactorAnalyzer(n_factors=n_components, rotation='varimax')
fa.fit(data)
```



```
rotated = pd.DataFrame(fa_rot.loadings_, index=cols)
rotated = rotated.applymap(lambda x: round(x, 3) if abs(x) >= 0.5 else '')
```

```
anova_table = sm.stats.anova_lm(model, typ=2)
print(anova_table)
```

```
X_const = add_constant(X)
```

```
# Step 5: Calculate VIF for each feature
```

```
vif_data = pd.DataFrame()
```

```
vif_data['Feature'] = X_const.columns
```

```
vif_data['VIF'] = [variance_inflation_factor(X_const.values, i) for i in range(X_const.shape[1])]
```

```
# Step 6: Display the VIF scores
```

```
print(vif_data)
```