文章版本：version1
Format: 1.5倍行距、12字、英文
撰寫時間：05/23 - 06/10
文章分工：

- abstract / introduction / discussion / Conclusion: Boa
- Literature Review: Phoebe ?
- Methodology:
  - data collection + how to measure variables: Roro ?
  - How to test hypo + Data Analysis Method:? 直接寫在 Data analysis 章節
- Data Analysis / Interpretation: Cindy + Leo

1. Abstract
2. A｜Introduction
3. B｜Literature review
   a. 參考們
      Roro: w6_110078512_邱若綺
      Boa: W6_109078517_林芝萱.pdf
      Phoebe: W6_110078503_張硯涵_Literature Review.pdf
      Leo: W6_110078509_施宇軒.pdf
      Cindy: W6_110078701_劉宇瑄.pdf
4. B｜Theoretical framework
5. methodology（實驗方法）
   a. B｜data collection + how to measure variables
   b. C｜How to test hypo + Data Analysis Method
6. C｜Result (Data analysis)
   a. Data cleaning
   b. Description Statistics - Random split vs Intention
   c. EDA (+Visualization)
   d. Normality & independence & Test
   e. H1 & H2
   f. Extra: Preference
7. A｜Discussion (特別是包含Limitation)
   a. 還有實驗 + 統計上的限制

Topic: Does Different Music Make People Willing to Pay More for Online Shopping?

Abstract ｜ Boa（完成🥰）

The purpose of this research is to explore whether different music creates different consumption behaviors in the e-commerce environment. The researchers invited 66 participants and formed them into 4 groups to listen to different music playlists, including fast, slow, classic, and popular music. While listening to music, participants were doing online shopping for 15 minutes on a fashion clothing website, and then feedback on the price they were willing for paying on this website. This research assumes that: In Hypothesis 1, groups receiving fast music had a significantly higher willingness for paying than groups receiving slow music under classic and pop types, and in Hypothesis 2, groups receiving classical music had a significantly higher willingness for paying than groups receiving pop music.

In order to test our hypotheses, the 4 groups were combined based on the hypotheses. That means, in H1, Group A and C were combined as the Fast Music Group, and so did Group B and D as the Slow Music Group. In H2, Group A and B were combined as the Classic Music Group, and so did Group C and D as the Pop Music Group. According to the results of the Wilcoxon test, it shows that both H1 and H2 are not significant. In addition to the limitation of the small number of samples and the experimental cost issues, the researchers also put forward discussions on this result to facilitate the iteration of future related research work.

Introduction ｜ Boa（資料來源待改🥰 - 還沒 **Gramarly**）

The use of e-commerce is increasing gradually. According to McKinsey's report, at least 2/3 of consumers in 9/13 countries surveyed had tried online shopping in 2020, and 65% of them said they intend to continue doing so. According to McKinsey's report (XXXX), in the U.S, it had been expected e-commerce penetration to grow 24% between 2019 and 2024, but it had reached 33% in July 2020. The e-commerce penetration in Latin America is also growing rapidly. One of the representative countries is Brazil. In the first half of 2020, the number of e-commerce app installations in Brazil increased by 99.3% compared with the same period in 2019 (Statista statistics, XXXX), and the size of the mobile e-commerce market grew to $13.43 billion, up 38.5% from 2019 (Euromonitor, XXX). As for Taiwan, the Ministry of Economic Affairs (XXXX) pointed out that the annual growth rate of online sales in the proportion of total retail turnover is 24.5%, which is 3.3% higher than the annual growth rate of overall retail turnover in 2021, which it's a record high.

However, under the fierce competition, we need to actively explore ways to improve the user's online shopping experience. The McKinsey report (XXXX) mentioned that consumers who shop online clearly lack brand loyalty. But, only 60% of businesses said they are basically ready for e-commerce. In addition, "trust" is more important in the e-commerce field than in traditional shop stores. It might

be because the current regulations on e-commerce in various countries are generally insufficient, and products purchased online are usually not immediately verified. (Gefen, D., & Straub, D. W., 2004) So, the goal of all e-commerce enterprises is to make good online shopping experiences and build strong confidence in their customers' minds.

While brick-and-mortar stores can manipulate five senses to influence consumers' behaviors, in the e-commerce field, music and colors are the strongest stimulus for consumers, and are also factors that companies can better grasp(Cheng, F. F., Wu, C. S., & Yen, D. C., 2009). Some studies have also pointed out that modern people are more able to receive "visual and sound" stimuli at the same time when consuming. There have been many studies presenting "Users' Interfaces (UI)" suggestions for online shopping websites to improve visual sense of consumers. But, the research of music stimulation in online shopping is relatively less(Klein, K., Melnyk, V., & Völckner, F., 2021). Therefore, we expect to take e-commerce as the theme, address a field study, and verify the influences of different music on participants' online shopping behavior changes. This research aims to test the willingness of participants to pay for online shopping of fashion clothes in fast, slow, classic, and popular music playlists, and to initiate suggestions for e-commerce website management.

Literature Review | Roro + Phoebe

The ambient difference between online and traditional shopping

Shopping online can significantly vary from the traditional pattern to manipulable factors of the shopping context, and Kotler(1973) has indicated conscious ambient design can positively influence consumer behaviors such as perception of store image, consumption, and time spending. Physical stores have atmospheric dimensions defined by Baker (1986), including in-store music(Milliman, 1982), ambient temperature and scents(Baker, 1986; Teller & Dennis, 2012), and interior design like lighting and furnishing(Golden and Zimmerman, 1986; Tantanatewin & Inkarojrit, 2016), the salesperson(Baker, 1986), and perceived crowding(Machleit et al., 2000). Online stores, in general, utilize e-commerce websites as a way to display products or services and interact with their customers. Given computer‐based environments of commercial sites, virtual experience, namely the design of user interfaces and interaction as well as certain sensory stimuli, is recognized as a critical cue for shopping (Huang, 2003; Cheng, Wu & Yen, 2007). Sensory, one module of experiential marketing, is found strongly associated with pleasant experiences (Schmitt, 1999). Specifically, Parsons (2002) found that auditory and visual stimuli are highly applicable to online shopping sites, and in the experiment, more than half of the participants perceived websites could provide visual stimuli for displaying products and auditory stimuli as the background stimulation. Additionally, music, a kind of aural stimulation, is one of the most cost-effective and

simple-to-customize factors to change the shopping environment(Yalch & Spangenberg, 2000).

The relationship between music and shopping

Music is viewed as an effective element to influence customers in physical stores(Milliman, 1982; Alpert & Alpert, 1990, Mattila & Wirtz, 2001; Wilson, 2003). Variations in music tempo, for instance, can have a considerable impact on the speed of in-store traffic flow(Milliman, 1982). And, when background music and the image of the store are congruent, purchase rate and time spent increase(Toldos et al., 2019).

Moreover, Mehrabian and Russell's (1974) stimulus-organism-response (S-O-R) model is widely used to study the relationship between stores' stimuli and customers' internal and external responses in mortar-and-brick stores. Then, Eroglu et al. (2001) created a conceptual model based on the S-O-R model to examine the impact of atmospheric conditions in online stores, in which there are affective and cognitive internal states in the period of the "O(organism)" stage. Affective response refers to three emotions that are Pleasure, Arousal, and Dominance (PAD), which are reactions influenced by atmospheric stimuli. However, due to Russell's (1979) study, many studies have decided to exclude the insignificant indicator, Dominance. A study found that the presence of music in the online shopping process significantly elicited customers' arousal valence (Chen et al, 2022), which shows music can also play a role on e-commerce platforms.

Responses to Fast Tempo Music

Music performed at a fast tempo is classified as "happy" songs, whereas those performed at a slow tempo are classified as "sad" songs, and high BPM music is not only considered more pleasing but more arousing than low BPM music (Bruner, 1990; Kellaris and Kent, 1991, 1994; Sweeney and Wyber 2002). According to Mehrabian and Russell's (1974) S-O-R model, when O (organism) is positive, S (stimulus) can directly affect R (responses). And, it has been identified that the response of purchase intent is positively influenced by emotion(Kim and Lennon, 2013). Aside from emotional perception, fast tempo music had a shorter perceived browsing time and a higher shifting frequency than slow tempo music when doing online shopping, and the effect of fast music tempo is recognized as an arousal inducer to motivate customers to process information (Chen et al., 2009). Hence, we can infer that

fast-paced music is more impactful than slow-paced music in increasing purchase intention. The following is our H1:

H1: Groups receiving fast music had a significantly higher willingness for paying than groups receiving slow music under classic and pop types.

Responses to Classical Music

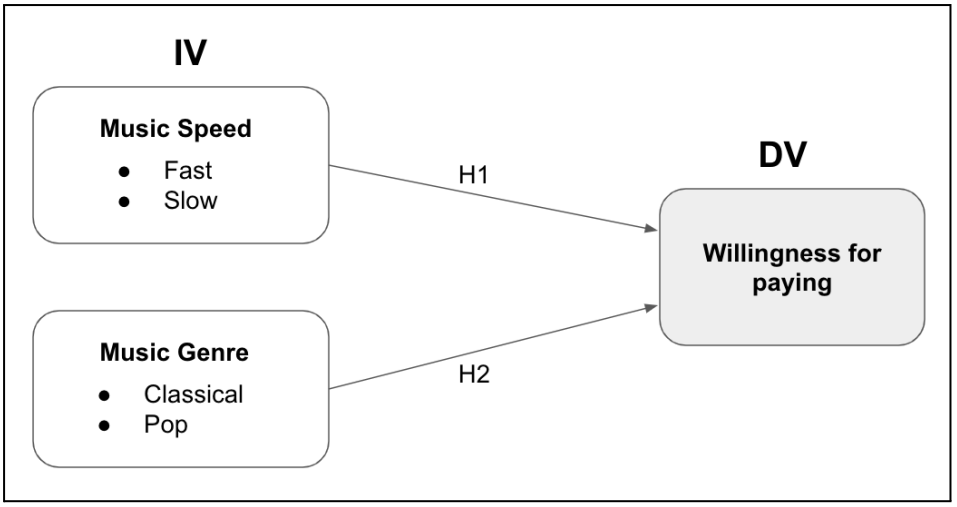
When it comes to the three dimensions of emotions of Mehrabian and Russell's (1974) S-O-R model, studies have shown pop music is more arousing, and classical music is more pleasing (Kellaris and Kent, 1994; Sweeney and Wyber 2002). However, classical music is associated with the concept of formality and high socioeconomic status (Stone, 1983). Yalch and Spangenberg (1990) found that classical music makes shoppers feel the price of the product should be higher, and consumers who were exposed to classical music had significantly increased sales (Areni and Kim, 1993). Plus, respondents perceive higher service quality when classical music is played alone or with soft lighting than pop music (Baker et al., 1994; Sweeney and Wyber 2002). Therefore, we reasoned that classical music could make participants pay more for shopping. The following is our H2:

H2: Groups receiving classical music had a significantly higher willingness for paying than groups receiving pop music.

- 其他會影響的
  網站的知名度、音樂的偏好 會對於購買情緒和認知有影響力。(Sweeney and Wyber, 2002; Chang and Chen, 2008; Hwang and Oh, 2019)

Contribution ***DONE*** 全部完成

In this paper, our contribution is divided into three parts. Unlike most previous literature discussing the relationship between brick-and-mortar stores and music consumption (Wilson, 2003; North, Shilcock, and Hargreaves, 2003; Dubé and Morin, 2001), we switch to an online website with background music so that managers can use different music in the booming e-commerce in the future. Furthermore, we evaluated the variety of music that could be used for future commercial purposes, since previous research had only looked at the presence or absence of music as a factor in the propensity to buy clothing online (Kim, Kim, and Lennon, 2009). Most importantly, unlike past literature that mostly used laboratory experiments to watch videotapes (Sweeney, and Wyber, 2002; Broekemier, Marquardt, and Gentry,2008), we used field experiments to increase the presence.



Research Design/Methodology

Methodology *DONE* 全部完成

The main purpose of this study is to design the willingness for paying influence by browsing online websites in different music contexts. And the field experiments are used to measure the effects of different variables.

Materials & Procedure

In this study, there were 66 participants. All of those people were between 20 and 30 years old. Participants were equally divided into four groups; those who had finished the experiment could receive a raffle prize worth NT2,000.

Before the experiment, they were asked to fill out a pre-questionnaire asking about their online shopping behavior, brand attitude, and music preferences. In terms of the music preferences survey, they were required to listen to eight 30-second songs and answer the questionnaires.

Then, participants were asked to complete a task. They were instructed to browse the same web for up to 15 minutes in a given shopping situation ("Recently, you're going to have a four-day and three-night graduation trip, so you come to Uniqlo's website to choose clothes to match."). In this section, we used Uniqlo's official website because the brand was well known and the style was simple. These features could reduce the impact while browsing the website. In addition, to make them more immersed in the music, each group had to listen to 3-4 songs within 15 minutes while browsing the website. Afterward, participants responded to a post-questionnaire that mainly asked about their willingness for paying.

Measurement DONE 全部完成

Pretest

The pre-questionnaire used for the experiment consisted of three parts. The first part focused on participants' online shopping behavior because Pahnila and Warsta (2010) discovered that e-commerce customer behavior was influenced by habit. The second part focused on brand attitude which was thought to influence brand preference and purchase intention (Chang and Liu, 2009). We adjusted the questions that contained six items (e.g., " I have positive opinions toward Uniqlo") to fit the experiments. And the questions were used by Chang and Liu (2009). The third part focused on music preference because many literature reviews show that it's an important measurement of the emotional evaluation of the service environment (Hui et al., 1997) and purchase intention. Broekemier et al (2008) found that liked/disliked music has a marginally significant direct effect on shopping intentions. To measure the music preference (Siebenaler, 1999), a five-point scale was used for respondents to express their level of likes. (1 = don't like it at all, 5 = like it very much)

Experimental design & Posttest

The experiment is a 2 (Speed: Fast / Slow) × 2 (Genre: Classical / Pop) between-subjects factorial design, as shown in Table ?.

The first independent variable was used to measure speed. According to Cheng and Yen (2009)'s research, we defined 72 BPM or lower as low speed and 94 BPM or more as high speed. The second independent variable was used to measure genre. According to Johnson (2002) classical music could be defined as follows," Musicologists use it to refer only to the music of the late eighteenth and early nineteenth centuries (Haydn, Mozart, Beethoven). " Therefore, we chose pieces played by Mozart as our classical music. In terms of pop music, the definition is close to that of Smith (2009) defined Top 40 songs in commercial positions as pop music. Therefore, we used the Taiwan Top 50 2022 song list on Spotify as our pop music. These two categories, speed and genre, matched each other. And all four different types of music have been used as background music.

The dependent measure collected in this experiment was the willingness for paying and to assess it, we used a post-questionnaire used by Babin and Attaway (2000). In addition, we adjusted the questionnaire (e.g., "Out of every NT3,000 I spend on Online Shopping, I spend _____ on the Uniqlo website.") to be more in line with our experimental situation.

6. Data Analysis and Hypothesis Testing | Cindy + Leo

According to the literature insights above, the homogeneity among the groups are vital.

By exploratory data analysis (section 6.1), we explore the homogeneity level of respondent's online shopping behavior, brand attitude, and music preferences among the groups.

The conclusion shows that

1. There is no significant difference among the treatment sets for the brand attitude and demographics. (section 6.1.1 & 6.1.4)
2. There are significant differences in the four treatments for online shopping habits. (section 6.1.2)
3. For the part of the music preference, we failed to provide the song with high-level homogeneity according to the questionnaire response. The correlation of the song in the same category provided an unsimilar preference score, which may have resulted from the possibility that people's music preferences consist of various viewpoints. (section 6.1.3)

In section 6.2, we combined the experimental treatments to form the groups in two hypotheses. And as section 6.1 indicated, there was a significant difference in online shopping habits, so to mitigate the between-group difference in it, we expelled the outliers from our data such that we can improve the homogeneity of participants.

After the group homogeneity improved, we discussed how we chose the method to testify the hypotheses in section 6.3. Then showed the results of hypothesis testing in section 6.4.

6.1 Exploratory Data Analysis

We received 66 samples in total. The quantities of each group are balanced. The statistics (mean, standard deviation, median) and the result of whether the variables varied between groups are shown in the separate sections below.

6.1.1 Demographics

Participants were asked about their age and gender. Due to the limited sample size, this study used the Kruskal-Wallis test to test if the groups have a significant age difference, and the Chi-squared test is applied to test if there were significant proportion differences in genders. The result showed no significant difference between the experimental groups in terms of age and gender. Table X summarizes the descriptive data.

6.1.2 Online Shopping Habit

Participants were asked "how much money do you spend on online shopping per month?". By using the Kruskal Wallis test between the four groups, we found a significant difference between them as shown in Table X. Simply speaking, Group A spends less, and Group B spends more on average. It was caused by the non-randomization of grouping. We thus combined our data according to the hypotheses and excluded outliers to alleviate this between-group difference which will be explained further in Section 6.2.

		Group A (Fast+Classical)	Group B (Slow+Classical)	Group C (Fast+Pop)	Group D (Slow+Pop)
Age	Mean	22.00	22.059	21.824	23.062
	S.D.	2.582	1.983	2.531	2.768
	Median	21.5	22	21	23.5
	P-value Kruskal-Wallis	0.4479			
Gender	Percentage	0.75	0.59	0.41	0.75
	P-value Chi-squared	0.1384			
Online Shopping Habits	Mean	578.125	2641.176	1085.294	1131.250
	S.D.	507.599	3750.843	1055.309	1245.642
	Median	500	1000	700	650
	P-value Kruskal-Wallis	* 0.01877			
Sample Size		16	17	17	16

- Gender is coded using a numeric scale, where male = 0 and female = 1. Meaning that 0.75 equals 75% of the participants were female.
- S.D stands for standard deviation.

6.1.3 Music Preference

The participants rated the four-song lists (Song List A, B, C, and D) by rating two songs from each song list. We averaged the ratings toward the four-song lists given by each experiment group in Table Y. After running Kendall's Coefficient of Concordance test separately on the four experiment groups, it showed that there was no agreement on which song list was favored more and less in Group A, B, and C. As for the participants in Group D, they had a low extent of agreement (W=0.301) on which song list was favored more and less. However, group D liked Song List B the most, which was not the song they listened to while undertaking the experiment.

	Group A (Fast+Classical)	Group B (Slow+Classical)	Group C (Fast+Pop)	Group D (Slow+Pop)
Song List A	3.22	3.21	3.18	3.38
Song List B	3.50	3.71	3.68	3.88
Song List C	3.41	3.53	3.18	3.44
Song List D	3.05	3.06	3.03	2.78
W	0.09	0.143	0.0838	0.301

P-value Kendall's Coefficient of Concordance	0.229	0.0628	0.234	*0.00236

Although Kendall's W showed that there were no agreements in Group A, B and C, we can't justify that the music preference of participants were controlled due to the lack of song homogeneity as Figure Z showed.

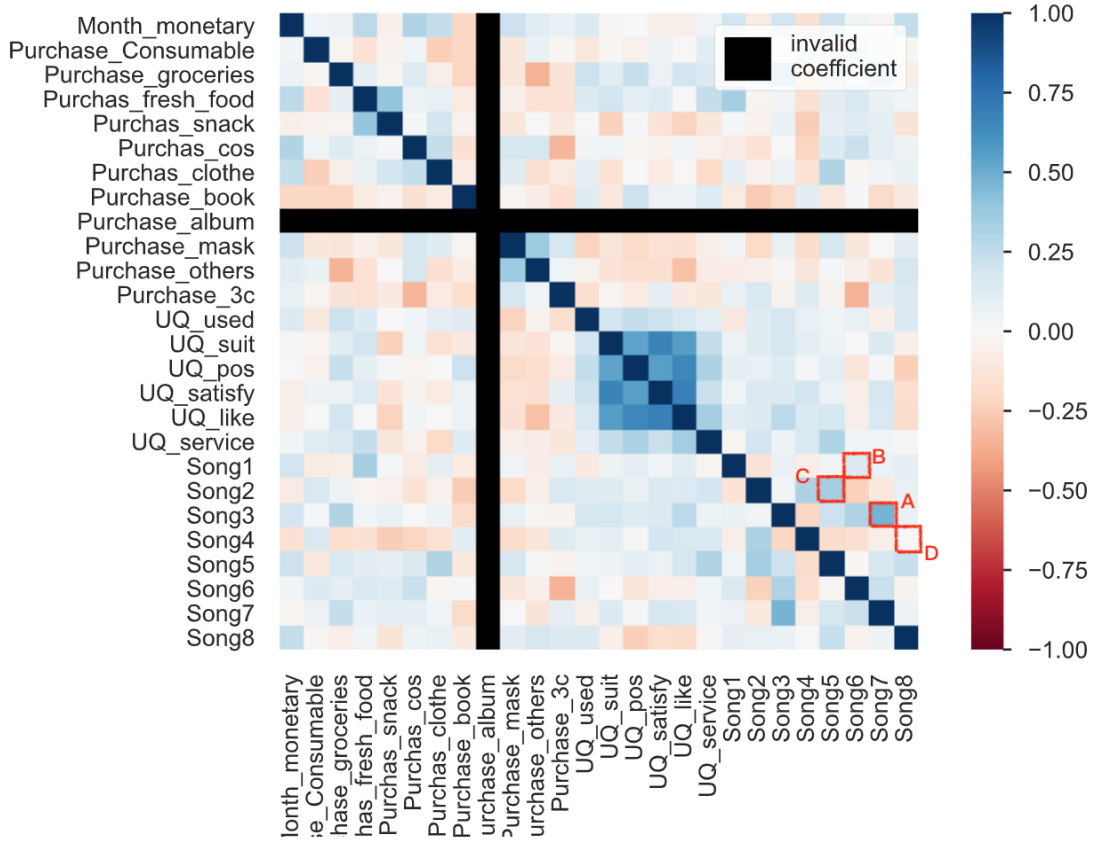


Figure Z

6.1.4 Uniqlo Brand Preference

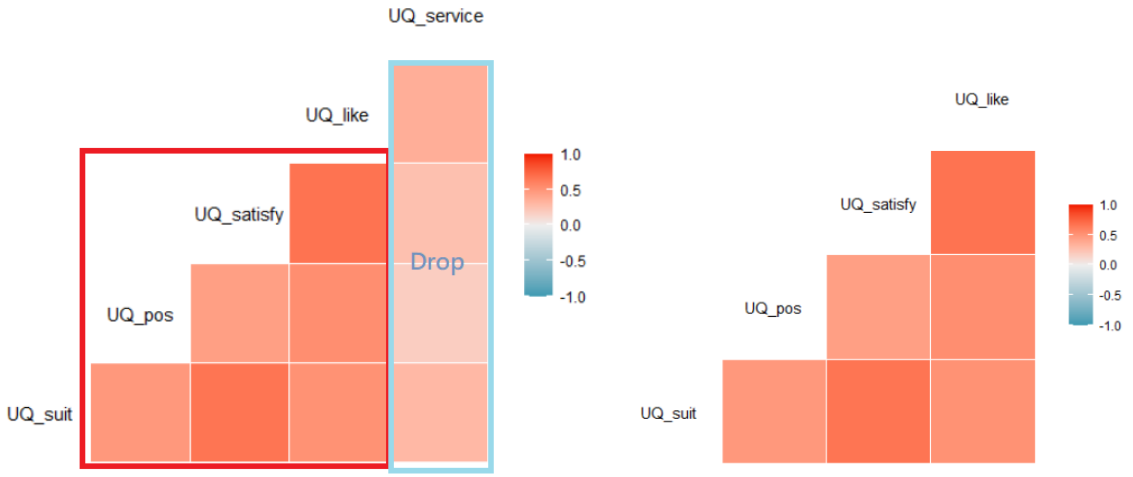


Figure 1: Dropping the UQ_service based on data driven viewpoint

From figure 1, it shows that the UQ_service is close to irrelevant, which resulted from the misleading questionnaire design. After excluding the UQ_service feature, Cronbach's α is 0.83, higher than 0.7. Therefore, we design to drop this feature.

After ordering the respondents in the same group aligned (Appendix: Figure 3), we can tell that there's no obvious light area in the chart. This can show that there is no serious effect between the experimental groups, which might result in a specific rating preference between the groups in response to the brand preference question.

According to Figure 4, the preference score is calculated as the sum-up score of the four questions, including UQ_suit, UQ_pos, UQ_satisfy, and UQ_like. Since there were several participants who had never used the Uniqlo website, their data were not included in the analysis as shown in Figure 4, where "Used Sample Size" is smaller than "Total Sample Size."

Group	Total	Average	Used Sample Size	Total Sample Size
A	169	15.363	11	16
B	172	14.333	12	17
C	224	12.286	14	17
D	203	13.230	13	16
Total	768	55.212	50	66

Figure 4: Grouping result

From Figure 5 (kindly refer to the appendix) - Density Plot of Grouping Total UQ_preference Score, to testify whether the means for each group were statistically significantly different, we would use the ANOVA.

Before estimating the ANOVA test, we made sure that the assumptions of the ANOVA are fulfilled.

Variance Test Summary: p-value = 0.3814 > 0.05 , we cannot reject H0 that the variance is the same among each group.

The p-value of the normality test is 0.04744 < 0..05, which suggests that the samples do not match the assumption of normality of ANOVA However, the one-way ANOVA is robust enough for data size (n > 30). And the normality increases after we do the further boosting to increase the data size. Hence, we can expect that the sample mean will reach normality based on the central limit theorem. Therefore, we conduct a One-way ANOVA as below:

	Degree of Freedom	Sum	Mean	F-value	P-value
Group	3	1923	6.410	1.016	0.394
Residuals	46	290.29	6.311		

Figure 6

The p-value=0.394 > 0.05, We are confident to say there isn't a statistical difference among the groups. With the testing result above, we can wipe the stains of uncontrol side-effects resulting from the brand preference in our case.

6.2 Data preparation

In order to test our hypotheses, we need to combine the groups because what we care about were fast versus slow music and classical versus pop music. Besides, if we group the samples into four groups according to the different song lists, there would be a significant difference in the online shopping habits. Hence, to test the hypotheses and mitigate this between-group difference, we combine them according to the hypothesized statement. That means, in Hypothesis 1, we combine Group A and Group C to form the Fast Music Group. We combined Group B and Group D to form the Slow Music Group. In Hypothesis 2, we combined Group A and Group B to form the Classical Music Group and combined Group C and Group D to form the Pop Music Group.

After we combined the groups, the difference was still large we then deleted outliers in each group. After examining the group differences using the Wilcoxon test, we can conclude that there were no significant between-group differences in terms of online shopping habits in the two hypothesis settings.

	H1	**H2**

		Fast Music (Group A+C)	Slow Music (Group B+D)	Classic Music (Group A+B)	Pop Music (Group C+D)
Online Shopping Habits **with outlier**	Mean	839.394	1909.091	1640.909	1107.576
	S.D.	862.478	2889.460	2872.538	1133.449
	Median	500	1000	700	700
Online Shopping Habits **without outlier**	Mean	667.742	1034.483	871.667	888.710
	S.D.	523.856	788.432	745.294	732.770
	Median	500	800	550	500
	Wilcoxon p-value	0.064		0.977	
Sample size		31	29	30	31

The outliers here were also expelled from the following hypothesis testing dataset to avoid their abnormal habit affecting how the dependent variable behaves.

At this point, we can finally say that there were no significant between-group differences in online shopping habits, brand preferences, and demographics. Hence these factors would not be the reason for the discrepancies in the dependent variable - the willingness for paying. As for the music preference, it should be improved in the future research.

6.3 Choose the statistical method for hypothesis testing

We chose to examine the median difference instead of the mean difference using the Wilcoxon two-sample test. (For a graphical explanation of the method choosing process, please kindly refer to appendix: Data Analysis - Flowchart.)

The reasons for this choice were:

1. Our data is not normally distributed, which disobeys the normality assumption of the ANOVA, which is used to examine whether there is a significant difference between two sample means.
2. The dependent variable is money-related, which would be strongly affected by the outlier. Therefore, the median is a better parameter to estimate.
3. The median in each group is much more homogeneous than the mean.

6.4 Hypothesis Testing

Here we estimate whether there is a significant difference in the median among the group (The median of Group A: 1000; B: 1000, C: 1200, D:1200. For density plot, kindly refer to the appendix- Figure: Median of WTP among the four groups.)

H1: Groups receiving fast music had a significantly higher willingness for paying than groups receiving slow music under classical and pop types.

Posttest Data		Fast (A＋C)	Slow (B＋D)
Willingness of consumers for paying	Mean	1188.065	1134.138
	Standard Deviation	783.375	690.505
	Median	1000	1000
Hypothesis Testing	Wilcoxon Test P-value (Fast - Slow)	0.4174	

For H1, the p-value of the Wilcoxon test is 0.4174 > 0.05. It indicates that we can not reject the null hypothesis that there is no significant difference between the median of the fast group and the slow group.(For density plot, kindly refer to the appendix- Figure: Median comparison in density for H1.)

H2: Groups receiving classical music had a significantly higher willingness for paying than groups receiving pop music.

PosTest Data		Classical (A＋B)	Pop (C＋D)
Willingness of consumers for paying	Mean	1089.333	1291.613
	Standard Deviation	718.110	807.226
	Median	1000	1200
Hypothesis Testing	Wilcoxon Test P-value (Classical - Pop)	Classical > Pop	0.8346
		Classical < Pop	0.169

As the p-value of H2 is 0.8346, we can not reject H0. It indicates that the group receiving classical music treatment had no significant difference in the willingness for paying compared to the pop music group. However, if we replace our original H2 with “Groups receiving pop music had a significantly higher willingness for paying than groups receiving classical music, the p-value (Classical < Pop) is 0.169, which is closer to the rejection region even though it’s still not significant. It goes against our previous assumption. (For density plot, kindly refer to the appendix- Figure: Median comparison in density for H2)

To summarize, both of the hypotheses are not significant. We provide no evidence for rejecting the null hypotheses.

H1	Groups receiving fast music had a significantly higher willingness for paying than groups receiving slow music under classical and pop types.	not support
H2	Groups receiving classical music had a significantly higher willingness for paying than groups receiving pop music	not support

Discussion/ Managerial Insights | Boa

7.1 Discussion of the Research Framework and the Experiment Design

7.1.1 The Reason of Insignificant of the Hypotheses

For H1, we assume fast music creates a higher willingness for paying than slow music while users' doing online shopping. The result of the pretest, which is set to measure the music preference of participants, is insignificant. It means that participants in this experiment had not distinguished different types of music. In controlling the homogeneity of music, this study lacks more precise or rigorous control. Therefore, the improvement should be a strict pilot test for a large number of participants in the homogeneity of music, and ensure that rhythm or type of music is distinguishable for them, then this hypothesis is built rigorously and meaningfully.

Furthermore, we also found the research mentioning that when classical music matches with BPM greater than 180, then this combination of music would bring an obviously great difference in people's emotions (Kellaris, J. J., & Kent, R. J., 1993). So, if the BPM of the music list used in the experiment is over 180, a more pronounced difference in participants' perceptions might be observed and highlighted.

For H2, classical music creates a higher willingness for paying than pop music while users' doing online shopping. Some researchers pointed out that "Classic music and classic goods will increase purchase intention" (Areni, C. S., & Kim, D., 1993). However, the price of classic products is usually higher and has larger differences than fashion products. We chose a fashion clothing website as the experiment field. The unit price of the products is low and the difference is small. That means it might not be compared with the classic commodities studied in the past. It is suggested that more detailed measurements are needed for improvements. For example, to test 2 kinds of websites, including both classical and fashion products, and test the differences between them. Also, the DV should include more dimensions for building a more balanced experimental framework and more precise assumptions.

7.1.2 The Improvement of the Dependent Variable

The dependent variable, "willingness for paying" was built based on related works (Yalch and Spangenberg, 1990), and used to measure the willingness of the participants to purchase on the fashion clothing website. Due to different experimental fields and frameworks, they may not be directly used in this research. Also, lots of researchers adopted PAD - including Pleasure, Arousal, and Dominance - to measure whether the hypotheses were supported, but it was not used in this research.

Based on the purpose of this research, exploring how to establish a better e-commerce environment to create a higher possibility of consumption, we believe that more diverse aspects can be used to

measure online shopping behaviors and willingness for paying. For example, creating a fake online shopping website could not only prevent brand preference from interfering with the results, but also allow participants to put the products into fake shopping carts, and to facilitate the subsequent calculation of the consumption amounts. This might be a more precise way to measure DVs and improve the accuracy of predictions.

7.2 Limitations

7.2.1 The Cost Issue

Due to the tight schedule, the number of participants in the experiment was lower than expected, and the random grouping was mainly performed according to the time available for participating in the experiment, but not in a rigorous randomized coding way. In addition, there's not enough time to test the real purchase behaviors of the participants. Under comprehensive balance and consideration, the experiment was designed to test "the willingness for paying" just by the questionnaire, not from real-life events.

Also, due to the device's limitations, the experiment was done remotely, not in person. So, it was difficult to ensure that all the participants were well following the regulations of the research team.

7.2.2 The Experiment Design

The experiment field was an official website of the fashion clothing brand, and it might lead to some potential brand preferences. It might be a better way to build a fake website that does not contain any brand image or color hints for relevant research to reduce the influence of brand preference. Also, the music type is another limitation of this experiment. The fast and slow music lists were limited to classical and pop music and did not include other types, like jazz, etc.

The most crucial factor is the gap between the experimental field and the actual situation. For example, (1) The assumption is that 15 minutes of music intervention could sufficiently influence purchase behaviors, but the reality may require longer-term intervention. (2) The natural environment might have more external influences, like friends' advice for purchasing, to interfere with users' decision-making.

- 八個指標

Conclusion｜Boa(完成🥰)

This study hypothesized that (1) fast music increases users' willingness for paying for online shopping more than slow music, and (2) classic music increases users' willingness for paying for online shopping more than popular music. The data was collected by a field experiment with 66 participants in 4 groups. The result shows that there's not enough evidence to verify that users' willingness for paying in the fast music online shopping environment is higher than that of slow music. Neither do users' willingness for paying in the classical music online shopping environment is higher than that of popular music. After data is analyzed, we still try to initiate integrated suggestions based on the contents of the literature review and this practical experimental experience. In addition to discussing the reasons for the insignificant, the results and suggestions are also useful for future works as an iterative reference.

References

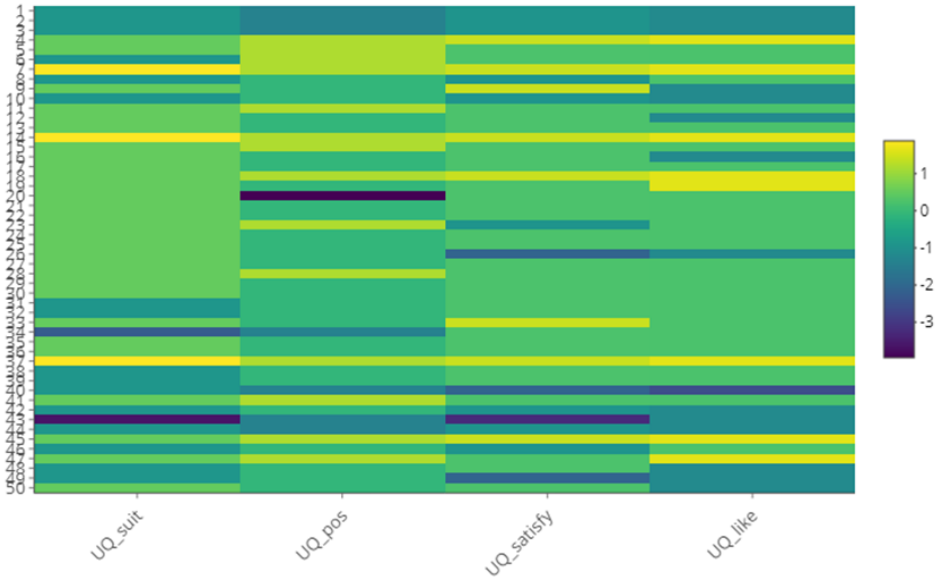
1. Areni, C. S., & Kim, D. (1993). The influence of background music on shopping behavior: classical versus top-forty music in a wine store. *ACR North American Advances*, 20, 336-340.
2. Babin, B. J., & Attaway, J. S. (2000). Atmospheric affect as a tool for creating value and gaining share of customer. *Journal of Business Research*, 49(2), 91-99.
3. Baker, Julie (1986), "The Role of the Environment in Marketing Services: The Consumer Perspective, " in John A. Czepeil, Carlole A. Congrarn and James Shanahan (eds.), The Service Challenge: Integrating for Competitive Advantage, Chicago, IL: *American Marketing Association*, 79-84.
4. Baker, J., Grewal, D., & Parasuraman, A. (1994). The influence of store environment on quality inferences and store image. *Journal of the academy of marketing science*, 22(4), 328-339.
5. Broekemier, G., Marquardt, R., & Gentry, J. W. (2008). An exploration of happy/sad and liked/disliked music effects on shopping intentions in a women's clothing store service setting. *Journal of Services Marketing*.
6. Bruner, G. C. (1990). Music, mood, and marketing. *Journal of marketing*, *54*(4), 94-104.
7. Chang, H. H., & Liu, Y. M. (2009). The impact of brand equity on brand preference and purchase intentions in the service industries. *The Service Industries Journal*, *29*(12), 1687-1706.
8. Chen, D., Zhang, X., Jiang, H., Meng, X., Zheng, J., Sun, L., & Zhang, K. (2022). How background music of shopping sites affects consumers during festival season. *Cognitive Computation and Systems*. 1-12
9. Cheng, F. F., Wu, C. S., & Yen, D. C. (2009). The effect of online store atmosphere on consumer's emotional responses–an experimental study of music and colour. *Behaviour & Information Technology*, 28(4), 323-334.
10. Dubé, L., & Morin, S. (2001). Background music pleasure and store evaluation: intensity effects and psychological mechanisms. *Journal of business Research*, *54*(2), 107-113.
11. Eroglu, S. A., Machleit, K. A., & Davis, L. M. (2001). Atmospheric qualities of online retailing: A conceptual model and implications. *Journal of Business Research*, 54(2), 177-184.
12. Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega*, *32*(6), 407-424.
13. Golden, L. G., & Zimmerman, D. A. (1986). *Effective retailing*. Houghton Mifflin.
14. Hul, M. K., Dube, L., & Chebat, J. C. (1997). The impact of music on consumers' reactions to waiting for services. *Journal of retailing*, *73*(1), 87-104.
15. Johnson, J. (2002). *Who needs classical music?: cultural choice and musical value*. Oxford University Press on Demand.
16. Kellaris, J. J., & Kent, R. J. (1993). An exploratory investigation of responses elicited by music varying in tempo, tonality, and texture. *Journal of Consumer Psychology*, 2(4), 381-401.

17. Kim, J., & Lennon, S. J. (2013). Effects of reputation and website quality on online consumers' emotion, perceived risk and purchase intention: Based on the stimulus-organism-response model. *Journal of Research in Interactive Marketing*, 7(1), 33-56.
18. Kim, J. H., Kim, M., & Lennon, S. J. (2009). Effects of web site atmospherics on consumer responses: music and product presentation. *Direct Marketing: An International Journal*.
19. Klein, K., Melnyk, V., & Völckner, F. (2021). Effects of background music on evaluations of visual images. *Psychology & Marketing*, *38*(12), 2240-2246.
20. Kotler, P. (1973). Atmospherics as a marketing tool. *Journal of retailing*, 49(4), 48-64.
21. North, A. C., Shilcock, A., & Hargreaves, D. J. (2003). The effect of musical style on restaurant customers' spending. *Environment and behavior*, *35*(5), 712-718.
22. Mattila, A. S., & Wirtz, J. (2001). Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of retailing*, 77(2), 273-289.
23. Mehrabian, A., & Russell, J. A. (1974). The basic emotional impact of environments. *Perceptual and motor skills*, 38(1), 283-301.
24. Milliman, R. E. (1982). Using background music to affect the behavior of supermarket shoppers. *Journal of marketing*, 46(3), 86-91.
25. Ministry of Economic Affairs. (2022, February 15). 網購市場順勢躍升新高, 成長率優於整體零售業. Department of Statistics. Retrieved April 8, 2022, from https://finance.ettoday.net/news/2218542
26. Pahnila, S., & Warsta, J. (2010). Online shopping viewed from a habit and value perspective. *Behaviour & Information Technology*, *29*(6), 621-632.
27. Siebenaler, D. J. (1999). Student song preference in the elementary music class. *Journal of Research in Music Education*, *47*(3), 213-223.
28. Stone, M. (1983). Some antecedents of music appreciation. *Psychology of Music*, 11(1), 26-31.
29. Sweeney, J. C., & Wyber, F. (2002). The role of cognitions and emotions in the music-approach-avoidance behavior relationship. *Journal of services marketing*, 7(2), 55-63.
30. Tantanatewin, W., & Inkarojrit, V. (2016). Effects of color and lighting on retail impression and identity. *Journal of Environmental Psychology*, 46, 197-205.
31. Teller, C., & Dennis, C. (2012). The effect of ambient scent on consumers' perception, emotions and behaviour: A critical review. *Journal of Marketing Management*, 28(1-2), 14-36.
32. Toldos, M. P., González, E. M., & Motyka, S. (2019). Exploring international atmospherics: The impact on shopping behaviors based on the language of the lyrics of music played in global apparel retailers' stores. *International Journal of Retail & Distribution Management*, 47(4), 368-383.
33. Trademag. (2021, November 15). 2021年「疫後全球電子商務發展趨勢」論壇暨市調系列發表會. 林月雲. Retrieved April 8, 2022, from https://www.trademag.org.tw/page/newsid1/?id=7852343&iz=6

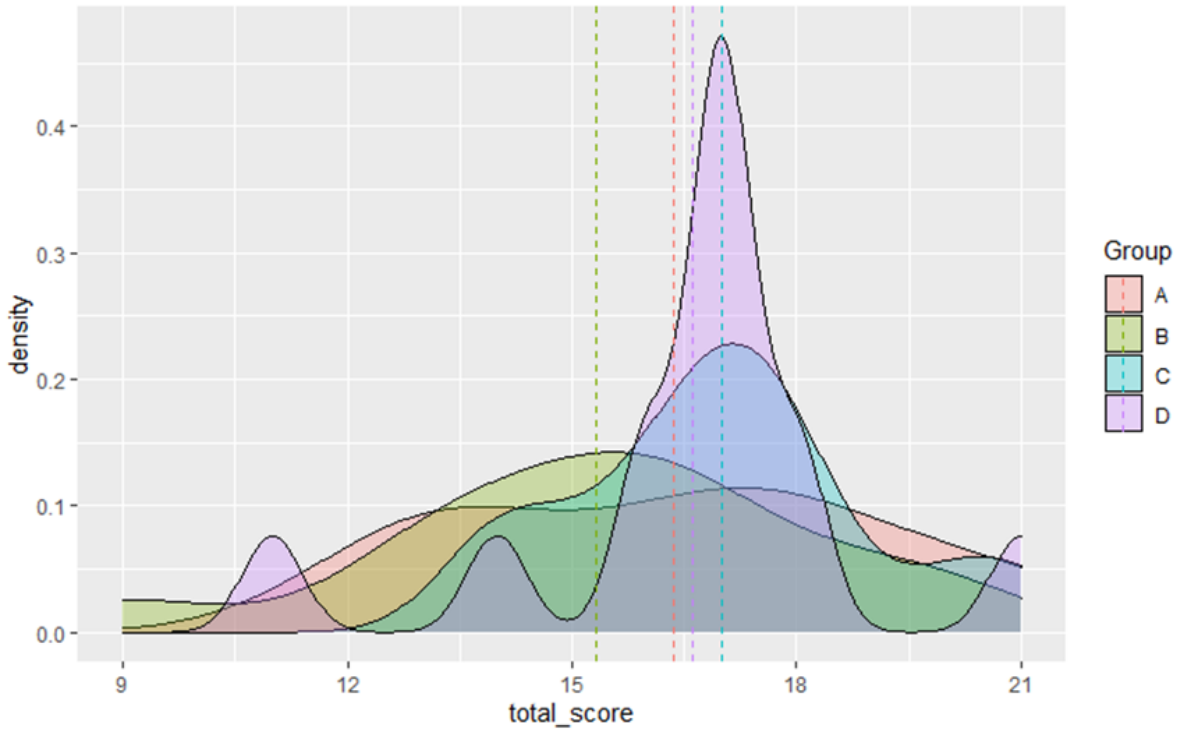
34. Russell, J. A. (1979). Affective space is bipolar. *Journal of personality and social psychology*, 37(3), 345.
35. Wilson, S. (2003). The effect of music on perceived atmosphere and purchase intentions in a restaurant. *Psychology of music*, *31*(1), 93-112.
36. Yalch, R., & Spangenberg, E. (1990). Effects of store music on shopping behavior. *Journal of Consumer Marketing*, 7(2), 55-63.

Appendix

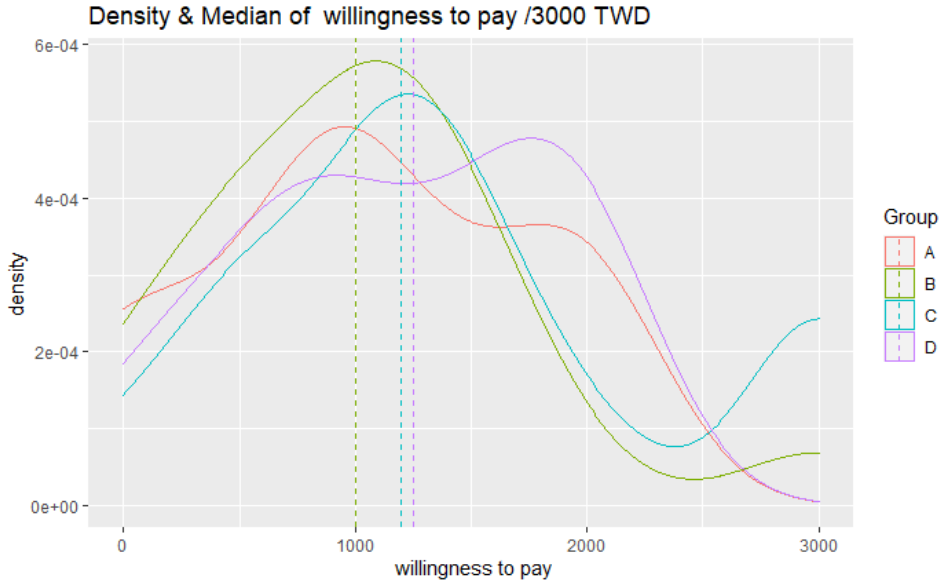
- **Figure 3**



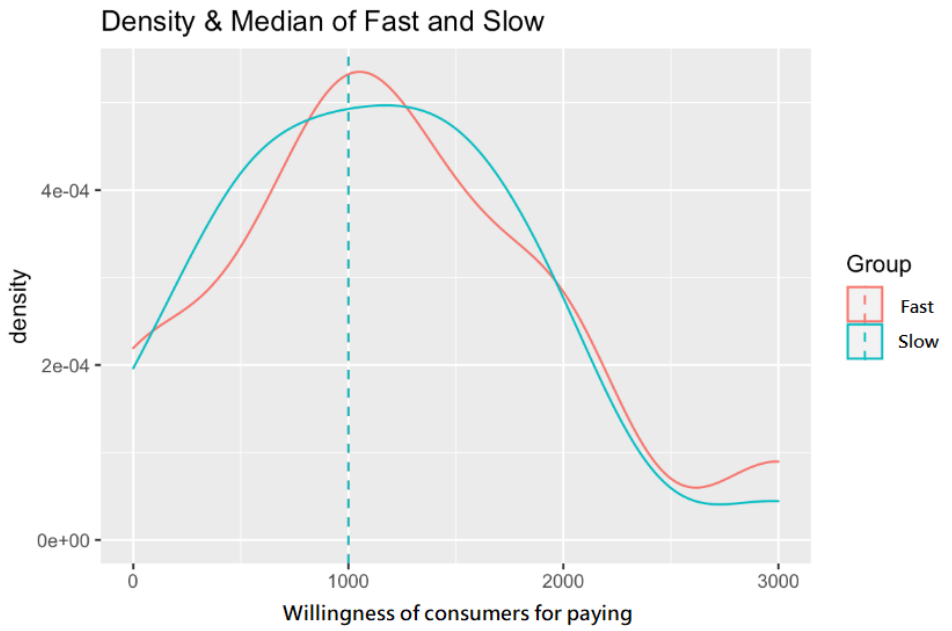
- **Figure 5: Density Plot of Grouping Total UQ_preference Score**



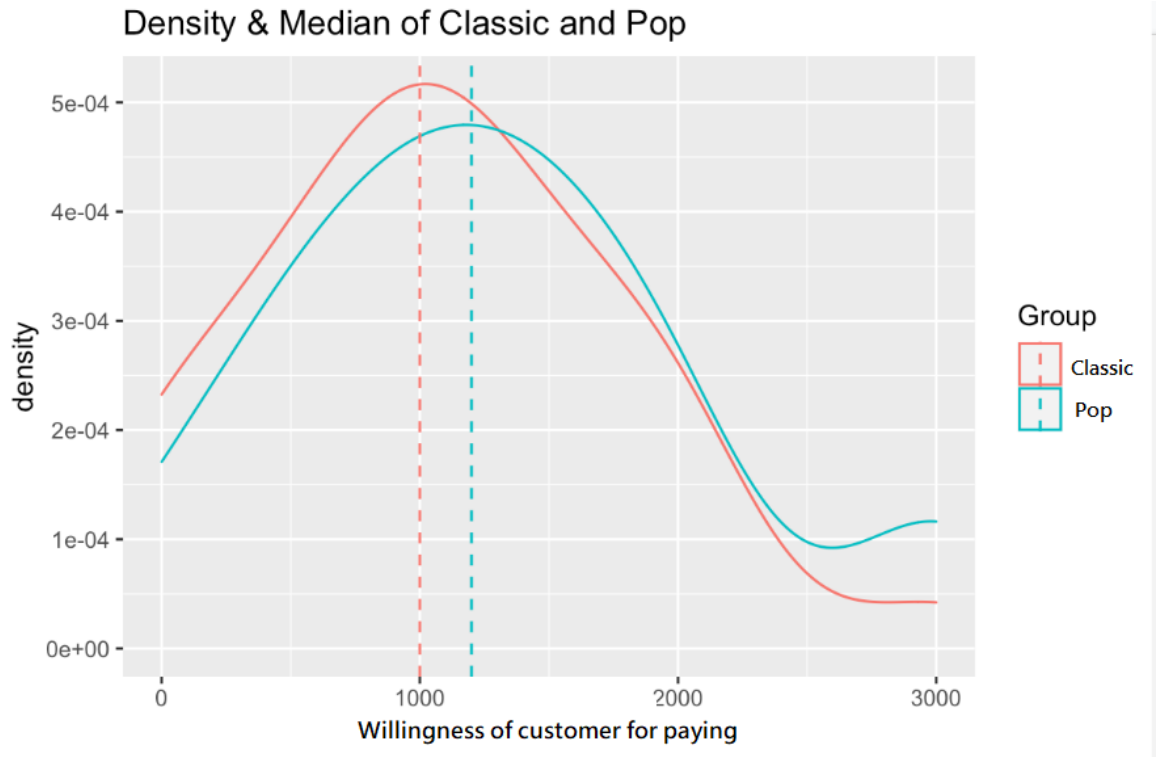
- **Figure: Median of WTP among the four groups**



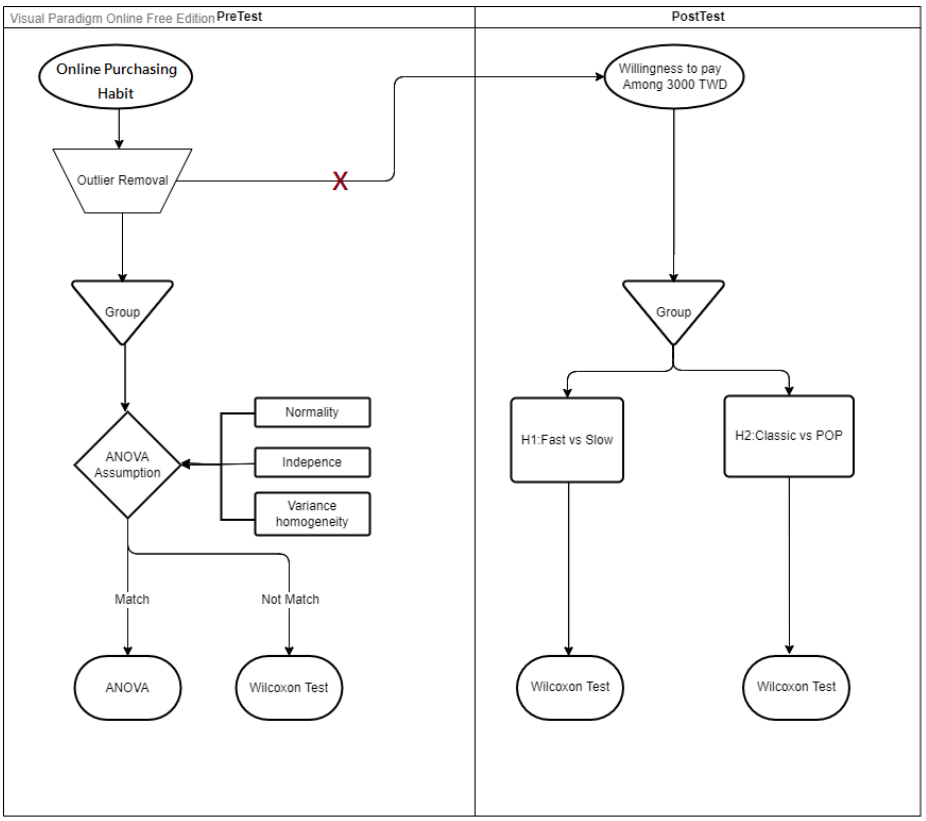
- **Figure: Median comparison in density for H1**



- **Figure: Median comparison in density for H2**



- **Data Analysis - Flowchart**



Final Presentation Q&A:

1. 怎麼確認真的是因為音樂影響想花多少錢？
   - **a.** 實驗更改後僅能做到音樂和願付價格的關係 **(**不同音樂是否有組間差異**)**
     - i. 第一版的實驗本來有加入時間 (Time Lag), 所以可以用推論 (inference) 的方式知道音樂是如何影響願付價格的;但在 proposal 跟老師討論過後, 因為時間、成本考量, 改了實驗方式, 改成 逛網購實驗當下聆聽音樂, 並在實驗後回答問卷。
     - ii. 同時因為實驗的時間成本考量下, 我們「沒有做到控制其他可能影響結果的變因, 僅有調查潛在會影響實驗結果的變因」，所以不能真的證明到是音樂「導致」願付價格的高或低。
   - **b.** 但我們在這樣的限制下, 有盡可能的調查會影響音樂對購物關係的潛在變項, 因此可以知道我們的結果是否會因為其他因素而被影響
     - i. 音樂偏好
       1. 雖然 A、B、C 組的組內音樂偏好沒有類似的傾向，D 組的組內音樂偏好有低程度的相似（W=0.301 ），而 D 組最偏好 B

歌單，但 **D** 組在實驗過程中不會聽到 **B** 歌單，所以並不會有因為有音樂的偏好影響到結果。

ii. 品牌偏好

1. 對 uniqlo 品牌偏好評分之加總於組間無顯著差異，因此我們也可以透過實驗後資料分析的方式 確認對 **uniqlo** 的偏好並不會成為影響實驗的原因。

iii. 消費習慣

1. 我們雖然在實驗後資料分析時發現四組的消費習慣差異很大，尤其以 A 組消費習慣金額最低（300元），B 組消費習慣金額最高（1000元），但在 Hypothese testing 時，我們是有進行併組的（H1: A+C vs. B+D; H2: A+B vs. C+D），我們發現併組後的消費習慣無組間差異，所以並 不會有因為有消費習慣的不同而影響到結果。

2. 音樂偏好？

a. 雖然第一題有寫到我們發現歌曲偏好不會影響到結果（即使 D 組對 B 歌單偏好，實驗也只有聽到 D 歌單不會聽到 B 歌單），但我們事後在資料分析上發現 來自相同歌單的兩兩歌曲之間 **correlation** 很低，也就是對於參與者來說即使兩首歌曲來自同一歌單，他們並沒有對於這兩首有一致的反應，換言之是我們在挑選各組歌單時，歌曲同質性沒有做得很好。

b. 我們事後針對音樂偏好沒有做好的部分，提供了三個可以改進之處：

i. 有文獻 (Kellaris & Kent, 1993) 表示當經典風格 (classical) 配上大於 180 的音樂速度 (BPM) 時，該組合的音樂對於人的情緒才會有足夠大的反應差異， 因此我們認為可以拉大音樂速度 **(BPM)**，才得以凸顯「快音樂」的感知。

ii. 在 Spotify 提供的歌曲特徵變數有高達九種 (ENERGY/DANCE/LOUD/VALENCE/LENGTH/ACOUSTIC/POP./RND)，因此我們認為不只是根據音樂速度 (BPM) 和歌曲風格而已，需要控制更多歌曲變數以確保其他特徵有一致的特性，避免影響到實驗結果。

iii. 在音樂偏好的調查上，我們認為因為一組歌單僅有挑選兩首歌曲可能是不足的，加上根據文獻 (Siebenaler, 1999) 僅讓受試者填寫一題 1-5 分以表明對於歌曲偏好，這可能也不足以了解一個人對於音樂風格的偏好，因此我們認為 可以增加同一風格的音樂數目以及增加調查音樂偏好的題目（如加入反向題）

3. 品牌偏好？

   a. 雖然在第一題有寫到我們發現品牌偏好不會影響到結果，但在事後資料分析上我們發現雖然在詢問品牌偏好的題題之間相關性高，Crunbach's alpha > 0.7，但是其實去除 Voting Bia 後，發現其實沒有測量得很好 **(crunbach alpha = 0.3)**。加上，即使我們發現有參與者在品牌偏好的問項中幾乎填答 3，但是 因為參與人數少（**n** 接近 **30**），因此無法直接刪除亂填答的人。

   b. 我們事後針對品牌偏好沒有做好的部分，提供了兩個可以改進之處：

      i. 因此我們認為 可以透過 **pilot** 的方式改善量表問的方式進行語句調整或是增加其他適合的問題，以避免受測者因為問題敘述不了解無法表明想法，所以直接填 3 的問題。

      ii. 我們發現如果想要盡可能地避免有品牌偏好的狀況發生，可以仿效菜單組使用「自製網站」以避免有品牌偏好的問題。

4. 中位數為甚麼都是整數？

   我們所詢問的問題為：請問您「每月」在電商平台購物的「平均消費金額」為何？，而我們所蒐集到的原始資料就是整數，即大家都回整數，如下圖所示：

B	C	D
實驗場次	請問您「每月」在電商平台購物的「平均次數」為何？	請問您「每月」在電商平台購物的「平均消費金額」為何？
C	2	500
C	2	700
C	2	3000
C	2	800
D	1	300
D	5	1500
D	3	500
D	4	2000
D	2	500
D	2	800
D	2	200
B	1	300
B	3	1000
B	4	2000
A	5	500
A	1	500
A	1	150
A	1	500
A	1	300
A	1	200
D	3	2000
D	2	2000
D	1	200
D	1	100
D	3	200