BA830 Team project

Christopher Chang, Yesol(Sally) Lee,Yongxian(Caroline) Lun, Linh To,

Introduction

In spite of the global recession and pandemics' influence on global economies, luxury brands are everywhere. According to Bain & Company, the luxury goods market returned to pre-pandemic growth in 2021. The sales are estimated to top 283 billion euros, which is $325 billion. Moreover, younger consumers are increasing. As reported by Boston Consulting Group, Gen Z and Millennials are set to represent more than 60 percent of the luxury market by 2025. In this regard, our team wanted to research how brand name affects the preference of products, especially amongst young consumers. Thus we used a survey to measure how people's preferences change among very similar designed clothing of high-end brands and inexpensive brands.

Methodology

Goal of the Experiment

The goal of the experiment was to observe the effect of exposing fashion brands on consumption decisions of young people who have graduated with the age range of 22 - 27, we decided to conduct the experiment on graduate students in various institutions. We used the blocking method with the gender of our participants so we also had to keep in mind the number of participants from each gender.

Survey Design

We designed the survey with Qualtrics platform and utilized the survey logic to create our treatment and control group and randomization for the participants. Each participant will see a survey with one question asking about their gender, and 10 questions asking them to choose among 3 fashion items. We divided our participants into male and female, control and treatment groups, and for those 4 groups, we conducted just one survey with information questions and 40 fashion brand questions.

Male and female questions assignment

There are 4 main components in the survey, which are: male_control, male_treatment, female_control and female_treatment. The survey will only display the right male and female questions according to the logic of the first question: "Which gender do you most closely identify with?". If the participants answer male, they will be moved into the male questions displaying male fashion brands, and vice versa.

Control and treatment assignment

Qualtrics options allowed us to randomly assign the participants into the control and treatment groups equally. If the participant answered male or female to the first question, they will be assigned randomly to the control or treatment group of their respective gender questions. The survey will equally display each gender to their control or treatment assignments to ensure the probability of the assignment.

Survey questions design

The survey is designed in 3 main parts: the introduction of our experiment, the information questions, and the main body part. In the survey introduction, we greeted the participants and briefly explained that this is the survey about fashion preference. In the information questions, participants were asked about their name and their gender identification between males and females. The main body of the survey is 10 questions asking the participants about their brand preference for 10 fashion items: sweater, boots, pants, scarf, hoodie, gloves, beanie, socks, coat, dress/suits. Each question is given 3 image choices from 3 different brands, 1 image from the high-end fashion brand and 2 images from the lower-range fashion brands. For the control group, the brand names were hidden from the pictures and the choices would only display as: “Option 1”, “Option 2” and “Option 3”. For the treatment group, the pictures show the brand names on the fashion items and the choices would display their respective brands, for example as: “Zara”, “Gucci” and “H&M”.

Randomization in question level

Qualtrics also gives us the options to randomize in both the question order and the question choices order. It resulted in the survey displaying the fashion item questions randomly; and in each question, the fashion brands would also be displayed randomly.

Scoring

The score of each respondent is calculated based on their preference for high-end fashion brands. The participant will be scored according to whether or not they will choose the high-brand items, the high-end brand will be assigned the score of 1, while other choices will be assigned the score of 0. As a result, for 10 fashion item questions, the minimum score would be 0 (the participant didn’t choose any of the high-end brands) and the maximum score would be 10 (the participant chose all high-end items). The Qualtrics option also allowed us to record the scores among all of our participants, given that we needed to assign and match the correct choices in each question to the scores. As a result, we could score consistently across all of our participants even though we used various randomization logic in our survey.

Procedure

Before sending out the survey, we first created a spreadsheet to keep track of each participant’s name, age, their gender and their respective experiment conductor so as to not have any duplicates in our end data result. Then we reached out to graduate students from Boston University and other schools to ask them to take our survey and track their survey completion progress. We aimed to have 15 participants in each of these 4 groups: male control group, male treatment group, female control group and female treatment group, so we had to survey a minimum of 60 people in total. The result of this experiment is 86 people answering our surveys, with 19 people in the male control group, 19 people in the male treatment group, 20 females in the control group and 23 females in the treatment group.

Data Analysis

Data Cleaning

Before we started data analysis, we firstly investigated the survey answers to verify if there are any not qualified response.We noticed that some of our participants haven’t completed all of the 10 survey questions and some of them accidentally answered the survey twice. We used Excel to drop these rows. We also removed irrelevant columns such as respondents_id, starting date, and more.

```r
# load the data
data <- fread('data_final_analysis.csv')
data$gender[data$gender==1] <- 0 # male
data$gender[data$gender==2] <- 1 # female
```

Exploratory Data Analysis

We firstly did some EDA to understand overview of the data.

```r
# Number of control(0) and treatment(1) samples
table(data$any_treatment)
```

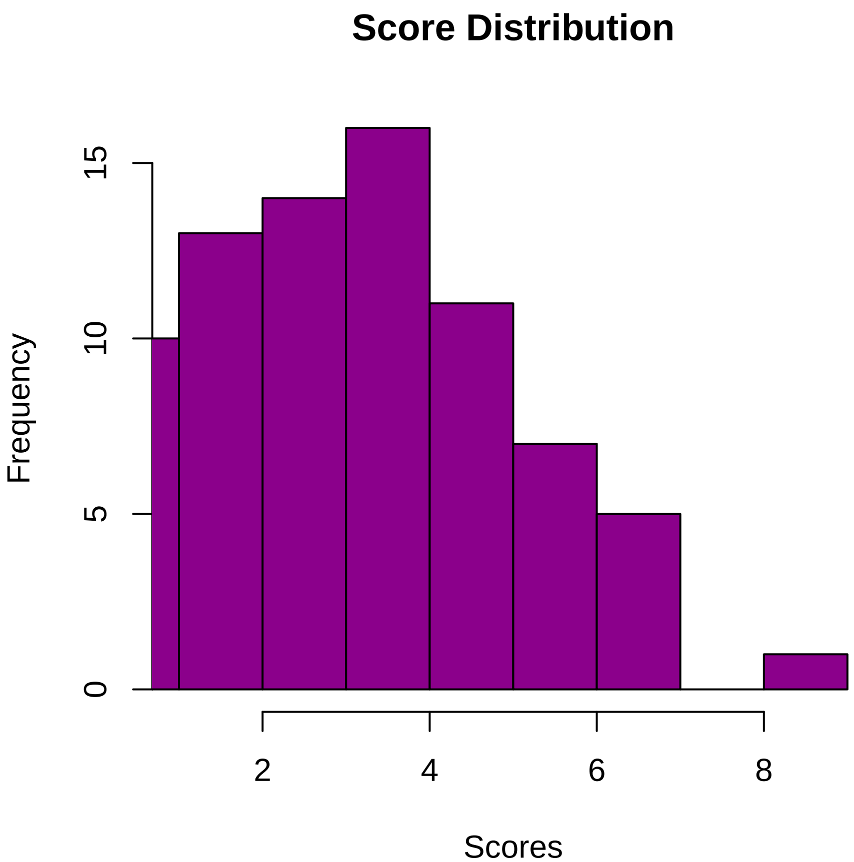
```
##
##  0  1
## 38 39
```

```r
# Number of control and treatment samples based on gender
data[,list(num_observation = .N),by=list(any_treatment, gender)][order(any_treatment, gender)]
```

```
##    any_treatment gender num_observation
## 1:             0      0              18
## 2:             0      1              20
## 3:             1      0              18
## 4:             1      1              21
```

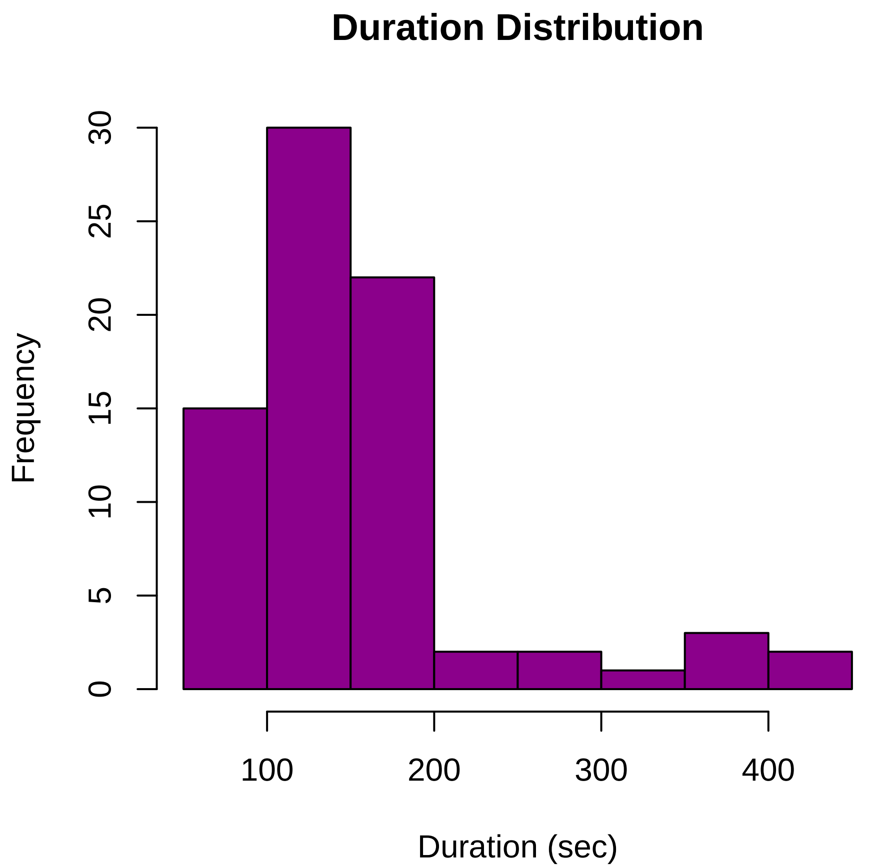
There are 38 individuals in the control group and 39 individuals in the treatment group including men and women. In the control group, there are 18 men and 20 women in the control group, and 18 men and 21 women in the treatment group.

```r
# Distribution of Scores
hist(data$score,
main="Score Distribution",
xlab="Scores",
xlim=c(1,9),
col="darkmagenta",
freq=TRUE
)
```




```r
# Distribution of Duration
hist(data$duration,
main="Duration Distribution",
```

```
xlab="Duration (sec)",
xlim=c(50,450),
col="darkmagenta",
freq=TRUE
)
```




```
median(data$score)
```

```
## [1] 4
```

```
median(data$duration)
```

```
## [1] 137
```

Both distributions are skewed to the right with medians of 4 and 137 for score and duration in seconds, respectfully.

```
pt <- PivotTable$new()
pt$addData(data)
#pt$addColumnDataGroups("gender")
pt$addRowDataGroups("any_treatment")
pt$addRowDataGroups("gender")
pt$defineCalculation(calculationName="Avg_Score", summariseExpression="round(mean(score, na.rm=TRUE),2)
pt$evaluatePivot()
pt
```

```
##                Avg_Score
## 0      0            3.44
##        1             3.4
```

```
##         Total       3.42
## 1       0           4.61
##         1           3.24
##         Total       3.87
## Total               3.65
```

The first column is any_treatment, and the second column is the gender, and the third one is the average score for different genders in different groups.

The average score in treatment group is higher than in control group, especially the male has higher score in the treatment group. On the other hand, the female in the treatment group has lower score than in the control group. This shows that the male in the experiment shows more tendency towards branded clothing than the female.

Estimated Average Treatment Effect

Simple Regression

```r
atehat <- data[any_treatment==1, mean(score)] - data[any_treatment==0, mean(score)]

simple_reg <- feols(score ~ any_treatment, data=data, se='white')

etable(simple_reg)
```

```
##                         simple_reg
## Dependent Var.:              score
##
## (Intercept)     3.421*** (0.2313)
## any_treatment     0.4507 (0.4266)
## _______________ _________________
## S.E. type       Heteroskeda.-rob.
## Observations                   77
## R2                        0.01451
## Adj. R2                   0.00137
```

We firstly computed the average treatment effect across all data points by using regression model. The estimate of average treatment effect was 0.451 and the standard error of this value is 0.4266. This means that respondents are more likely to select the high-end brand products when they know the product was made by a luxury brand. The p-value was 0.3 thus, it is not statistically significant. Confidence interval also includes 0, thus we cannot reject the null hypothesis.

Conditional treatment effect

```r
male <- data[gender==0]
female <- data[gender==1]

#CATE of Male
m_ate <- male[any_treatment == 1, mean(score)] - male[any_treatment == 0, mean(score)]

print(m_ate)
```

```
## [1] 1.17
```

```r
t.test(male[any_treatment==1, score], male[any_treatment==0, score])
```

```
##
##  Welch Two Sample t-test
```

```
##
## data:  male[any_treatment == 1, score] and male[any_treatment == 0, score]
## t = 2, df = 31, p-value = 0.06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.0704  2.4037
## sample estimates:
## mean of x mean of y
##      4.61      3.44
```

```
#CATE of Female
f_ate <- female[any_treatment == 1, mean(score)] - female[any_treatment == 0, mean(score)]

print(f_ate)
```

```
## [1] -0.162
```

```
t.test(female[any_treatment==1, score], female[any_treatment==0, score])
```

```
##
##  Welch Two Sample t-test
##
## data:  female[any_treatment == 1, score] and female[any_treatment == 0, score]
## t = -0.3, df = 33, p-value = 0.8
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.33  1.01
## sample estimates:
## mean of x mean of y
##      3.24      3.40
```

We thought the average treatment effect may differ by gender of respondents. So we also checked the conditional treatment effect by gender. CATE for male respondents was 1.17. CATE for males had a p-value of 0.06, therefore it is statistically insignificant in significance level of 0.05. Females, on the other hand, showed a CATE of -0.162, and the p-value was 0.8. Therefore, CATE for females is not statistically significant. It was interesting because our team initially expected that female respondents would have a bigger treatment effect.

Controlling a Covariate

```
# gender as a covariate
cov_reg <- feols(score ~ any_treatment + gender, data=data, se='white')

etable(cov_reg)
```

```
##                           cov_reg
## Dependent Var.:             score
##
## (Intercept)     3.798*** (0.3385)
## any_treatment     0.4594 (0.4213)
## gender          -0.7163. (0.4237)
## _______________ _________________
## S.E. type       Heteroskeda.-rob.
## Observations                   77
## R2                        0.05099
## Adj. R2                   0.02534
```

We also tried controlling gender as a covariate to increase the precision of the estimate. In this case, the estimate of average treatment effect slightly increased to 0.4594. Standard error of the value changed to 0.4213. Thus, the model found the more precise treatment effect compared to the previous regression.

Randomization Check

Balance Check using Regression Model

A randomization check is a regression where the dependent variable occurs before the experiment. It should be very unlikely that there are substantial differences in before experiment variables if the experiment was done properly.

```
regression_pre_effects_gender <- feols(gender ~ any_treatment,
                                       data = data)
summary(regression_pre_effects_gender)
```

```
## OLS estimation, Dep. Var.: gender
## Observations: 77
## Standard-errors: IID
##               Estimate Std. Error t value   Pr(>|t|)
## (Intercept)     0.5263      0.082   6.418 1.1199e-08 ***
## any_treatment   0.0121      0.115   0.105 9.1633e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.498908   Adj. R2: -0.013183
```

In this data, we only had gender as a variable occurs before the experiment. The result proves that the estimate of any_treatment is not statistically significant, which means the covariate is not affected by the treatment. Thus, it proves that the randomization is done properly.

Balance Check using prop.test

Our team also used the prop.test function to check whether the randomization proportion was intended.

```
num_obs_treat <- dim(data[any_treatment==1])[1]

num_obs_all <- nrow(data)

proportion_treatment <- (100/2) * 0.01

prop.test(num_obs_treat, num_obs_all, p = proportion_treatment)
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  num_obs_treat out of num_obs_all, null probability proportion_treatment
## X-squared = 0, df = 1, p-value = 1
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.391 0.621
## sample estimates:
##     p
## 0.506
```

Based on these results, the 0.5 is in the confidence interval and p-value of 1 is high, we can't reject the null of a 50% treatment assignment probability at the 5% level.Thus, the randomization was done properly.

Statistical Power

```
cohens_d <- atehat/sd(data[, score])
cohens_d
```

```
## [1] 0.239
```

The Cohens' D measures the effect size of the difference between the mean of treatment and the mean of control. The Cohen's D of 0.239 is a small effect.

```
num_control<- nrow(data[any_treatment==0])

pwr.t2n.test(n1 = num_obs_treat, n2 = num_control, d = cohens_d,
             sig.level = .05, power = NULL)
```

```
##
##      t test power calculation
##
##              n1 = 39
##              n2 = 38
##               d = 0.239
##       sig.level = 0.05
##           power = 0.179
##     alternative = two.sided
```

The power is 0.179 which is a bit less, meaning that we are a bit likely to detect the effect. This is because we have small Cohen's D and our experiment only had about 80 observations.

```
pwr.t.test(n = NULL, d = cohens_d,  sig.level = .05, power = 0.8)
```

```
##
##      Two-sample t test power calculation
##
##               n = 275
##               d = 0.239
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

If we want to have higher power of 0.8, we need 275 observations in total when the Cohen's D does not change.

Limitation

We found there are several possible ways to improve the project and get more insightful information.

Adding more products

We used only 1 product for each product category. There might be some characteristics such as colors, length of clothing, and style that can possibly affect the amount of treatment effect. So if we choose a specific product category and have more diverse styled products in the questionnaire, we could understand more about how knowing brand name would affect people's preference.

Expand Age Range of Respondents

Our team only concentrated on people in their mid 20's. People in other age groups, like teenagers and junior professionals, can have different perceptions of luxury brands. There can be a certain pattern of treatment effect by age group. For example, the treatment effect might be small within teenagers, because they cannot afford luxuries, and as they grow up and get a job, the interests on high-end brands increase and possibly the treatment effect also gets bigger.

Adding more brands

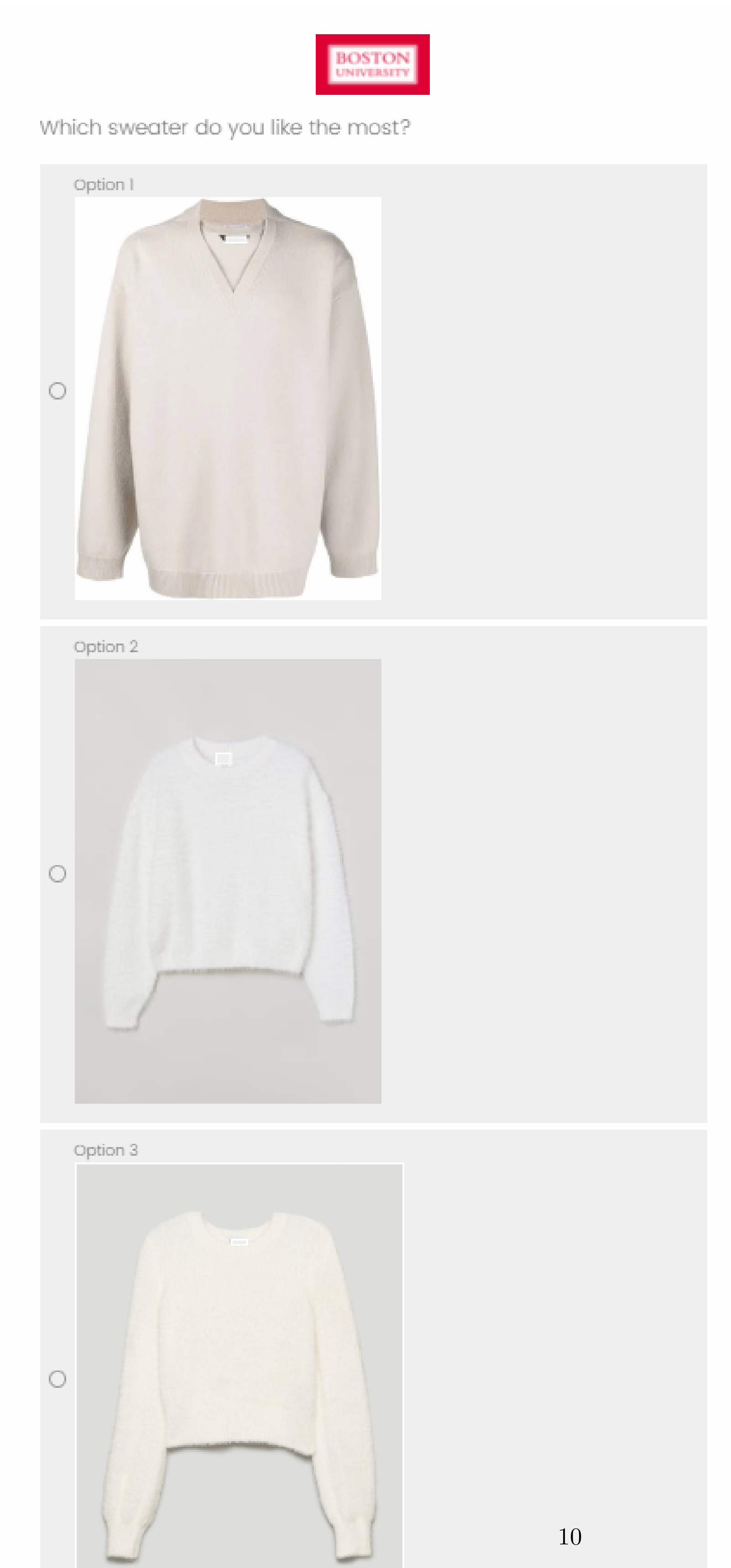
In this study, we used high-end brands like Gucci, Dior, and Balenciaga. But there are so many mid-luxury brands such as Michael Kors, DKNY and so on. If we include those brands, we can get more insight about brands' power to people's preference.

Conclusion

We designed the experiment to analyze the effect of displayed brand names on people's preference for the nearly same designed fashion items, specifically those who are in their mid 20's. We computed the score for each respondent by counting the number of luxuries they chose from the survey. We detected a positive estimated treatment effect, but this value is not statistically significant. There was a difference between ATE of men and that of women. Surprisingly, males were more likely to choose high-end brands when they were seeing the picture of product and brand name together. Despite our estimated average treatment effect being not statistically significant, we confirmed that our experiment was done properly based on the randomization check. The results of statistical power indicates that we need around 275 observations to reach higher power of 0.8 that can increase the likelihood to detect the effect. For further studies, we can expand the product categories and attributes to see how the average treatment effect can vary by the product categories or design. We can also collect more data from people of different ages.

Appendix

Preview of Survey Form



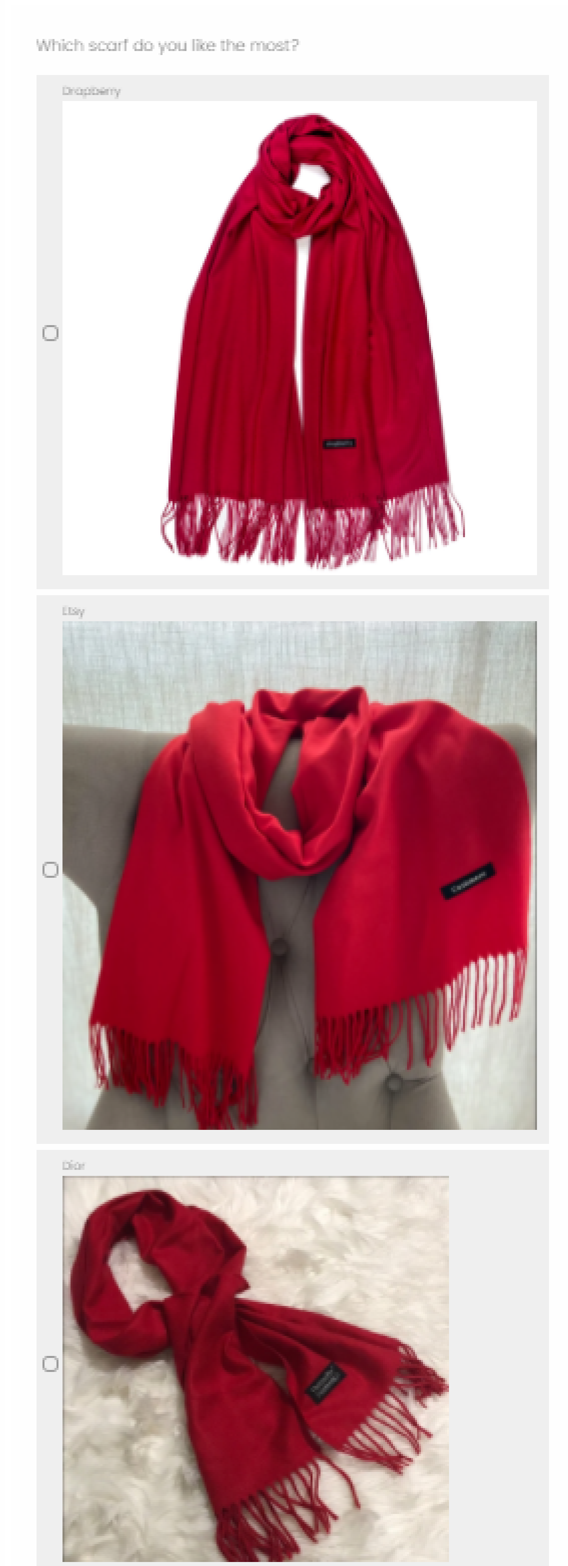
BOSTON UNIVERSITY

Which sweater do you like the most?

Option 1

Option 2

Option 3



Which scarf do you like the most?

Dropberry

Etsy

Dior

Logic of the Survey Platform

Qualtrics question display logic: https://www.qualtrics.com/support/survey-platform/survey-module/question-options/display-logic/

Qualtrics question randomization: https://www.qualtrics.com/support/survey-platform/survey-module/block-options/question-randomization/

Qualtrics choice randomization: https://www.qualtrics.com/support/survey-platform/survey-module/question-options/choice-randomization/

Qualtrics scoring system: https://www.qualtrics.com/support/survey-platform/survey-module/survey-tools/scoring/