

# BA830 – Report

## Impact of Advertisement Vibrancy on Consumer Appeal

Mar 6, 2025

Section A1, Team 10

Aastha Surana, Amisha Kelkar, Pratik Mahajan,

Raskirt Bhatia, Soham Sarvade



## **Introduction and Motivation:**

In digital advertising, ad design significantly influences consumer attention, perception, and engagement. Marketers strive to optimize visual elements like vibrancy, yet their impact remains uncertain. Ads vary from minimalist designs (clean, simple layouts) to vibrant ads (bold colors, dynamic imagery). Call-to-action (CTA) elements further encourage interaction, but their effectiveness alongside vibrancy is unclear.

This study examines whether vibrant ads enhance appeal and whether adding a CTA increases engagement. The findings will help businesses craft data-driven, visually compelling ads to boost user engagement and conversions.

## **Research Question:**

Does the vibrancy of a print advertisement influence consumer appeal, and does the addition of a call-to-action further enhance engagement?

## **Hypothesis Tested:**

- **H0 (Null Hypothesis):** There is no significant difference in consumer appeal between minimalist and vibrant advertisements, with or without a call-to-action.
- **H1 (Alternative Hypothesis):** Vibrant advertisements lead to higher consumer appeal compared to minimalist advertisements, and adding a call-to-action further increases engagement.

## **Treatment and Randomization Approach:**

The experiment involved three groups:

- **Control Group:** Shown print advertisements with a minimalist design.
- **Treatment 1:** Shown the same brands but with a more vibrant design.
- **Treatment 2:** Shown the same brands and vibrant designs as Treatment 1, but with an additional call-to-action button.

**Unit of Randomization:** Randomization was conducted within subjects at the survey question-respondent level, meaning each respondent was randomly assigned to view one of the three ad variations for each brand in the survey.

**Randomization approach:** We employed **simple randomization** at the respondent-question level, ensuring that each participant was randomly assigned to view one of the three ad variations for each brand. This approach distributed treatments evenly, minimizing bias and ensuring a fair comparison across conditions.

## **Outcome Measurement and Data Collection:**

**Data Collection:** A Qualtrics survey was conducted with 66 respondents, ensuring randomized exposure to one of three ad variations (Control, Treatment 1, or Treatment 2) for unbiased distribution.

**Population:** The survey was shared with BU students, friends, and international participants to ensure diversity. All questions were mandatory, ensuring complete responses.

**Brands:** We selected six widely recognized and popular brands, each representing a distinct industry: technology (Apple), fast food (McDonald's), beverages (Pepsi), footwear (Nike), alcohol (GreyGoose), and airline (Emirates)

The survey recorded:

1. Brand Interaction: Whether the respondent had purchased from the brand before.
2. Appeal Rating: A 1-10 scale measuring how appealing they found each advertisement.
3. Click Likelihood: How likely they would be to engage with the ad if it appeared in their feed.
4. Demographic Data: Age, gender, and country of residence.

### Techniques Used

1. Use of Covariates: Demographic data (age, gender, country) were used as covariates in regression analysis to control for respondent differences, improving accuracy and ensuring precise treatment effect estimation.

### Data Cleaning:

To ensure consistency and accuracy, the dataset was cleaned through:

- Standardization: Categorical responses were categorized, and binary (Yes/No) values mapped to 1 and 0.
- Column Removal: Personal identifiers (RecipientEmail, IPAddress) and redundant fields (Status, Progress) were dropped.
- Formatting: Numeric values in text responses were extracted, the first two rows removed, and key columns renamed.
- Reshaping & Standardization: The dataset was transformed into long format, aligning treatments with responses (Click Likelihood, Appeal Rating, Past Buyers), and Frequency of Ads was grouped into 1-5, 5-10, 10+, and Never.

### Exploratory Data Analysis:

#### Sample Composition and Demographics

The dataset includes respondents from both male and female demographics, helping us understand the sample composition and potential differences in advertisement engagement based on gender. Additionally, the geographical demographics of respondents reveal a diverse participant set, including Boston University students and international participants. Understanding the country-wise distribution aids in assessing how cultural differences might influence advertisement engagement. Furthermore, the demographic breakdown by age provides insight into how different age groups perceive advertisement vibrancy and engagement, which is crucial for tailoring advertising strategies effectively.

#### Advertisement Engagement Analysis

The likelihood of respondents engaging with an advertisement was assessed on a scale from 1 to 10, showing that engagement increased from 1 to 8, peaking at 7 and 8. This suggests that most respondents found the advertisements moderately to highly appealing. However, ratings at 9 and 10 slightly declined, indicating that while the advertisements were engaging, they did not fully convince all respondents to take action. Lower ratings (1-3) were relatively rare, indicating minimal negative reception. Additionally, respondents rated advertisement attractiveness on a scale from 1 to 5, with the majority rating them at 4,

signifying a generally favorable reception. However, ratings of 1 and 2 were also common, suggesting that a significant portion of respondents found the ads less appealing. The lowest count was observed at rating 5, indicating fewer respondents considered the ads highly appealing.

Note: All relevant visualizations supporting this analysis are included in the appendix for reference.

### **Randomization Check:**

This analysis assesses whether the random assignment of respondents to treatment groups across different brands was effective in balancing key demographic characteristics: Gender, Age Group, and Country. Ensuring balance across these variables is crucial to isolating the effects of treatment and minimizing confounding biases in subsequent analyses.

### **Methodology**

- Proportions Z-test for Gender (binary variable), conducted pairwise between treatment groups.
- Chi-Square Test for Age Group and Country (multi-category categorical variables).
- Tests were conducted separately for each brand, with  $p > 0.05$  indicating that the randomization was successful..

### **Key Findings**

1. Randomization Appears Successful for Most Brands:
  - For Apple, Pepsi, McDonald's (McD), Nike, and Emirates, all p-values were greater than 0.05, indicating no statistically significant differences in demographic distributions across treatment and control groups.
  - This suggests that randomization was effective in balancing Gender, Age Group, and Country across these brands, ensuring comparable groups for analysis.
2. Potential Gender Imbalance in GreyGoose:
  - The Gender comparison between Treatment\_1 and Treatment\_2 yielded  $p = 0.007$ , indicating a potential imbalance in gender distribution.
  - While this imbalance does not necessarily invalidate results, it may require additional consideration, such as including Gender as a control variable in downstream analyses to mitigate any potential bias.

### **Implications**

- Largely Successful Randomization: Across most brands, the absence of significant differences confirms that randomization was effectively implemented, reducing the likelihood of demographic-driven biases in results.
- Addressing McD's Gender Imbalance: The observed imbalance in Gender for McDonald's could be handled by:
  - Controlling for Gender in further statistical analyses.
  - Running additional robustness checks to assess the impact of this imbalance.
  - Applying a multiple-comparison correction to verify if the significance holds after adjustments.

### T-Test:

Independent t-tests were conducted to compare means between the control group and the two treatment groups for two key outcome variables **Click Likelihood** and **Appeal Rating**. The t-test results are summarized in the appendix.

The comparison between Minimalist ads (Control) and Vibrant ads (Treatment\_1) for Click Likelihood yielded a statistically significant result, indicating that vibrant ads increased the likelihood of respondents clicking the advertisement. The Minimalist vs. Vibrant + CTA (Treatment\_2) for Click Likelihood test was not significant, suggesting that adding a CTA did not significantly alter click likelihood compared to minimalist ads.

Neither of the Appeal Rating comparisons (Minimalist vs. Vibrant, Minimalist vs. Vibrant + CTA) showed statistical significance, implying that neither vibrancy alone nor vibrancy with a CTA significantly affected how appealing respondents found the advertisement.

### Cohen's D Analysis

Cohen's d was computed to quantify the effect size of treatment differences for Click Likelihood and Appeal Rating. Effect sizes provide a measure of practical significance beyond just statistical significance. A small effect is indicated by Cohen's d = 0.2. A medium effect corresponds to Cohen's d = 0.5, while a large effect is represented by Cohen's d = 0.8.

### Interpretation

- **Click Likelihood:**

- Treatment\_1 had a moderate positive effect (Cohen's d = 0.348), meaning users were somewhat more likely to click compared to the Minimalist design.
- Treatment\_2 had a small positive effect (Cohen's d = 0.140), suggesting a weaker influence on click likelihood.
- The combined effect (Treatment\_1 + Treatment\_2) was negative ( $d = -0.217$ ), indicating a potential counteracting effect when both treatments were applied together.

- **Appeal Rating:**

- Across all comparisons, effect sizes were small, meaning the differences in appeal ratings were minimal and likely not practically significant.
- The negative Cohen's d values for Treatment\_2 and the combined treatments suggest that these versions may have slightly reduced appeal compared to the Minimalist design, though the small effect size implies a negligible real-world impact.

Outcome	Comparison	Value	Effect Size
<b>Click Likelihood</b>	Control vs Vibrant Ads	0.35	Medium

	Control vs Vibrant Ads w/ CTA	0.14	Small
	Vibrant vs Vibrant Ads w/ CTA	-0.22	Medium
<b>Appeal Rating</b>	Control vs Vibrant Ads	0.02	Small
	Control vs Vibrant Ads w/ CTA	-0.18	Small
	Vibrant vs Vibrant Ads w/ CTA	-0.2	Small

## Implications

The analysis shows that Click Likelihood was more affected than Appeal Rating, with Treatment\_1 having a moderate effect, Treatment\_2 a small effect, and their combination resulting in negative effects, possibly due to design conflicts. In contrast, Appeal Rating showed minimal differences, suggesting that the treatments did not significantly impact perceived appeal. The negative effect sizes warrant further analysis to identify potential overcomplication or unintended design flaws. Overall, Treatment\_1 appears to be the best option for increasing clicks, but none of the treatments meaningfully enhance appeal, highlighting the need for further design refinement.

## Average Treatment Effect (ATE) and Conditional Average Treatment Effects (CATE):

### Methodology

To assess the overall impact of ad vibrancy and the addition of a call-to-action (CTA), we computed the Average Treatment Effects (ATE) by comparing the mean outcomes between each treatment group and the control group. Two key outcomes were examined: Appeal Rating and Click Likelihood.

- ATE Calculation:
  - Vibrant Ads (Treatment 1) vs. Control: The difference in mean outcomes indicates the net effect of a more vibrant design.
  - Vibrant Ads with CTA (Treatment 2) vs. Control: Similarly, this reflects the additional impact of a CTA element.
  - In addition, we estimated Conditional ATEs (CATE) by segmenting the sample into subgroups based on demographic characteristics (e.g., Gender, Country, and Age Group) to explore potential heterogeneity in treatment effects.

### Key Findings

Treatment	Click Likelihood	Appeal Rating
Vibrant Ads	0.91	0.027
Vibrant Ads with CTA	0.371	-0.236

- **ATE Results:**
  - Appeal Rating: Vibrant Ads exhibit a negligible positive effect (+0.027), whereas Vibrant Ads with CTA reduce appeal slightly compared to the control.
  - Click Likelihood: Vibrant Ads increase click likelihood by +0.91, while Vibrant Ads with CTA offer a moderate increase of +0.37 relative to the control.
- **CATE Insights:**
  - Subgroup analyses reveal that treatment effects vary by demographic segment. For example, among females, Vibrant Ads significantly boost click likelihood compared to the CTA-enhanced version.

## **Interpretations**

Overall, the results suggest that while the visual enhancement of ads (i.e., using vibrant colors) has a positive impact on consumer engagement—especially in terms of click likelihood—the addition of a CTA does not consistently enhance appeal and, in some cases, may reduce it. The heterogeneous effects observed in the CATE analysis imply that demographic factors, such as gender and regional differences, moderate the treatment's impact. These insights underscore the importance of tailoring ad design strategies to specific audience segments rather than adopting a one-size-fits-all approach.

## **Regression Analysis:**

### **Methodology**

We estimated ordinary least squares (OLS) regressions to assess the impact of ad treatments on two outcomes: Click Likelihood and Appeal Rating. For each outcome, two models were estimated:

- Model (1): Includes key predictors—treatment indicators and core demographic/ad exposure variables.
- Model (2): Adds additional covariates (e.g., detailed demographic factors and ad frequency) to improve the model fit and control for confounding influences.

This dual-model approach enables us to examine both the baseline treatment effects and the adjusted effects after accounting for respondent characteristics.

## **Key Findings**

### **Click Likelihood:**

- In Model (1), older respondents (55+ or “Prefer not to say”), frequent ad viewers (5–10 times/day), and males exhibit higher click likelihoods. However, both treatment conditions reduce click likelihood compared to control (Vibrant Ads: -2.09 points; Vibrant Ads with CTA: -2.34 points). The coefficients, however, were not significant at the 95% Confidence Interval.
- In Model (2), with additional covariates, the overall fit improves ( $R^2 = 0.496$ ). Here, Vibrant Ads show a significant reduction (-2.29 points) in click likelihood, while the effect of Vibrant Ads with CTA (-1.86 points) is not statistically significant.

### **Appeal Rating:**

- Model (1) reveals that, relative to control, both Vibrant Ads and Vibrant Ads with CTA significantly reduce appeal by 1.70 and 2.20 points, respectively ( $R^2 = 0.434$ ).
- In Model (2), after controlling for demographics and ad frequency, positive influences from older age, U.S. residence, and male gender are observed, yet both treatments continue to lower appeal by approximately 1.46 and 1.43 points, respectively. The model explains around 67% of the variance.

## **Interpretations**

The findings suggest that while demographic factors and higher ad exposure boost engagement, Vibrant Ads without a CTA reduce click likelihood compared to minimalist controls. Despite engaging certain subgroups, increased vibrancy may sometimes deter clicks. Both treatments also lower overall appeal, indicating a trade-off between engagement and attractiveness.

Overall, the regression analyses highlight the complex effects of ad design. While ad frequency and demographics drive engagement, modifications like vibrancy and CTAs may negatively impact appeal. These insights stress the need to balance visual impact with consumer preferences in digital ad design.

## **Limitations**

The study's findings are constrained by a small sample size of approximately 66 respondents, which limits generalizability. A larger sample could enhance statistical power and detect smaller yet meaningful effects. Additionally, while the study included international participants, the majority were likely from Boston University and their extended networks, reducing demographic diversity. The experiment was conducted in a digital format via an online survey, which may not fully capture how advertisements perform in other mediums such as print, television, or outdoor billboards.

Moreover, consumer appeal and engagement are inherently subjective, influenced by personal preferences, biases, and mood at the time of response—factors that were not explicitly controlled. Another limitation is the reliance on self-reported engagement intentions rather than real-world behavioral data, such as clicks, purchases, or brand interactions in actual advertising settings.

## **Conclusion**

This study explored how ad vibrancy and CTAs influence consumer appeal and engagement. Findings suggest that while vibrant ads attract more clicks, they don't always enhance appeal. CTAs may not consistently boost engagement and, in some cases, can reduce ad attractiveness. Our randomized design ensured fairness, though a slight gender imbalance was noted for Grey Goose, which may influence results.

## **Key Takeaways:**

1. Vibrant ads increase clicks but don't always enhance appeal.
2. CTAs may not always improve engagement and can sometimes lower appeal.
3. Consumer responses vary, influenced by factors like gender and ad exposure frequency.

These insights emphasize the need for customized ad designs over a one-size-fits-all approach. Future research with a larger, diverse sample and real-world tracking could offer deeper behavioral insights.

## References

1. Karmarkar, Uma R., Brian Knutson, Michael Platt, and Moran Cerf. "Neuromarketing: What You Need to Know." Harvard Business Review, January 2019.  
<https://hbr.org/2019/01/neuromarketing-what-you-need-to-know>
2. "Multisensory Marketing: Engage Your Audience on a Deeper Level." Growth by Design. Accessed March 4, 2025. <https://growth-by-design.co.uk/multisensory-marketing/>

## Appendix

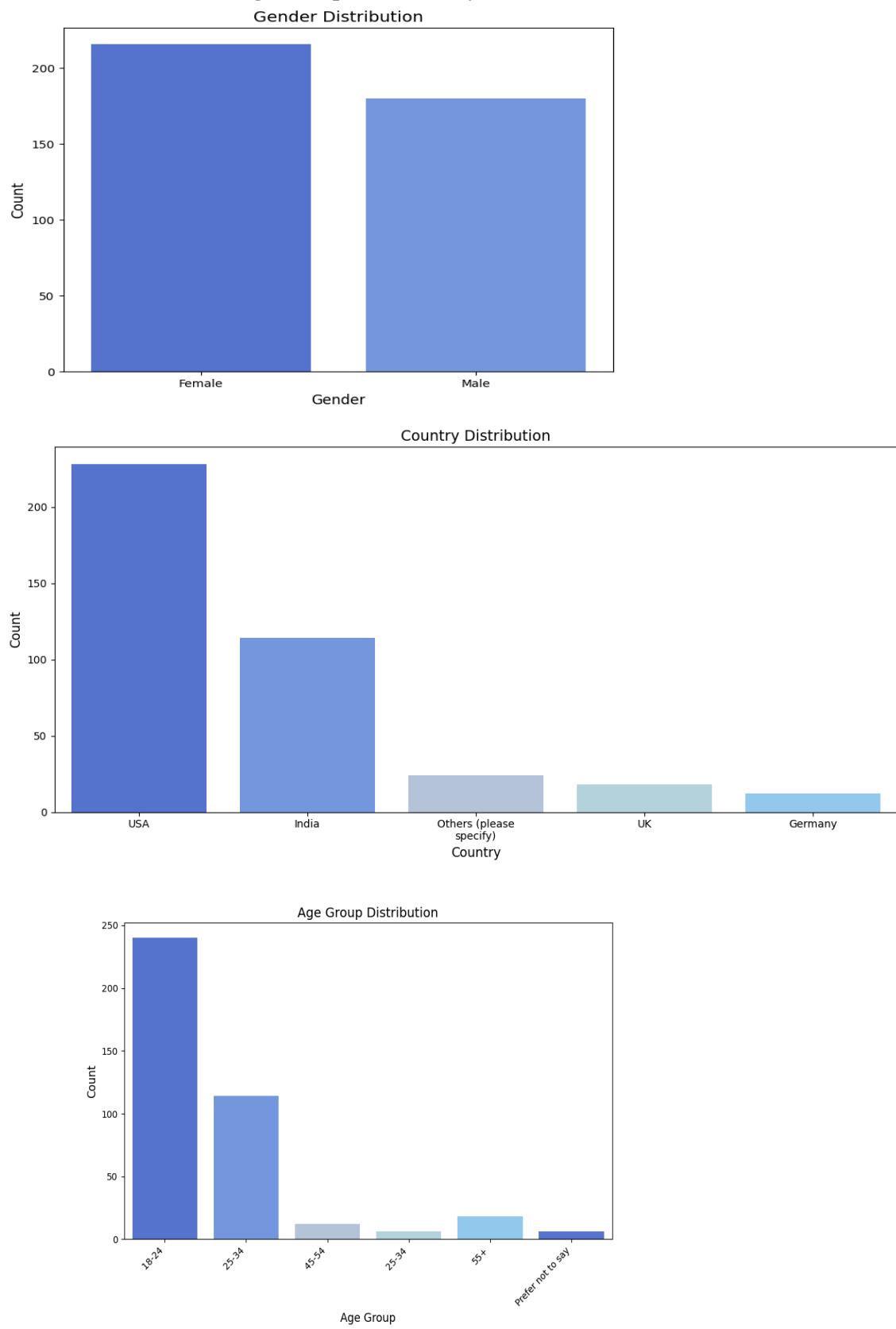
1. **Link to the Colab Notebook:** [Colab Notebook](#)
2. **Brand Images Used in the Experiment**

Below, we provide the set of brand images that were used in our survey across the three experimental groups—**minimalist, vibrant, and vibrant with a call-to-action (CTA)**—to assess variations in consumer appeal and engagement.

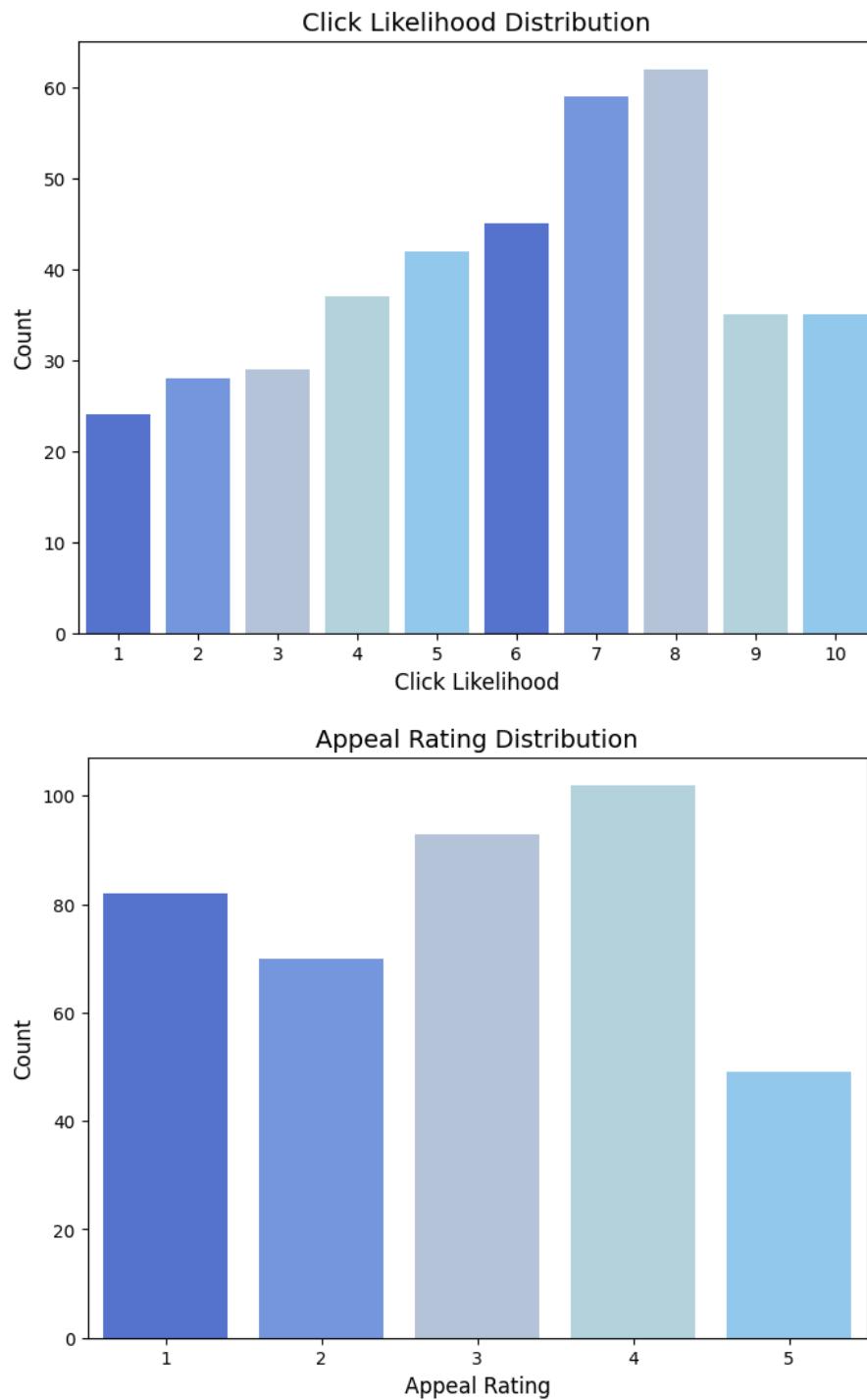
Industry - Brand	Control (Minimalist design)	Treatment 1 (Vibrant design )	Treatment 2 (Vibrant design + call to action)
Tech - Apple			
Beverage - Pepsi			
Fast Food - McDonald's			

Footwear - Nike			
Alcohol - Grey Goose			
Airline - Emirates		<p>Earth is 71% water</p> <p>we cover the rest</p>	<p>Earth is 71% water</p> <p>we cover the rest</p> <p>BOOK NOW &gt;</p>

### 3. Distribution of Gender, Age Group and Country of Residence Distribution



#### 4. Click Likelihood and Appeal Ratings Recorded in our Experiment



## 5. Randomization Test Results

Brand	Gender (C vs. Vibrant Ads)	Gender (C vs. Vibrant Ads w/CTA)	Gender (Vibrant Ads vs. Vibrant Ads w/CTA)	Age_Group (p)	Country (p)
Apple	0.258	0.663	0.464	0.617	0.173
Pepsi	0.106	0.464	0.342	0.806	0.695
McDonalds	0.124	0.415	0.554	0.710	0.318
GreyGoose	0.175	0.161	0.007	0.443	0.377
Nike	0.545	0.464	0.247	0.252	0.911
Emirates	0.895	0.622	0.702	0.371	0.956

### T-test Results:

#### Click Likelihood

Comparison	t-statistic	p-value
Control (Minimalist) vs. Treatment_1 (Vibrant) -	-2.92	0.004
Control (Minimalist) vs. Treatment_2 (Vibrant + CTA)	-1.13	0.26

#### Appeal Rating

Comparison	t-statistic	p-value
Control (Minimalist) vs. Treatment_1 (Vibrant) -	-0.17	0.86
Control (Minimalist) vs. Treatment_2 (Vibrant + CTA)	1.45	0.15