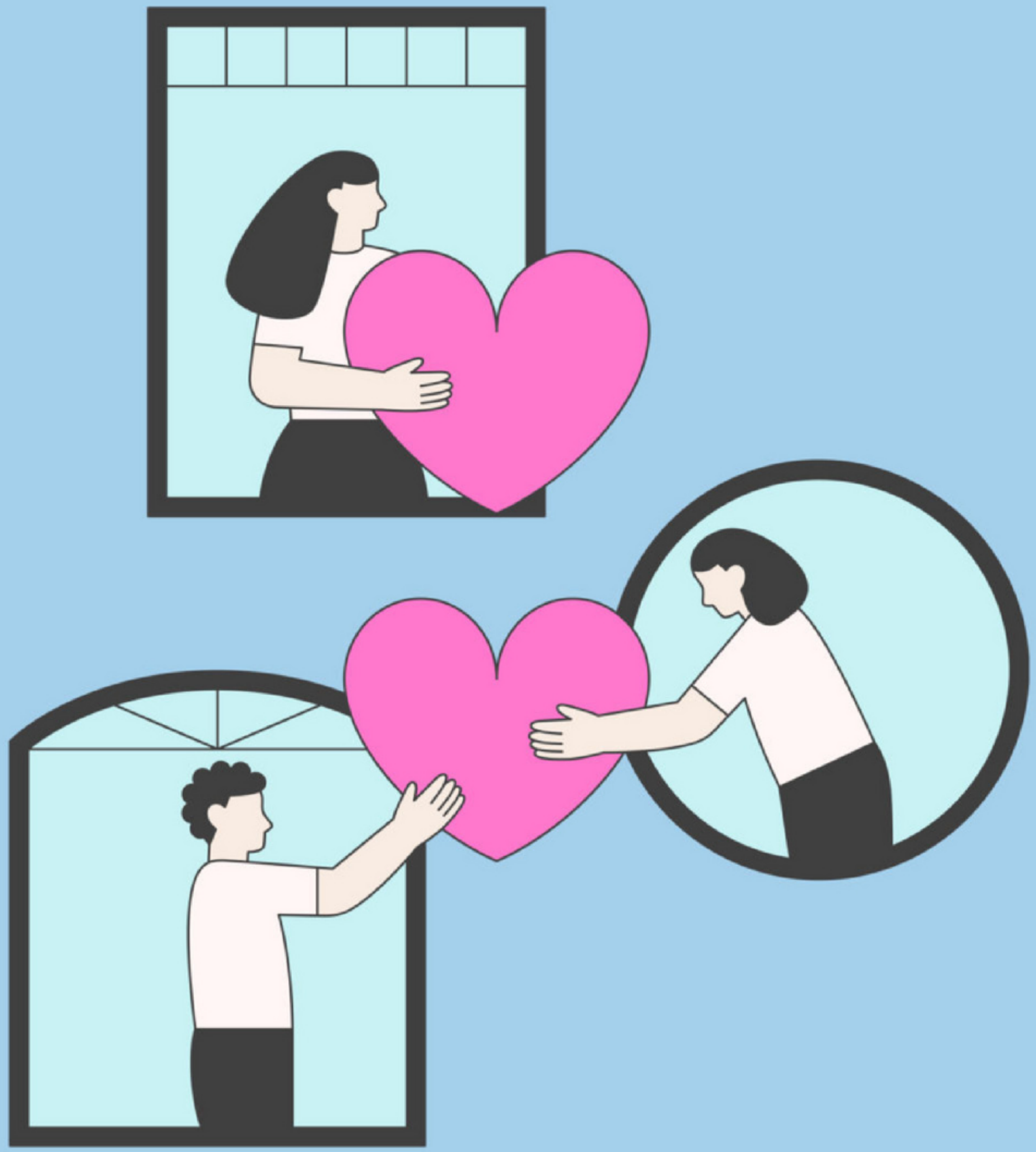
Online marketing campaign effectiveness

For Charitable Organisations



Redacted for privacy reasons

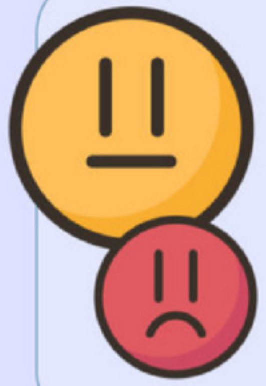
Our research highlights a number of key insights on the success of text-based campaigns to aid managers of charities.

Part 1

Emotional states, such as anger and sadness, in text-based campaigns are a driver for clicks. These states promise exponential positive effects, but should be used with caution as their effectiveness is influenced by charity context.

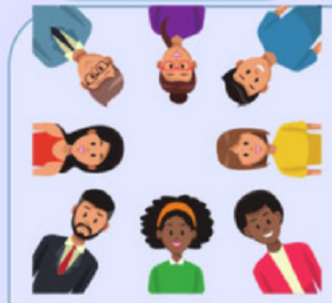
Text-based campaigns' clicks will benefit from a moderate number of words. We have found a positive effect on clicks up to a threshold value of 40 words.

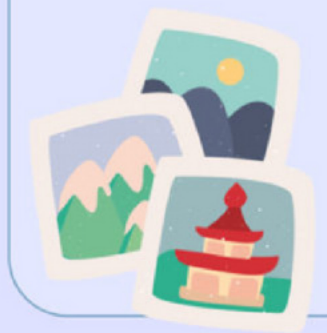
The contextual focus in charity text-based campaigns is crucial for success. We have found certain topics to enhance clicks, while others to do the opposite. Therefore, topic framing is essential for managers to get right based on their target.



Negative sentiment in text-based campaigns is a driver for clicks. Strategic use of negative connotation may enhance campaign success.

Part 2





Including images in an email, a campaign has a higher chance to get open, click or donate rather than one that doesn't include any images. We found out that images that portray people had slightly more opens and clicks.

People, they do not give attention to the format of the letters so much. Since people are more visually oriented, they were less influenced by the headers.

Data Exploration 1.1

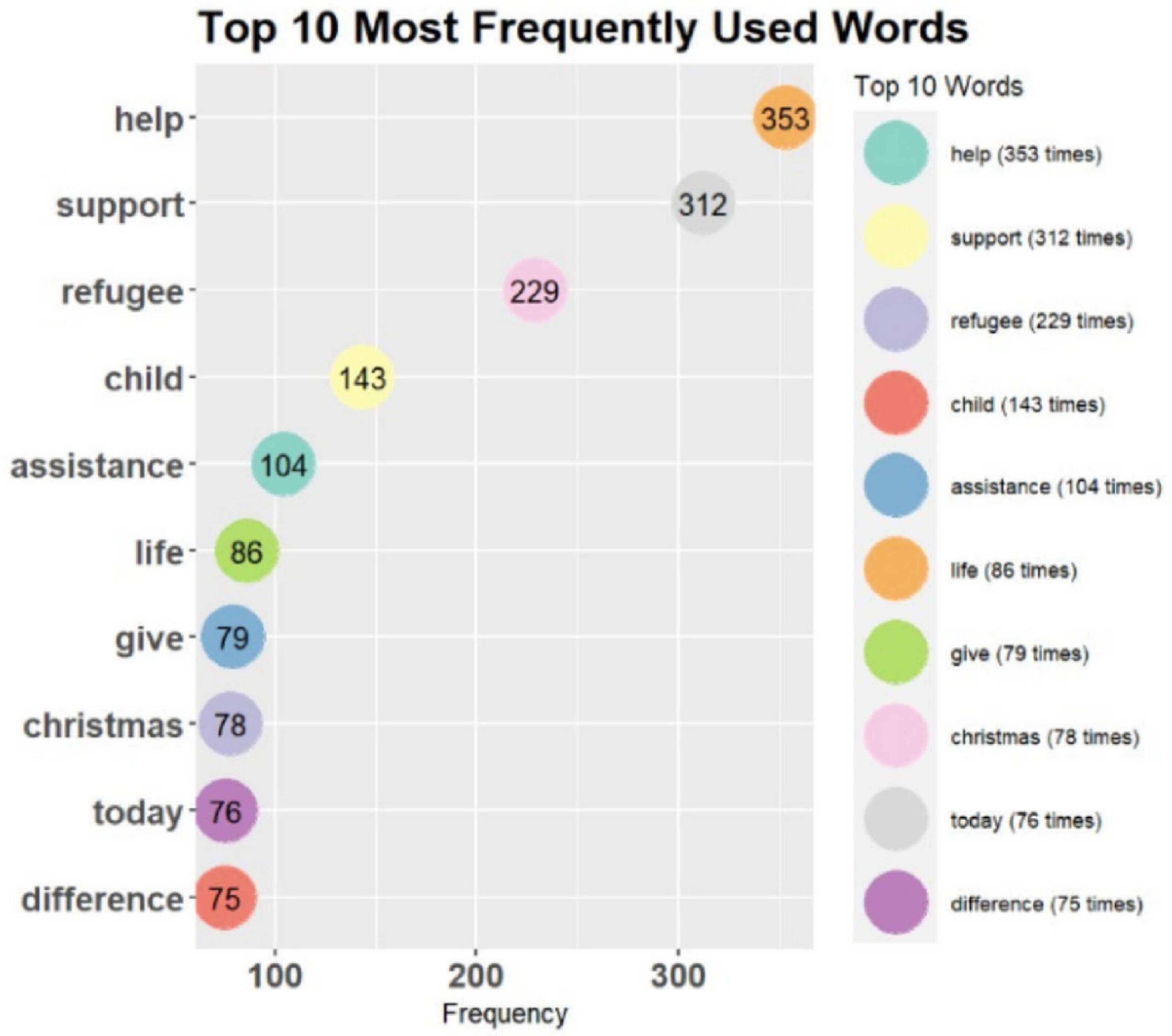
Median

The median number of clicks generated by the campaign is **233.80**

Mean

The average number of words contained in each text is **19**

Top 10 words

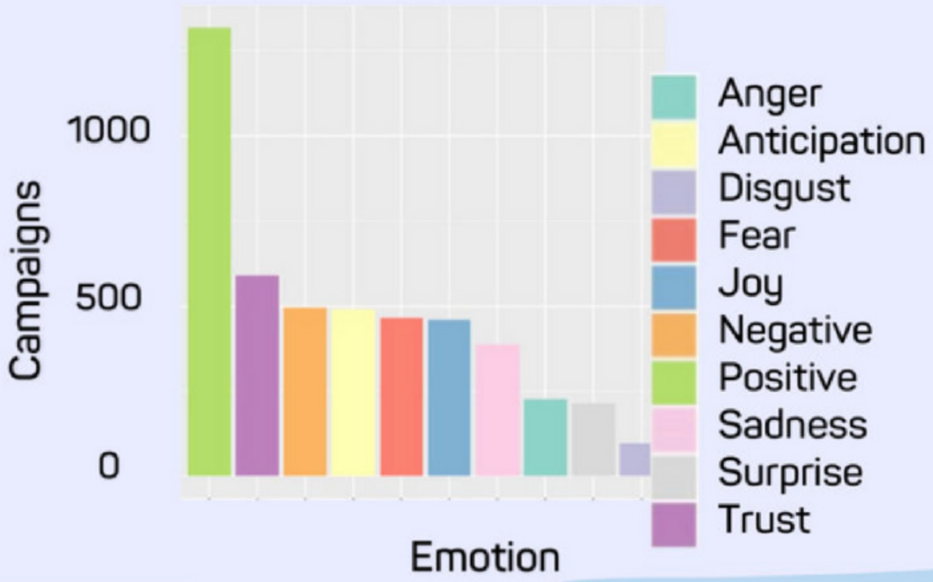


Overall Distribution of Words

Variable Generation 1.2

Distribution of emotions throughout campaigns

- Throughout the campaigns, Lexicon NRC assigned approximately 1500 words to the "positive" emotion
- "Trust" is the emotion with the second-highest frequency, with 500 words being matched to it
- The pattern is discontinued with "negative", being the third emotion with the highest frequency

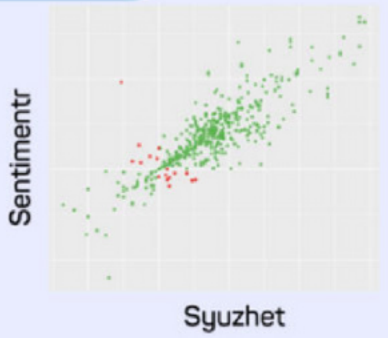


How well do the methods perform?

Syuzhet and Sentimentr agree on

96%

of the cases



- The green dots signify agreeableness between the two methods . Since most dots are green, there is consensus on the majority of observations. The number of sentences the methods agree on is **543**.
- The red dots, on the other hand, point to misclassifications. The number of sentences the methods disagree on is **22.**

Top 10 words by sentiment

Negative	Positive
Vulnerable	Support
Flee	Work
Emergency	Protect
Distress	Recommendation
Disaster	Contribution
Wild	Rewarding
Threaten	Protection
Disorder	Well
Cruelty	Sustainable
Crisis	Supporting

The number one ranked negative word has a count of 35, while the number one ranked positive word has a count of a little above 300

There are 10 times more positive words than negative on average.

Topic modelling

5

is the optimal number of topics

The value of the following algorithms is minimized:
- Arun2010
- CaoJuan2009

The value that minimizes them sits at 5

The value of the following algorithms is maximized:
- Griffiths2004
- Deveaud2014

The value that maximizes them sits at 5

Topic Exploration, Interpretation & Probabilities

We used the top ten terms per topic to come up with the following topic names:

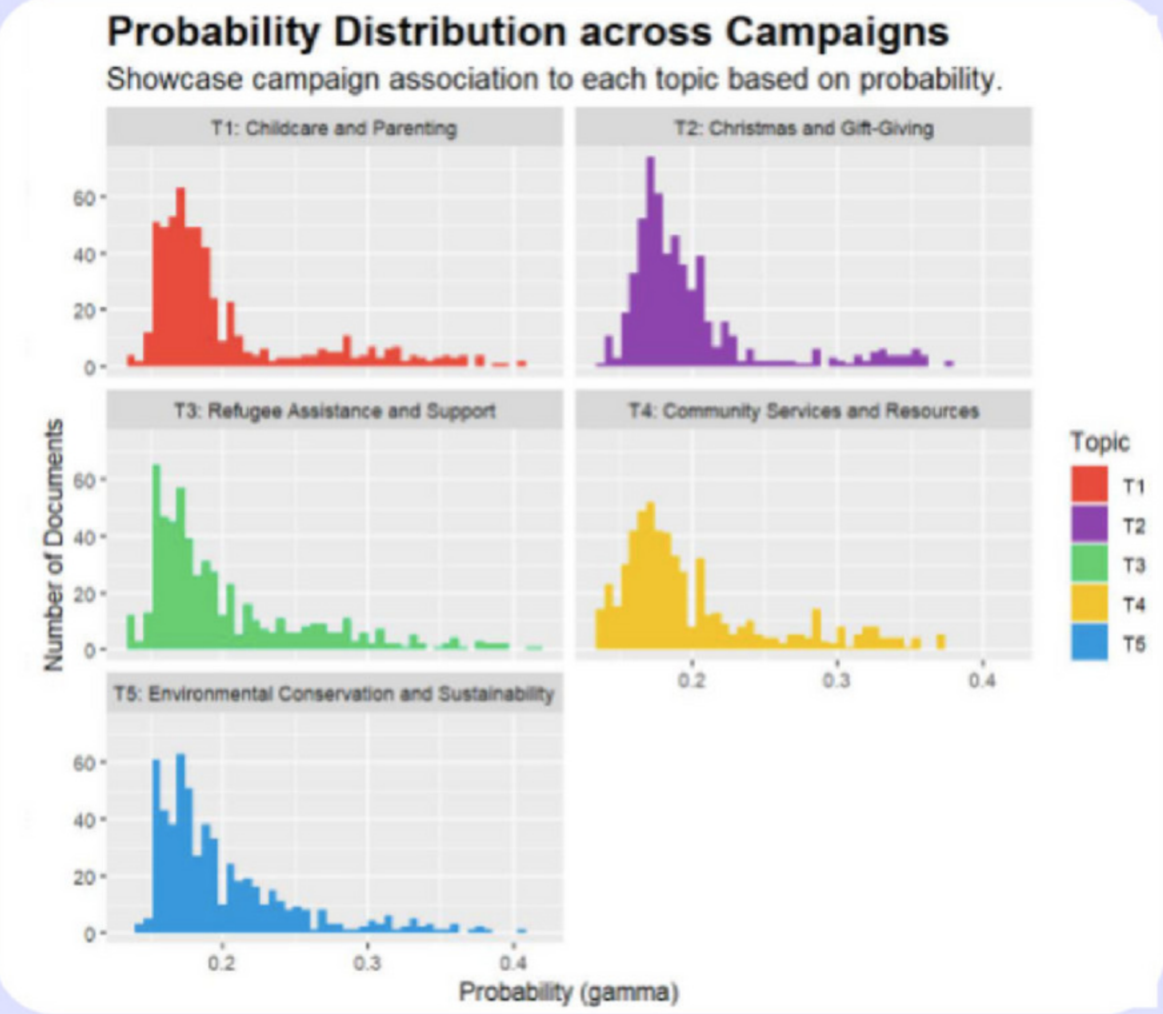
TOPIC 1
Childcare & Parenting
child · parents · care

TOPIC 2
Christmas & Gift-Giving
christmas · give · family

TOPIC 3
Refugee Assistance & Support
refugee · assistance · shelter

TOPIC 4
Community Services and Resources
services · counseling · charity

TOPIC 5
Environmental Conservation & Sustainability
world · save · nature



Findings

1.3

Methodology: Road to the final model

❶ Correlations

Investigating correlations of our variables with clicks, we already see that **negative sentiment and emotions are correlated with more clicks.**

Overview of Sig. Correlations

\+ disgust, anger, sadness, fear, nr. of words, topic 2, topic 5

− anticipation, trust, joy, afinn, sentimentr, syuzhet, topic 1

❷ Full Linear Model

Fitting a linear model with all sentiment variables, topic probabilities, and word counts resulted in **multicollinearity issues**, presenting a need to **reduce the number of variables**.

❸ **Ensemble methods:** Lasso & Elastic net

After performing the two ensemble techniques, we find them to have similarly low RMSE for out-of-sample prediction (~67). To build our final linear model we prioritize Elastic net's structure as it has one less variable (AFINN sentiment).

Based on this, we choose the following **predictors for clicks:**

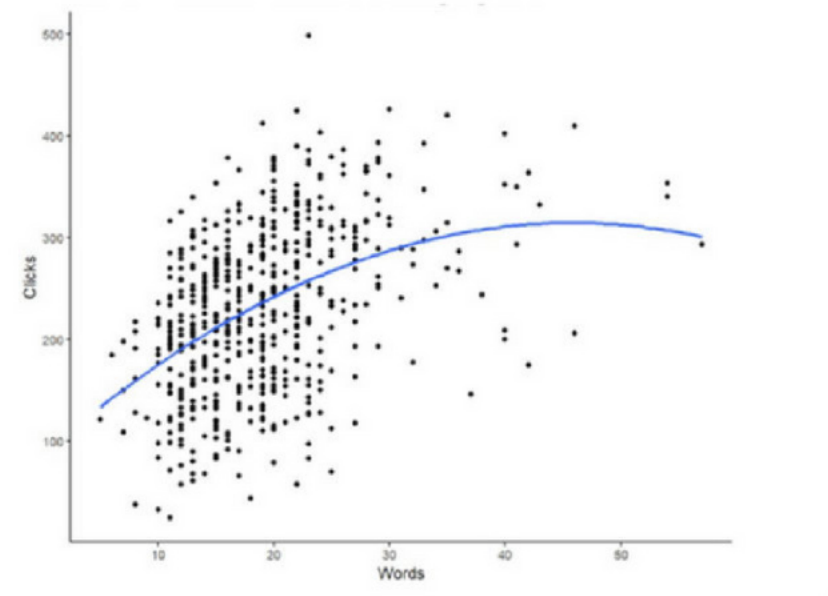
topic 2, topic 3, nr. of words, syuzhet, bing, anger, joy, highclick-outlier dummy, sadness, anticipation, fear

! Non-linear effects in: nr. of words, bing, anger, sadness

We fit a model including the quadratic term separately for each of the variables that are relevant for our final model. Afterward, we perform ANOVA's. We found the quadratic term to be significant for **number of words** ($p < .001$), **bing** ($p < .001$), **anger** ($p = 0.04$) and **sadness** ($p = 0.02$), indicating a non-linear relationship with clicks. Although we keep these variables in our linear model we do not interpret them linearly. Instead, we visualize the relationships and base our findings on that.

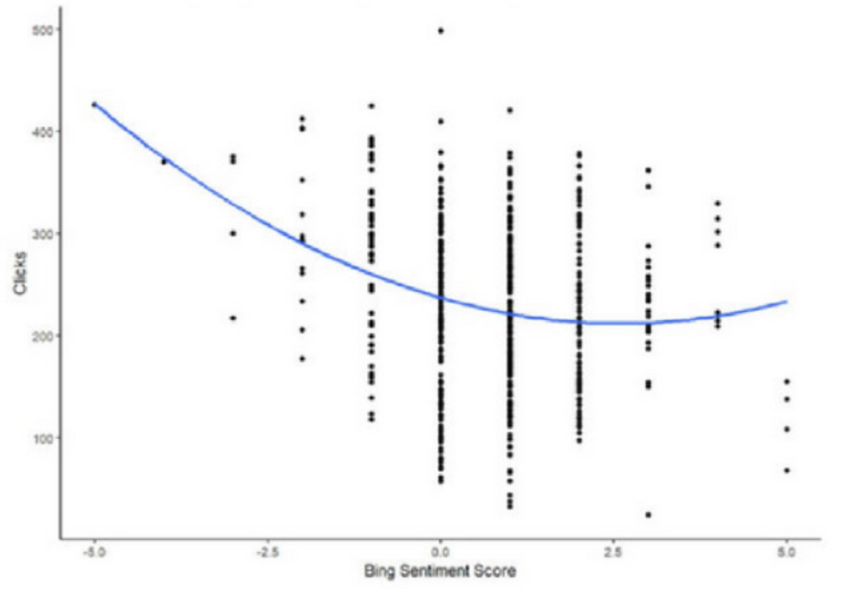
nr. of words

Initially number of words have a positive effect on clicks. This turns negative after around 40 words.



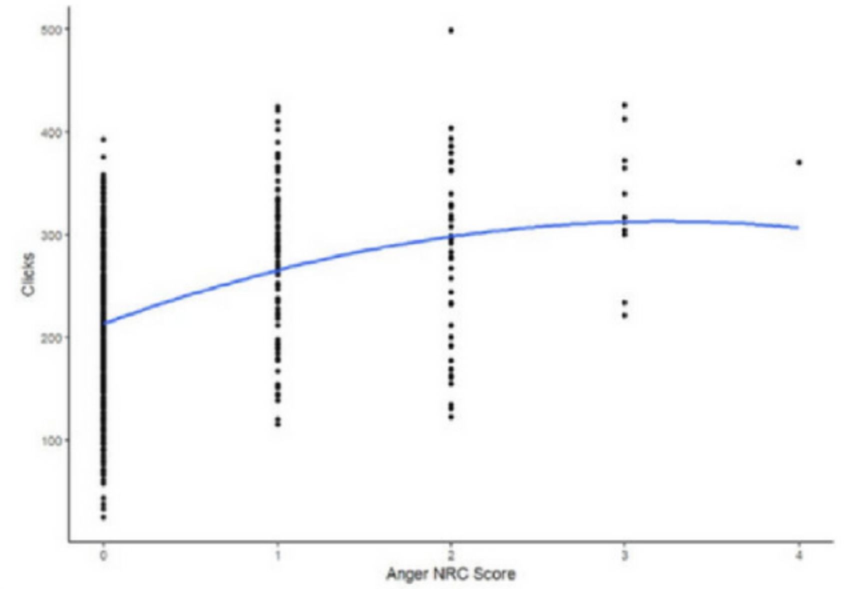
bing

As Bing sentiment turn more and more positive, clicks decrease exponentially. At 2.5 it appears to switch.



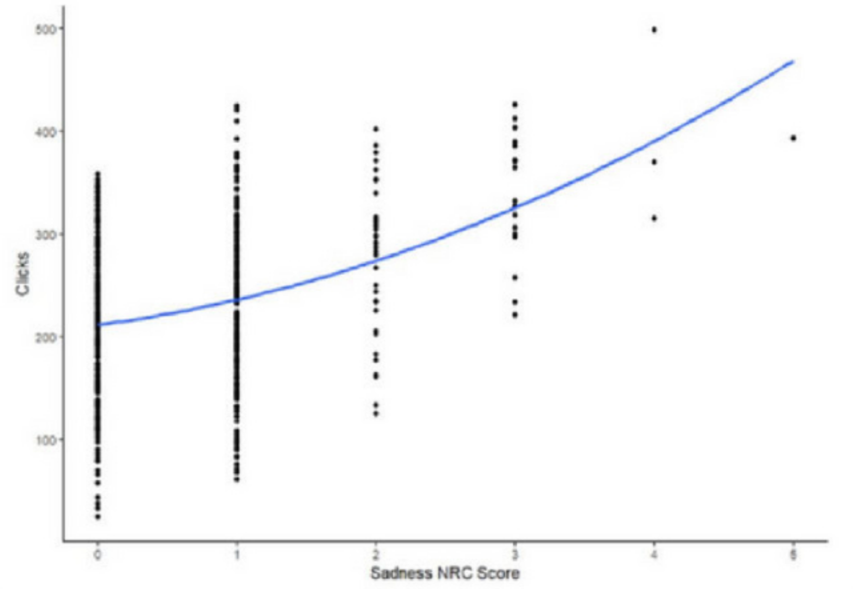
anger

Anger has a growing positive relationship with clicks, however too much anger turns it negative.



sadness

There appears to be an exponential positive relationship between sadness and clicks.



Findings

1.3

Results: Final linear model

Overall, the model is significant, and our independent variables explain 37% of the variance in clicks [$p < .001$; Adj. $R^2 = 0.37$].

We explore the significant predictors below:

Topic 2: Christmas & Gift-giving

We found topic 2 to have a significant and positive effect on clicks [$\beta = 163.69$; $p = .009$]. Therefore, holiday-themed text should result in more clicks.

Topic 3: Refugee Assistance

Topic 3 has a significant negative effect on clicks [$\beta = -125.49$; $p = .049$]. Therefore, refugee-themed text may impact clicks negatively.

Syuzhet & Bing Sentiment

Syuzhet has a significant negative effect on clicks [$\beta = -7.51$; $p = .022$]. Therefore, as sentiment decreases, clicks increase. Text with negative sentiment should result in more clicks. Similarly, Bing has an exponentially negative effect, however up to a point.

NRC: Anger & Sadness

Sadness has a significant non-linear exponential positive effect on clicks, while anger appears to have a positive relationship before hitting plateau, after which the relationship may turn negative or decrease its positive magnitude.

When we consider the above results, we can see that people in our sample are drawn to negativity and tend to pay attention to it. Further perspective is given thanks to NRC emotions, which show that anger and sadness generate clicks. Finally, Christmas-themed text should also help with clicks, however we should also consider the possibility that in the holiday period there might be more activity and generosity in general.

ML Approach: Tree models

For a more robust insight, we utilized two ML approaches, namely, a CART tree and a Bagging model.

1 **CART Tree:** confirmation of our LM results

Overall, the CART model confirms our findings. We see that the highest click branches are ones related to negative sentiment and sadness. It is especially interesting to see the most important node is anger.

We highlight the most impactful findings below:

Number of Words: for text with low anger sentiment (<0.5)

- Without taking into account sentiment: text with fewer words has higher clicks (<13).
- With sentiment, text with fewer words (<21.5) and sad sentiment (>.5) bring more clicks.
- Text with more words (≥21.5) have most clicks if they are less related to Topic 5 (<0.258).

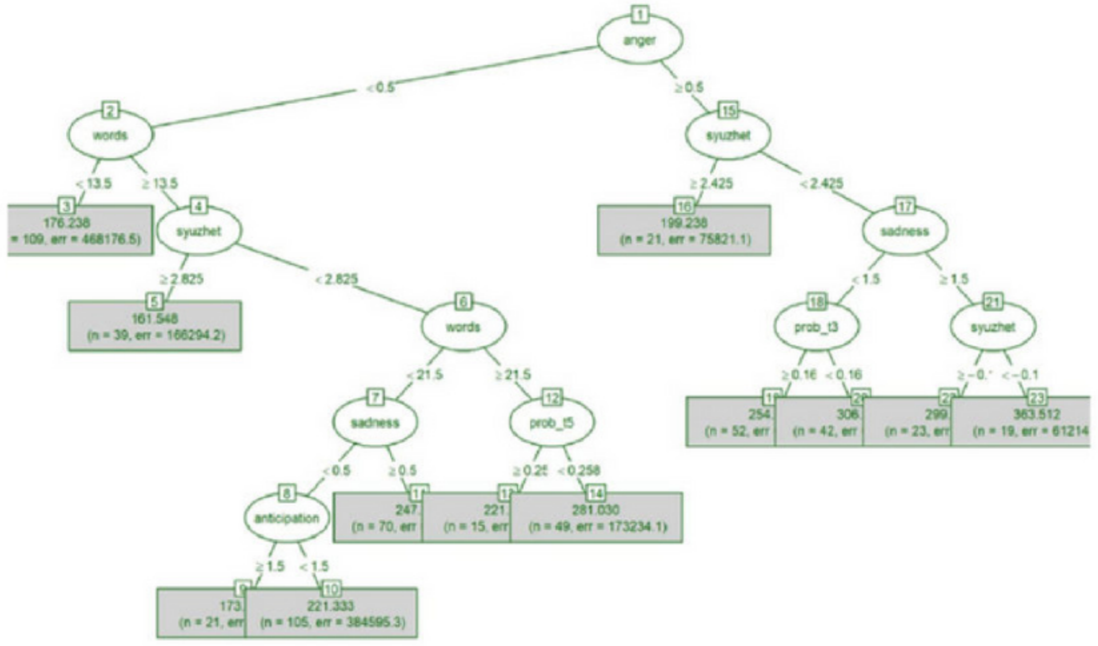
Sentiment: for text with higher anger sentiment (>0.5), without taking into account number of words.

We find the **highest click branch** here. It consists of texts with negative Syuzhet sentiment (<-0.1) and sadness (≥1.5) resulting in 364 clicks. This click number is 20% higher than the second-highest clicks branch and more than 100% higher than the lowest clicks branch.

2 **Bagging:** variable importance

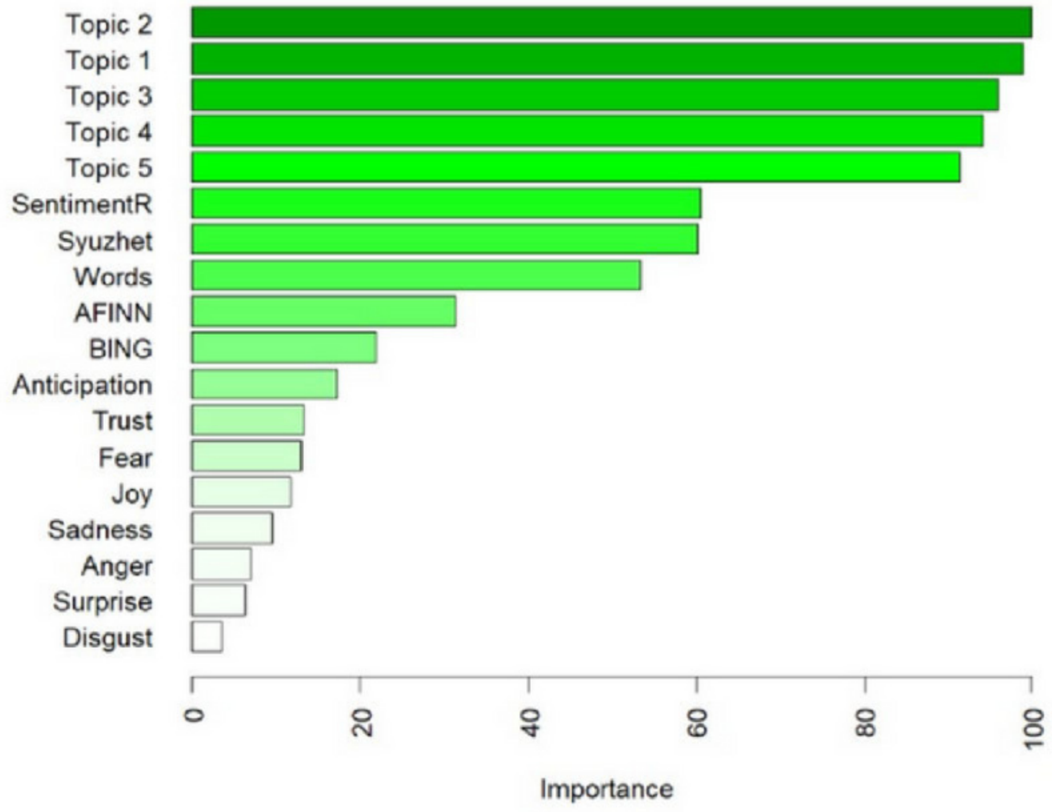
Finally, our bagging model with 500 bags, found that the topic variables are most important, with sentimentr/syuzhet sentiments and number of words next.

CART Tree Model Visualization



Variable Importance in Predicting Clicks

Based on a 500 bag model.



Findings 1.3

Discussion: based on previous literature

The Role of Emotions in Viral Content

Negative emotions

It has been found that emotions play a critical role in the likelihood of online content going viral. There is a rich body of literature on this, but especially relevant are the findings of Berger & Milkman (2012), who demonstrated that content that elicits high-arousal states such as anger or anxiety tends to be more viral than low-arousal states such as sadness. These findings partially align with those of our linear regression, as we have also found that anger has a positive influence on clicks, however, we also find sadness to have this effect. Anger is further confirmed to be a key in driving clicks by our CART, which has it as the root node.

Positive emotions

In contradiction, to our findings, Berger & Milkman (2012) find that positive emotions like awe and amusement are most likely to make content go viral. However, it is mentioned that this relationship is complex and context-dependent.

The Role of Social Influence

Another interesting finding is that successful online campaigns tend to leverage social influence and mimicry (Dafonte-Gómez, 2015). Therefore, it might be beneficial for charity organizations to portray people being kind and giving. This also aligns with the significant positive effect Topic 2, Christmas & Gift-giving, has.

Context & Beyond

Recent research has found that campaign success of charities depends on factors such as the type of charity and the gender of the consumer, in relation to which emotional appeal is more effective. For example, charities about animals benefit from sadness, while those for children benefit from joy. (Kwon et al., 2022)

	Expected	Unexpected
STRONG	• **Anger** has been shown to have a positive relationship with likelihood of virality. We also see this, albeit non-linearly. • **Social influence** has been shown to have a positive relationship with likelihood of virality. We also see that campaigns related to **gift-giving** have a positive relationship with clicks, showcasing the power of social influence.	• **Sadness** has been shown to have an exponential non-linear positive relationship with clicks. This goes somewhat against the findings of previous literature, however not entirely. Namely, latest literature finds the important role of context where sad emotions may actually drive clicks, for example in the case of charitable campaigns related to children (Kwon et al., 2022). We also further demonstrate this in our CART model.
WEAK	• **Context,** or in our case topics, have been shown by our bagging model to be most important in predicting clicks. This although not conclusive may be showcasing the importance of context aligning with existing literature. • **Number of words** has an expected non-linear relationship, where up to a threshold (~40 words) value words positively affect clicks, however turn negative after.	• **Joy and other high-arousal positive states** have been shown throughout literature to be significant and positively driving likelihood of virality and campaign success. In the case of our results, we did not find a significant effect of joy or other positive high-arousal states. • **Refugee-related** topic has a negative relationship with clicks, this may be due to poor context-emotion match.

Data Exploration 2.1

Randomization

Number of days

All three groups have a similar mean number of days since the last email was sent, at **135** days.

Past donations

There are differences between the groups; the image_landscape and image_person groups have similar mean values (around **262-263**). The no image group has a lower mean value, at **182.**

Past visits

All three groups have similar mean values (around 5.94-5.95).

Age & Gender

The mean in age is very similar across the three groups, ranging from 45.0 to 45.1.

The mean gender (which is likely a binary variable indicating male or female) is also very similar across all three groups, ranging from 0.498 to 0.500.

age and gender of the recipients in each group are not significantly different from each other, and that these variables are not likely to have a strong impact on any differences observed in the previous code block.

Past donations greater than 0

All three groups have a mean proportion of past donations greater than zero of 1, which suggests that all recipients in each group have made at least one past donation.

Past visits greater than 0

A high proportion of recipients in each group have made at least one past visit, with mean proportions ranging from 0.998 to 0.999. This suggests that the recipients in each group are likely to have some level of engagement with the organization or cause represented by the email campaign.

Randomization was **partially successful.**
The groups are similar in most respects; except for the differences spotted in the means of past donations.

Descriptive statistics

MEAN	opened emails	clicks	donations
image_landscape	0.652	0.148	51.7
image_person	0.718	0.185	51.2
no_image	0.229	0.0586	24.8

Do images and headers drive the rates of opened emails, clicks and donations?

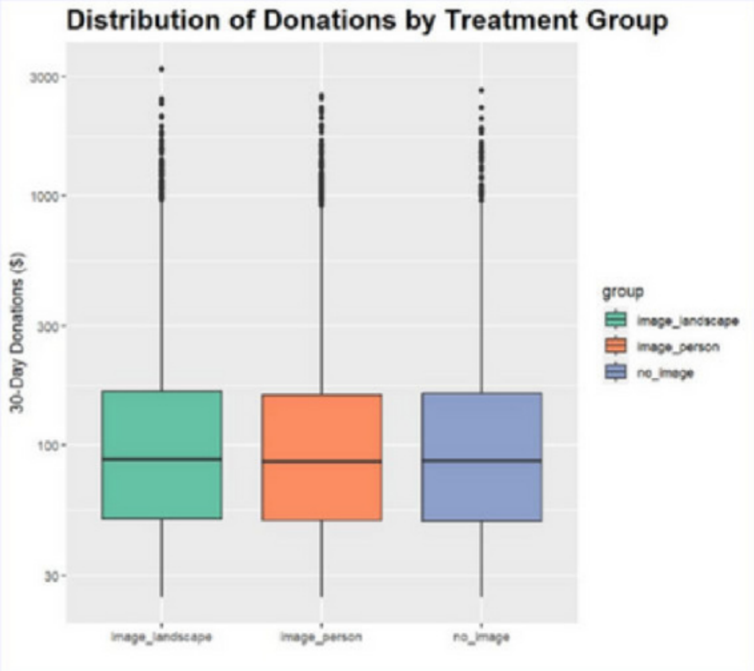
Based on the descriptive statistics, there is a difference in the means of opened emails, clicks and donations between the image and the non image groups. To investigate whether the difference in means is significant as well as the difference in the means of the header/no header groups, we conducted a T-test.

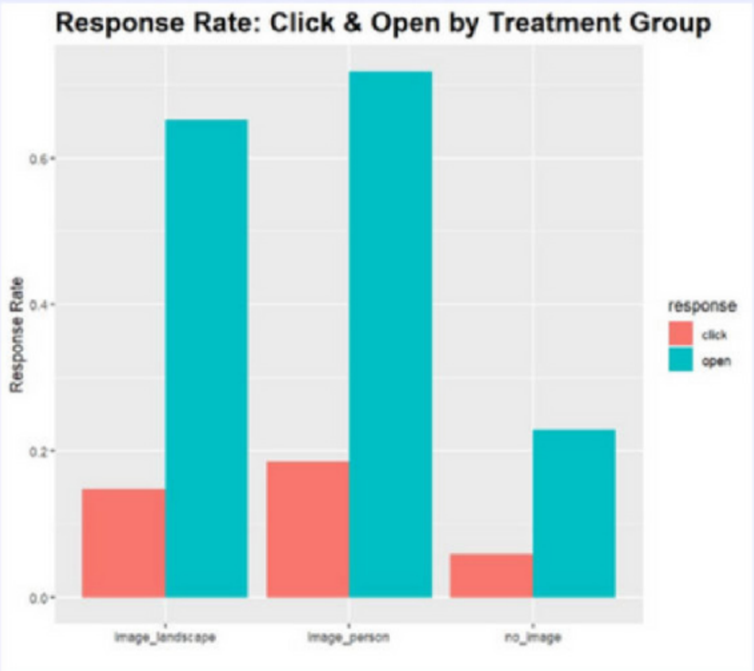
	open	clicks	donations
image vs no image	p=2.2e^-16	p=2.2e^-16	p=2.2e^-16
header vs no header	p=0.936	p=0.812	p=0.03064

1. The results indicate that the difference between the means of the image and no image groups is significant. Therefore we can conclude that emails with images increase the number of opened emails, clicks and donations.

2. The results indicate that the differences between the means of the header and no header group is not significant for open and click rates but it is significant for donations. Therefore we can conclude that emails with headers do not increase the number of opened emails and clicks but they do incresse the amount of donations.

To investigate our findings (1) even further, we conducted a proportion test to to analize the effectiveness of different email designs (images vs no images) in terms of email open rates, click rates and donation rates. The output of this test suggests that the difference in proportions of the two groups is statistically significant, with a very low p-value (less than 2.2e-16).





Are there are any interaction between headers and images?

In order to test for interaction effects between the two independent variables (headers and images), a linear regression model is conducted. We include an interaction term to the model and test its significance.

	Estimate	p-value
Header	-0.0023226	<2e-16 ***
Image	0.4557718	<2e-16 ***
Header: image	0.0005824	0.915

1. As seen by the significance levels (p-value.> 0.05), the null hypothesis that there are no interaction effects between the two variables cannot be rejected. Therefore, there is no significant interaction between headers and images in terms of email open rates. The effect of image on email open rates is different at different levels of headers, and viceversa.
2. However, as seen by the high positive coefficient (0.4557718), the presence of an image is strongly associated with higher open rates. Therefore, it is recommended to use emails with images to maximize open rates. It is worth experimenting with different types of images to determine which ones are most efective at capturing recipients' attention.

Do the effects vary depending on donor characteristics?

To investigate whether the effects of emails containing headers or images varied depending on donor characteristics, we executed a linear model

age	gender	age:gender
p=0.4102	p=0.0831	p=0.2066

Based on the p-values, we can conclude that the age and gender predictors are not significant at the 5% significance level, as their p-values are greater than 0.05. The p-value for the interaction term is also not significant, indicating that the relationship between age and donation does not vary significantly by gender.

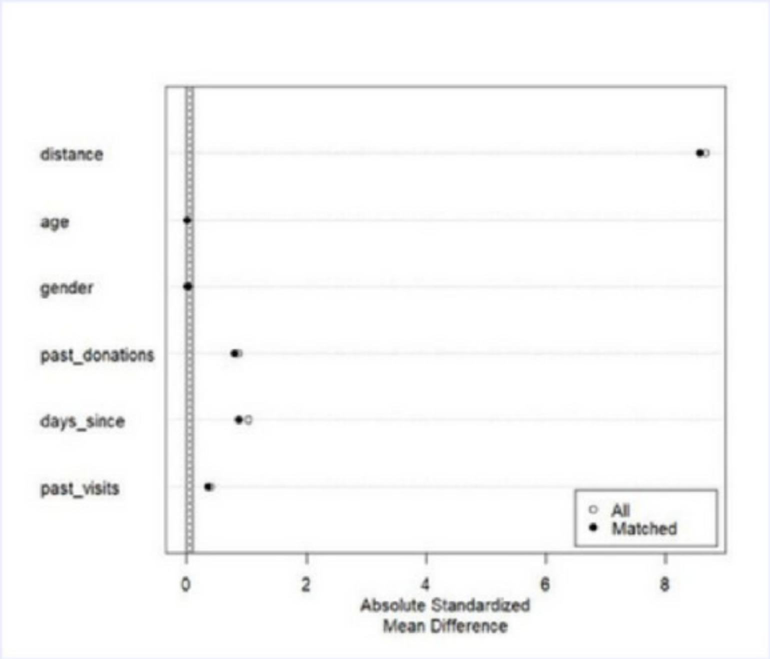
Investigating the effect of individual profiles in donation behavior

To investigate whether individuals who did a donation have different profiles than those who did not make a donation we conducted the following steps:

1. We create dummies to see who donated and who didn't.
2. We ran a binomial logit model to predict the probability of donating based on different individual's profiles. The results indicate that past donations and past visits are significant in predicting the probability of donating.

Since the results indicate that different profiles have an effect on donation behavior, we conducted a propensity score matching to correct for such differences.

In contrast, with the pre-propensity score matching, only past visits appear to be significant in predicting donation behavior post matching



We then examine how the effect of images/ headers on donation amount change **post matching**, by conducting a linear model.

header	image
p=0.04756	p=0.06361

After correcting the effects of different profiles in donation behavior, we found that only headers have statistical significance to predict donation amount.

References

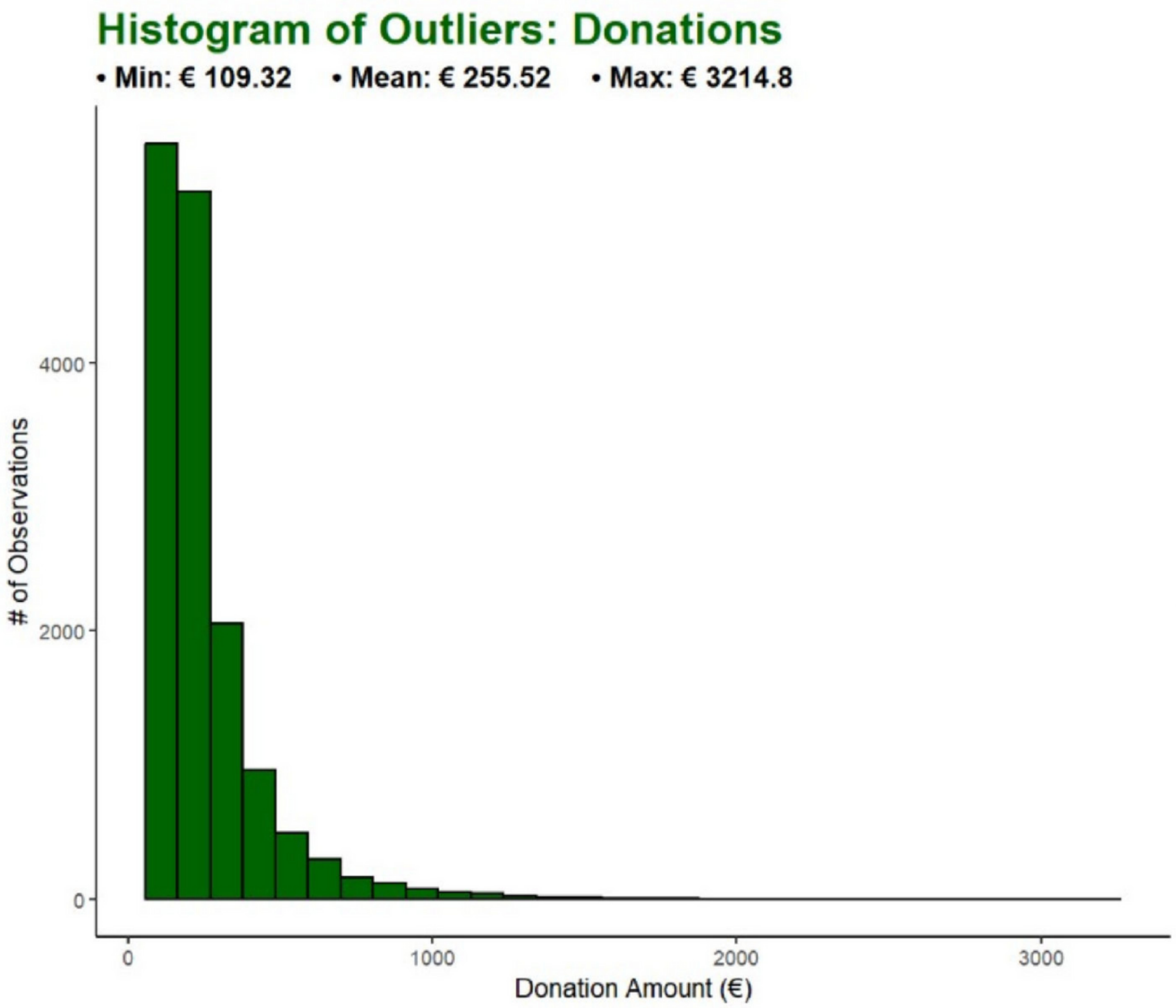
- Berger, J., & Milkman, K. L. (2012). What makes online content viral? Journal of Marketing Research, 49(2), 192-205.
- Dafonte-Gómez, A. (2015). The Key Elements of Viral Advertising. From Motivation to Emotion in the Most Shared Videos. Comunicar, 45, 139-146.
- Hlavac, M. (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. Central European Labour Studies Institute (CELSI). https://CRAN.R-project.org/package=stargazer
- Kwon, J., Lin, H., Deng, L., Dellicompagni, T., & Kang, M. Y. (2022). Computerized emotional content analysis: empirical findings based on charity social media advertisements. International Journal of Advertising, 41(7), 1314-1337.

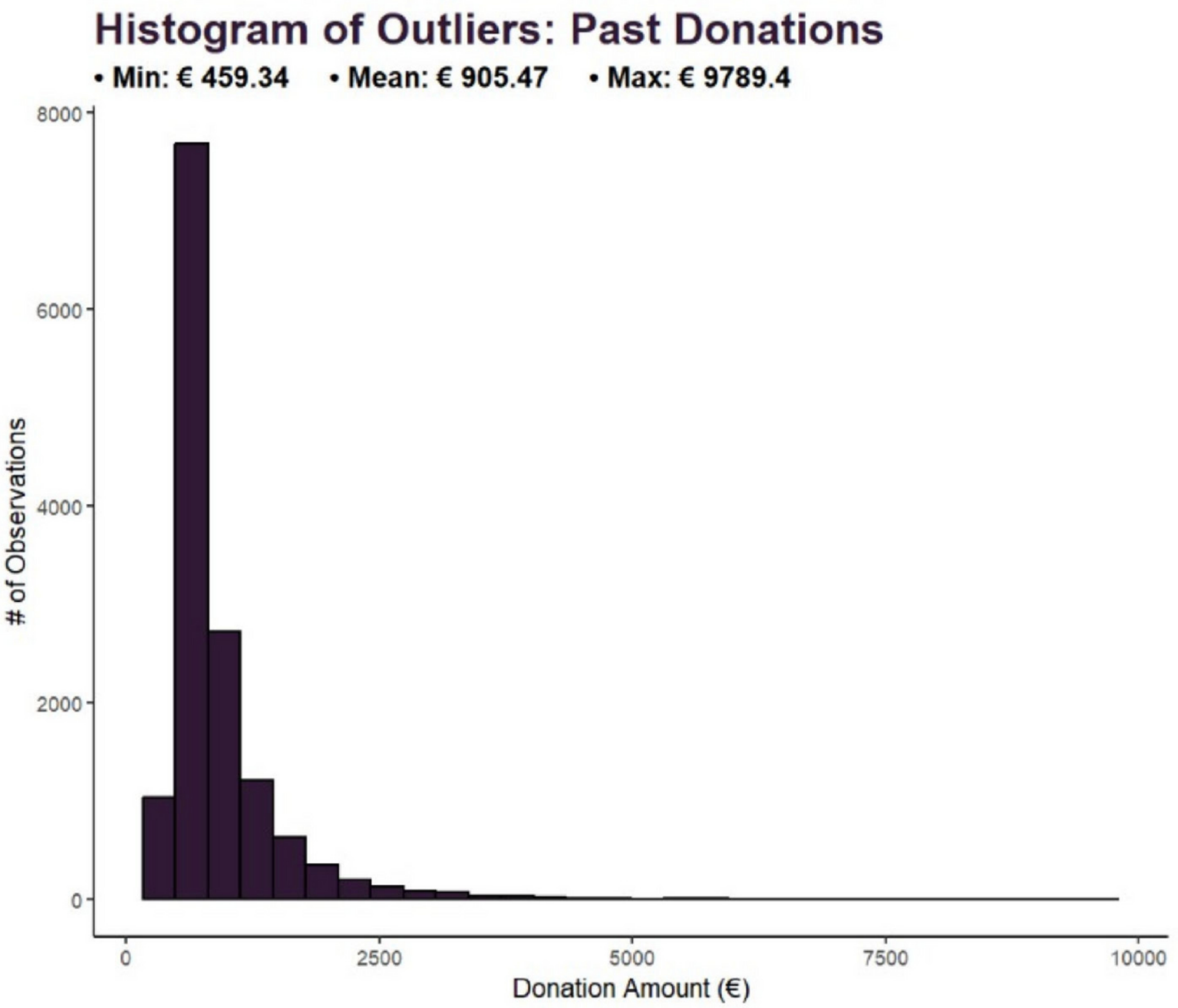
APPENDIX A

The **plots** behind the insights.

APPENDIX A-1.1

Histograms of observations classified as outliers.



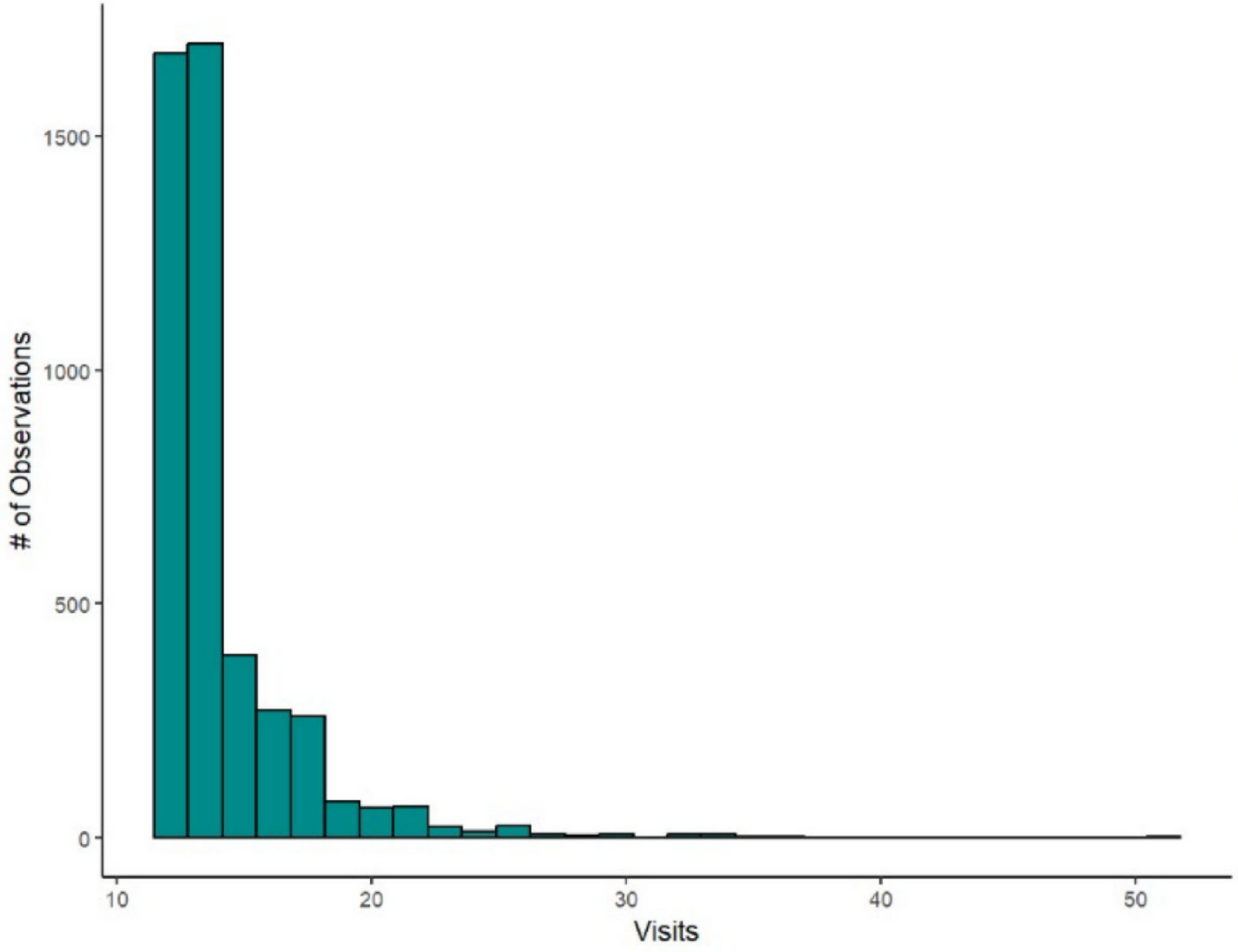


APPENDIX A-1.2

Histograms of observations classified as outliers.

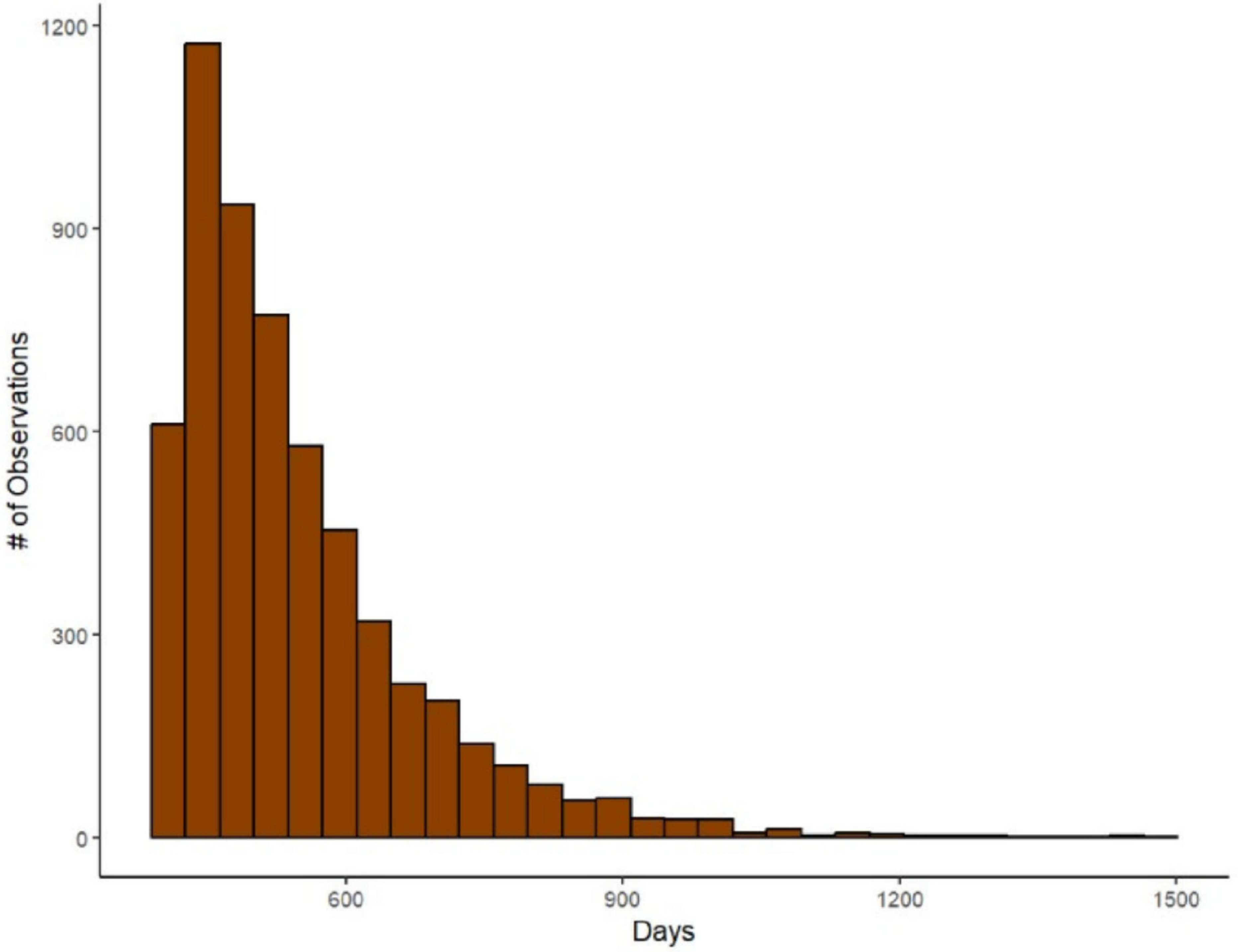
Histogram of Outliers: Past Visits

• Min: 12 visits • Mean: 14 visits • Max: 51 visits



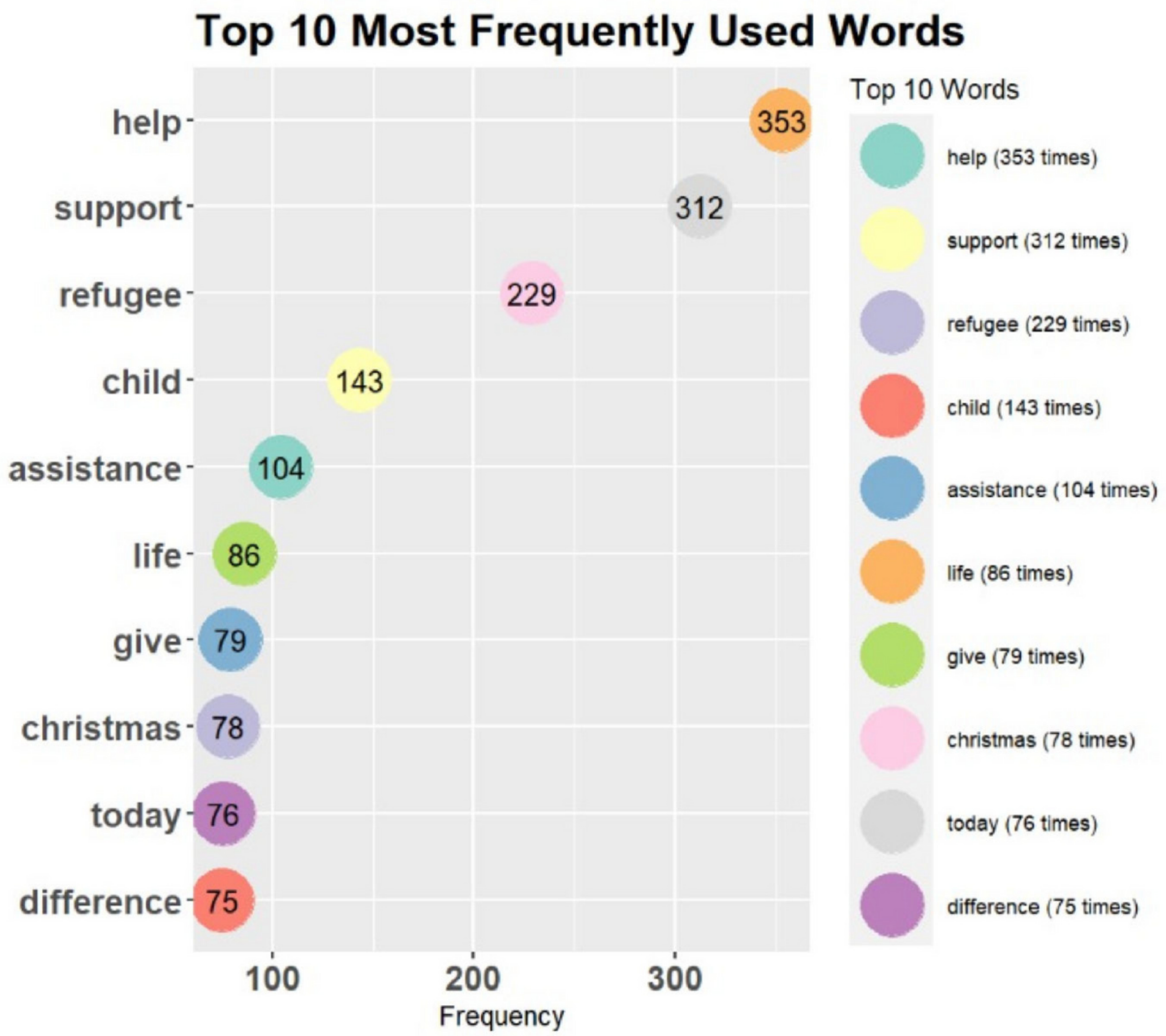
Histogram of Outliers: Days Since

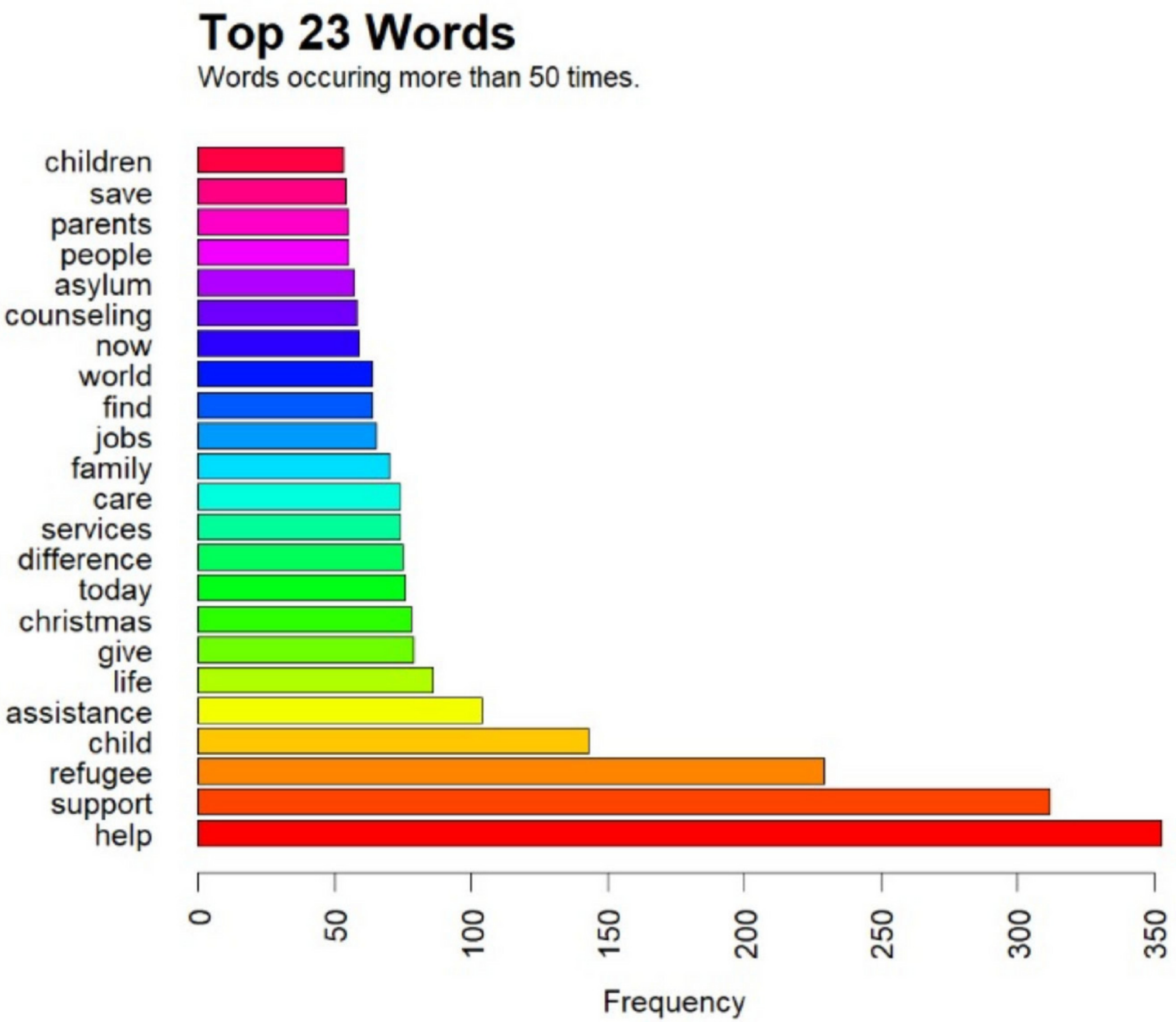
• Min: 413 days • Mean: 548 days • Max: 1488 days



APPENDIX A-2.1

Exploring word frequency through text mining methods.



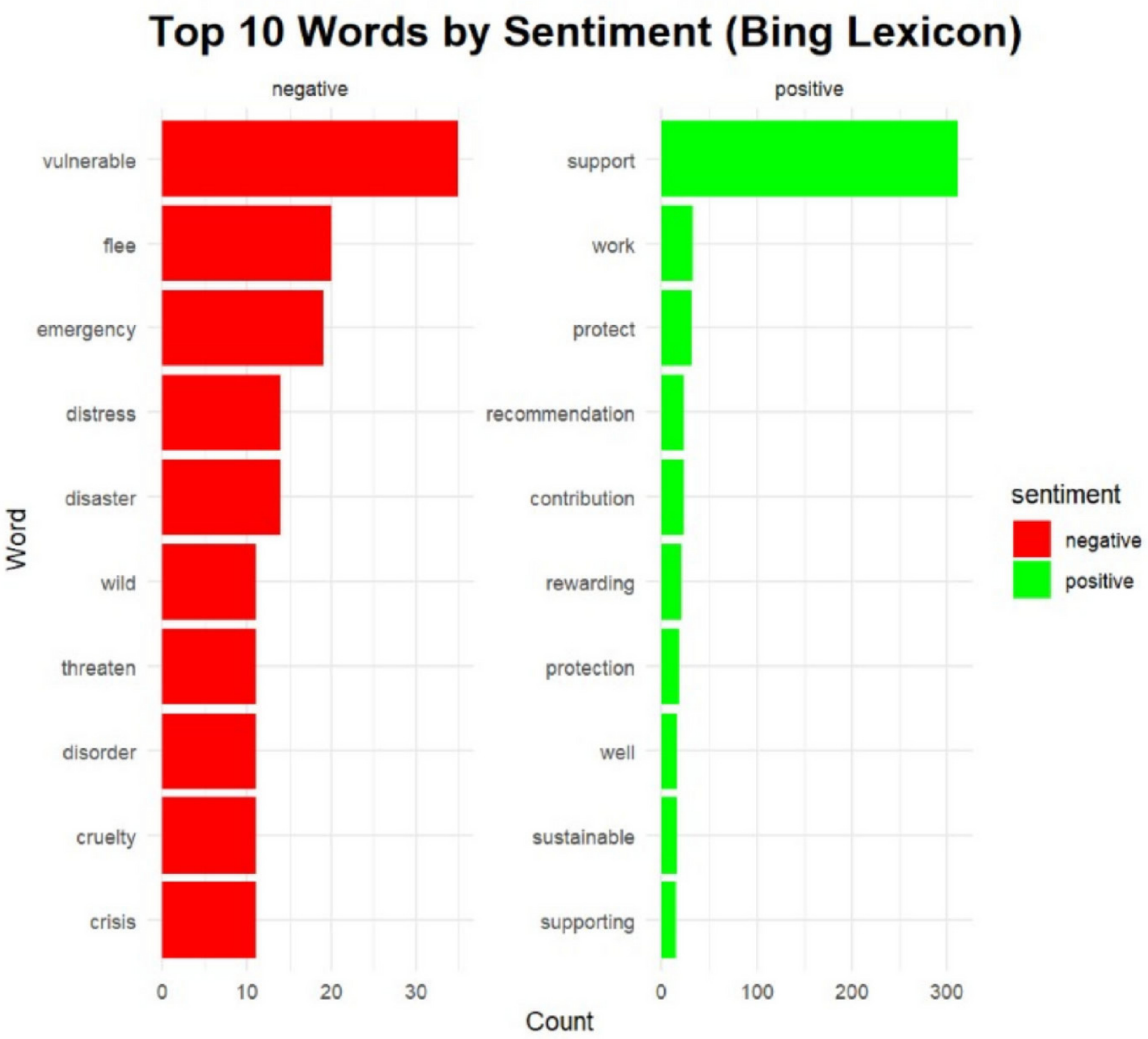


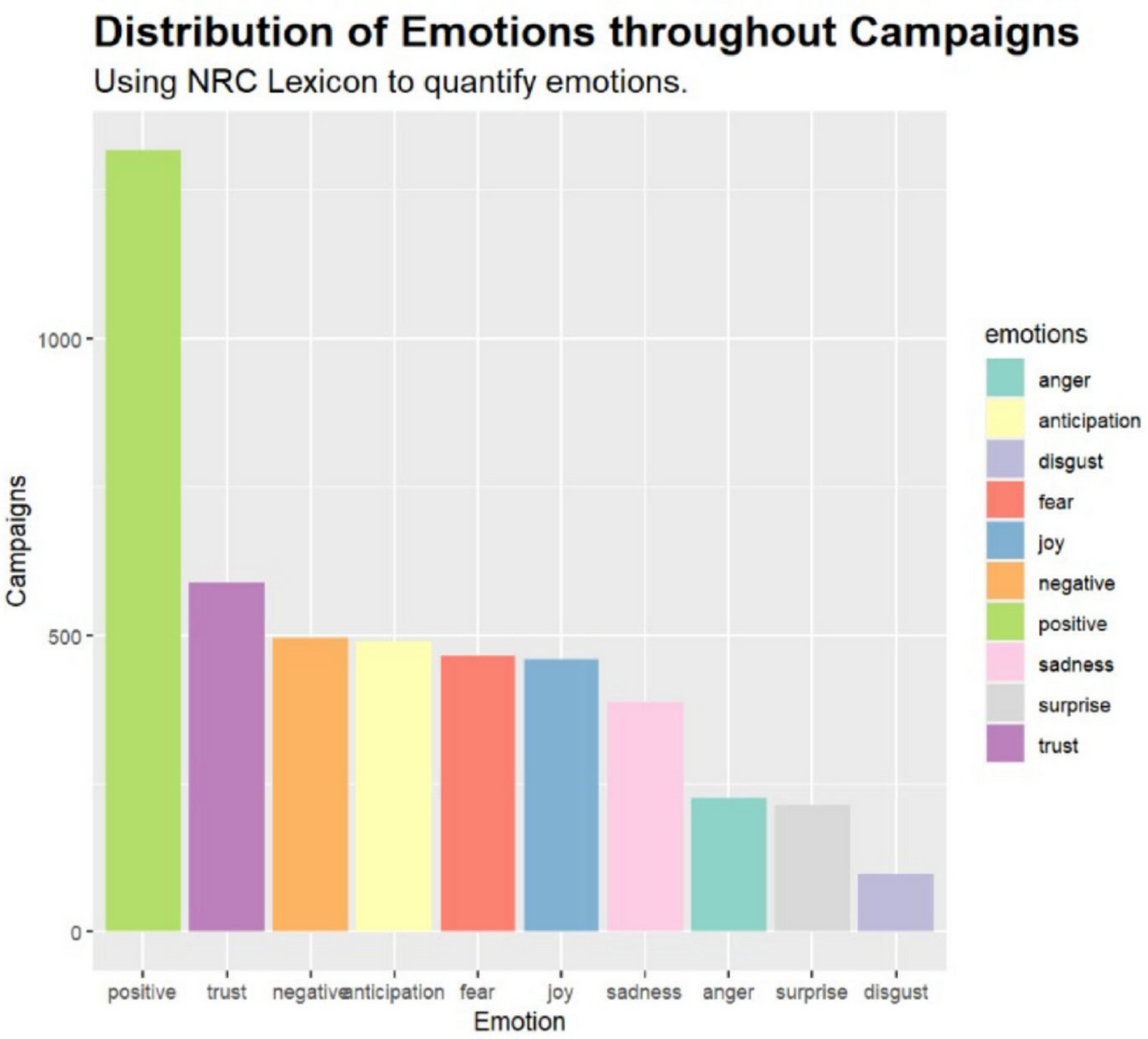
APPENDIX A-2.2

Exploring word frequency through text mining methods.

APPENDIX A-3

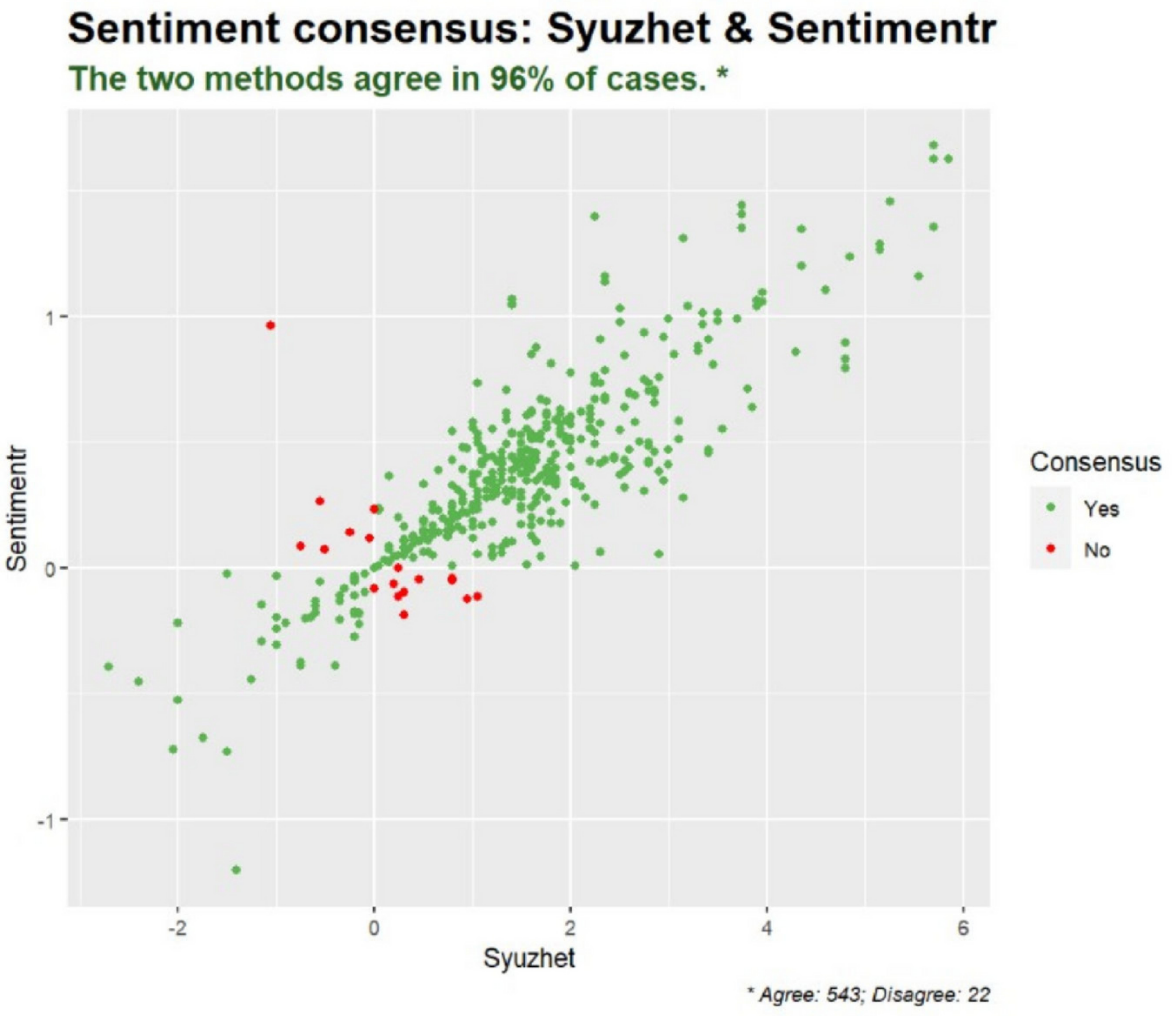
Employing Lexicon-based methods in combination with text mining to better understand our data.

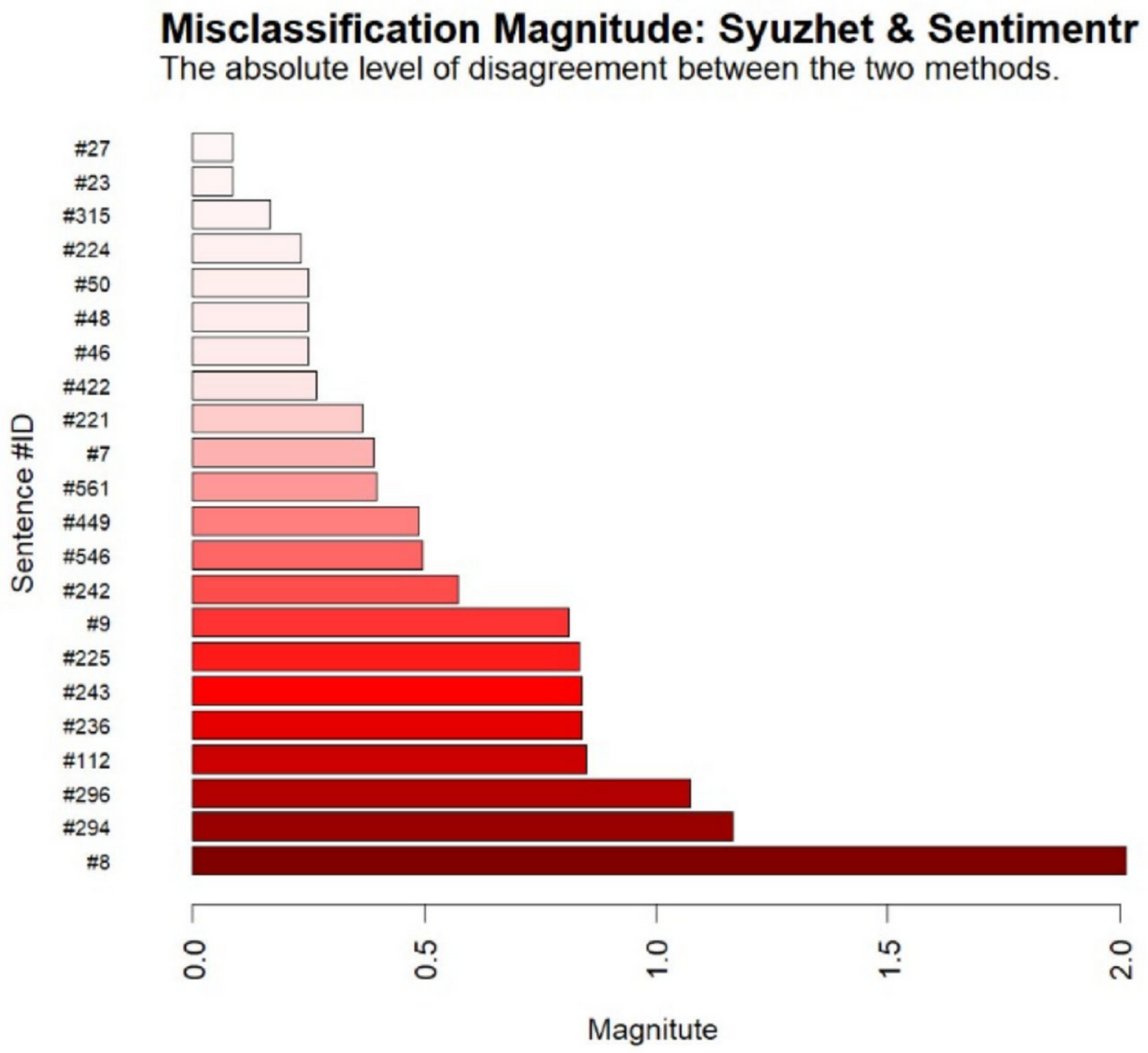




APPENDIX A-4

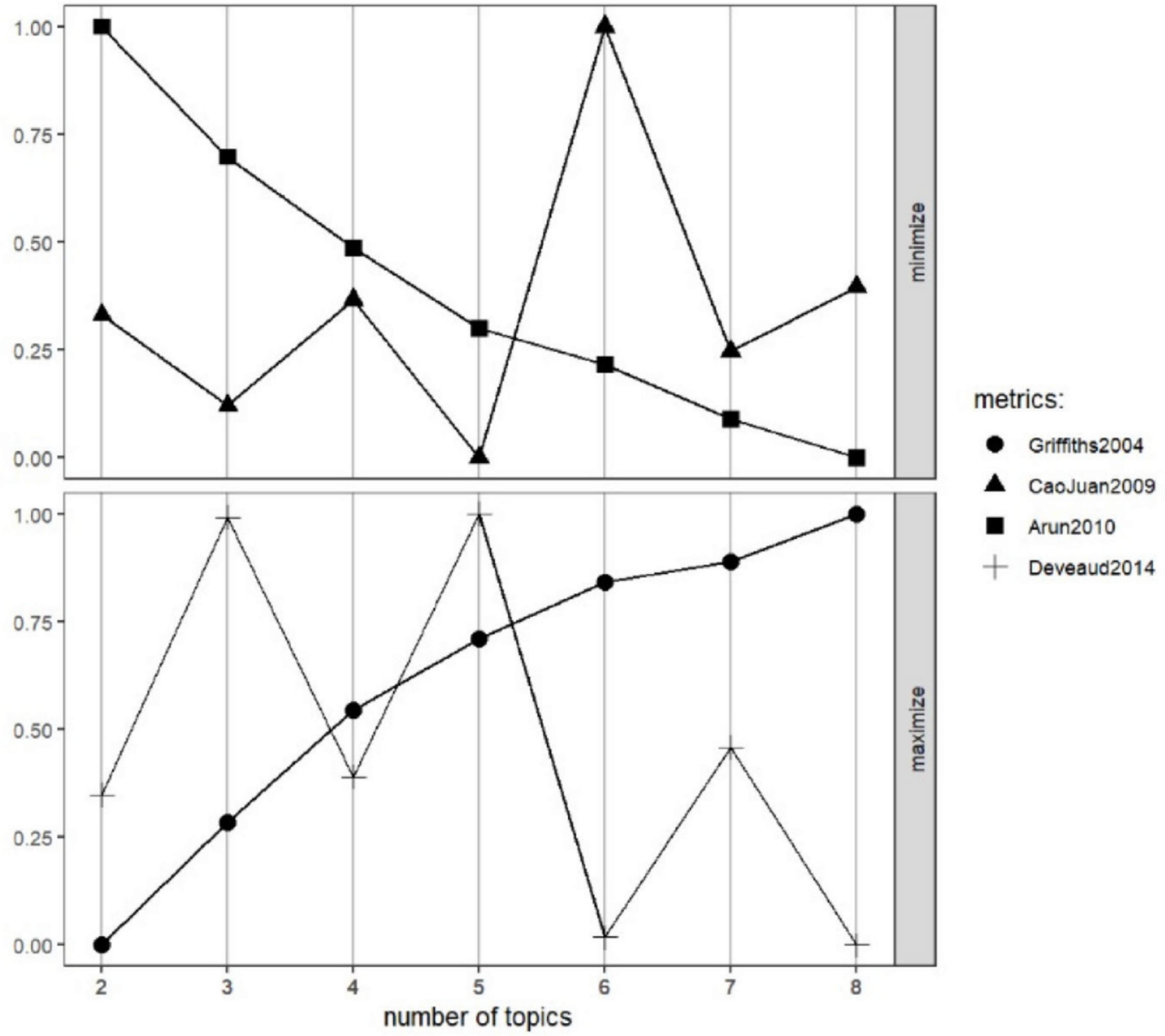
Assessing the agreement between the Syuzhet and Sentimentr models to understand the quality of their predictions.

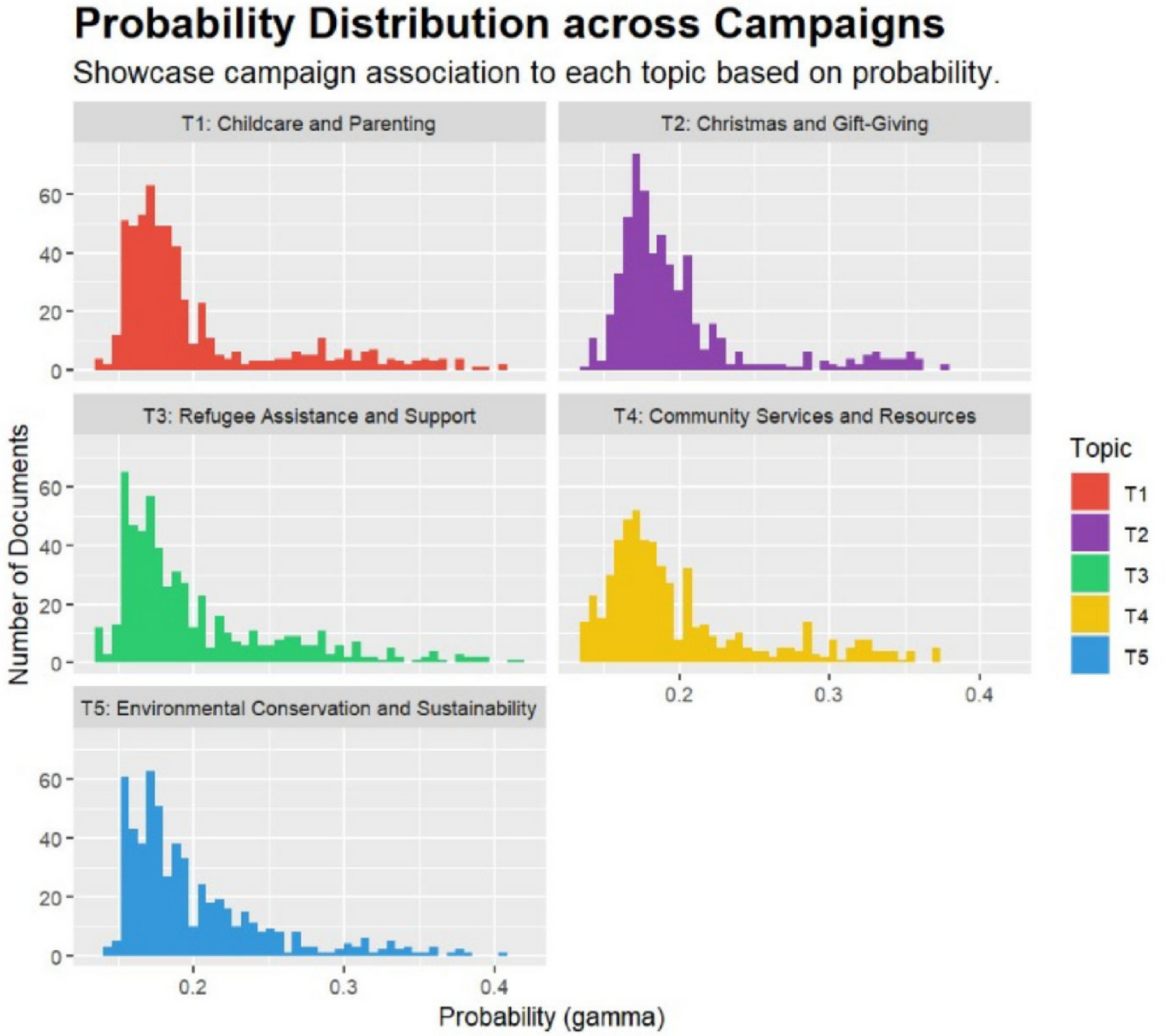




APPENDIX A-5

Performing topic modeling, firstly choosing the number of topics, then exploring the distribution of theta across documents.

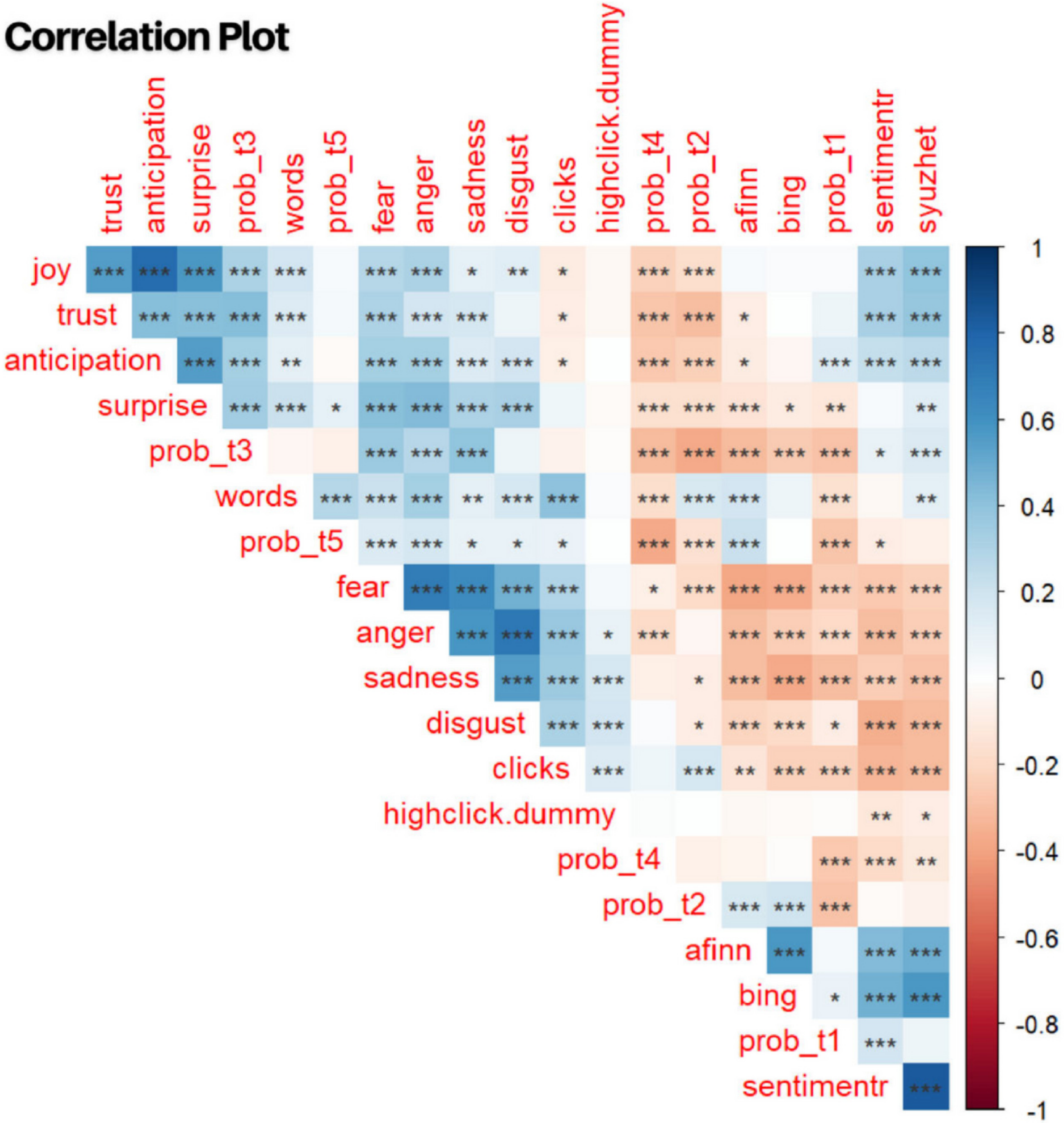




APPENDIX A-6.1

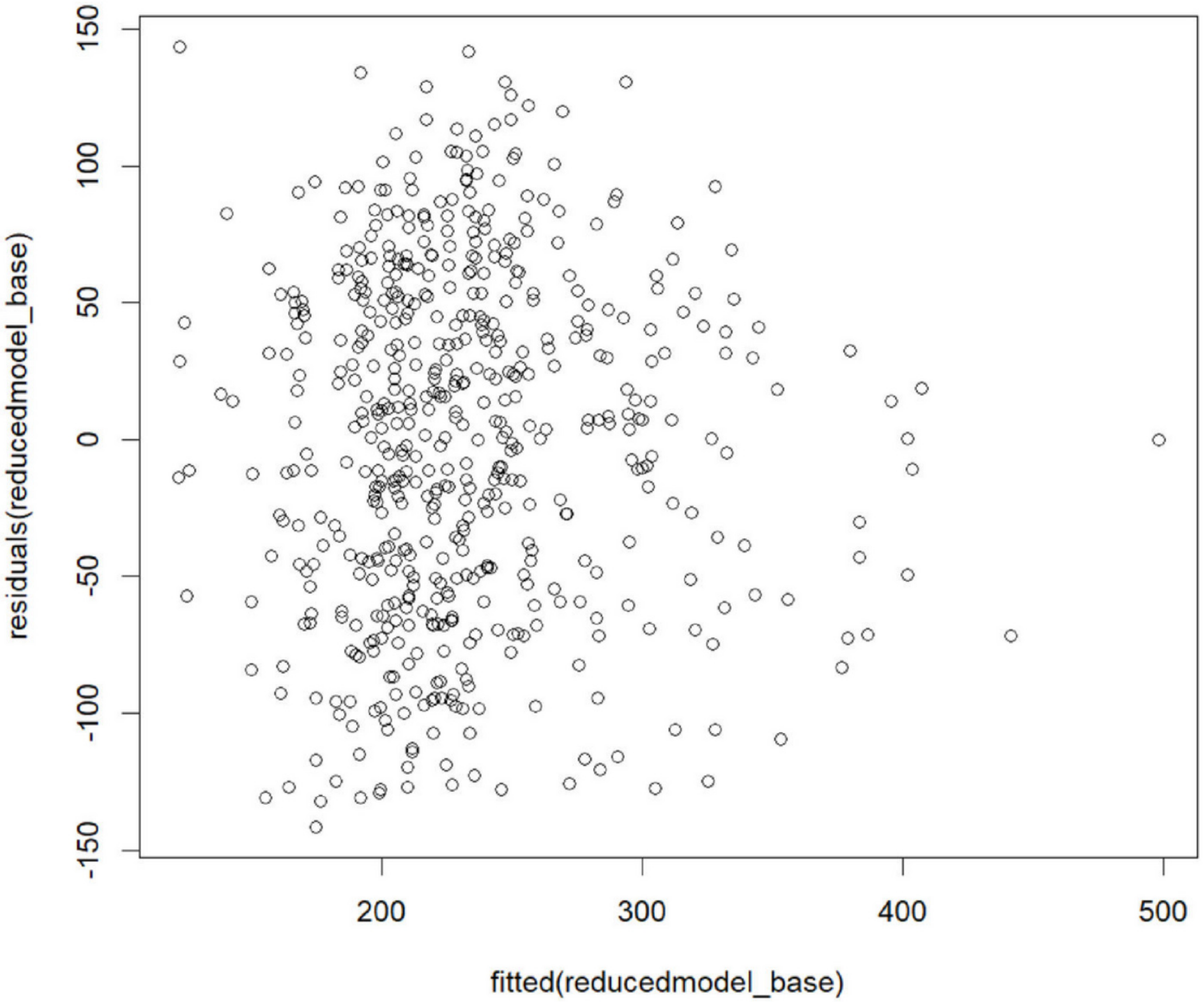
Building a final linear regression model and exploring non-linear effects.

Correlation Plot



Residuals Plot

Visual checks for homoscedasticity and linearity.

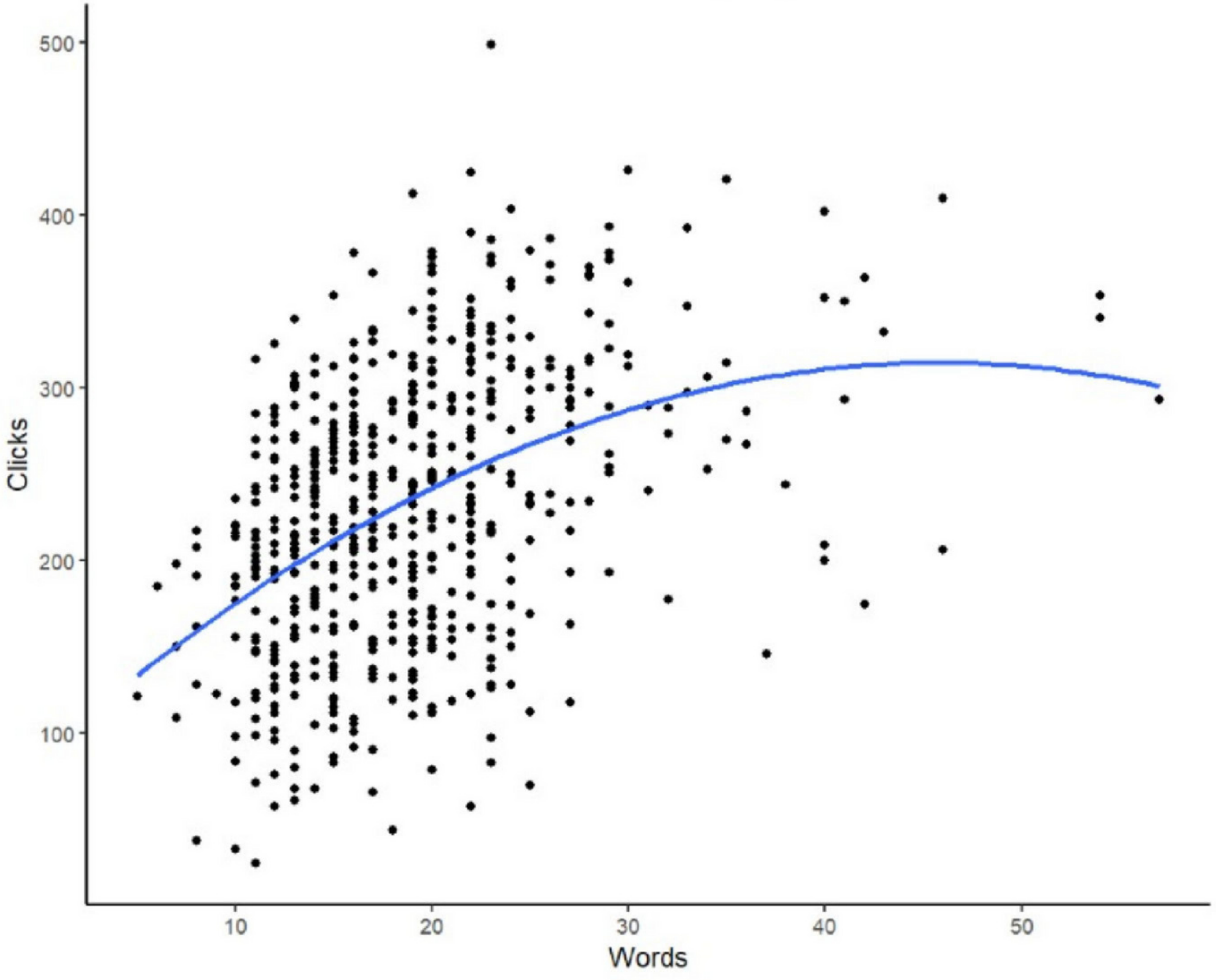


APPENDIX A-6.2

Building a final linear regression model and exploring non-linear effects.

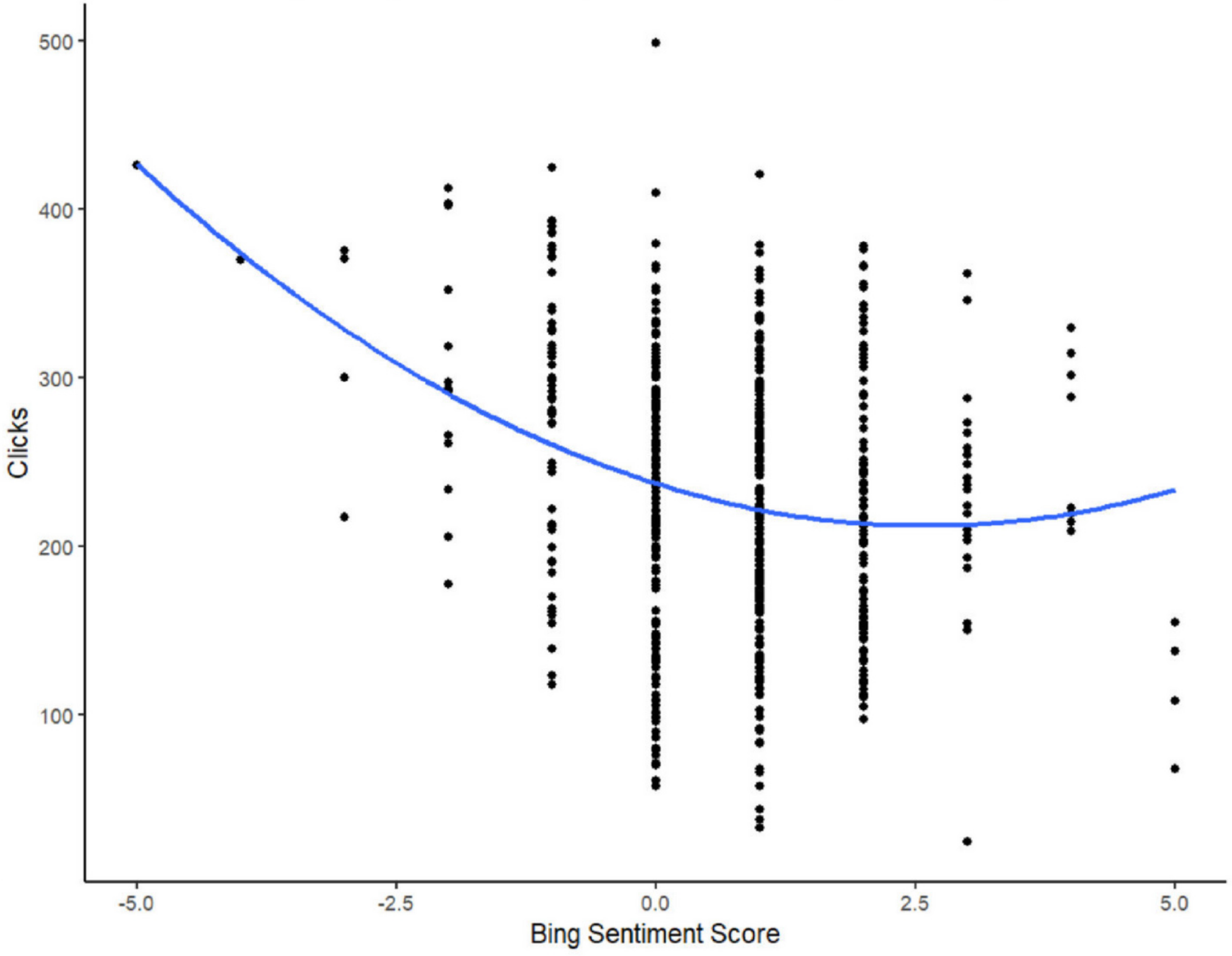
Non-linear Relationship: Clicks ~ Words

After 40 words the number of clicks start going down.



Non-linear Relationship: Clicks ~ Bing

As sentiment goes up clicks fall up to 3.0, then they start rising.

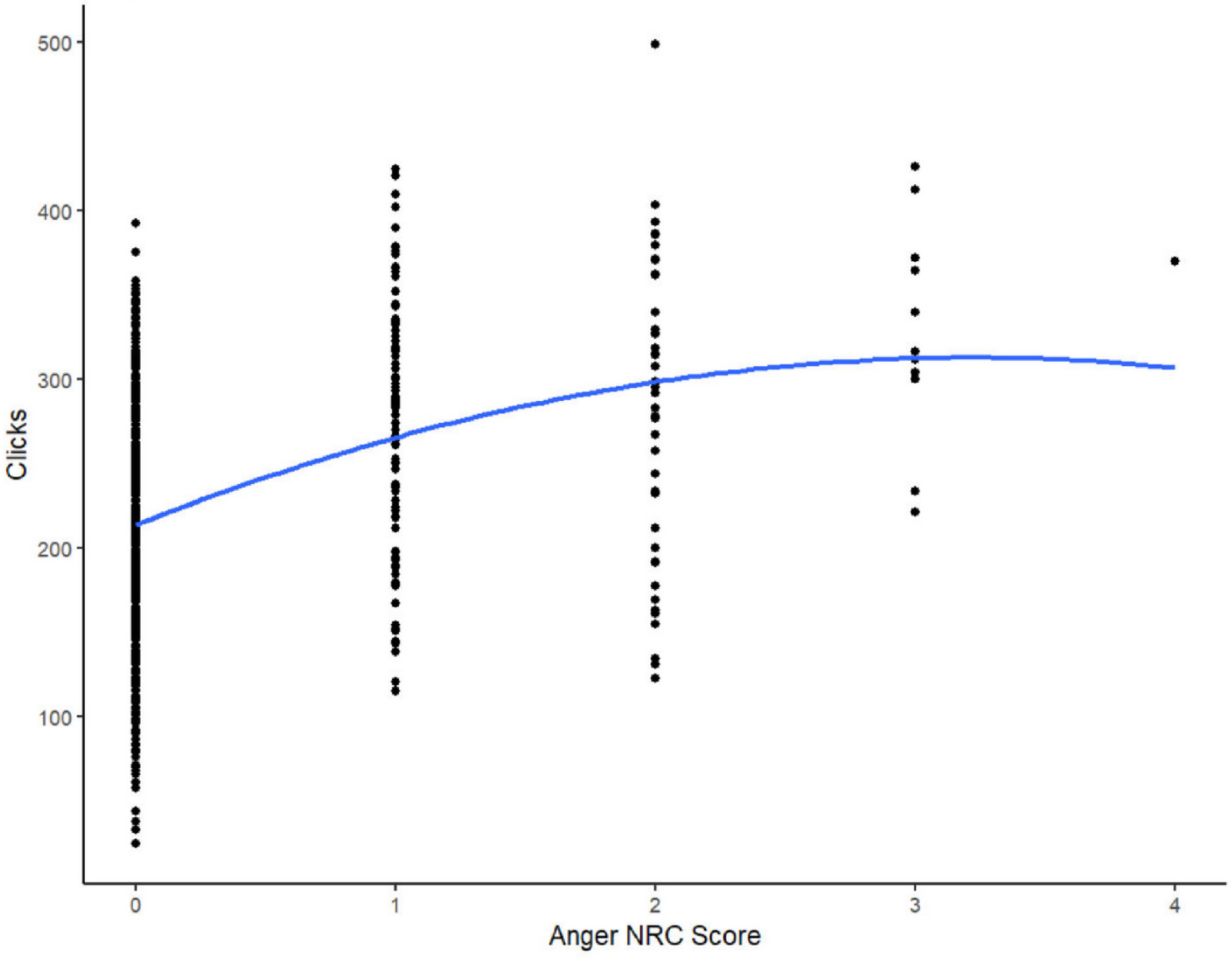


APPENDIX A-6.3

Building a final linear regression model and exploring non-linear effects.

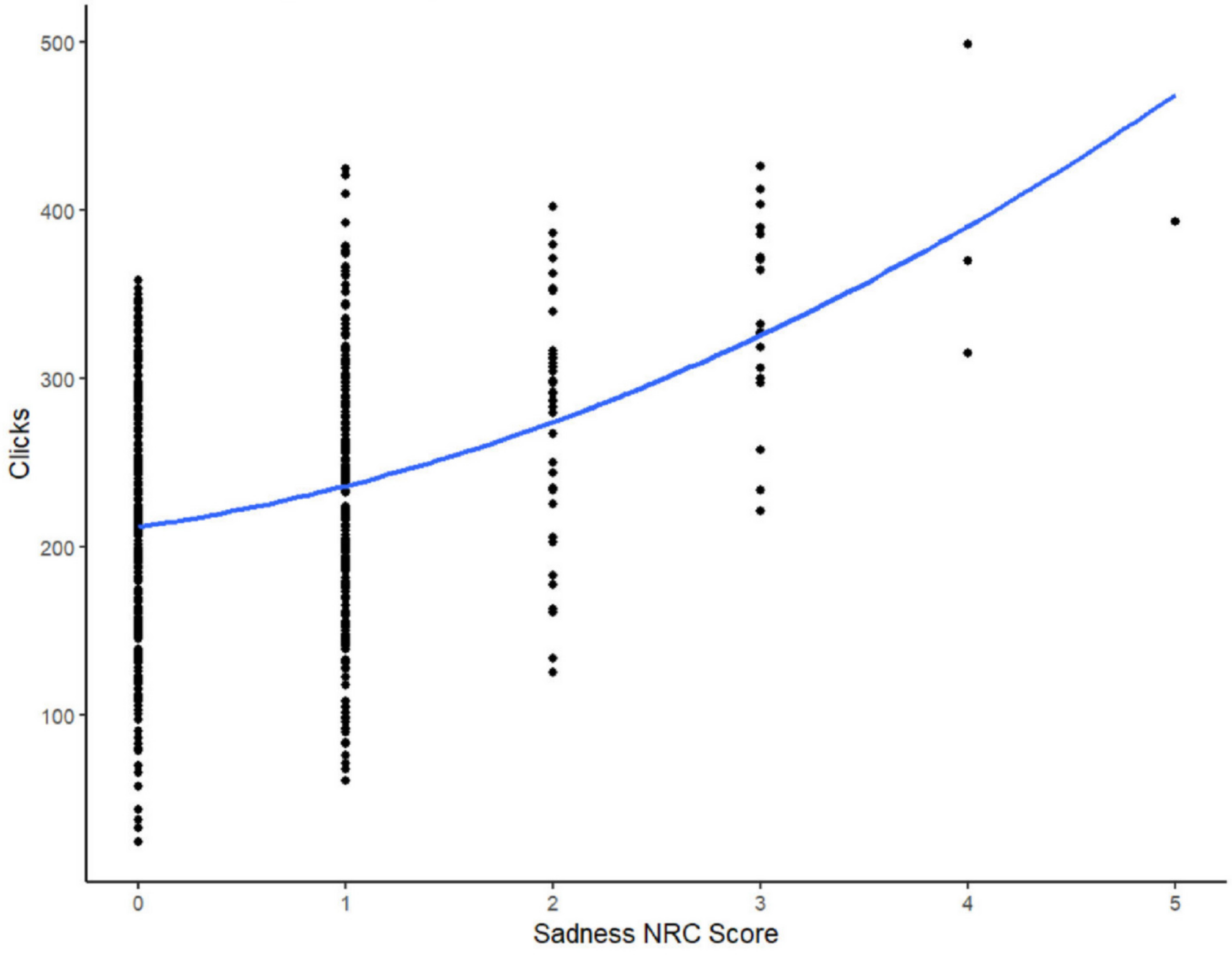
Non-linear Relationship: Clicks ~ Anger

Angry sentiment's clicks peak at 3 then fall.



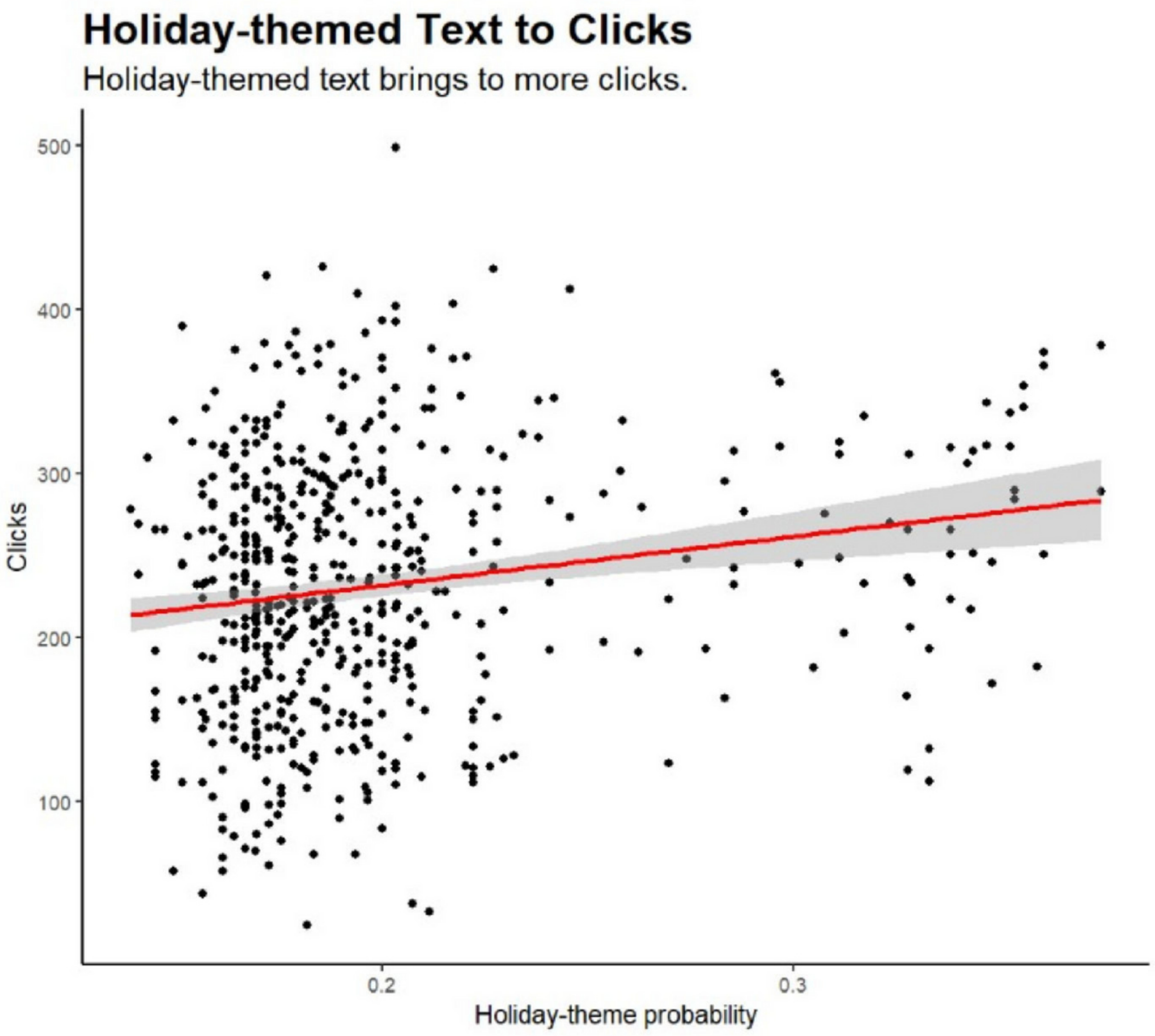
Non-linear Relationship: Clicks ~ Sadness

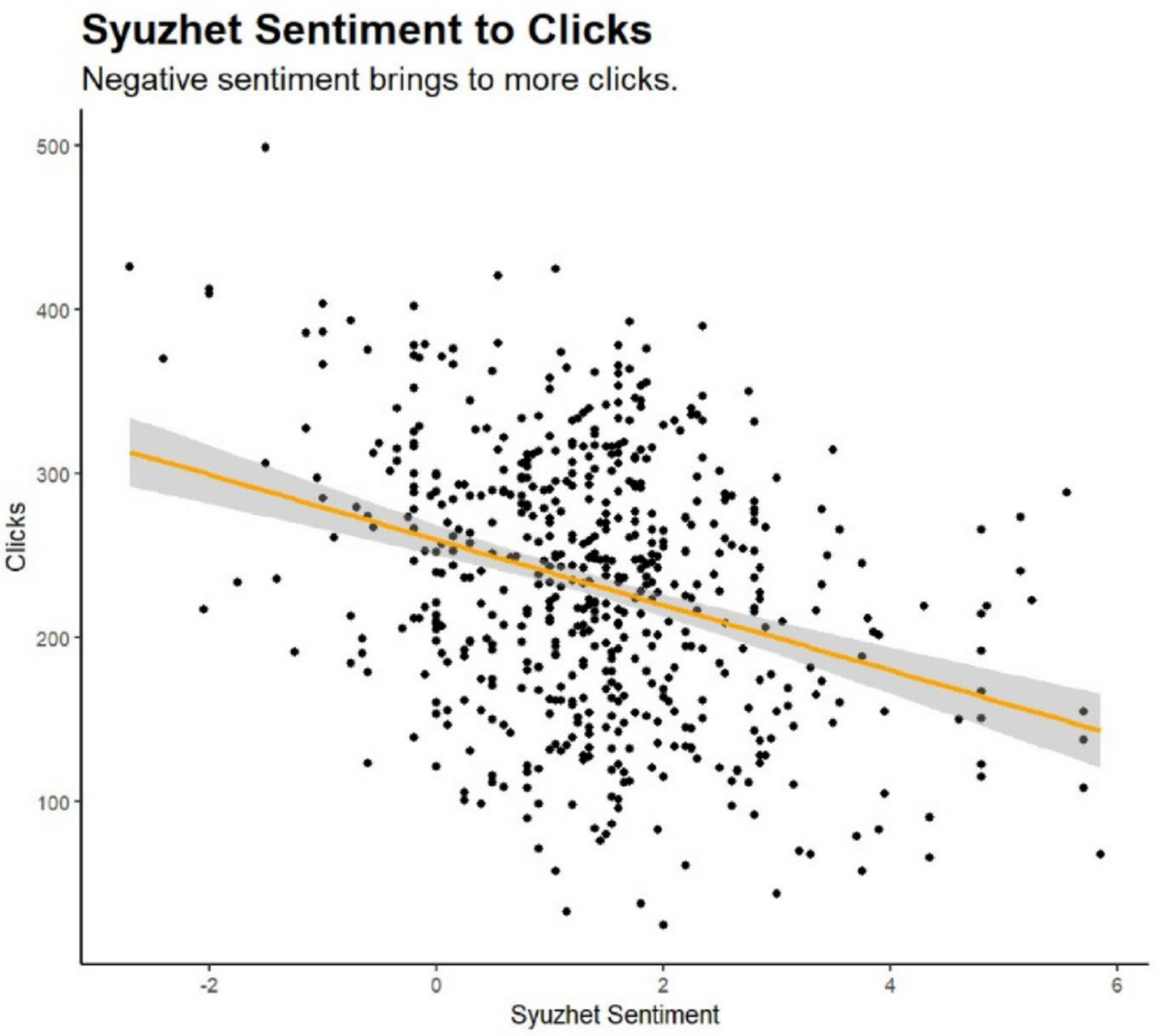
Clicks rise exponentially as sad sentiment rises.



APPENDIX A-6.4

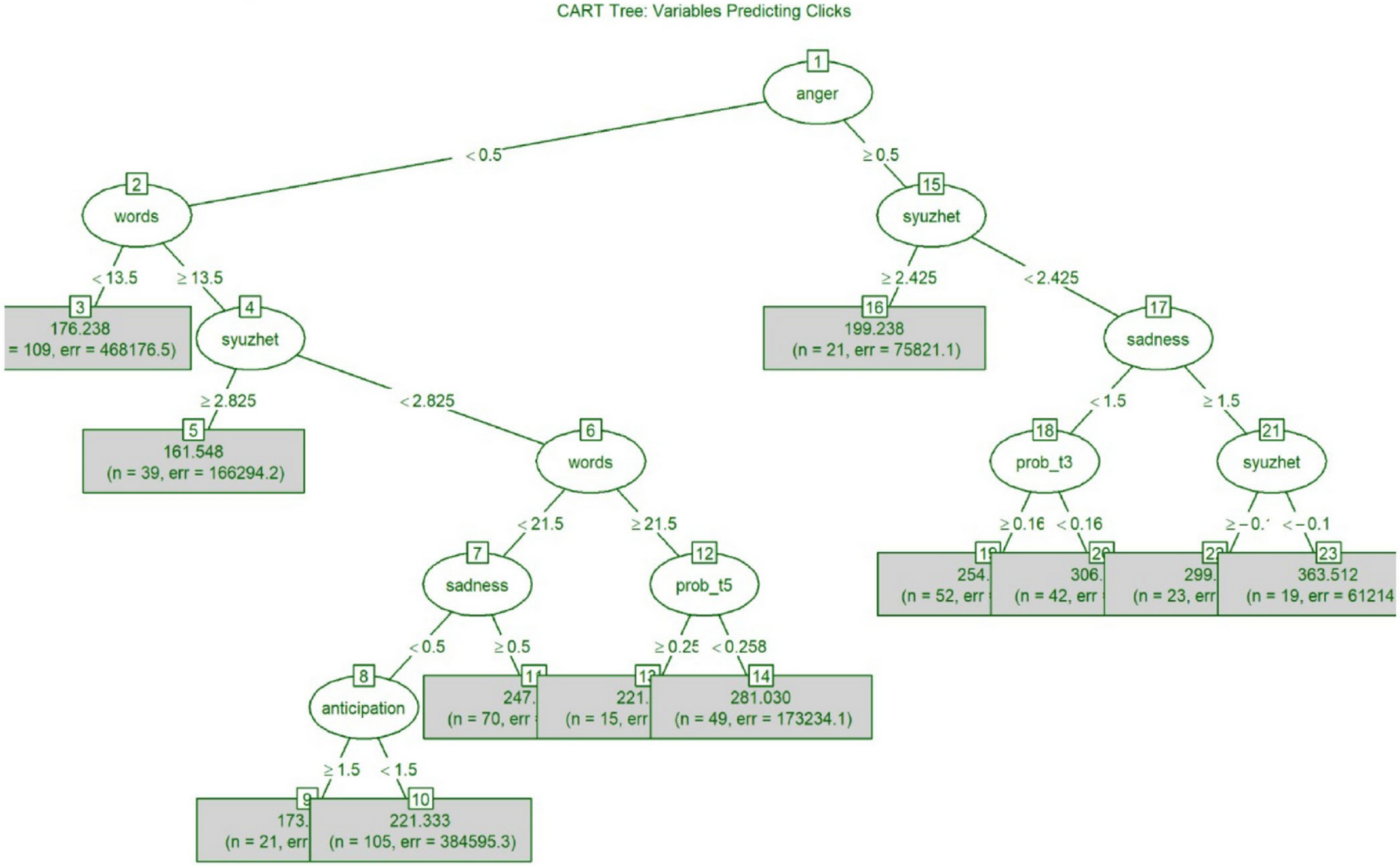
Building a final linear regression model and exploring non-linear effects.





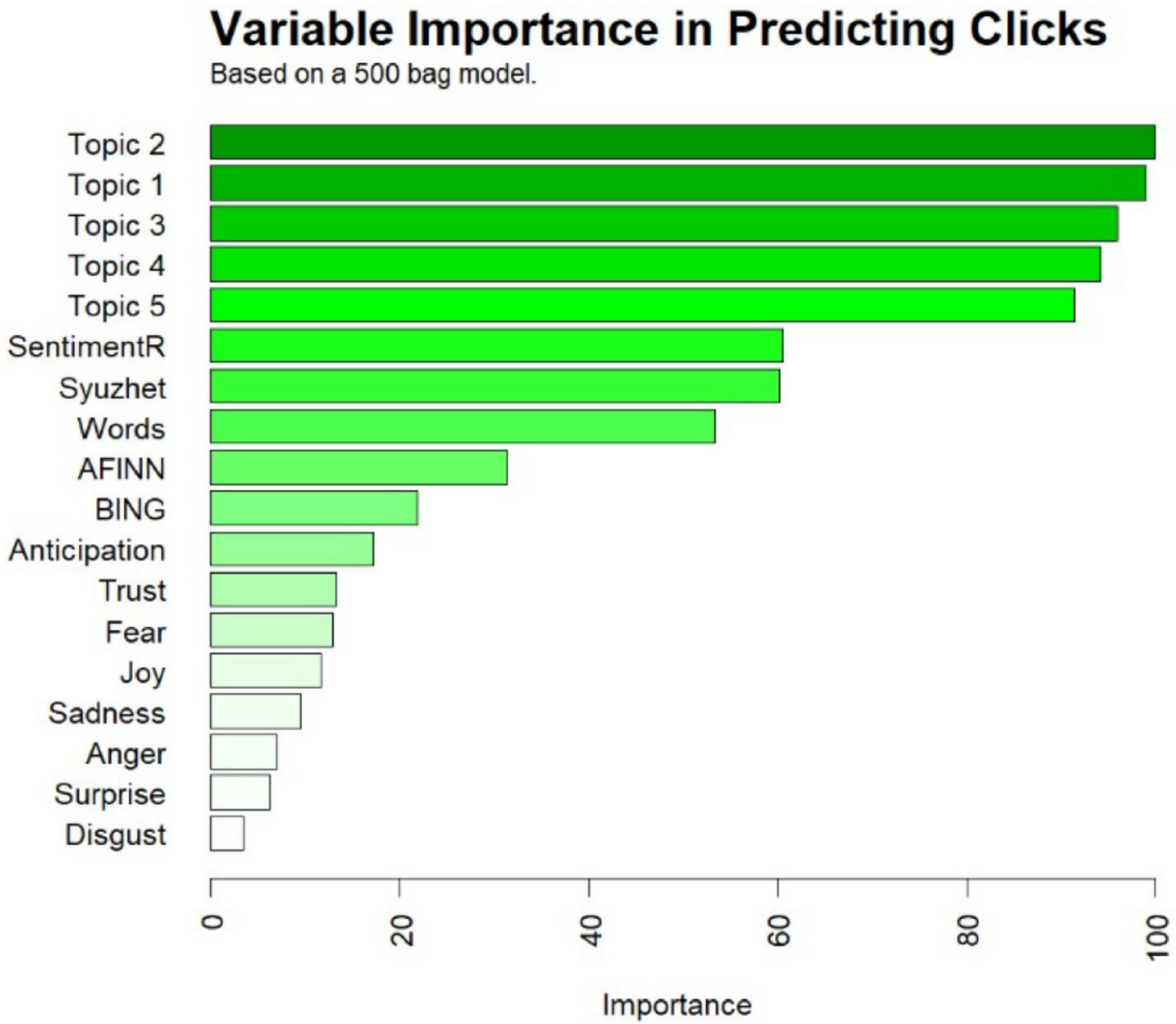
APPENDIX A-7.1

Employing machine learning models to further understand the relationship between clicks and our independent variables.



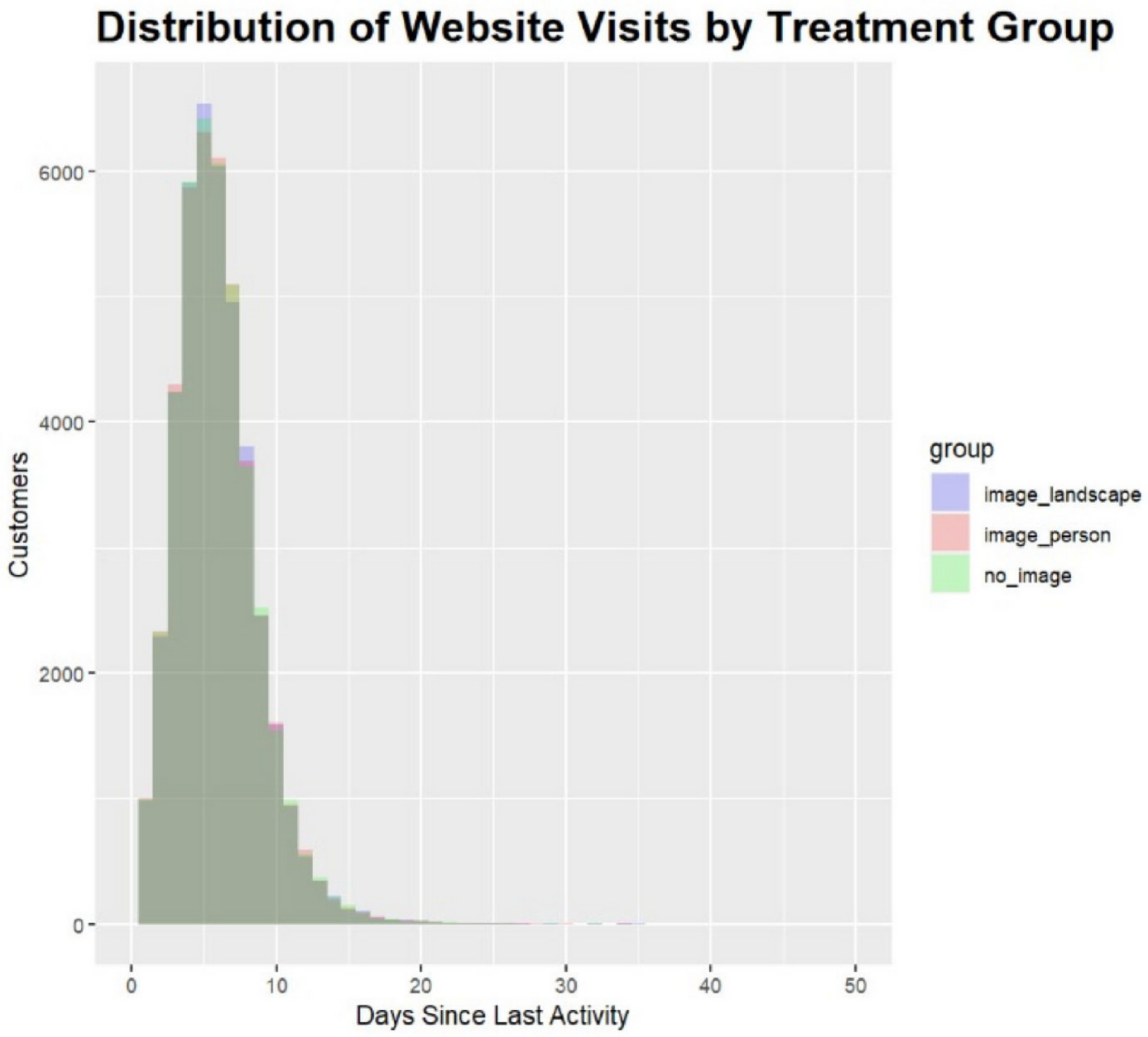
APPENDIX A-7.2

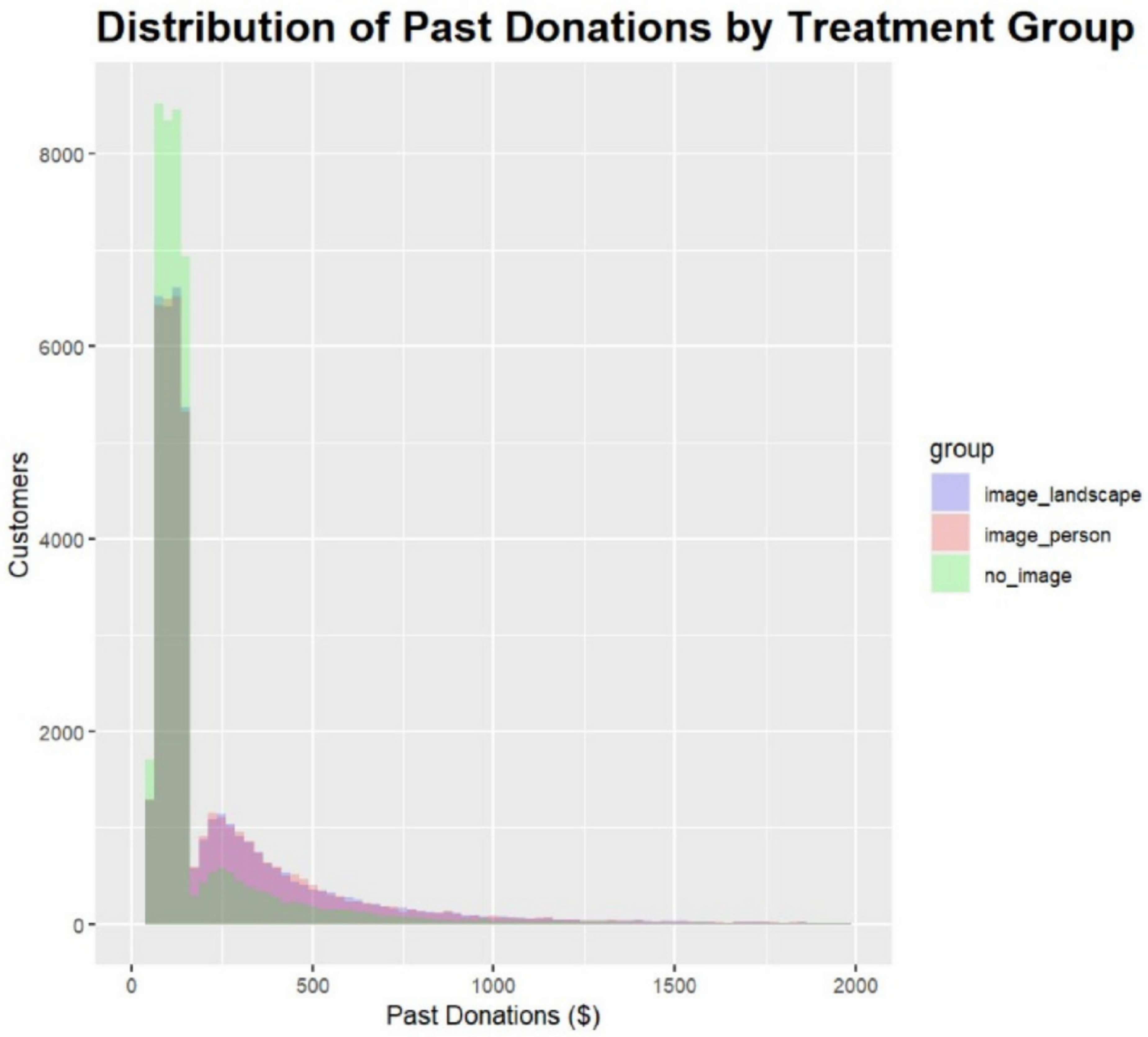
Employing machine learning models to further understand the relationship between clicks and our independent variables.



APPENDIX A-8.1

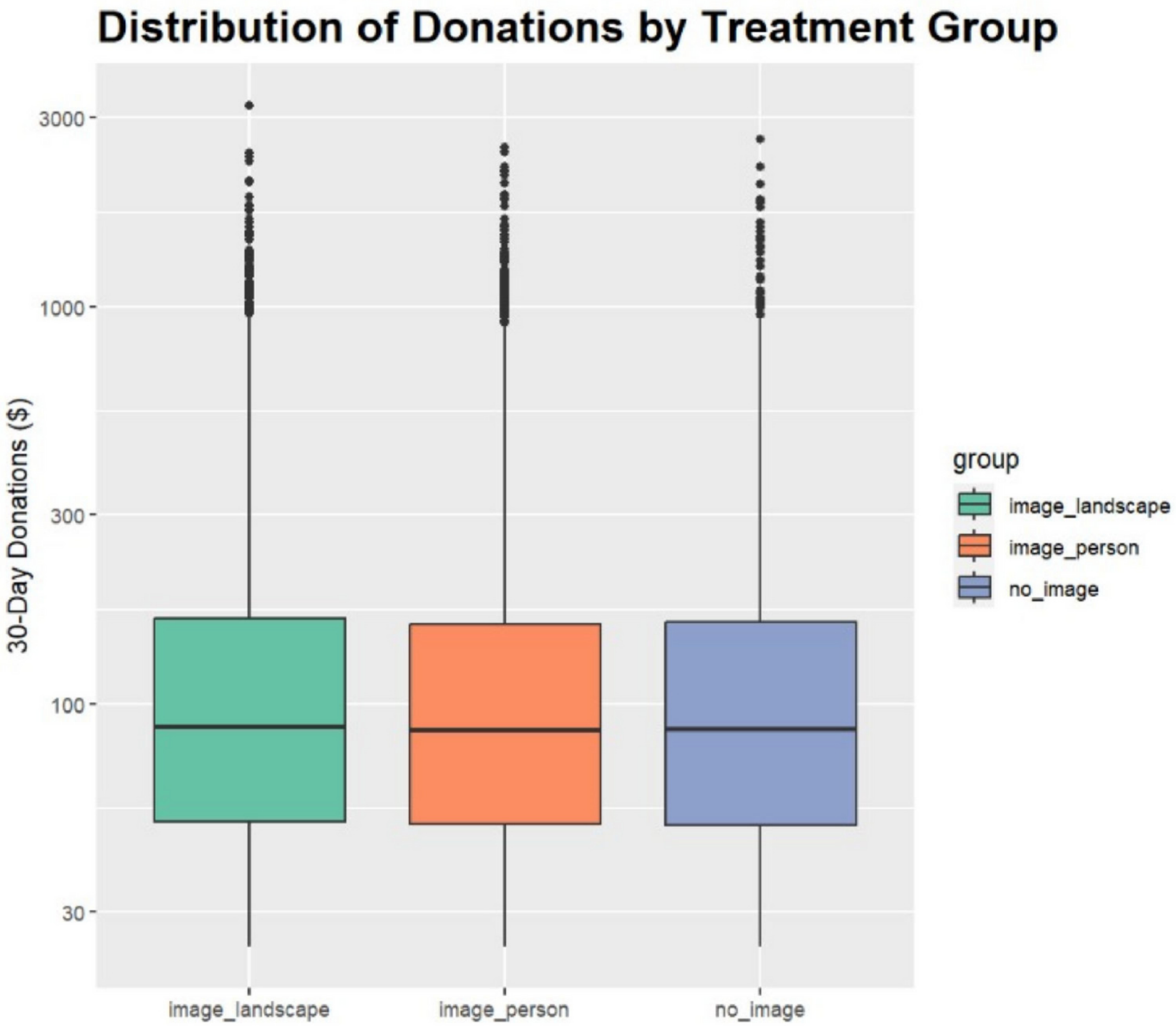
Exploring the distribution of website visits, past donations and donations by treatment group.





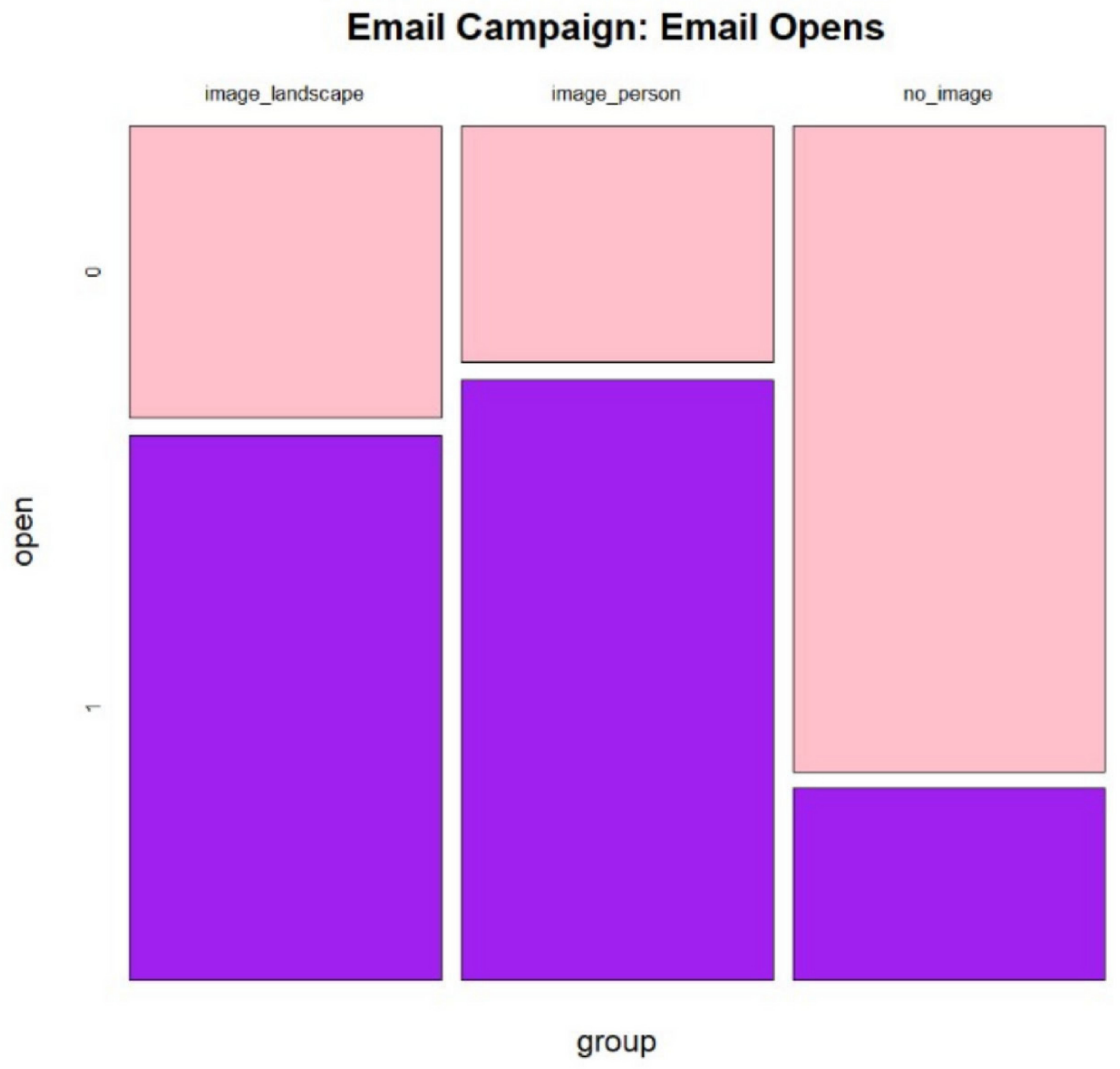
APPENDIX A-8.2

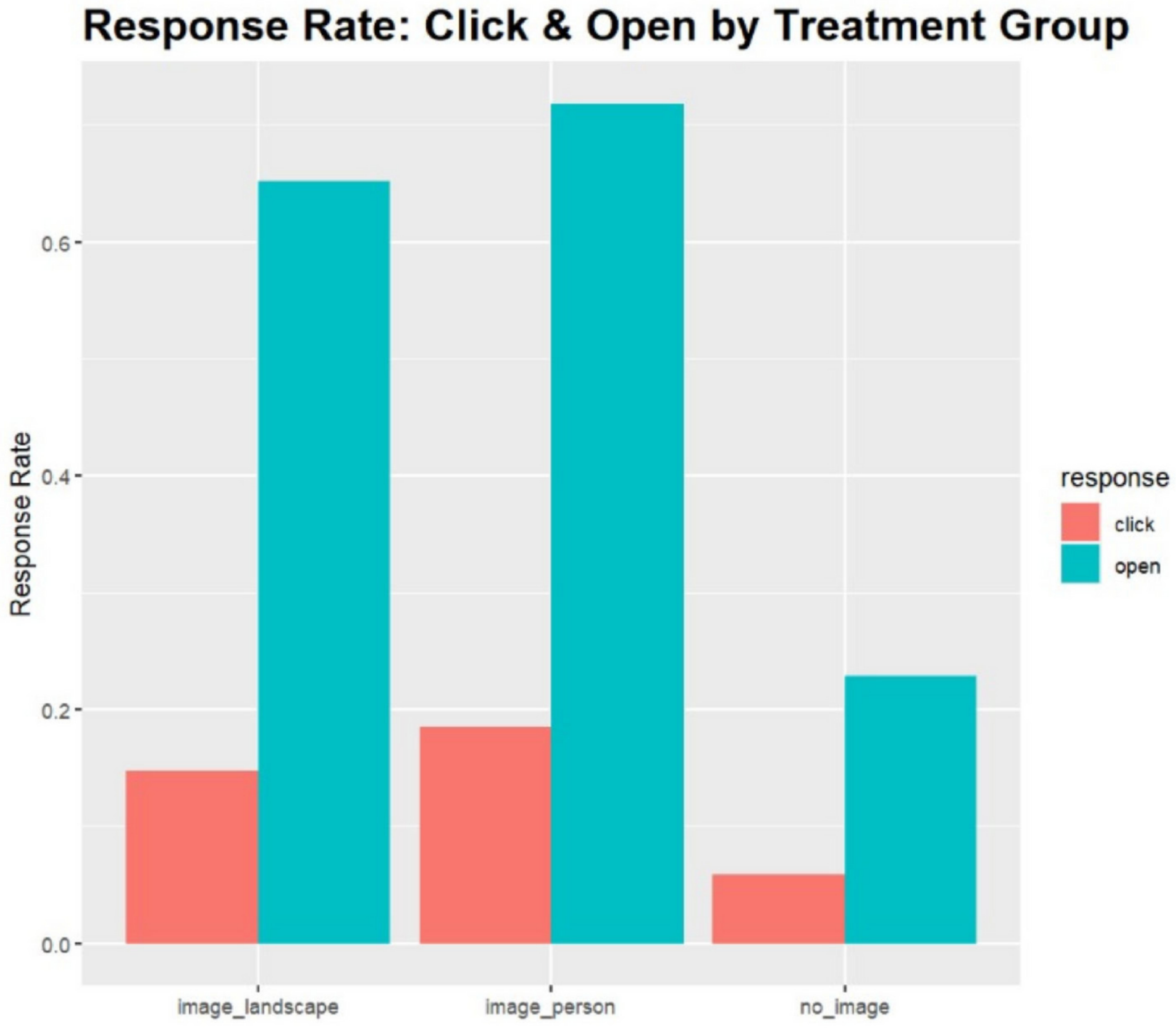
Exploring the distribution of website visits, past donations and donations by treatment group.



APPENDIX A-9

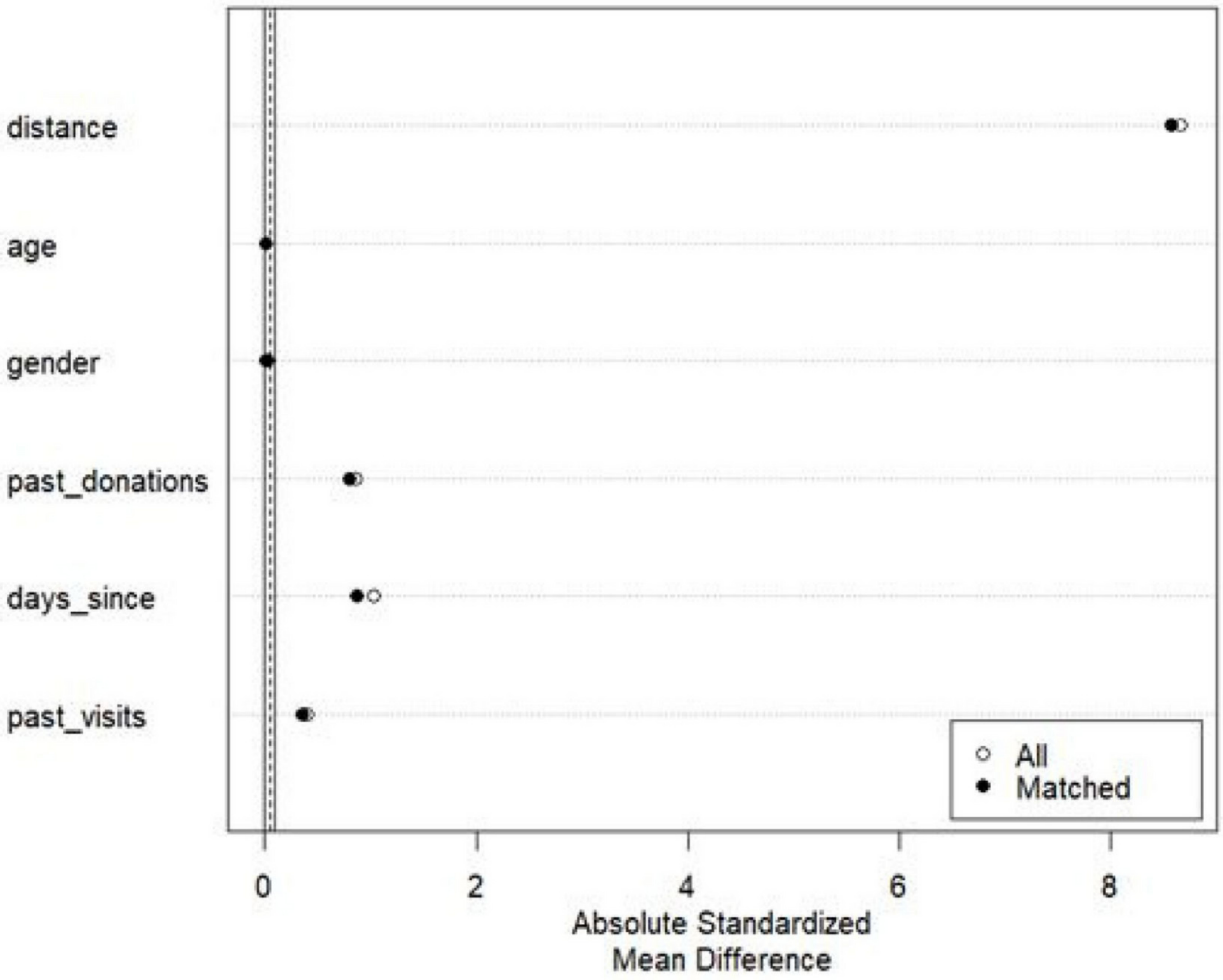
Exploring email opens and response rate (clicks and opens) per treatment group.





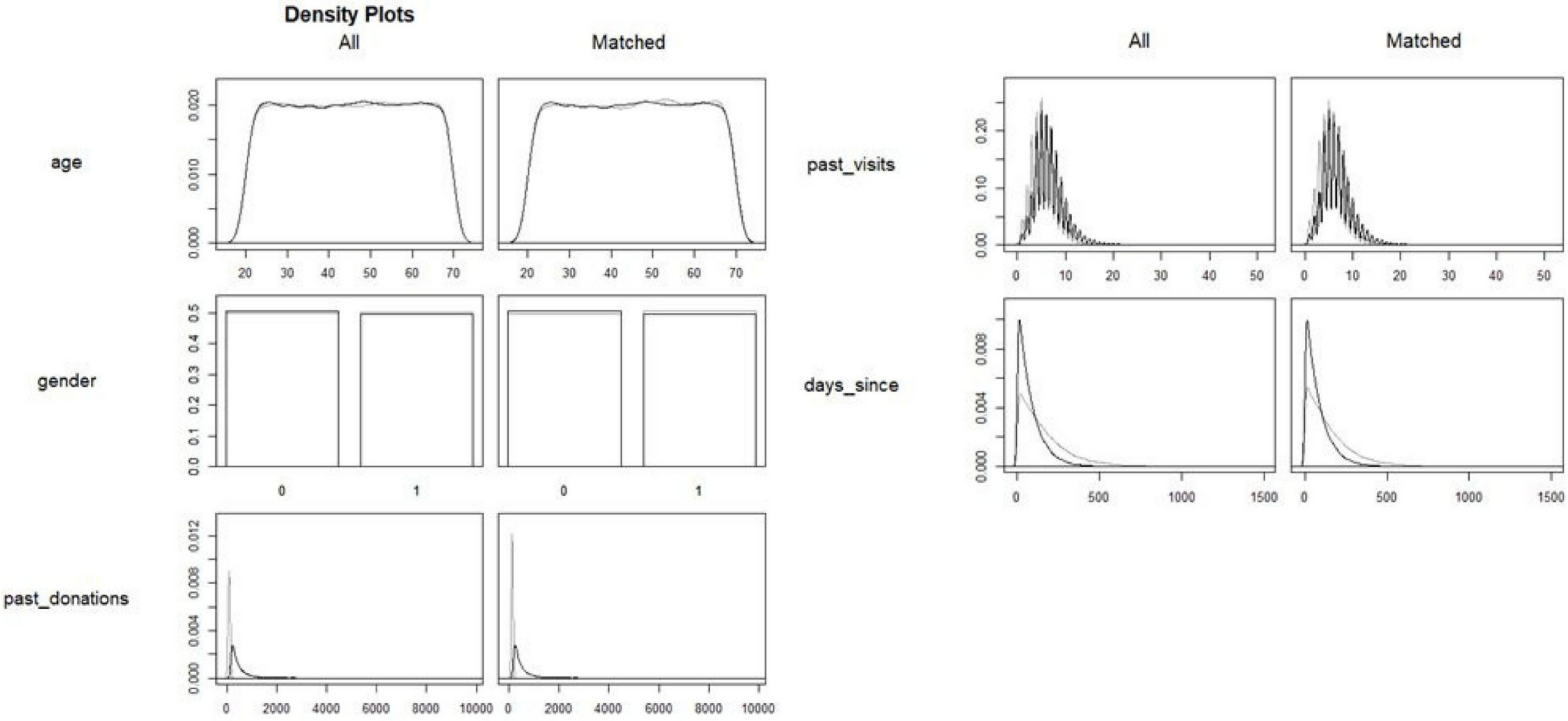
APPENDIX A-10.1

Performing propensity score matching.



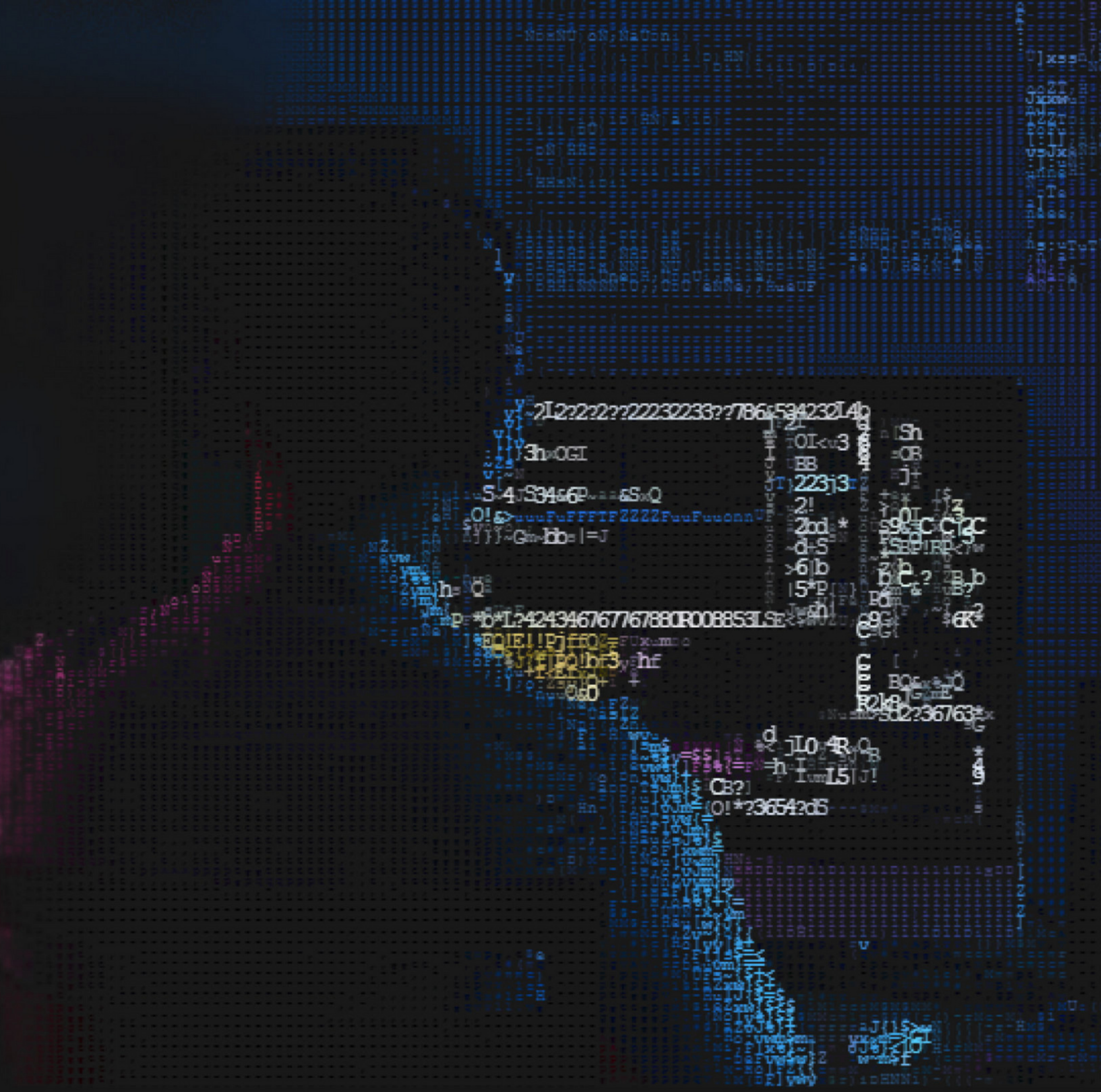
APPENDIX A-10.2

Performing propensity score matching.



APPENDIX B

The **tables** behind the insights.



APPENDIX B-1

Building a final linear regression model, accounting for non-linear effects.

Table 1: Final Linear Model Results

	Dependent variable:
	clicks
Topic 2 (prob_t2)	163.686**
	(62.483)
Topic 3 (prob_t3)	−125.486*
	(63.527)
Number of Words (words)†	4.119***
	(.428)
Syuzhet sentiment (syuzhet)	−7.751*
	(3.373)
Bing sentiment (bing)	−6.639*
	(2.990)
NRC: Anger (anger)	14.339*
	(5.972)
NRC: Anticipation (anticipation)	−3.370
	(3.973)
NRC: Fear (fear)	2.531
	(3.996)
NRC: Joy (joy)	−8.758
	(4.494)
NRC: Sadness (sadness)	18.832***
	(4.788)
Outlier: High clicks (highclick.dummy)	125.758
	(65.217)
Intercept	150.673***
	(20.298)
Observations	565
R^2	.385
Adjusted R^2	.373
Residual Std. Error	63.554 (df = 553)
F Statistic	31.481*** (df = 11; 553)

Note: *p<0.05; **p<0.01; ***p<0.001

† Squared; due to non-linear effects, it is not possible to interpret linearly.

APPENDIX B-2

Investigating interraction between headers and images.

Table 2: Investigating interraction between headers and images.

	Dependent variable:
	open
header	$-.002$
	(.004)
image	$.456^{***}$
	(.004)
header:image	.001
	(.005)
Constant	$.230^{***}$
	(.003)
Observations	123,988
R^2	.186
Adjusted R^2	.186
Residual Std. Error	.450 (df = 123984)
F Statistic	$9,424.045^{***}$ (df = 3; 123984)
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Main Finding:	There is no significant interaction effect.

APPENDIX B-3

Investigating effect variance based on donor characteristics.

Table 3: Investigating effect variance based on Donor characteristic

	Dependent variable:
	donate
age	−.025
	(.030)
gender	−3.537
	(2.041)
age:gender	.054
	(.043)
Constant	44.274***
	(1.439)
Observations	123,988
R^2	.00004
Adjusted R^2	.00001
Residual Std. Error	109.652 (df = 123984)
F Statistic	1.541 (df = 3; 123984)
Note:	*p<0.05; **p<0.01; ***p<0.001
Main Finding:	There is no significant variance in the effect based on donor characteristics.

APPENDIX B-5

Fitting models before and after performing propensity matching.

Table 4: Binomial Logit Model (Pre-matching)

	Dependent variable:
	donate_people
age	−0.00000
	(.0001)
gender	−.001
	(.002)
past_donations	.001***
	(0.00000)
days_since	−.001***
	(.00001)
past_visits	.023***
	(.0004)
Constant	.085***
	(.004)
Observations	123,988
R^2	.388
Adjusted R^2	.388
Residual Std. Error	.362 (df = 123982)
F Statistic	15,751.680*** (df = 5; 123982)
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 5: Linear Model for Header (Post-matching)

	Dependent variable:
	donate
header	.137
	(.292)
pscore1	389.657***
	(.449)
Constant	−86.769***
	(.272)
Observations	76,826
R^2	.907
Adjusted R^2	.907
Residual Std. Error	40.404 (df = 76823)
F Statistic	375,863.900*** (df = 2; 76823)
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 6: Linear Model for Image (Post-matching)

	Dependent variable:
	donate
image	5.095***
	(.320)
pscore1	388.909***
	(.451)
Constant	−89.981***
	(.309)
Observations	76,826
R^2	.908
Adjusted R^2	.908
Residual Std. Error	40.338 (df = 76823)
F Statistic	377,228.800*** (df = 2; 76823)
Note:	*p<0.05; **p<0.01; ***p<0.001

APPENDIX C

The **R script** behind the insights.