Instagram 2-Way Factorial Analysis

Prepared by Viktor Shaumann and Paul Thompson

Contents

Executive Summary

The purpose of our analysis is to investigate the results of a multivariate experiment conducted by Instagram. The experiment is designed to measure the relationship between users' engagement time and both type and prevalence of the ads. The two way factorial analysis allows us to quantify the interaction effects of multiple types of variables (type and prevalence) on the response (user engagement time). Therefore, we can estimate which specific types of ads, ad display frequencies or combinations of different types and frequencies reduce or increase the time user spends with the app.

We would typically expect a higher proportion of ads to negatively impact the user experience. Analysis of experiment results indicates that Instagram could increase the user engagement with their app by including ads that contain video and decreasing prevalence of the actual ads.

Overview

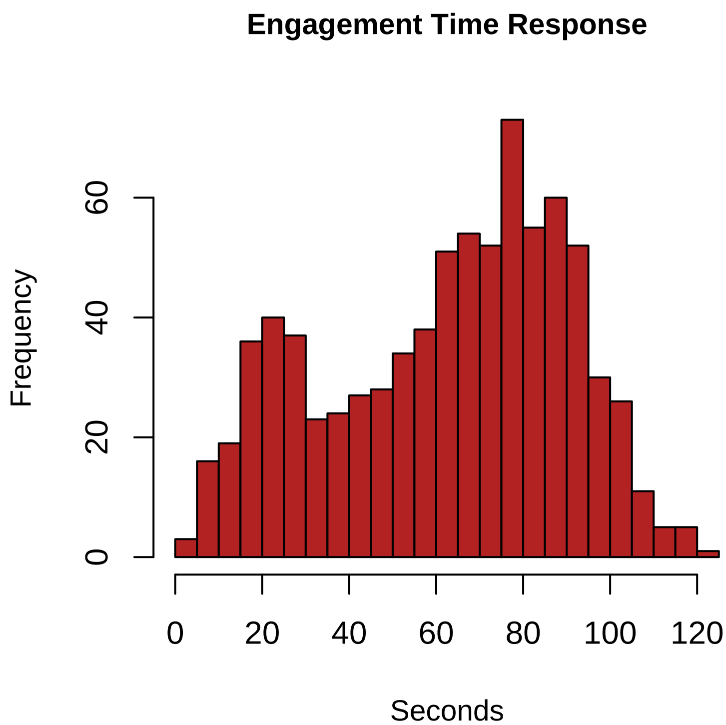
Mock experiment data was generated to simulate the relationship between user engagement time and both type (picture, video) and prevalence (quantified as a proportion of the posts user sees such as 0, 0.1, 0.2, or 0.5) of the ads.

800 observations were collected, where each possible combination of ad type and ad prevalence was replicated 100 times (runs). For example, one combination would be a video ad that was displayed with 0.5 prevalence (half of the items user saw were ads).

Exploratory Data Analysis

Before diving into the analysis, we look at the distributions of each of the variables to familiarize ourselves with the data and get a preliminary idea of differences between levels.

A histogram of the response approximately resembles a bimodal distribution. This could due to the effect that the two factors have on different groups of users.



Analysis

The first step in our analysis is to determine whether individual factors or their interactions impact time. We can access these relationships both visually and quantitatively.

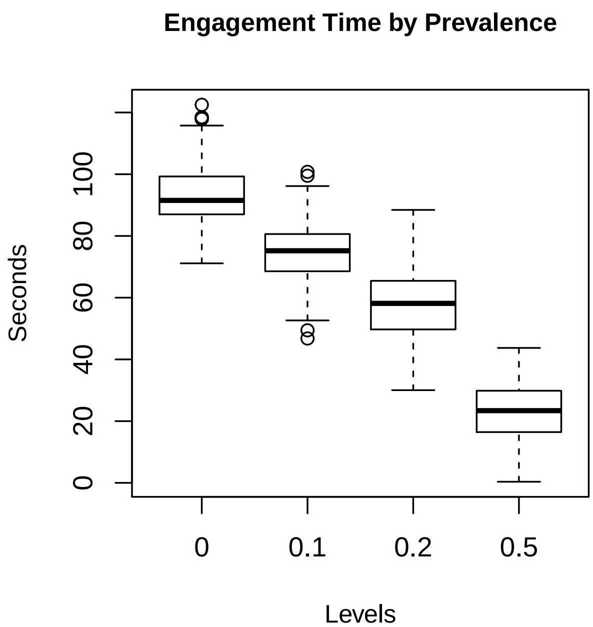
Effect Plots

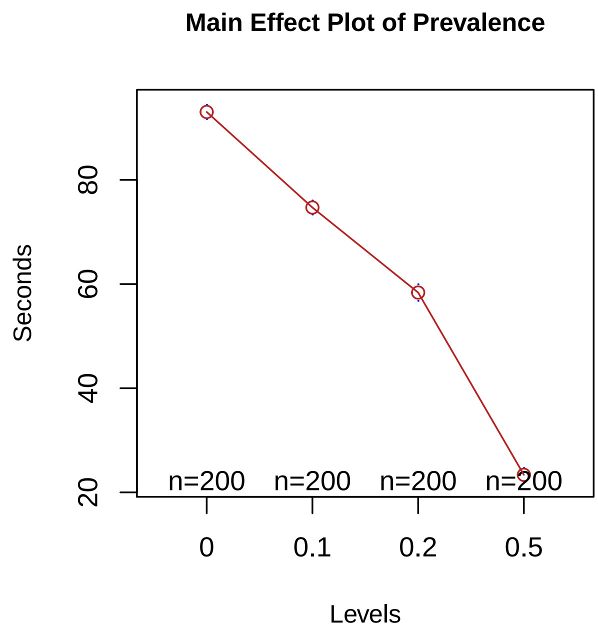
We can visualize effects of individual factors on response variable (Box plots and Main Effect Plots) as well as the combined effects of every possible combination of factors' levels (Interaction Plots).

Box and Main Effect plots allow us to visually inspect how statistical measures and mean effects vary per level of each factor.

Prevalence

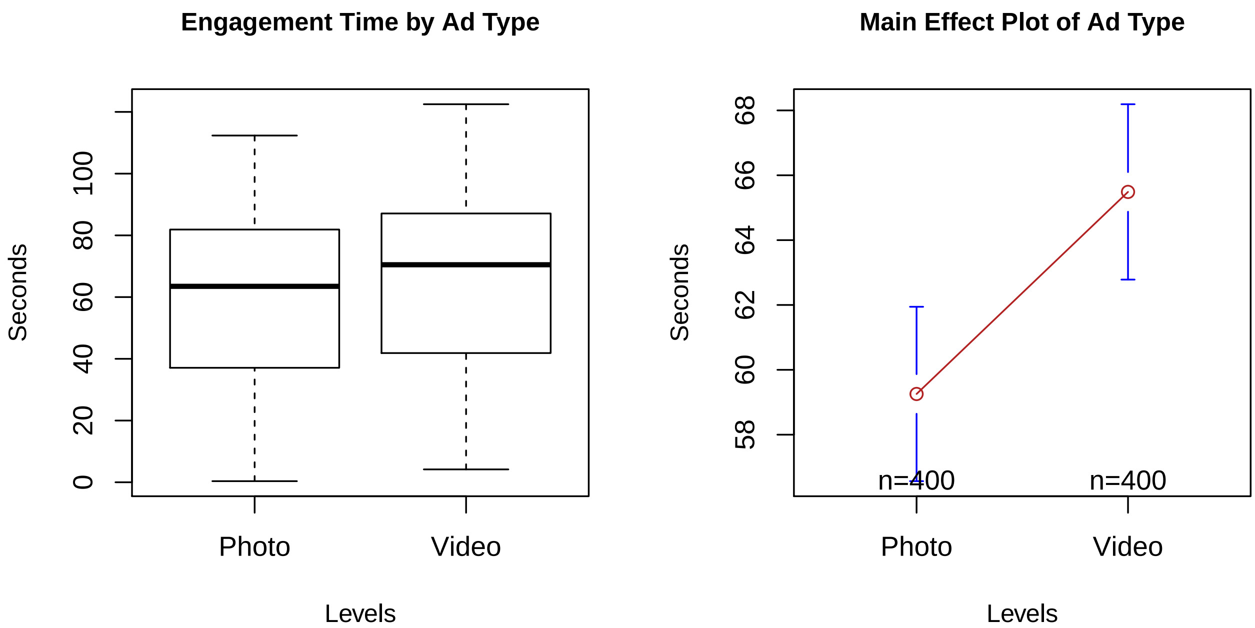
As expected, both a box plot and main effect plot indicate that mean and median engagement time falls dramatically as ad prevalence in the user's feed is increased from 0% of items shown to 50% of items shown. The box plot also shows that the variance differs little between levels.





Type

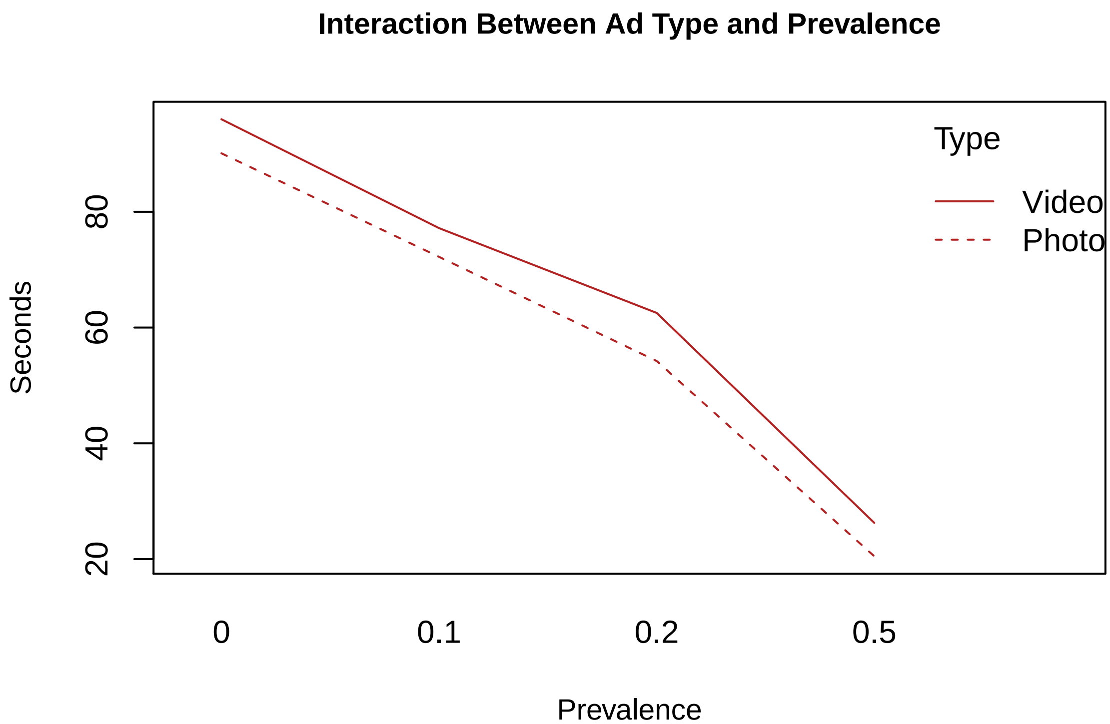
We look at the box plot of engagement type by ad type and main effect plot by ad type. There is a smaller difference in median and mean responses between the two ad types than there were between different prevalence levels. We can see from the Main Effect Plot that the difference in mean response between Photo and Video ads is approximately 5 seconds. There doesn't seem to be a difference in variance.



Pairwise Interaction Plot (Type and Prevalence)

A Pairwise Interaction Plot allows us to inspect whether an interaction effect between levels of factors is present. For example, we can see whether mean response varies by a specific combination of Type and Prevalence or is constant per each variable. Each point on the plot represents a mean response per combination of two levels, one from each factor. When interaction effects between combinations of different levels are present, we expect to see lines that have different slope and direction. When no interaction effect is present, we can expect to see lines that are approximately parallel to each other (same slope and direction).

The Interaction Plot below shows that lines follow approximately the same slope at all points. Therefore, we can conclude that there are no interaction effects present and we should focus on examining effects of the individual factors.



Quantitative Analysis

Multiple Regression Model

In order to quantify the relationships that were evident from the plots above, we can also perform a formal analysis to quantify an exact effect of different factors and their levels on the response variable.

First step in this analysis is to fit a Multiple Regression Model with each level and all possible interactions of the levels. We performed a residual analysis and verified that assumption of normality and identical distribution are not violated. Please refer to the Appendix for more information. Model was fit using R as follows:

```
model = lm(Time ~ Type + Prevalence + Prevalence*Type)
```

The results of the model output are shown in the Table 1 below.

Table 1: Time   Type + Prevalence + Prevalence*Type

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	90.1075	0.9677	93.11	0.0000
TypeVideo	5.8823	1.3686	4.30	0.0000
Prevalence0.1	-17.8942	1.3686	-13.07	0.0000
Prevalence0.2	-35.8900	1.3686	-26.22	0.0000
Prevalence0.5	-69.6244	1.3686	-50.87	0.0000
TypeVideo:Prevalence0.1	-0.9169	1.9355	-0.47	0.6358
TypeVideo:Prevalence0.2	2.4263	1.9355	1.25	0.2104
TypeVideo:Prevalence0.5	-0.1125	1.9355	-0.06	0.9537

ANOVA

After we fit the model, we proceed with ANOVA to identify the variables that are responsible for the most variance in the response. These variables are Type, Prevalence and an interaction variable of Type and Prevalence. Table 2 below shows the output of the ANOVA model.

Table 2: ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Type	1	7766.29	7766.29	82.93	0.0000
Prevalence	3	526050.50	175350.17	1872.34	0.0000
Type:Prevalence	3	312.63	104.21	1.11	0.3430
Residuals	792	74172.94	93.65		

The F-test statistics in the ANOVA table for Type and Prevalence are statistically significant while the F-test statistic for the interaction between Type and Prevalence is not. This means we reject the hypothesis that the effect of different levels of Type and Prevalence on the response is the same. However, we fail to conclude that different levels of Type influences how different levels of Prevalence affect the response and vice versa. Further, this confirms the visual relationship we saw from the interaction plot in previous section. These results indicate that response is impacted by individual levels of each variable rather than their interactions.

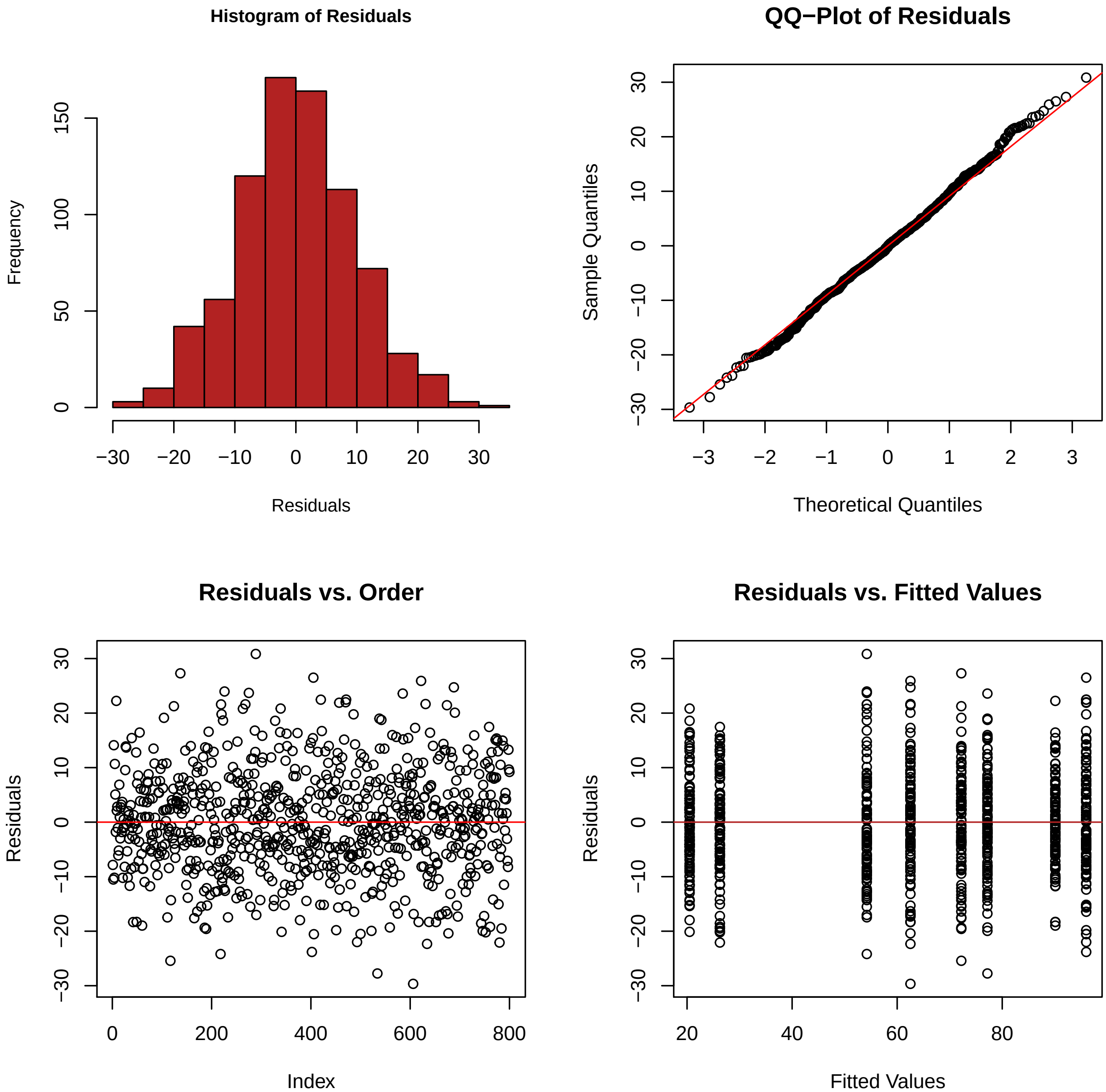
Since we reject the null hypothesis that altering the levels of prevalence and type doesn't change the response, we can do t-tests to determine if the individual levels of the factors effect on the response. We will use the lm model summary in Table 1 above for the t-statistics and their corresponding p-values. Unlike with pairwise t-tests between different levels, we are testing the null hypothesis that the estimated coefficient is equal to zero ($H_0 : \beta_j = 0$). Therefore, we are testing if the estimated effect of the factor being at that level is significant. Since the p-values for TypeVideo, Prevalence0.1, Prevalence0.2, and Prevalence0.5 are significant, we reject the null hypothesis that these levels don't effect the response differently than if they weren't at these levels.

In order to quantify specific effects of each level on the response we examine results of our Multiple Regression Model in Table 1. Model output indicates that Video and Prevalence 0.1 through 0.5 are statistically significant variables based on their respective p-values. Based on the beta coefficients we quantify effects of each of the levels on the response. For instance, we can see that video ad increases user engagement time by 5.8 seconds compared to the image ad on average, holding all other variables constant. We can also see the individual effects of each of the Prevalence levels on the response. For example, Prevalence of 0.5 decreases the user engagement by 69.6 seconds on average, holding all other variables constant.

Conclusion

Analysis of experiment results indicates that Instagram would increase their user engagement with their app by including ads that contain video and decrease prevalence of the actual ads. Based on this information, Instagram can make a better informed decision on the trade-off of the the user experience and monetization of their app with advertisements.

Appendix



Histogram of Residuals
Frequency
0
50
100
150
−30
−20
−10
0
10
20
30
Residuals
QQ−Plot of Residuals
Sample Quantiles
−30
−20
−10
0
10
20
30
−3
−2
−1
0
1
2
3
Theoretical Quantiles
Residuals vs. Order
Residuals
−30
−20
−10
0
10
20
30
0
200
400
600
800
Index
Residuals vs. Fitted Values
Residuals
−30
−20
−10
0
10
20
30
20
40
60
80
Fitted Values