

Assessing Color, Subjects, and Popularity on Instagram

Research conducted by:

Brandon Levesque
633 Main Street
Burlington, VT 05405
Contact: balevesq@uvm.edu

December 9, 2020

ABSTRACT

Achieving popularity on Instagram, a popular social media and photo-sharing platform, is quite a difficult task. Users may be wondering what the best method is for gaining a large number of, “likes,” and after creating various models that would shed a light on what the secret might be, we believe that we have come to some conclusions for what to do (and not do); based on our data. It has been found that photos of animals, on average, gain over 1.5 times more, “likes,” than both portraits and landscapes. Gray is the most often used colors around the outer edges of photos, with brown being most often in the center. The top-center section of an image most significantly affects the number of, “likes,” that post receives; when no model-intercept/constant is included. An account’s, “Follower,” count is the most significant predictor of success. Finally, the color brown, in the center-right section of an image, is the most significant predictor of, “likes,” compared to other colors.

Keywords

Colors, analysis, Instagram, social media, modelling

1. INTRODUCTION

Around 7 years ago, I started photography. Seeing the world through my camera lens, and being able to document all of my adventures, enticed me to continue with this hobby, and thus allowed me to expand my skills. Now, I hope to expand my photography-audience on platforms such as Instagram, but do not know what types of images would be successful. With this project, I intend to answer the question, “**What colors and subjects most significantly contribute towards an Instagram post being popular?**”

Aside from gaining, “likes,” on the internet, this research could be used in a number of fields, such as advertising/marketing, media creation, and education. Essentially, this data shows what colors and subjects attract a person’s attention, so that they choose to leave a, “like,” on the post, which would be quite valuable for any organizations that rely on enticing a human to do some task (buying a product, enjoying a movie/show, staying focused in school, etc).

The mind is a beautiful, yet misunderstood thing. It can easily be influenced in some cases, while completely ignorant of others. In this fairly simple analysis, we plan to generate useful information that will help artists, content-creators, and marketers more appropriately tailor their content towards audiences. Thus, allowing them to become more successful in their respective field.

2. RELATED WORK

Photography has a concept called, “the rule of thirds,” when composing your image. This concept involves evenly dividing an image into thirds both horizontally and vertically (Fig 1, Dunlop, J) and having the subject of the image being places on the intersection of those lines. In addition to coloring, I also want to investigate whether the area in the image where that color if present affects its attractiveness (i.e. red being present in the bottom, left corner of the image compared to the top, right).



Fig 1

A study, published in 2011 by the University of Toronto and Adobe Systems, Inc., conducted by Peter O'Donovan, Aseem Agarwala, and Aaron Hertzmann looked to explore color preferences in humans; using large datasets. In their studies, they found that people preferred simple color combinations and themes that only relied on 2 or 3 colors (O'Donovan). In the study, computer programs, Adobe Kuler and COLOURLovers, allowed online-users to create their ideal color-palette (using a maximum of 5 colors), which would then be submitted to the research team, in order to determine the preferences of how many different colors a palette should have, as well as what those colors should be.

Additionally, an article published by Emma Norris, in sponsorship with Plann X Canva, explored the relationships between colors and emotions; detailing what each of the primary and secondary colors provoked as an emotional response from audiences (Norris, E).

Studies have also been conducted on the influence of photographs, such as a thesis by Jeffrey L. Hall, which found the category of photograph to be the most significant factor in determining appeal. The most important were ranked as: the subject in still-life and landscape, and depth of field in people and landscape photography (Hall, J). This study utilized a sample of people who rated what aspects of a photo where the most significant to them, and then processed the data using a MANOVA method at a significance level of 0.05

Although these studies looked into a person’s response to an individual aspect of photography (color verses subject or color and emotion), this report will combine these components, to give future Instagram-influencers the knowledge that they need to post images with the most-attractive subject and coloring. We will determine which sections of an image are most significant, the colors present in them, as well as which photography subject is most associated with higher, “like,” counts.

METHODS

In this study, 30 images were collected from 5, popular, Instagram photography accounts listed on an online-article (“Famous Photographers On Instagram”). Each image was then cut into 9 equally-sized sections, and labeled 1 through 9 (from left to right, and top to bottom). These sections were then processed through a color-analysis tool from geotests.net, to find the most often occurring pixel color in section of the original image. The color results, along with number of likes, account-followers, type of image (animal, landscape, or portrait), post-date, and collection data, were recorded in a spreadsheet, and then exported as a CSV file; to be imported into Python. In total, approximately 2,250 pieces of data were manually collected for this project.

Multiple linear regression models were then built; using various numbers of predictors/sections of images, in order to find the most significant area of a post, along with the most significant color used. This type of modeling was chosen because our data consisted of multiple predictors (the sections of an image), numerous color options (10), the type of subject (3), and was designed to predict the number of, “likes,” that a post would receive given these parameters, so MLR felt like the most appropriate method to proceed with.

In addition, bar plots were conducted to portray the counts of color occurrences in each section of the images, as well as the counts of the three different types of images. These help the audience better understand the number of instances of each color or subject occurring in the dataset.

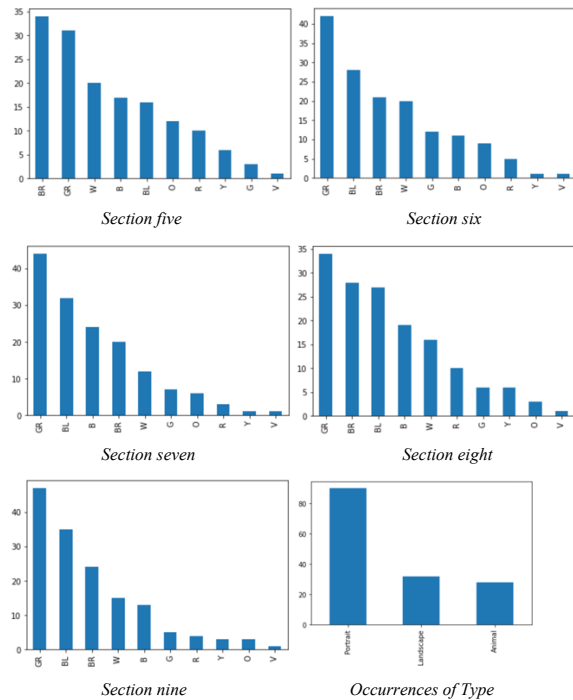
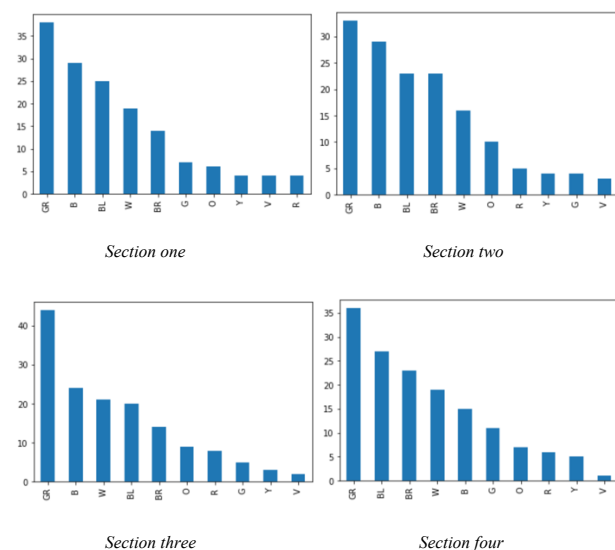
Other types of analyses, such as Random Forest and Lasso regression were conducted, but the results were not found to be useful.

2.1 Subsection 1: Data cleaning

Due to this dataset being created by hand, little to no data cleaning was required to proceed with our analysis. After importing my CSV file into Python, I executed the, “.dropna(),” command on my dataframe, in order to remove any NA (missing) values from my dataframe. Due to the way that my CSV file was created by Excel, rows 151 to 300 were empty, and so I removed them from the dataset by creating a subset of my original dataframe and saving it over itself.

2.2 Subsection 2: Bar plots

As previously mentioned, it was found that the color gray was the most often used color in all of the 8 outer-sections of images, with the center’s most often used color being brown (section five). Below you will find the bar plots of the nine image-sections, with the bars labeled by color. The following legend is used to identify each bar plot label: B = blue, BL = black, BR = brown, G = green, R = red, O = orange, Y = yellow, V = violet, W = white, and GR = gray. In the final bar plot, the occurrences of the, “Type,” of images are displayed. We can see that portrait photos are posted more than the combination of both landscape and animal ones.



2.3 Subsection 3: Regression and P-Values

Multiple linear regression was used to determine the most significant predictors of popularity of Instagram posts. The highest R-squared value (0.785) model that was produced utilized the predictors of, “Followers,” and, “Section (one through nine),” to determine the number of, “Likes,” that a post received. With a p-value of less than 0.000, it was found that the, “Followers,” predictor was most significant, followed by, “Section seven,” (lower-left) at a p-value of 0.100.

Another model, with R-squared of 0.665 that used the nine image sections and the, “Type,” predictor, found the image, “Type,” to be the most significant predictor (p-value of 0.018), followed by, “Section seven,” again (p-value of 0.028).

To determine the most significant color and where it was located in an image, the categorical labels found in each section, as well as the labels for the, “Types,” were replaced by, “dummy,” variables, so that regression could be performed (before the colors were relabeled as values 0 – 9). In a model observing the nine sections compared to number of, “Likes,” it was found, at a p-value of 0.037, that the color brown in, “Section six,” was most significant, when an intercept was included. A model that did not contain the intercept found also found brown in, “Section six,” to be the most significant, with a p-value of 0.024.

Models that included the different, “Types,” of images then showed that, “Animal,” images (with a p-value of less than 0.00), were most significant, along with the model’s intercept. Interestingly enough, “Section four,” and, “Section six,” has equal significance when they were the color violet at a p-value of 0.030. With the intercept removed, “Animal,” photos still were the most significant, while the color and area changed to white in, “Section five,” (p-value of 0.023).

Finally, the image sections were grouped into rows (sections one-three, four-six, and seven-nine). The top row’s most significant

predictor was gray in, "Section three." The middle row was found to be green in, "Section four." The bottom row showed white in, "Section seven."

The color that brought the most positive increase in, "Likes," if present, individually, was red in, "Section nine," with a value of $7.644e+04$.

The color that brought the most positive increase in, "Likes," if present, in the top row, was violet in, "Section two," with a value of $6.199e+04$.

The color that brought the most positive increase in, "Likes," if present, in the middle row, was green in, "Section four," with a value of $4.945e+04$.

Finally, the color that brought the most positive increase in, "Likes," if present, in the bottom row, was violet in, "Section seven."

3. RESULTS AND FUTURE WORK

Overall, it is very interesting how color, and where it is placed in an image, as well as the type of image, affects its popularity on Instagram.

For future work, I would like to explore how accurate my model is at predicting a photography account's number of, "likes," on newer posts that they make, as well as applying this to my own account (@asalenssees). From my own knowledge, though, I foresee some discrepancies occurring between the predicted and actual values of the model. For example, if an account, "likes," a post (typically more than one) on another account, there is a chance that that account will, in-turn, give, "likes," to your posts. This color and subject related model would not account for such factors, unfortunately.

Another way that I hope to expand upon this study would be to include more specific, "Types," of labels for images, such as: astrophotography, nightscape, urban, rural, sunrise/sunset, daytime, sponsored, and others, as I only utilized broad terms in my dataset. Potentially, this could become a more significant predictor of popularity than it already is.

4. REFERENCES

- [1] Dunlop, J. (2020, July 14). What is the Rule of Thirds? (And How to Use it in Photos!). Retrieved October 10, 2020, from <https://expertphotography.com/improve-your-composition-the-rule-of-thirds/>
- [2] O'Donovan, P., Agarwala, A., & Hertzmann, A. (2011). Color Compatibility From Large Datasets. Retrieved December 8, 2020, from <http://www.dgp.toronto.edu/~donovan/color/colorcomp.pdf>
- [3] Norris, E. (2020, March 11). The psychology of color in Instagram marketing. Retrieved December 8, 2020, from <https://www.planthat.com/psychology-of-color/>
- [4] Hall, J. (2007, May). ELEMENTS THAT INFLUENCE VISUAL APPEAL IN PHOTOGRAPHS. Retrieved December 8, 2020, from https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=bgsu1174794370&disposition=inline
- [5] Github Repository: <https://github.com/blevesque749/stat287>

About the author:

Brandon Levesque: a photographer and video editor that grew up in Vermont, but has been fortunate enough to travel to China, Hong Kong, and South Korea, in order to view the world through his camera lens. This project was largely inspired by my love for photography, and desire to expand my brand on Instagram.

**C
o
l
u
m
n
s**

**o
n**

**L
a
s
t**

**P
a
g
e**

**S
h
o
u
l
d**

**B
e**

**M
a**