# Trends Across Famous Artists

Esha Maheshwari

Feb 22, 2024

CMPLXSYS 351

## Project Goals

The primary objective of this project is to conduct a color and image analysis of specific artists’ works, exploring various quantifiable aspects of their artistic expressions. This investigation aims to deepen our understanding of the relationships between artists, genres, nationalities, and the visual elements within artworks. By leveraging advanced color and image analysis techniques, this project seeks to uncover patterns, trends, and unique characteristics that define the artistic expression of selected artists.

The parts of image and color analysis that I will analyze are dominant colors, entropy, and HSV (hue, saturation, value) values. Dominant color refers to the color that is most prevalent in the piece. I will use these colors to further analyze the hue, saturation, and value components. The entropy of a piece refers to the complexity contained in a given neighborhood, typically defined by a structuring element. Higher entropy means higher ‘randomness’ present in the work, which can often appear through texture, contrast, or blurriness.

Ultimately, the findings will contribute to a more nuanced comprehension of the visual language employed by artists, enriching our appreciation of their creative outputs.

## Data

I used the ‘Best Artworks of All Time’[[1]](#footnote-0), retrieved from Kaggle, uploaded by the user ‘Icaro’ in 2019. It is a collection of paintings of the 50 most influential artists of all time. The dataset contains three files/directories, one for the dataset of information for each artist, one that contains full sized images, and one that contains resized images (compressed versions of the full sized images).

From the artists file, I looked at the following variables; name, genre, nationality, and paintings (number of their paintings in the dataset). To reduce the processing runtime for the original 8000+ images that were in this dataset, I decided to use the resized image files instead of the full sized images. The resized images directory contains only .jpg images of each file.

## Implementation

### Preprocessing

I began with completing some preprocessing on the dataset. I noticed immediately that one artist had poorly-named images and another artist had pieces that were missing in the database. I removed these artists and their respective pieces from consideration in my analysis. To reduce processing runtime for the functions I would have to eventually complete, I randomly chose 20 images per artist and ended with a total of 860 images for my analysis. I copied these images to a new folder to keep them separate from the 8000+ images that were initially in the resized directory.

### Dominant Colors

I used the ColorThief[[2]](#footnote-1) package to extract the dominant color in each image and stored that value.

### Calculating Delta E

I used the CIELAB color space, also referred to as L\*a\*b\*, which is designed to approximate human vision. The L\* component closely matches human perception of lightness, while the a\* and b\* components measure the four unique colors of human vision (red, green, blue, yellow).

The Delta E value of a color describes its color difference or distance from another color and is measured on a scale from 0 to 100, where 0 is less color difference, and 100 indicates complete distortion. Essentially, I will use the Delta E value to analyze the similarities and differences in colors across various artists, genres, and nationalities.

I converted each dominant color found previously to a CIELAB value. I separated my calculations into artists, genres, and nationalities, to compare similarities within each subset. Then, for every other color in that group (artists/genres/nationalities), I found the Delta E value and stored it. I calculated the average of these values and stored it for each artwork, for each group. This value essentially tells us how similar a color is to other dominant colors of the same group. For example, for an artwork ‘Amedeo\_Modigliani\_33.jpg’, the average Delta E value for the artist (Amedeo Modigliani) will tell us how similar or different the dominant colors in this artwork are to others with the same artist. For the same artwork, the average Delta E value for the genre (Expressionism) will tell us how similar or different the dominant colors in this artwork are to others with the same genre. The same goes for nationality.

### Calculating Entropy

Entropy refers to the complexity contained in a given neighborhood, typically defined by a structuring element. More entropy means more ‘randomness’ present in the image. This can come through texture, how blurry/defined a picture is (more blurred means less entropy), and several other factors.

I used the scikit-image libraries' entropy[[3]](#footnote-2) method , which takes in an input image in grayscale. For each artwork, I converted the image to grayscale, then calculated and stored the entropy.

### Calculating HSV Values

The HSV (hue, saturation, value) color model that remaps the traditional RGB color model into dimensions that are more intuitive for humans. The hue is the color portion of the model ranging from 0 to 360 degrees (where red falls between 0-60 degrees, and magenta falls between 301-360 degrees, with all other colors falling between this range). The saturation describes the amount of gray in a particular color, sometimes appearing from 0 to 1. Reducing this component towards 0 introduces more gray and produces a faded effect. The value describes the brightness or intensity of the color, from 0 to 100 percent, where 0 is completely black, and 100 is the brightest and reveals the most color.

For each artwork, I converted its respective dominant color to HSV and stored the individual hue, saturation, and value values.

### Comparing By Artist

I grouped the artworks by artist and for each value I calculated above (artist Delta E values, entropy, hue, saturation, value), I found the average for each artist and sorted it as such.

### Comparing By Genre

I grouped the artworks by genre and for each value I calculated above (genre Delta E values, entropy, hue, saturation, value), I found the average for each genre and sorted it as such.

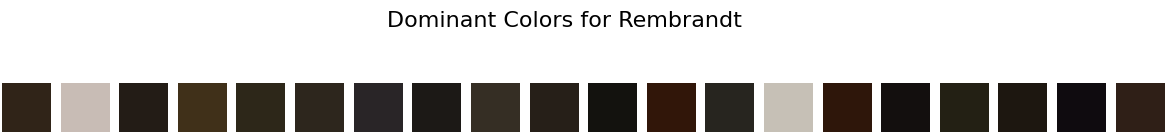
### Comparing By Nationality

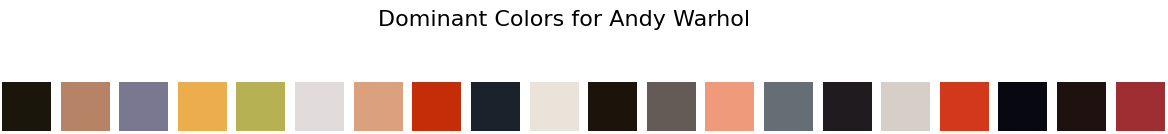
I grouped the artworks by nationality and for each value I calculated above (nationality Delta E values, entropy, hue, saturation, value), I found the average for each nationality and sorted it as such.

## Results and Discussion

### Comparing By Artist

Rembrandt and Michelangelo have the lowest Delta E values ranging from 19.41 to 19.85 in that order respectively. Andy Warhol and Joan Miro have the highest Delta E values ranging from 51.72 to 49.99 in that order respectively. The artists with the lowest Delta E values tend to use similar dominant colors within their work. The artists with the highest Delta E values tend to use a wider variety of colors within their work. Visually, we can see this through a display of a subset of the dominant colors for these artists in Figure 1.





*Fig 1. Dominant colors for Rembrandt and Andy Warhol side by side.*

Rembrandt and Gustav Klimt have the lowest entropy values ranging from 6.40 and 6.42 respectively. Michelangelo and Diego Rivera have the highest entropy values ranging from 7.63 to 7.62 respectively. Artists with low entropy values indicate less prominent strokes, a more ‘blurry’ effect, or fewer structural elements. Artists with high entropy values indicate more defined strokes and more structural elements. Visually, we can see this in figure 2.



*Fig 2. Michelangelo (first and second) and Rembrandt (third and fourth) artworks*

Giotto di Bondone and Titian have the lowest hue values ranging from 0.07 to 0.08 respectively. Piet Mondrian and Salvador Dali have the highest hue values ranging from 0.36 to 0.35 respectively. I did not see a significant result when comparing hue values, which makes logical sense, because hues only describe the undertone color used, not how saturated or intense the color is (which is easier to pick up on). Visually, I personally did not see a difference in artists based on their hue values (Figure 3). That being said, quantitatively, there is a relationship between different artists and average hues.





*Fig 3. Dominant colors for Giotto di Bondone and Piet Mondrian side by side.*

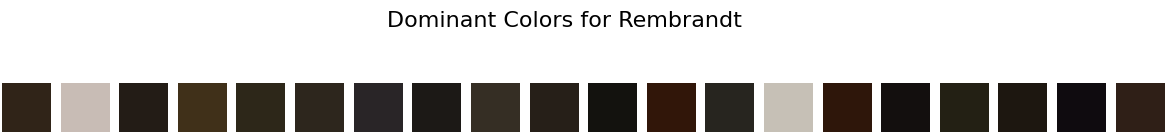
Alfred Sisley and Claude Monet have the lowest saturation values ranging from 0.14 to 0.17 respectively. Hieronymus Bosch and Frida Kahlo have the highest saturation values ranging from 0.59 to 0.58 respectively. Visually, this is easy to see in the dominant colors for the artists as shown in Figure 4, where low saturation artists seemed to have more grays and more gray-toned dominant colors while high saturation artists seemed to have less grays and less gray-toned dominant colors.





*Fig 4. Dominant colors for Alfred Sisley and Hieronymus Bosch side by side.*

Rembrandt and Eugene Delacrois have the lowest value components ranging from 0.21 to 0.29 respectively. Leonardo da Vinci and Michelangelo have the highest value components ranging from 0.69 to 0.68 respectively. Low value component artists have less brightness or intensity in their dominant colors, having lots of ‘black-toned’ colors present. High value component artists have more brightness and intensity, without many ‘black-toned’ colors present. Visually, we can see darker (closer to black) colors for artists with low value components, and lighter (closer to white) colors for artists with high value components as shown in Figure 5.





*Fig 5. Dominant colors for Rembrandt and Leonardo da Vinci side by side.*

### Comparing By Genre[[4]](#footnote-3)

Proto Renaissance and Baroque have the lowest Delta E values, ranging from 22.60 to 24.50 respectively. Pop Art and Primitivism, Surrealism have the highest Delta E values, ranging from 51.72 to 47.67 respectively.

Symbolism, Art Nouveau and Pop Art have the lowest entropy values, ranging from 6.42 to 6.52 respectively. Social realism, Muralism and Abstract Expressionism have the highest entropy values, both around 7.62.

Proto Renaissance and High Renaissance, Mannerism have the lowest hue values, ranging from 0.07 and 0.08 respectively. Neoplasticism and Symbolism, Expressionism have the highest hue values, ranging from 0.36 to 0.34 respectively.

Symbolism and Symbolism, Art Nouveau have the lowest saturation values, both around 0.20. Primitivism, Surrealism and Social Realism, Muralism have the highest saturation values, ranging from 0.58 to 0.55 respectively.

Baroque and High Renaissance, Mannerism have the lowest value components, ranging from 0.28 to 0.31 respectively. Symbolism, Art Nouveau and Suprematism have the highest value components, both around 0.66.

### Comparing By Nationality

French, British and Austrian have the lowest Delta E values ranging from 24.90 to 26.24 respectively. Mexican and German, Swiss have the highest Delta E values ranging from 44.90 to 44.65 respectively.

Austrian and Spanish have the lowest entropy values ranging from 6.42 to 6.97 respectively. Mexican and Norwegian have the highest entropy values ranging from 7.46 to 7.43 respectively.

British and Austrian have the lowest hue values, both around 0.14. Norwegian and French, British, Belarusian have the highest hue values, both around 0.34.

French, British and Austrian have the lowest saturation values ranging from 0.14 to 0.20 respectively. Mexican and Dutch have the highest saturation values ranging from 0.56 to 0.40 respectively.

Spanish, Greek and Flemish have the lowest value components, ranging from 0.35 to 0.41 respectively. Austrian and British have the highest value components, ranging from 0.66 to 0.65 respectively.

### Problems Encountered

Due to the dataset being geared towards the 50 most influential *artists* of all time, there is a discrepancy in the number of available paintings for different genres and nationalities. For example, many genres/nationalities only have one associated artist with them with 20 of that artists’ respective paintings to analyze. Although we can normalize our values for the total paintings available in the dataset across genres/nationalities, the fact remains that there is not a wide enough variety of paintings from different groups to make an effective conclusion. So, although similar analyses were completed across artists, genres, and nationalities, the only analysis that can effectively be used to make decisions on relationships are those completed across artists.

Due to time constraints, a further analysis into genres/nationalities was unable to be made.

## Conclusion

Knowing artist trends in colors and image analysis can be valuable for several reasons. For one, artists often reflect and respond to cultural and societal trends of their time. Understanding these artistic trends can provide insights into the prevailing themes, emotions, and issues in society. Different colors and visual elements also evoke specific emotions and responses and understanding these trends can help communicators choose visuals that effectively convey their intended messages. This can be especially helpful in fields such as marketing, advertising, and graphic design. Art trends are also a subject of study in academic and artistic institutions and understanding these trends helps art students and enthusiasts contextualize their work within broader artistic movements.

1. <https://www.kaggle.com/datasets/ikarus777/best-artworks-of-all-time> [↑](#footnote-ref-0)
2. <https://github.com/fengsp/color-thief-py> [↑](#footnote-ref-1)
3. Used the scikit-image library to calculate this: <https://scikit-image.org/docs/stable/auto_examples/filters/plot_entropy.html#sphx-glr-auto-examples-filters-plot-entropy-py> [↑](#footnote-ref-2)
4. The general analysis of what distinct components mean stay similar whether they are compared by artist, genre, or nationality. For the remaining results and discussion section, I will simply state the components results, assuming the same analysis as described above. [↑](#footnote-ref-3)