|  |
| --- |
|  |

|  |
| --- |
| Budget Collector  GLOBAL DATA ARTISTS (GROUP 7) |

|  |  |
| --- | --- |
| Lilian Chepng'eno;  Hojun Chan;  Joao Pedro Martins Morand |  |

Final Project Report

|  |
| --- |
| How Color has Shaped the Visual Language of Art  *Join us in the process of creating an interactive medium for artists, art collectors, and art enthusiasts as we visualize and get to understand the color changes influenced by changes in art movements, globally, using data analytics.*  Summary  Our team has partnered with Budget Collector with a goal to investigate whether there is any correlation between the colors used in a painting, when the painting was created and where the painting was created. This will involve the extraction of the dominant colors of an image, using image processing techniques, and thereafter statistical methods and time series visualization to establish correlation. The successful achievement of the project objectives will be beneficial to art collectors, art enthusiasts, and artists in different aspects. |

Introduction

Understanding color has long been an unsolved problem. Johann Wolfgang von Goethe’s book "Theory of Colors" was one of the earliest formal explorations of the subject. Goethe studied the relationships among the different colors and how they seemed to interact with each other. He stated that different colors have different psychological impacts on a person’s emotions and mood. Minor differences in the color can evoke a different emotion.

The purpose of this project is not to give a strict answer to how colors were used in an artwork’s time period, style or origin. Just as the artists wanted their pieces to be thought-provoking and engage the audience in a way that is open for interpretation, we want our interactive medium to allow the user to visualize correlations and trends on how colors were used.

Chart, sunburst chart

Description automatically generated

Johann Wolfgang von Goethe's Theory of Colour

THE DATASET

Our dataset of choice is the Budget Collector Inc dataset. We chose this dataset for the following reasons.

* Relevance: The dataset contains all the variables that we would like to analyze, i.e., pictures of the painting, and corresponding metadata related to period, location and style. It also contains information on the artists, media and tags related to the picture
* Credibility: Since the data was prepared and provided by Budget Collector, we can trust that the information is credible, and any subsequent analysis done using the data can be trusted.
* Completeness: The most important features (Period, Location, Image, style) required for our analysis have no missing values.

Data Exploration & Analysis

The visuals below give a representation of the 396 images of the raw dataset based on different attributes. As a quick summary of the raw dataset, most of the images were:

* Painted in Europe & North America
* Created in 19th & 20th Century
* In the style of Expressionism, Romanticism, Baroque, Impressionism, and Abstract Expressionism
* Painted on canvas with oil
* Contained humans, nature, and abstract images

Graphical user interface, application

Description automatically generated with medium confidenceMap

Description automatically generated

Graphical user interface

Description automatically generated with medium confidenceGraphical user interface, application, Teams

Description automatically generated

Extract, Transform & Loading

1. Extraction: The dataset was stored in AirTable, containing 396 images. To load the data into python we used API calls. We then transformed the loaded data into a Pandas data frame.
2. Data Cleansing: We confirmed that there were no duplicates of images in the data within the datasets. However, there were small typos in the dimension names which we later cleaned up such as duplication in time periods, medium, tags etc. We checked for null values and found 7 features out of 14 had missing data. To address this, we used imputation to address the missing datapoints.
3. Image Loading: We accessed the images through the image links stored in the Data frame using the Python Imaging library together with the HTTP requests library.

Limitations of Analyzing a Small Sample Set of Digital Artwork

Analyzing a small sample of digital artwork can be limiting for several reasons:

* Lack of representation: A small sample size may not be representative of the population of digital artwork for a specific style, period, region, etc. leading to biases and inaccurate conclusions about the overall characteristics or trends of digital artwork.
* Limited scope: The image itself does not give a comprehensive understanding of the context in which the artwork was created. For example, it may not capture the environmental, social, political, and cultural influences on the artist's creative process.
* Sampling error: A small sample size may cause sampling errors, such as random variation or outliers.

In general, looking at a small sample of digital art can be helpful for coming up with first insights and theories, but it is crucial to be aware of its limitations and possible biases. To get more accurate results, we can increase the sample size or use multiple data sources in order to get more trustworthy and valid results.

OUR APPROACH

Data Modeling using K-Means

After Data extraction into python and basic cleansing of the data, we embarked on extraction of the dominant colors for each image in the dataset. To do this, we decided to use Kmeans clustering algorithm using 5 clusters. To run Kmeans, we converted each image to a numpy array then we reshaped the array obtained into a format that could train a machine learning model based on the features of the image (-1,3).

Challenges

After reshaping the image and running the algorithm, we encountered several issues:

* Some images were not in the format of RGB (-1, 3) but in RGB-A (-1,4). To resolve this issue, we converted these images to the RGB (Red, Green, Blue) format. We used the python imaging library to perform this
* We initially tested our algorithm, with a few images and it ran perfectly, however, when we ran the algorithm on the entire dataset, we found that the performance was quite slow, processing 10 images in 300 mins. To resolve this, we resized all the images while maintaining quality and the performance improved significantly to running the entire dataset within 25mins. We used Python's OpenCV library to resize the images

Results

* With the issues resolved, we managed to run the algorithm on the entire dataset, whereby we obtained the cluster centers for each image.
* With the cluster centers obtained, we converted these to the RGB color representation. We also ranked each of the colors in terms of dominance by counting the labels for each cluster, the cluster with the most labels being the most dominant (Dominant color 1) and the one with the least labels being the least dominant (Dominant color 5).
* We created a simplified pallet of 252 colors, instead of the millions of possible RGB combinations then calculated the distance in the RGB values of the image’s color versus the RGB values of every color in our color palette. We then assigned a color name, Hex Code, and the Color shade to the five dominant colors for each image based on the shortest distance.
* In our color palette we assigned each color to a shade (11 shades), This is because in as much as colors can be varied, they are derived from a certain shade. This was also to be beneficial in the subsequent analysis of creating and interactive medium.
* We created 4 python dictionaries for the color name, color dominance, color hex code, and color shade whereby the dictionary keys were the image IDs and the values were the respective dominance, name, hex code and shade.
* We then converted the dictionaries into 4 different Pandas dataframes. We then merged the four dataframes into one final dataframe which we used to merge to our original dataset using the Image ID.
* This formed our final dataset which we exported into a Comma delimited file that we used to create on interactive medium.

A picture containing text

Description automatically generatedA picture containing text

Description automatically generated

*Original Resized image, versus the clustered image*

Chart, waterfall chart

Description automatically generated

*The dominant colors from clustered image above*

Chart, bar chart

Description automatically generated

*Bar plot of the dominant colors, by percentage of the color occurrence*

Color names from Pallet: Sienna; Turquoise; Lavender; Matte Black; Citrine

Color shades from Pallet: Brown; Blue; Purple; Black; Yellow;

Preface

Flaws of Analyzing Digital Artwork Through a Computer

There are several potential flaws in analyzing digital artwork through a computer, including:

* Lack of Detail: A computer screen might not be able to display a digital artwork's entire spectrum of colors and details. This could cause the examination of the artwork to lose information and nuance.
* Technical Limitations: The tools and equipment used to examine digital artwork may be restricted in their ability to zoom in or out or in their support for particular file types. These restrictions may compromise an analysis's precision and comprehensiveness.
* Inability to Evaluate Texture: A computer may be able to analyze the colors and shapes in a digital artwork, but it might fail to do so when evaluating the texture or other physical aspects of the piece.

A picture containing text, different, items, several

Description automatically generated

*Example: Limitations of Analyzing Digital Artwork Through a Computer*

*Apparatus and Hand from Google Images*

Key findings

Despite the limitations of analyzing digital artwork through a computer, the key findings were significant enough to allow us to analyze the correlation and trends with the color of artworks throughout time. Our original hypothesis was that there would be a distinct, single color for each period. However, this was not the case. In fact, we found that artists used the same several colors to symbolize different messages, evoke a range emotions, and create depth. From the results obtained, the most dominant colors are black, brown, and gray.

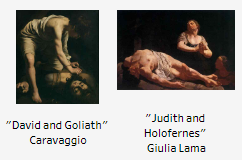
Black

Graphical user interface, application, website

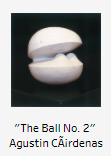
Description automatically generated

The color black is timeless and versatile, and it has been the most dominant color found in all of the images we analyzed. Black has been used to represent a wide variety of styles, such as Baroque, Expressionism, and Romanticism, and has been popular in many regions, including Europe, North America, and Asia. Black has been prevalent throughout the 17th to 21st century, and through looking at the images and tags, we can see themes of evil, sorrow, rebellion, and power. For example, in the images below, we can see things like “death”, “man”, “fear”, etc.

Since black has a wide range of symbolism, it may be a common choice for artists as it can depict different messages and elicit different emotions. Black can convey a sense of drama and intensity in a painting, which is often used by artists that want to depict strong emotions, such as grief or anger such as in “David and Goliath” and “Judith and Holofernes”.



In addition, black is frequently used as a background color to make the center image pop. Black can create a stark contrast with lighter colors, making them stand out more and adding depth to the artwork such as in “The Ball No. 2”. For example, in a portrait painting, using black to define the shapes of the face can make the features more prominent such as in “John Quincy Adams” portrait.



Although some images in our dataset do not have a lot of black, black was still returned as the most dominant color or among the top five colors. We therefore did research to understand why this is the case. One reason is that carbon black charcoal has been used by artists to draw their initial designs before beginning a painting since prehistoric times till the present day. These initial charcoal sketches are an essential step in creating art because they help define the composition and establish the relative values of the elements depicted. This implies that the shade may still appear black even if additional colors are added on top of the charcoal sketch.

Additionally, from our color palette of 252 colors, we noticed that some of the green shades were being classified as black (in the image below). We think this might be due to some of the limitations of analyzing art through a computer or due to our palette of colors not having a wider array of colors.

A picture containing timeline

Description automatically generated

*Example: Colors classified to black*

Our interactive medium showed that black was widely used in the 19th and 20th centuries in the Europe and North America regions.

Graphical user interface, application

Description automatically generated

We researched to understand how black was used in those periods and regions, and we found that black was a first choice for many European art academies as it could be used to create strong contrasts and to define the contours of the objects in paintings. This is consistent with some of the art pieces in our dataset whose dominant color was black. In addition, the realist movement emerged in the 19th century, and artists started using black to depict drab, industrial landscapes of the time, as well as the somber clothing worn by many working-class people as well as priests and judges. Improvements in the manufacture of synthetic pigments in the 19th and 20th centuries allowed artists to produce more powerful, dark hues than ever before, and black started to appear more frequently in paintings around this time. Black was associated with drama, mystery, and depth, and it was a powerful tool for artists who wanted to evoke strong emotions or convey a sense of darkness or melancholy.

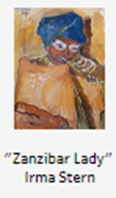
Overall, black is a versatile and powerful color that has been used to convey a wide range of messages and emotions throughout art history, and it continues to be a popular choice for artists today.

Brown

A picture containing text, screenshot, different, several

Description automatically generated

Brown has been a color used over time to represent earthiness, ruggedness, health, and nature. While it may be considered drab in comparison to other colors, it allows artists to create a sense of realism on canvas. Brown is commonly used in landscape paintings to depict natural elements like rocks, trees, and soil, and in still-life paintings to create the illusion of wood or leather textures such as in “Claude Monet Painting by the Edge of a Wood”. It can also be used to add shading and depth to skin tones in portraits, as seen in "Zanzibar Lady."



Brown is easily mixed with other colors to create a range of earthy tones and shades, and it can create depth, texture, and warmth in a painting.

Graphical user interface, application

Description automatically generated

Darker shades of brown can create a somber or melancholy atmosphere, while warmer shades can create a sense of comfort. However, the limitations of using computers to analyze digital art and the narrow palette used for analysis may lead to misclassifications, as seen with some green colors being classified as brown due to the reaction of copper-based paints to light and oxygen in paintings made from the 15th to 17th centuries. In the 19th and 20th centuries, many of the brown pigments used in painting, such as raw umber, burnt umber, and sienna, were naturally occurring earth pigments that were readily available and affordable. This made them a popular choice for artists who wanted to create a naturalistic or earthy color palette.

Brown is a versatile color that can be used to create a broad variety of tones and tints, from deep and rich to subdued and delicate. It was a popular choice for artists who wanted to add depth and dimension to their works, particularly in paintings that emphasized still life or natural landscapes. Additionally, brown's symbolic associations with stability, earthiness, and tradition may have made it a popular choice for artists who wanted to convey a sense of rootedness or historical continuity.

Gray

Graphical user interface, application

Description automatically generated

Gray is a dominant color in many paintings, often used to convey neutrality and balance. This could be because gray is the shade between black and white. Our research indicates that gray is commonly used for two reasons: First, artists use gray to create shadows and highlights. By mixing different shades of gray with other colors, they can add depth and dimension to their work. For example, in the painting "Butterfly," gray was used multiple times for shading. In "Haystacks (Effect of Snow and Sun)," darker tones create the effect of implied shadows. Second, gray can be used to convey a melancholy, reflective mood in a painting.



Artists like Rembrandt and Velázquez often use gray tones to evoke feelings of reflection and melancholy. For instance, Rembrandt used gray to depict strong water waves in one of his shown below paintings.



The Storm on the Sea of Galilee

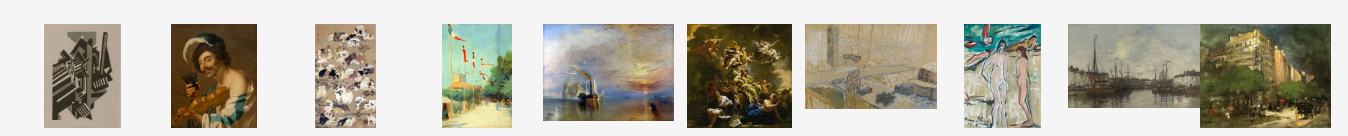
Rembrandt

Graphical user interface, application

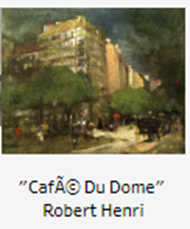
Description automatically generated

During the 19th and 20th centuries in Europe and North America, gray became a popular color in art. The Industrial Revolution of the 19th century led to the mass production of paint, and the range of available colors expanded greatly. Gray was one of the new colors available, and artists were eager to experiment with it. The art movements of the time, such as realism and impressionism, sought to depict the world in a more objective and realistic manner. Gray was useful for this because it is a more neutral and realistic color that could capture the subtleties of light and shadow. Photography also became popular during this time, and artists may have been influenced by it. Photographs were produced in grayscale and sepia tints, and artists may have been trying to mimic these qualities in their paintings. Lastly, gray may have been used for its symbolism. It is a subdued and somber color that conveys a sense of melancholy and contemplation. This was a common theme during the 19th and 20th centuries. Some artists in our dataset, such as Rembrandt, used gray in this way.

Green



Green has traditionally been connected to growth and new beginnings. It also denotes abundance and rejuvenation. In contrast, green can also stand for lack of experience, jealousy, or envy.. Even though green did not appear among the most dominant colors in our previous analysis, we believe it is widely used for different reasons. One of the reasons is to represent nature, as green is often associated with grass, trees, and foliage. In many landscape paintings, it creates a sense of tranquility and harmony with the environment. For example, in "Café Du Dome," green gives peace to the hustle and bustle of city life.



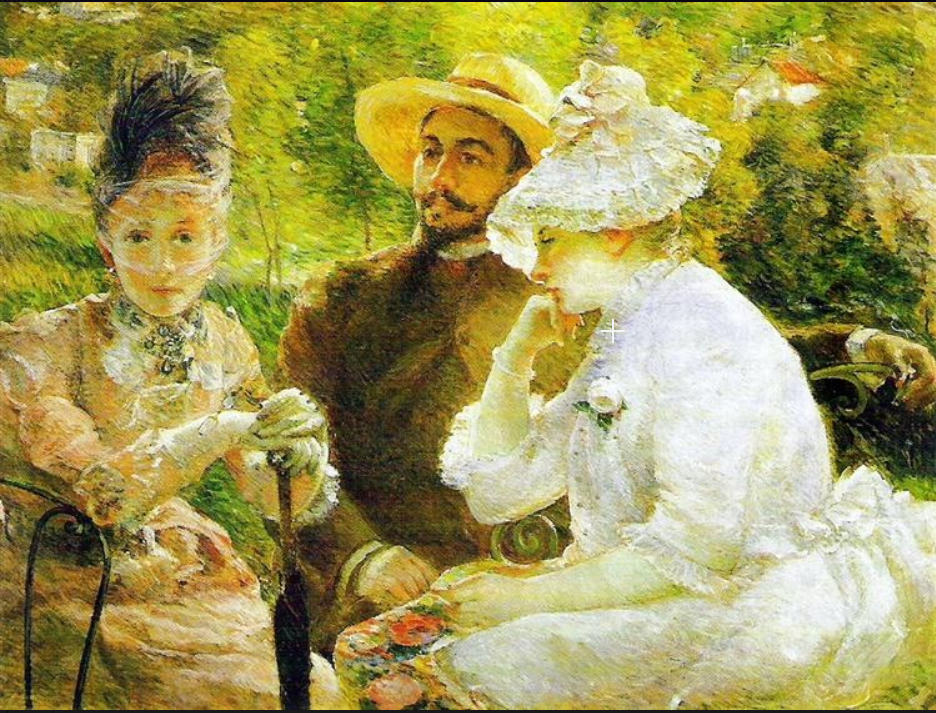
Similarly, in "Monk by the Sea" by Caspar David Friedrich, shades of green depict the horizon, emphasizing the monk's frail form before the majesty of nature, this is highlighted by the wide expanses of sea and sky. Green is also used to convey emotions and moods, representing jealousy, envy, growth, renewal, or hope, among others. In the painting "Wisdom and Virtue" by Luca Giordano, shades of green might have been used to symbolize growth. Using allegories and legendary creatures to represent many virtues and qualities, Luca Giordano created the artwork below with the concept of elevating people through wisdom and virtue.



Monk by the Sea Wisdom and Virtue

Caspar David Friedrich Luca Giordano

Moreover, in some cultures, green has special cultural or historical significance. In the 19th and 20th century Green was used to reflect the influence of the Impressionist movement which emerged in the mid-19th century. It was renowned for capturing the transitory effects of light and environment with the use of vivid, luminous colors. Green was a particularly important color for many Impressionist painters, as it was associated with the lush countryside and the shimmering reflections of water. The below image On the Terrace at SÃ¨vres by Marie Bracquemond from our dataset had the impressionism style.



On the Terrace at SÃ¨vres

Marie Bracquemond

Similarly, during that time, artists experimented with new and intense color combinations to convey a range of emotions and psychological states. Green was often used in this context to represent jealousy, envy, or a sense of unease or anxiety. Furthermore, artists in the 19th and 20th centuries experimented with new techniques and materials, including nontraditional pigments and dyes, which allowed them to create bold and striking greens that were previously impossible to achieve.

Graphical user interface, application

Description automatically generated

CONCLUSION

Ultimately, the analysis of artwork is subjective and heavily relies on human interpretation. While a computer can help use simplify the loading and processing of the images, it may not be able to fully capture the emotional or cultural significance of a piece of digital artwork. When viewed through a digital medium, it lacks the same context the artist may have originally intended the artwork to viewed in. Context, in terms of how the artwork interacts with its environment and subjects, such as its size, lighting or location is very important.

Through this project, we found that color usage in artwork has varied widely depending on the region and culture in which the artwork was created. Different regions and time periods have developed their own unique color palettes that reflect their artistic traditions and cultural values.

From our analysis, we saw common themes as to why some colors are chosen, these are: for symbolism, to create contrast, due to artistic movements or due to technological advancement or new materials and techniques. This, therefore, shows that there might be correlation between the color choices of an art piece, the period it was painted and where it was painted.

We, however, cannot conclusively establish correlation due to how small our dataset was. Additionally, the dataset was skewed in that, most of the images were painted in Europe and North America and during the 19th and 20th centuries. This therefore meant we were not able to understand whether there was any influence in color choices of art due to art movements or where the painting was made for the rest of the regions. Lastly our color palettes were not very wide and hence this might have also provided unconclusive results in that some colors might have been classified into the wrong shades.

Recommendations

From our analysis, we saw that there might be correlation of the color choices of an art piece to where there painting was made and the period it was created, due to the reoccurring themes we saw for the four-color shades we analyzed in the findings section above. We however were not able to conclusively establish this due to the limitations stated above and as such we would recommend the below:

* Further analysis to be done on a bigger dataset containing images from all regions and ensuring the data is balanced and not skewed in terms of regions or periods or any other features.
* A wider palette of colors and shades to be used to classify the colors to avoid wrongful classification of colors.
* To get the optimum number of clusters for each image using a statistical method such as the Elbow method. In our case we used 5 clusters for all images, which might not be the optimum number for all images.

REFERENCES

Douma, M., curator. (2008). *Pigments through the Ages*. Retrieved April 12, 2023, from <https://www.webexhibits.org/pigments>.

Mayfair Gallery Guides. (2018) *19th Century European Painting: Key Styles & Movements.* Retrieved April 12, 2023, from <https://www.mayfairgallery.com/blog/19th-century-european-painting-styles-movements/>

PressPacs. (2019). *Why some greens turn brown in historical paintings.* Retrieved April 12, 2023, from <https://www.acs.org/pressroom/presspacs/2019/acs-presspac-october-2-2019/why-some-greens-turn-brown-in-historical-paintings-.html>

Maria Popova. (2012). *Goethe on the Psychology of Color and Emotion.* Retrieved April 12, 2023, from <https://www.themarginalian.org/2012/08/17/goethe-theory-of-colours/>