

The background is a dynamic, abstract composition of swirling, smoke-like or ink-like patterns in vibrant colors. On the left, there are deep purple and magenta hues. On the right, there are bright cyan, pink, and yellow tones. These colors blend and swirl together against a dark, almost black, background, creating a sense of movement and depth. The overall effect is reminiscent of liquid art or a microscopic view of colored particles in motion.

Color in VOGUE

Callie Bouzane
August 27, 2019

VOGUE

- Monthly magazine covering the beauty and fashion industries:
 - Both are multi-billion dollar industries in the US alone
- 127 years of covering these topics
- Vogue Archive goes back to first issue (1892)
- Strong presence in the fashion industry with tons of data to explore

1910



1960



2010

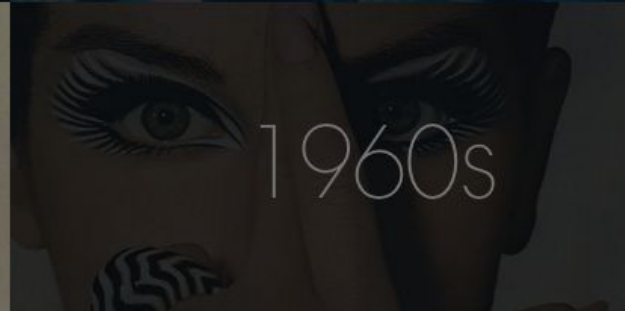


OBJECTIVE

Pull and analyze color in Vogue cover images
to identify any trends

AGENDA

1. Data Collection (Web-scraping Vogue's Archive)
2. Data Processing
3. OpenCV (Extracting Color from an image with OpenCV and KMeans)
4. Model (Binary Classification - Predicting Editor in Chief)
5. **Conclusions**
6. Next Steps



DATA COLLECTION

2,078 images

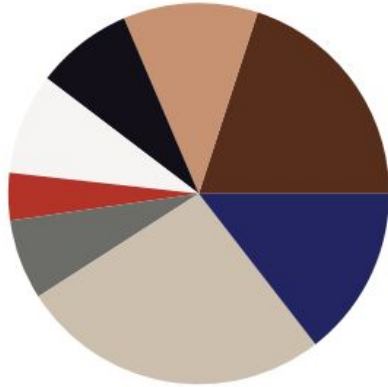
- Vogue Archive
- Beautiful Soup
- Pandas DataFrame
- Download images

DATA PROCESSING: COLOR TO DATAFRAME



STEP 1

Image



STEP 2

OpenCV

```
array([[131, 131, 120],  
       [164, 153, 126],  
       [102, 83, 52],  
       ...,  
       [160, 143, 134],  
       [155, 147, 143],  
       [ 79, 82, 41]])
```

STEP 3

RGB Array

	r7	g7	b7
1059	164	100	94
1060	144	71	54
1061	188	173	139
1062	85	70	62
1063	86	77	76

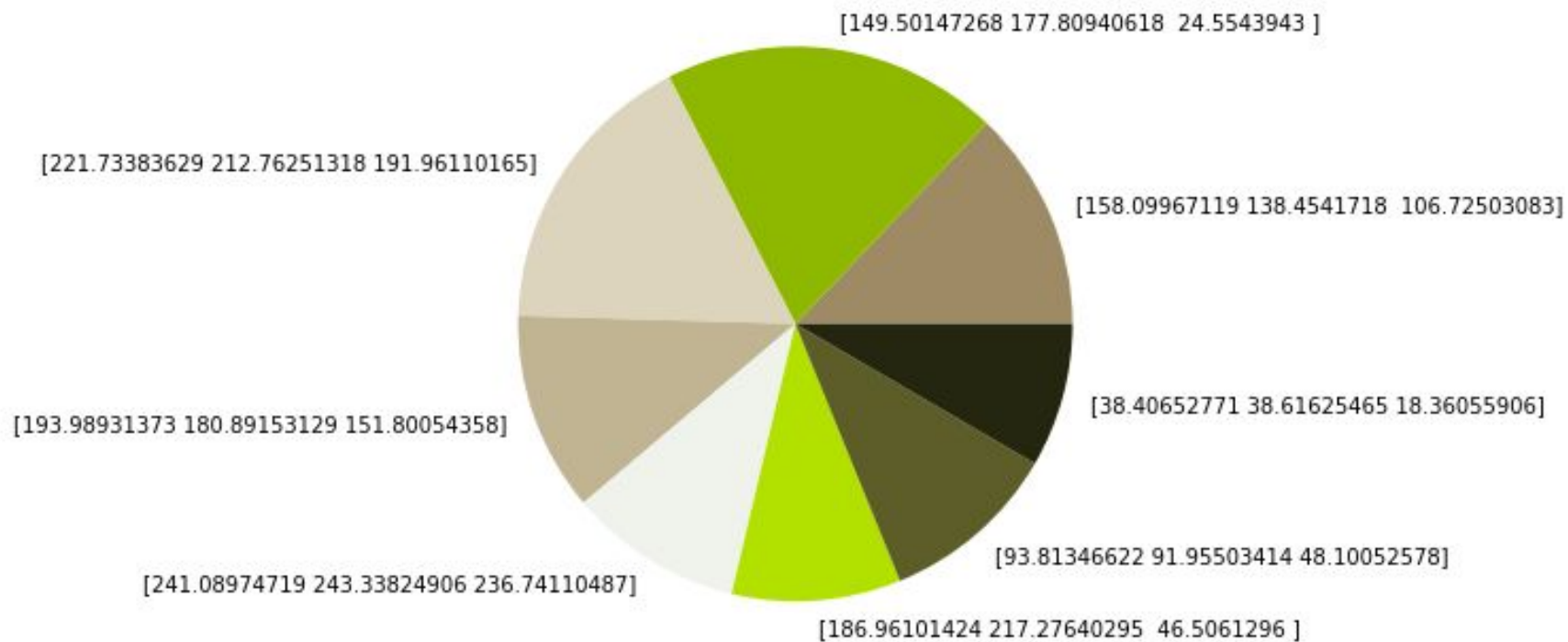
STEP 4

DataFrame

OPEN CV

1. Open Source Computer Vision
2. Uses Numpy to extract RGB values in an image
 - a. OpenCV extracts color from the reverse order, BGR
 - b. You need to re-order so that you are not looking at BGR Meryl Streep





FINAL DATAFRAME

Feature	Description
Date	Datetime format of the cover's publication date
Month	Month of the cover's publication date
Year	Year of the cover's publication date
Name	.jpg name of the cover's image
f#	frequency (in pixels) of an RGB value in a cover's image
r#	value of R (red) in an RGB value
g#	value of G (green) in an RGB value
b#	value of B (blue) in an RGB value
color#	the RGB array of r#, g#, b#
xkcd#	the xkcd study color name that most closely matches the color#
distance#	the distance between the xkcd# RGB array and the color# RGB array



BINARY CLASSIFICATION MODEL

- 1950's to present
- Predicting whether Anna Wintour was Editor in Chief for a given cover
- Features: colors (dummy variables)
- Logistic, Random Forest, and **Bayes**
- Baseline accuracy: 62%
- Train score: ~80.9%
- Test score: ~72.5%

CONFUSION MATRIX

	Actual Positive	Actual Negative
Predicted Positive	50	23
Predicted Negative	61	172

- Model ran better than the baseline
- There are trends worth exploring
- OpenCV less computationally taxing compared to a Convolutional Neural Network
- Exploring Color and Sales

CONCLUSIONS

NEXT STEPS

1. Develop a time series
2. Identify whether there are any trends in seasons and/or decades
3. Create a multi-classification model for seasons and or/decades
4. Create a KMeans Clustering model to identify any other potential trends

SOURCES

<https://www.mckinsey.com/industries/retail/our-insights/the-state-of-fashion>

<https://www.reuters.com/brandfeatures/venture-capital/article?id=30351>

<https://blog.xkcd.com/2010/05/03/color-survey-results/>

<https://frankturnerv.com/portfolio/fashion-models-from-fashion-models/>

<https://towardsdatascience.com/color-identification-in-images-machine-learning-application-b26e770c4c71>

https://scikit-image.org/docs/dev/auto_examples/transform/plot_histogram_matching.html#sphx-glr-auto-examples-transform-plot-histogram-matching-py

<https://www.pyimagesearch.com/2014/08/04/opencv-python-color-detection/>



Thank you!

Does anyone have any questions?