



Spotting Color Trends in Singapore's Fashion Scene

An unsupervised learning approach

By: Lim Shi Min



TABLE OF CONTENTS

O1.

Motivation and
Problem Statement

O2.

Data Collection,
Processing, EDA

O3.

Modelling

O4.

Visualizing a use-case

O5.

Conclusion & Future
Work

O1.

MOTIVATION & PROBLEM STATEMENT

SINGAPORE'S FASHION SCENE



Love, Bonito, one of Singapore's most well-known fashion retailer, clocked S\$31.4mn in annual sales in 2020

Most blogshops have weekly launches; important to keep up with the trend. Besides material and cut, color is also an important aspect

The top 30 fashion influencers have between 130K to 1.2M followers

Current Situation



It's a new month/season coming up. What colors should the shop pick for production?

Current Situation & Problem Statement



Manually scroll through Instagram, fashion magazines for trending colors to get inspiration?

Information cannot be stored and referred to systematically, making it difficult to really discern trends

Problem Statement

Help Singapore fashion retailers systematically spot trending color themes from Instagram photos of top fashion influencers, through unsupervised learning

ENVISIONED END PRODUCT

Select Month

Jan 2021

Feb 2021

Mar 2021

Apr 2021

May 2021

Jun 2021

Jul 2021

Aug 2021

Sep 2021

Oct 2021

Nov 2021

Color themes of the month

Select color to view relevant photos



As seen on:

- Influencer X
- Influencer Y
- Influencer Z

MEASURES OF SUCCESS

Metrics

1. Photos contributing to a cluster make sense

- Subjectively look through photos belonging to a cluster and see if there is a common color theme that can be spotted

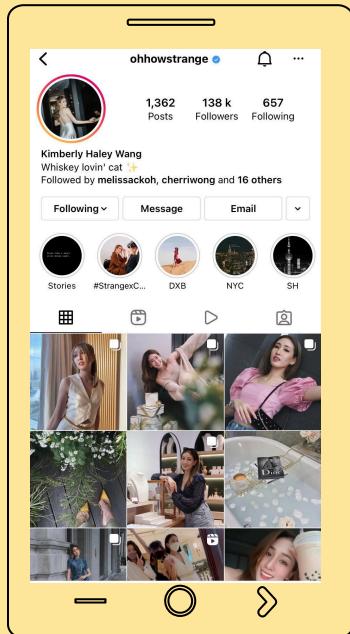
2. Able to offer new and interesting insights to trending colors

- Can install a click count on the color theme selection to track usage
- Gather feedback from users

3. Quality of clusters are good

- Use silhouette score to ensure that the quality of clustering is not too bad (minimally above 0.3)

O2. Data Collection & Processing



DOWNLOADING INSTAGRAM POSTS

STEP 1

List of top fashion influencers in Singapore from Starngage¹



Female Fashion Influencers x 34



Follower Count > 120k/influencer



Singapore

STEP 2

Remove influencers with topics such as food and pets tagged to them, and those that are tagged wrongly

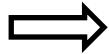
STEP 3

Use Instaloader (python library) to download posts from the selected accounts, from Jan-Nov 2021



>12000 photos

1. Starngage is an influencer marketing platform that has worked with >5,000 advertisers



PROCESSING THE PHOTOS

STEP 1

Remove background
using Python library
`rembg`

STEP 3

Resize images and
extract relevant RGBA
data using Python
Imaging Library

STEP 2

Remove irrelevant
photos (e.g. no
humans)



STEP 4

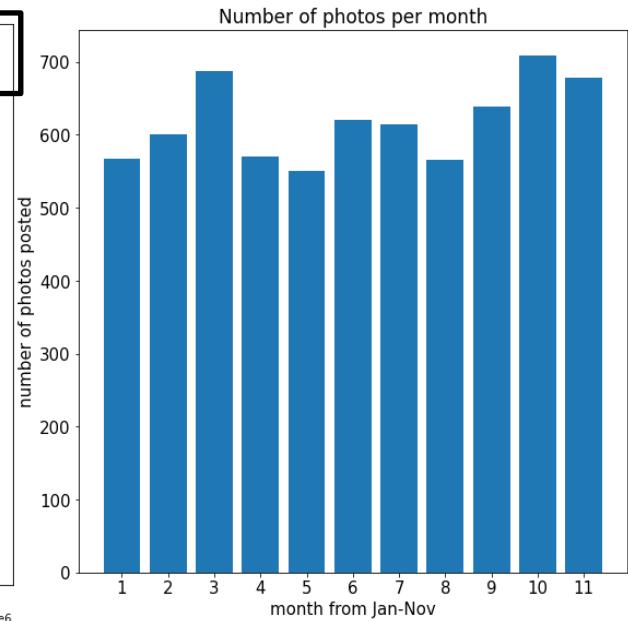
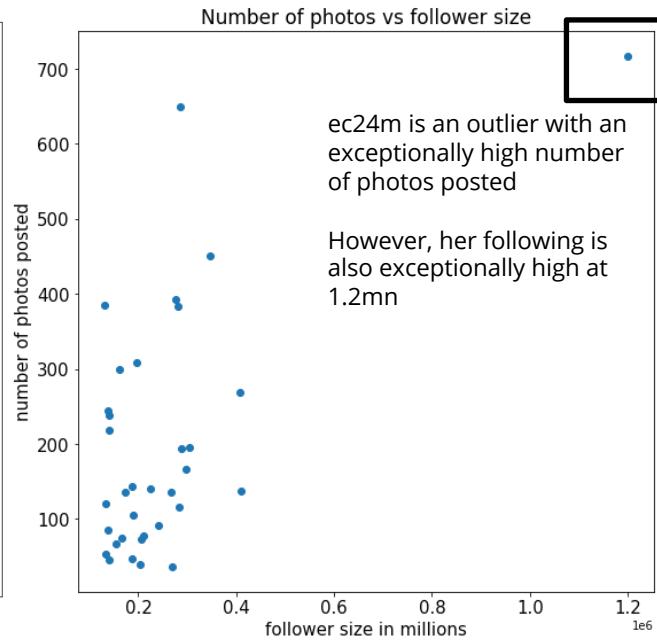
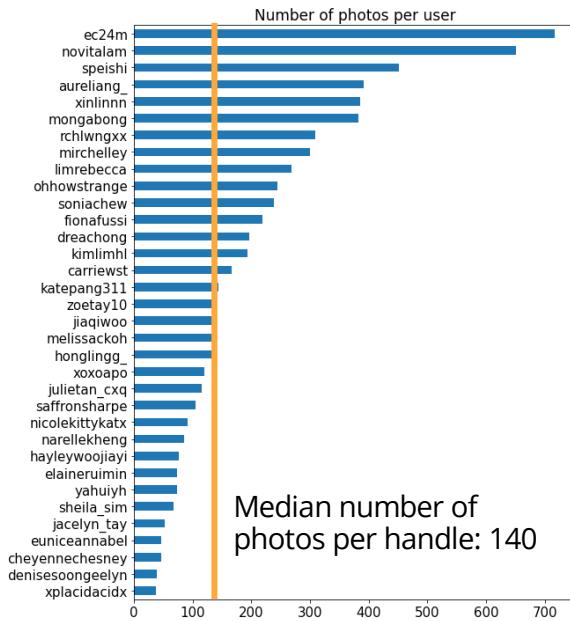
Save RGBA
information in csv

- Manual exercise at this stage
- But this is made easier to spot after background removal
- Left with 6800 images

- Data storage is a non-trivial issue as each photo has $40,000 \text{ pixels} \times 3 (\text{R}, \text{G}, \text{B}) = 120,000$ data points
- Data had to be transformed and saved by batches to csv

UNDERSTANDING THE DATA

Is there a good representation of photos from the different Instagram accounts, and across the months?



Quite consistent no. of photos per month, with each month having >500 photos

03.

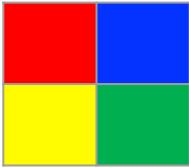
MODELLING

CHOICE OF MODEL

WHY UNSUPERVISED LEARNING

- Unlabeled Instagram images
- Color themes are not as straightforward as we think

Labelling color groups are easy only in extreme cases



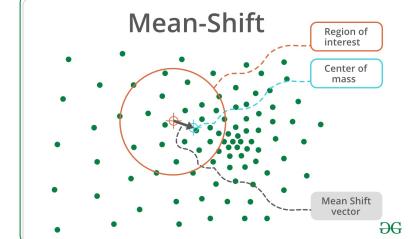
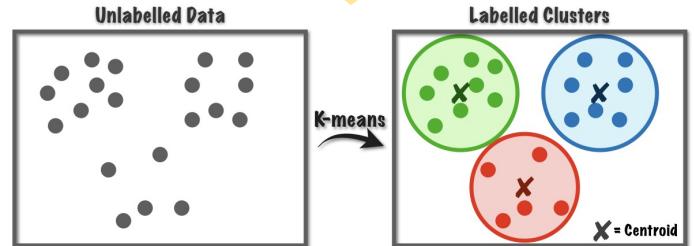
Can we label these as easily?

Red/
Purple?
Yellow/
Brown?
Grey/
Blue/
Green?
Blue/Purple?

MODELS CONSIDERED

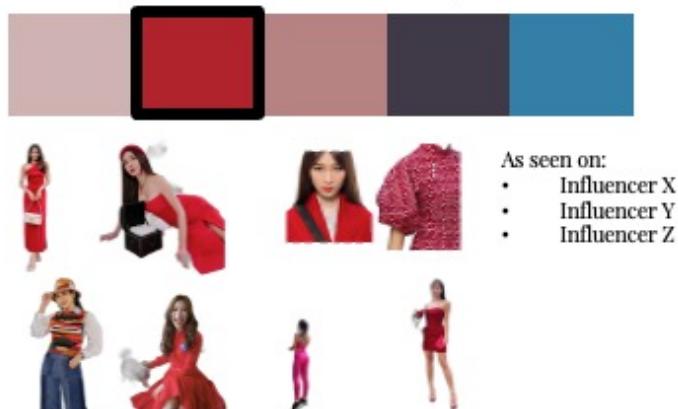
- K-means
- Mean-shift

} Centroid-based



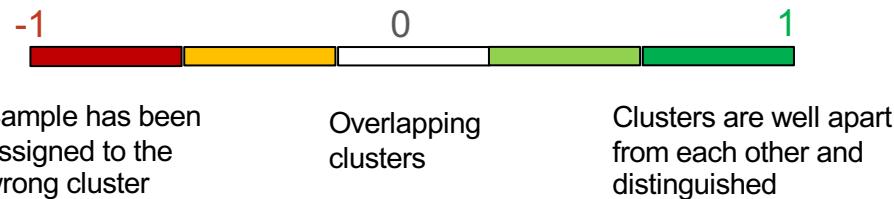
EVALUATION OF MODEL

SUBJECTIVE CALL: DO THE CLUSTERS MAKE SENSE?



OBJECTIVE MEASURE: SILHOUETTE SCORE

- Measures how cohesive each cluster is
- And how far apart clusters are from each other



CLUSTERING RESULTS ON INDIVIDUAL PHOTOS



K-means: K=2 Silhouette Score: 0.79



K-means: K=3 Silhouette Score: 0.72



Mean-shift Silhouette Score: 0.62



K-means: K=2 Silhouette Score: 0.66



K-means: K=3 Silhouette Score: 0.57

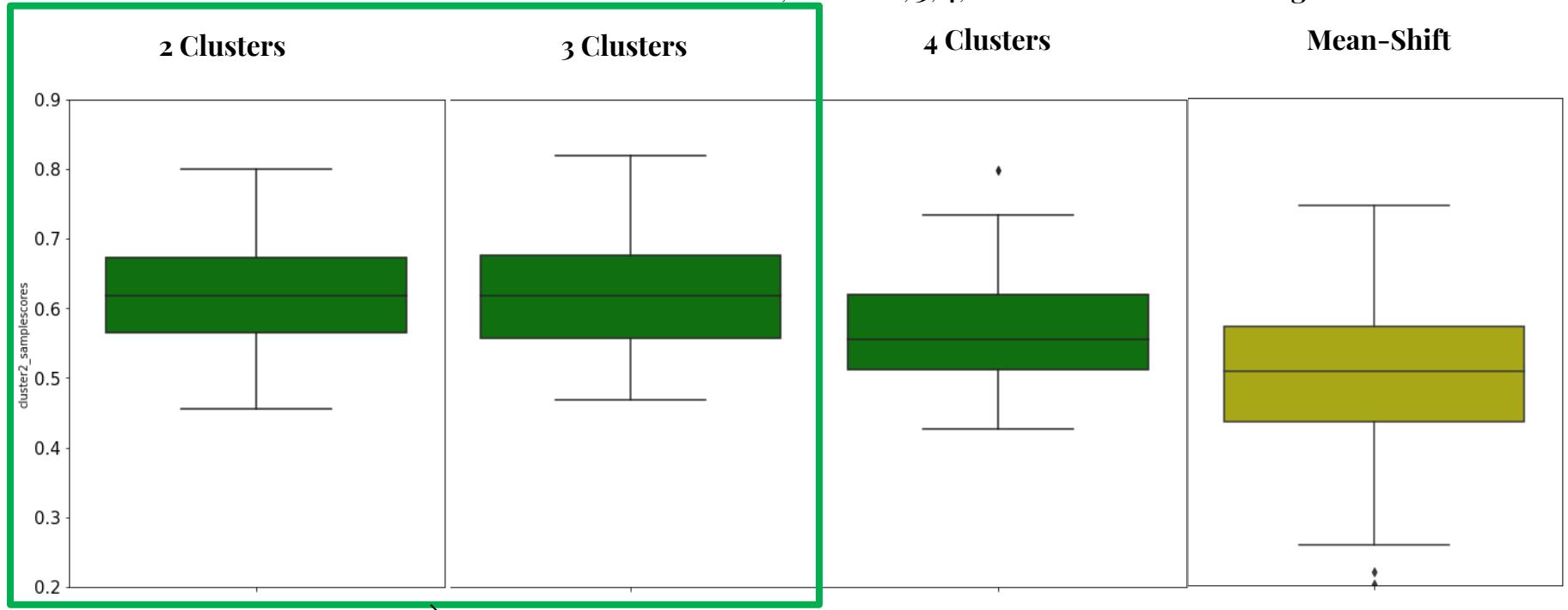


Mean-shift Silhouette Score: 0.58



DETERMINING NUMBER OF CLUSTERS FOR EACH PHOTO

Distribution of Silhouette Scores¹ for Each Photo, for K = 2, 3, 4, and Mean-shift Clustering



Since silhouette scores are similar for 2/3 clusters, I have decided to use 3 clusters so as to potentially capture more colors

1. Scores are calculated for a random sample of 100 photos instead of all 6800 photos, as score calculation takes up a lot of computing power

CLUSTERING RESULTS FOR EACH MONTH'S PHOTOS

K=3 clustering for each photo



x 6800 photos

Consolidating the centroids of each photo

Month	Photo	Centroids
1	1	a, b, c
1	2	d, e, f
	3	g, h, i
2	4	j, k, l

Cluster by month

Mean silhouette score
for each month
0.50

Silhouette scores do not differ too much for 3/5/7 clusters.
The user can toggle across the number of clusters to see if new inspirations
can be gathered from increasing the number of clusters

O4.

Visualizing a use-case: Color inspirations for CNY 2022

CLUSTERING RESULTS ON MONTHLY PHOTOS

USE CASE: INSPIRATION FOR COLOR THEMES FOR UPCOMING CNY

FEB 2021

Silhouette Score: 0.48



Some interesting patterns (differing hues of red as patterns) emerged – could be an inspiration for this year?

CLUSTERING RESULTS ON MONTHLY PHOTOS

USE CASE: INSPIRATION FOR COLOR THEMES FOR UPCOMING CNY

FEB 2021

Silhouette Score: 0.48



Muted pink shades
seem to be a popular
choice too

CLUSTERING RESULTS ON MONTHLY PHOTOS

USE CASE: INSPIRATION FOR COLOR THEMES FOR UPCOMING CNY

FEB 2021

Silhouette Score: 0.48



Surprisingly, bright blue/teal seemed to be popular during CNY last year

CLUSTERING RESULTS ON MONTHLY PHOTOS

USE CASE: INSPIRATION FOR COLOR THEMES FOR UPCOMING CNY

NOV 2021

Silhouette Score: 0.42



Bright orange-red hues seem to be appearing more in Nov, and could be also appropriate for CNY

05.

Conclusion & Future Work

CONCLUSION

1.

Photos contributing to a cluster make sense



- Color themes of the clothes were generally clear from the photos in a common cluster

2.

Able to offer new and interesting insights to trending colors



- CNY case study: bright blue and orange hues surfaced as possible colors to consider for upcoming CNY

3.

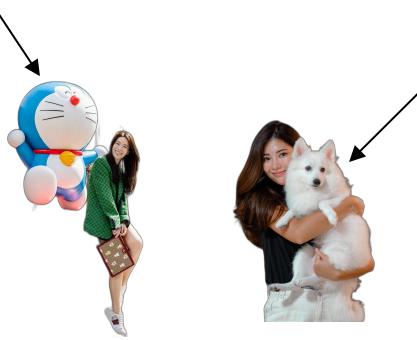
Quality of clusters are good



- Silhouette score at the monthly consolidated level is at 0.5

Overall, I will recommend the use of the K-means clustering method to discover color trends, to help Singapore fashion retailers with deciding on the color palette of their future launches

FUTURE WORK



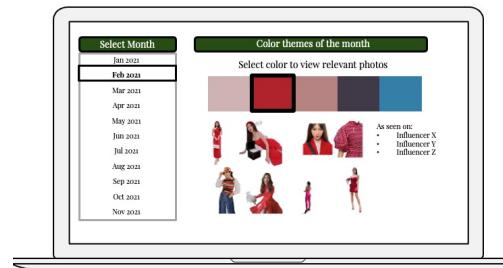
Issue: How to filter out irrelevant photos systematically

- **Current tactical solution:** Manually remove
- **To explore:** People counter with OpenCV (only keep photos where human count is 1).
- **To further think about:** How about photos that have one human, but have other dominating objects?



Issue: Nude/brown/black appears frequently as a cluster due to hair/skin

- **Current tactical solution:** Remove centroids that are close to nude/brown/black for each photo, before clustering at the monthly level
- **To explore:** Object segmentation to detect clothes before doing color clustering



And of course, to deploy into a webpage/app for easy access



THE END

References used for the project

- Medium article on K means clustering for dominant color detection [[link](#)]
- Youtube tutorial on interactive color detection
- Background removal library [[link](#)]
- Function to apply image pre-processing to whole folder [[link](#)]
- Instaloader documentation [[link](#)]
- Plotting clusters [[link](#)]

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by Freepik.

Please keep this slide for attribution.