FASHION

# Trend Clustering Tool ↗

Joyce Ishikawa & Shivani Dedhia

# TREND CLUSTERING

## ACTIVITY

You are a merchandiser for Tory Burch Women's Accessories

Women's Accessories include Socks, Belts, Scarfs, Bag Charms, Umbrella, Jewellery and Hats

Pick any **one category** from the Women's Accessories to analyse current trends

Take any 3 competitors and screenshot major trends that you spot

Activity Google Docs

TORY BURCH

FENDI

BURBERRY

LOEWE

# TREND CLUSTERING

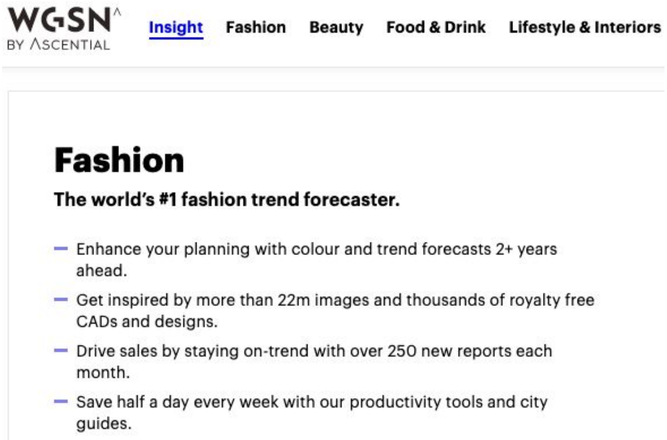
## WHY DO WE NEED IT?

Our user:
Merchandisers/Trend forecasters

Problem:
Users need to collect images of trends from various sources through manual labor screenshotting.

Fashion forecasting websites like WGSN, Fashion Snoops, Mintel etc are used to identify major trends across fashion. Too few image examples given.

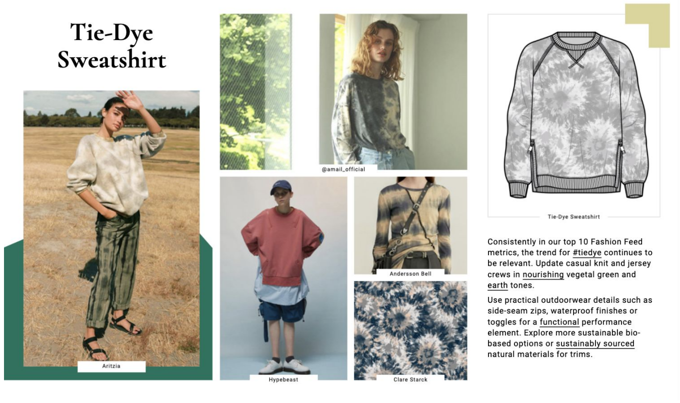
*Trend : *The most repeated pattern, silhouette, material, color across brands are trending*



**Fashion**

**The world's #1 fashion trend forecaster.**

- Enhance your planning with colour and trend forecasts 2+ years ahead.
- Get inspired by more than 22m images and thousands of royalty free CADs and designs.
- Drive sales by staying on-trend with over 250 new reports each month.
- Save half a day every week with our productivity tools and city guides.



Consistently in our top 10 Fashion Feed metrics, the trend for #tiedye continues to be relevant. Update casual knit and jersey crews in nourishing vegetal green and earth tones.

Use practical outdoorwear details such as side-seam zips, waterproof finishes or toggles for a functional performance element. Explore more sustainable bio-based options or sustainably sourced natural materials for trims.

# TREND CLUSTERING

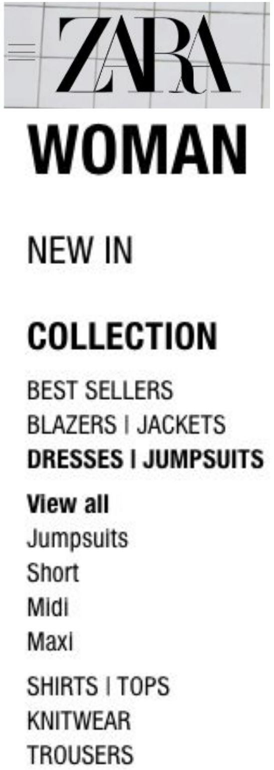
## WHY DO WE NEED IT?

*Why can't we just filter by image tags?*

All brands have different names to separate their categories.

**Zara** and **H&M** have different naming conventions to separate their 'Dress' category as seen on the left

Making it difficult and tedious for merchandisers to spot trends across many sites.



ZARA

WOMAN

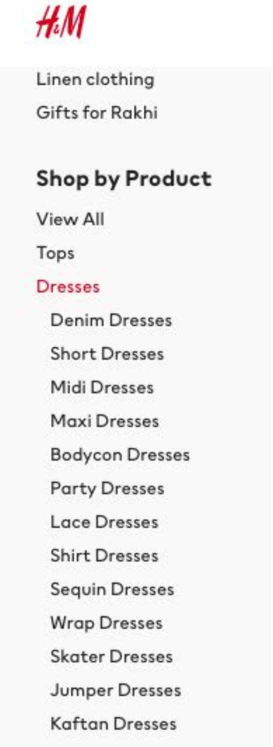
NEW IN

COLLECTION

BEST SELLERS
BLAZERS | JACKETS
**DRESSES | JUMPSUITS**

**View all**
Jumpsuits
Short
Midi
Maxi

SHIRTS | TOPS
KNITWEAR
TROUSERS



H&M

Linen clothing
Gifts for Rakhi

**Shop by Product**

View All
Tops
Dresses
- Denim Dresses
- Short Dresses
- Midi Dresses
- Maxi Dresses
- Bodycon Dresses
- Party Dresses
- Lace Dresses
- Shirt Dresses
- Sequin Dresses
- Wrap Dresses
- Skater Dresses
- Jumper Dresses
- Kaftan Dresses

# TREND CLUSTERING

## HOW DOES THE TOOL WORK?

**OVERVIEW**

1. Image scraped from desired fashion websites (Used Fatkun chrome extension)
   100 images from 4 websites(ZARA, H&M, ASOS, TopShop)
2. Rename files by source-number.jpg “ZARA-1.jpg” into training data (400 images)
3. Fit the training data set calculating feature vectors and train the model to find similar images by cosine similarity
4. Use the similar items function to access images grouped by similarity
5. Output: A list of folders of images that are ranked by similarity to the target image

# TREND CLUSTERING

## HOW DOES THE TOOL WORK?

Fitting our dataset:
For each image, their feature vectors will be calculated and stored in a sqlite database

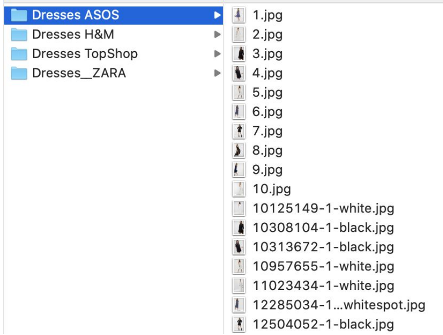
Features will be calculated for the original image, as well as the image flipped left-right and up-down, and the image rotated by angles of 90, 180, and 270 degrees.

Similarity model:
The library loops through all the items, and finds the items in the rest of the dataset which look the most similar.

Similarity is based on the cosine distance between the feature vectors calculated using the model provided (‘VGG19’ and ‘Inception_Resnet_V2’).

# TREND CLUSTERING

## HOW DOES THE TOOL WORK?



Dresses ASOS
Dresses H&M
Dresses TopShop
Dresses__ZARA

1.jpg
2.jpg
3.jpg
4.jpg
5.jpg
6.jpg
7.jpg
8.jpg
9.jpg
10.jpg
10125149-1-white.jpg
10308104-1-black.jpg
10313672-1-black.jpg
10957655-1-white.jpg
11023434-1-white.jpg
12285034-1...whitespot.jpg
12504052-1-black.jpg

1. Web scrape our images

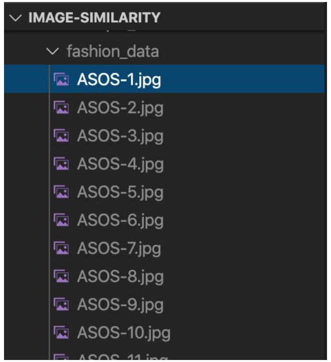
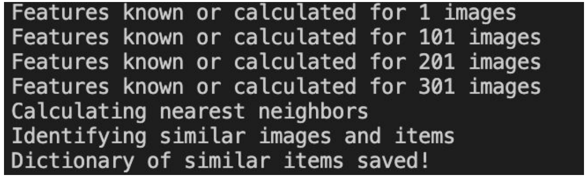


IMAGE-SIMILARITY
fashion_data
ASOS-1.jpg
ASOS-2.jpg
ASOS-3.jpg
ASOS-4.jpg
ASOS-5.jpg
ASOS-6.jpg
ASOS-7.jpg
ASOS-8.jpg
ASOS-9.jpg
ASOS-10.jpg
ASOS-11.jpg

2. Rename each file by retailer-number

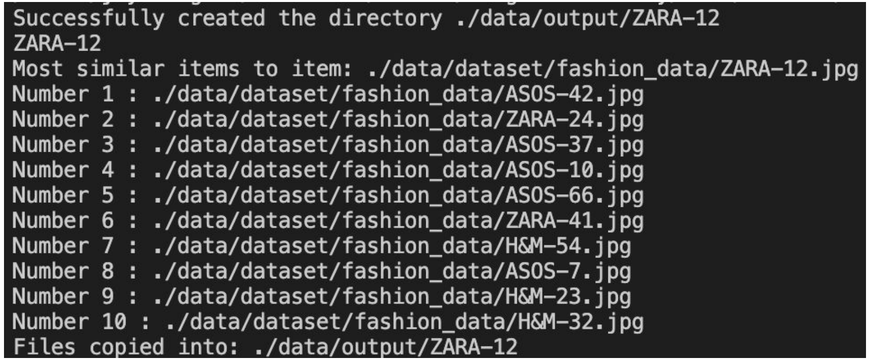


```
Features known or calculated for 1 images
Features known or calculated for 101 images
Features known or calculated for 201 images
Features known or calculated for 301 images
Calculating nearest neighbors
Identifying similar images and items
Dictionary of similar items saved!
```

3. Fit data and  by calling:
python main.py --task fit --dataset fashion_data

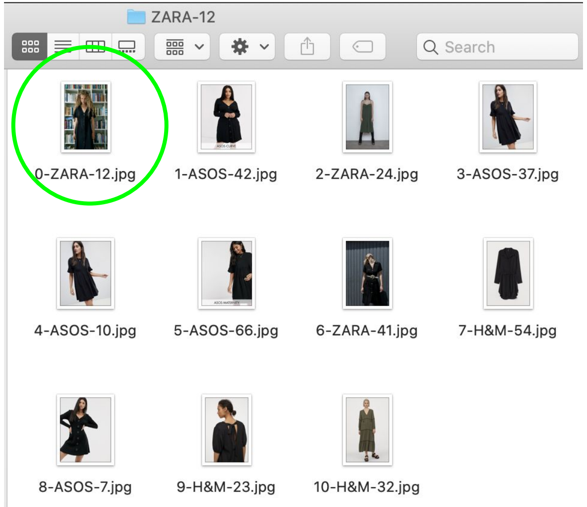
# TREND CLUSTERING

## HOW DOES THE TOOL WORK?



```
Successfully created the directory ./data/output/ZARA-12
ZARA-12
Most similar items to item: ./data/dataset/fashion_data/ZARA-12.jpg
Number 1 : ./data/dataset/fashion_data/ASOS-42.jpg
Number 2 : ./data/dataset/fashion_data/ZARA-24.jpg
Number 3 : ./data/dataset/fashion_data/ASOS-37.jpg
Number 4 : ./data/dataset/fashion_data/ASOS-10.jpg
Number 5 : ./data/dataset/fashion_data/ASOS-66.jpg
Number 6 : ./data/dataset/fashion_data/ZARA-41.jpg
Number 7 : ./data/dataset/fashion_data/H&M-54.jpg
Number 8 : ./data/dataset/fashion_data/ASOS-7.jpg
Number 9 : ./data/dataset/fashion_data/H&M-23.jpg
Number 10 : ./data/dataset/fashion_data/H&M-32.jpg
Files copied into: ./data/output/ZARA-12
```

4. Run a call to show similar items:
python main.py --task show_similar_items --dataset fashion_data



5. Randomly select one image as a target image
and find 10 similar images (ranked in order of similarity)

# TREND CLUSTERING

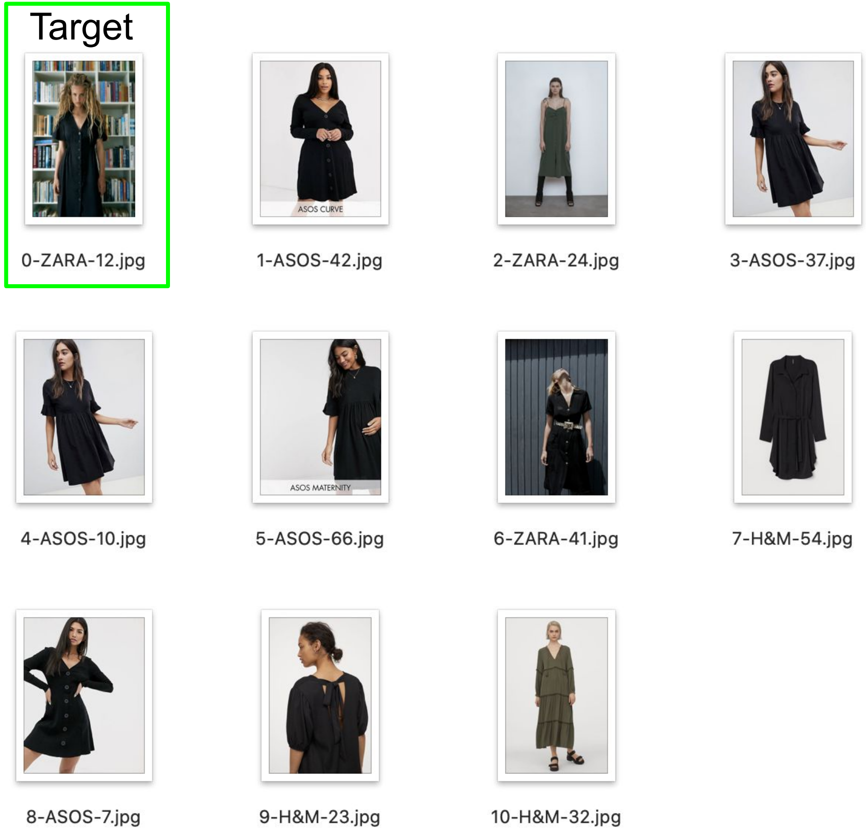
## OUTPUT

It creates a folder with 10 similar images to the target image.

These are ranked by similarity and stored in a new folder.

This folder automatically has the same name as the target image making it easy to identify.

E.g Folder name: ZARA-12



# TREND CLUSTERING

## LIMITATIONS

While training the model, if there is a corrupted image it breaks the code and does not run until that image is removed, making it a long process to train the model.

Duplicate images found in the dataset result in inaccurate results.

The time to train the model with more than 1000 images will take hours.
Approx 30 mins to train under 500 images.

Currently the function outputs only one folder at a time with 10 ranked images.

We can’t tell the difference in distance (similarity) between image rank 1 and 2

# TREND CLUSTERING

## NEXT STEPS

- Automatically make multiple folders and group similar images;
- Increase accuracy by letting the output be limited by the distances in the kNN;
- Increase our dataset and to check if accuracy increases;
- Create a script that will scrape web images from various sources.