StreetStyle: Exploring world-wide clothing styles from millions of photos

Kevin Matzen, Kavita Bala, and Noah Snavely
Experiments by: Angela Lin and Lakshay Narula

StreetStyle Dataset Statistics

Size:

- Total: 27084 entries
- 23908 Distinct Instagram Posts

Fields Associated With Each Image:

- City ID
- Month ID
- Upload Time
- Upload Location
- Bounding Box Coordinates
- Image Size
- Attribute Annotations

Original Image

Cropped Image

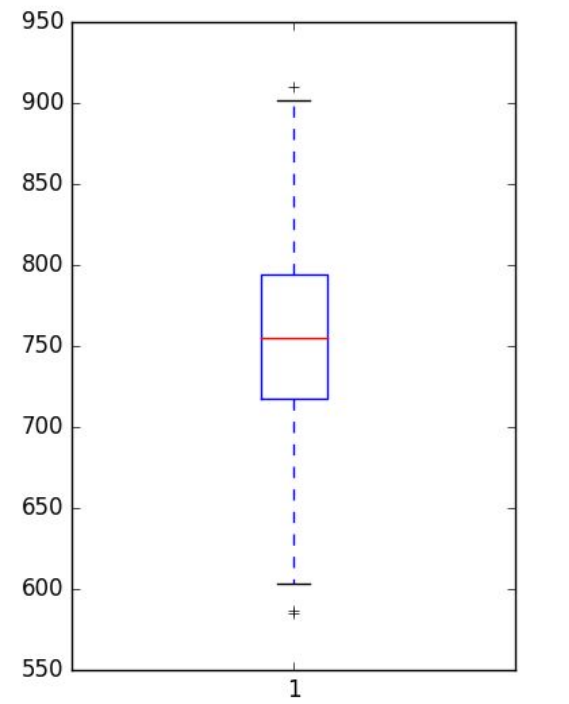
Cities in the dataset

- Bangkok
- Beijing
- Berlin
- BogotÃ¡
- Budapest
- Buenos Aires
- Cairo
- Delhi
- Dhaka
- Guangzhou
- Istanbul
- Jakarta
- Karachi
- Kiev
- Kolkata
- Lagos
- London
- Los Angeles
- Madrid
- Manila
- Mexico City
- Milan
- Moscow
- Mumbai
- New York City
- Osaka
- Paris
- Rio de Janeiro
- Rome
- SÃ£o Paulo
- Seoul
- Shanghai
- Sofia
- Sydney
- Tianjin
- Tokyo

Missing Cities:

- Austin
- Chicago
- Johannesburg
- Nairobi
- Seattle
- Singapore
- Toronto
- Vancouver

Number of Samples in Each City



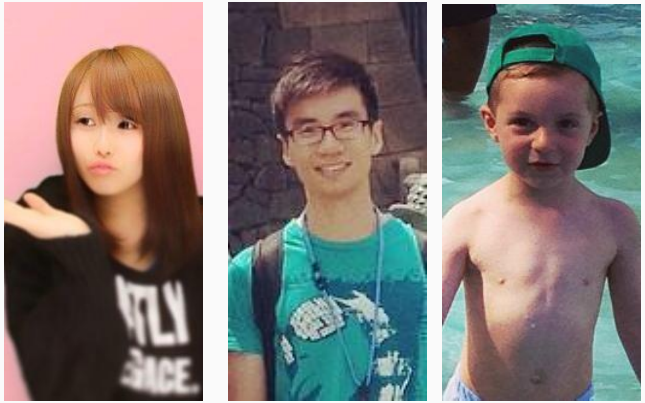
Median: 755 Samples

Examples from the dataset

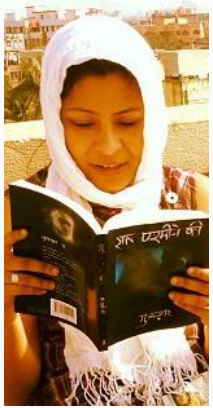
Clothing Category Shirt Outerwear T-shirt Dress Tank top Suit Sweater

Neckline Shape Round Folded V-shape

Sleeve Length Long Sleeve Short Sleeve No Sleeve



Style In The Wild

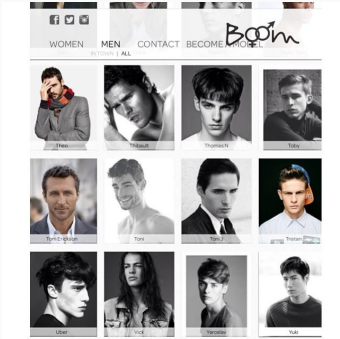


Multiple People & Occlusions

Variation in Pose

Varying Image Quality

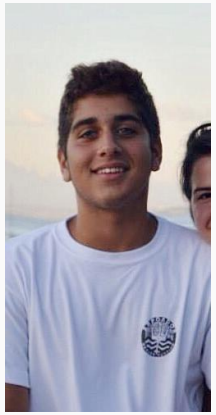
Cultural Holidays



Variety of Sources

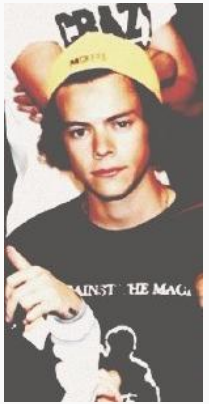
Co-occurrence Trends Using Conditional Probabilities

Style based



Functionality based

Reasonable Dependencies



Attributes Occur Outside of Bounding Boxes

Sleeve Length: Short Sleeve

Sleeve Length: Long Sleeve

Label: Dress Category

Label: Floral Clothing Pattern

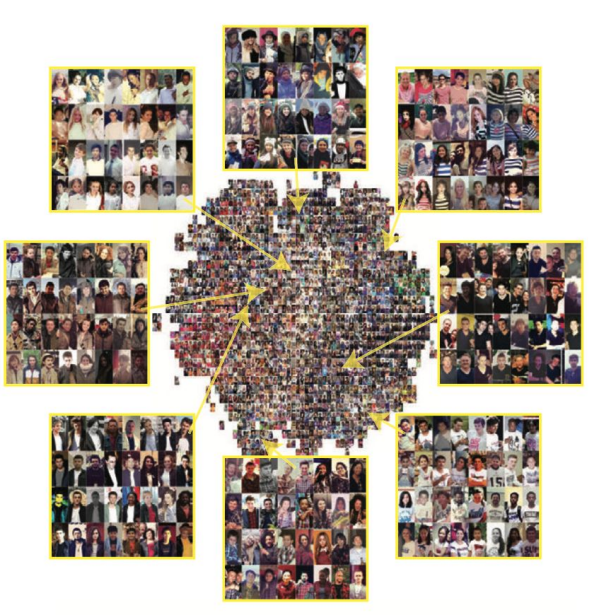
Purpose

The paper has the following theme:

1. Need large data to find trends
2. Large data tough to label
3. Label 27K images
4. Train a CNN to generate features, cluster to find trends

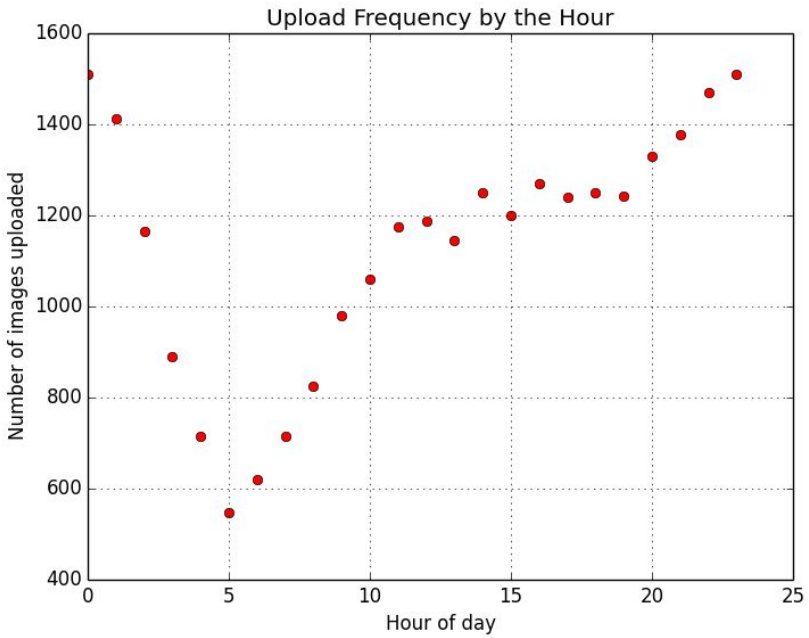
However…

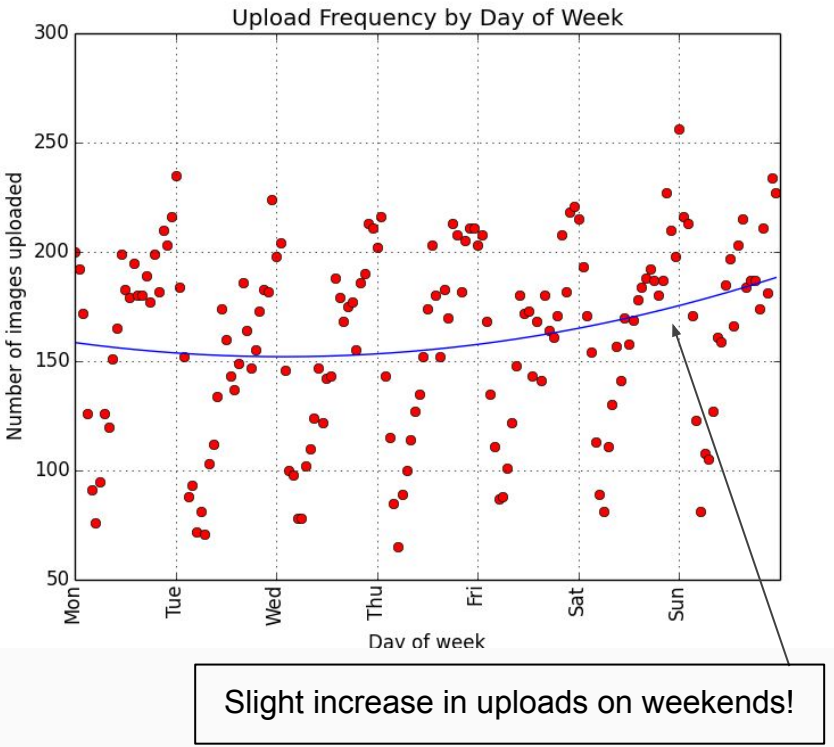
1. We don't have access to code
2. We don't have access to the large dataset (>10 million images)
3. We do have access to the labeled 27K images



These set of experiments attempt to expose possible interesting trends in the dataset that could be explored!

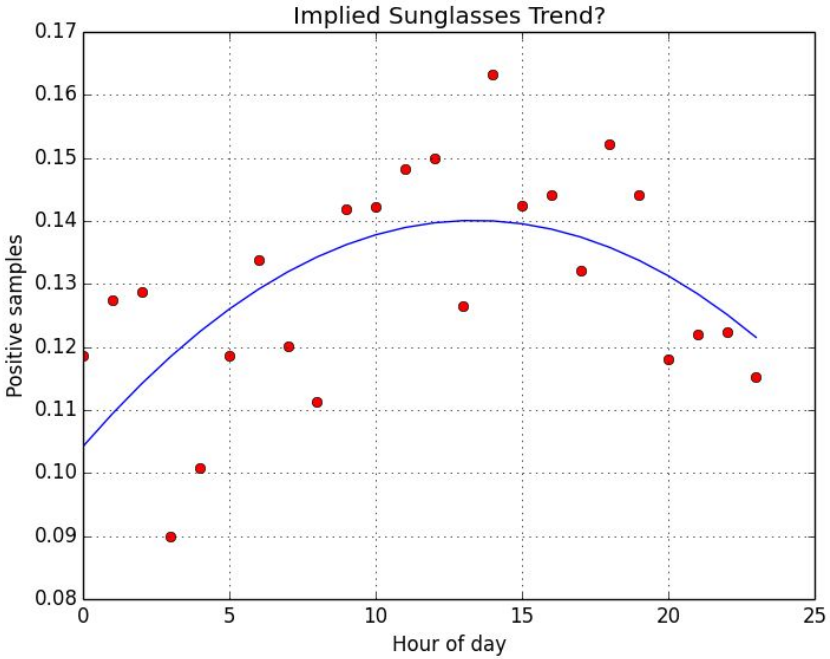
Image Upload Trends

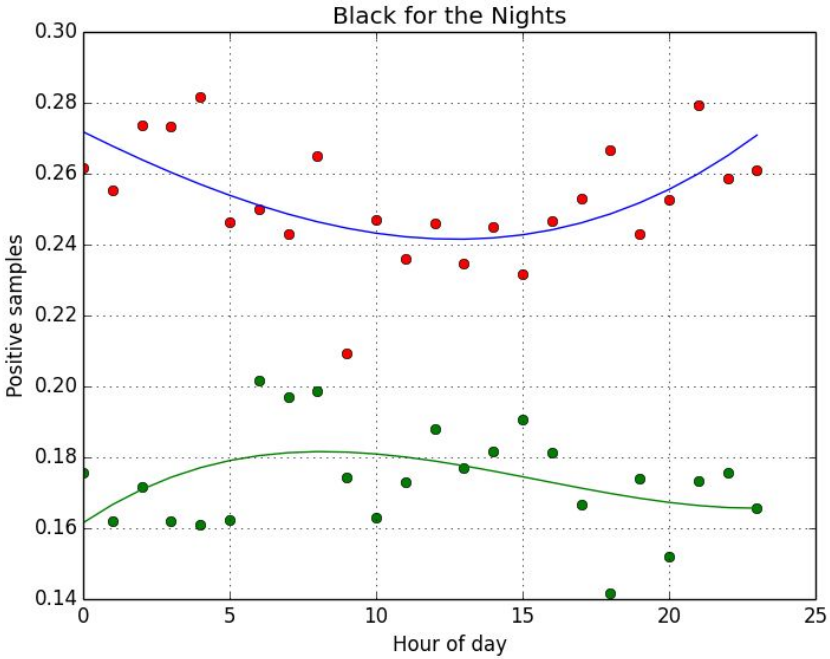




We normalize all results by the average upload frequency to observe true trends.

Daily Trends

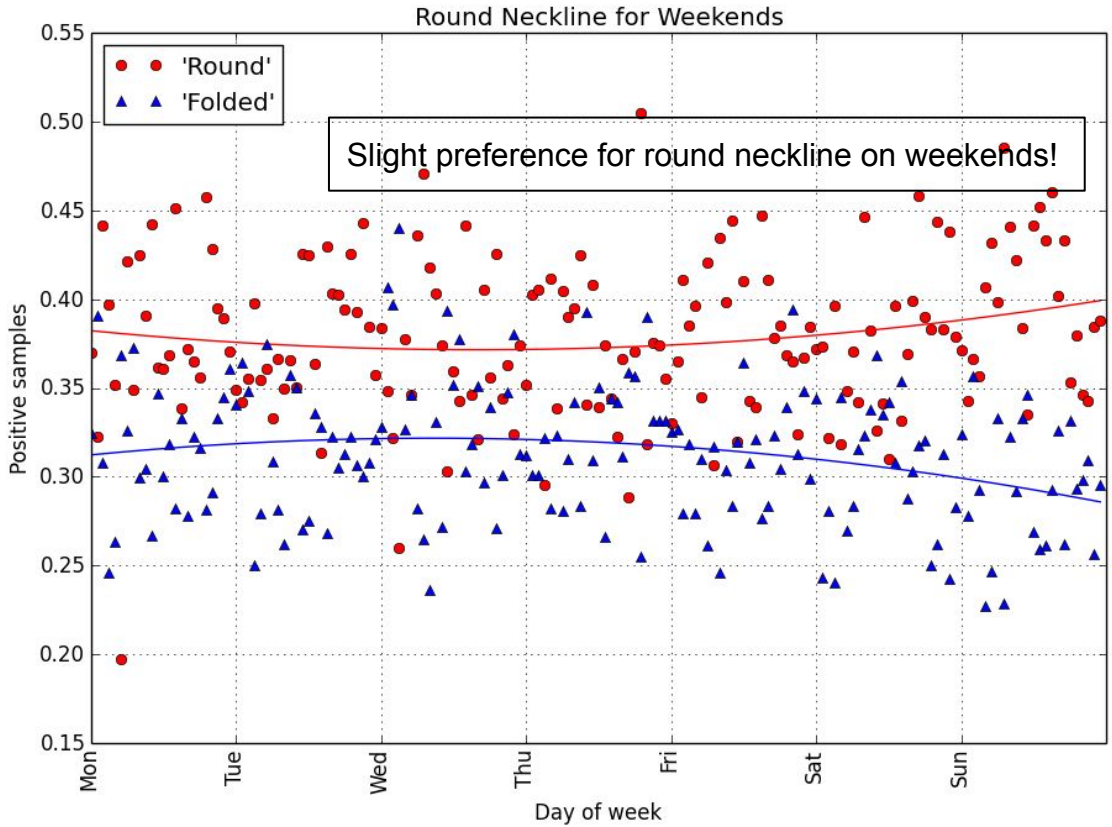




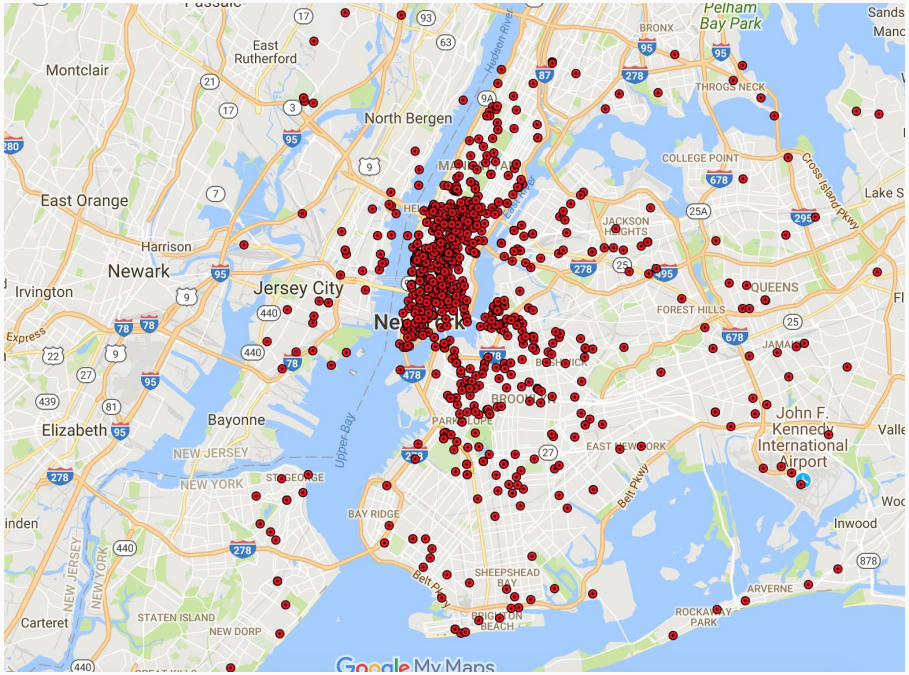
Higher fraction of glasses during day time might suggest about 3% presence of sunglasses!

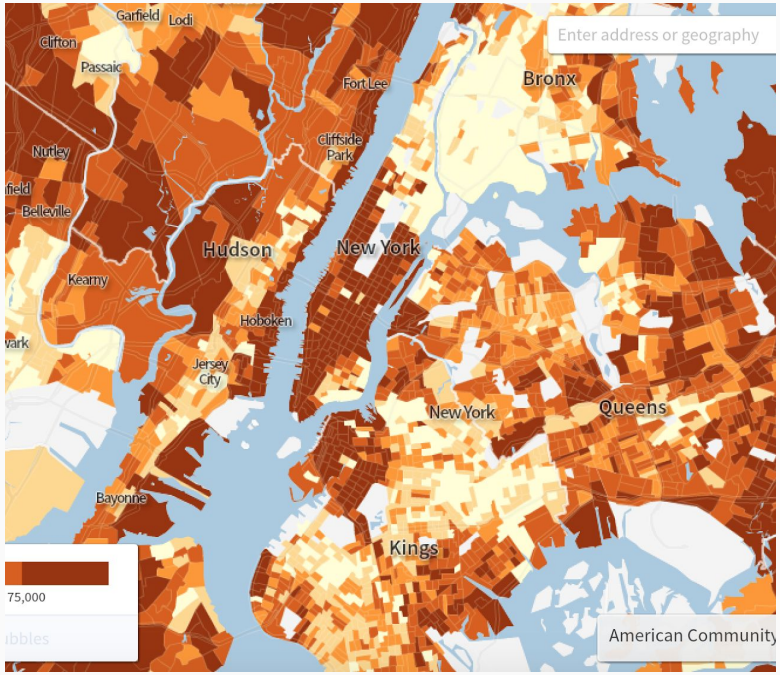
Black color appears to be more popular at night.

Weekly Trends



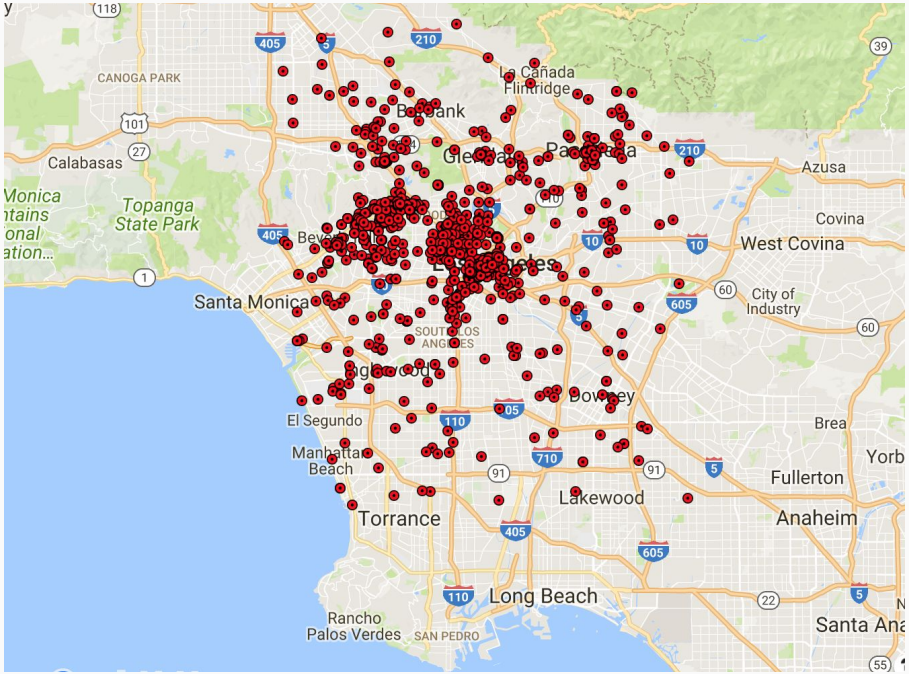
Instagram and Demographics (NYC)

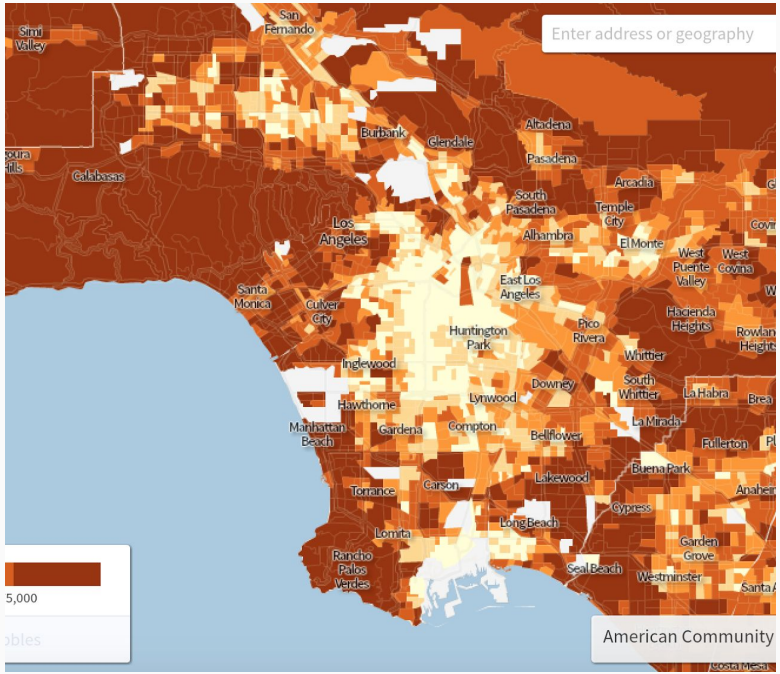




Instagram uploads, population density, and median income are all positively correlated.

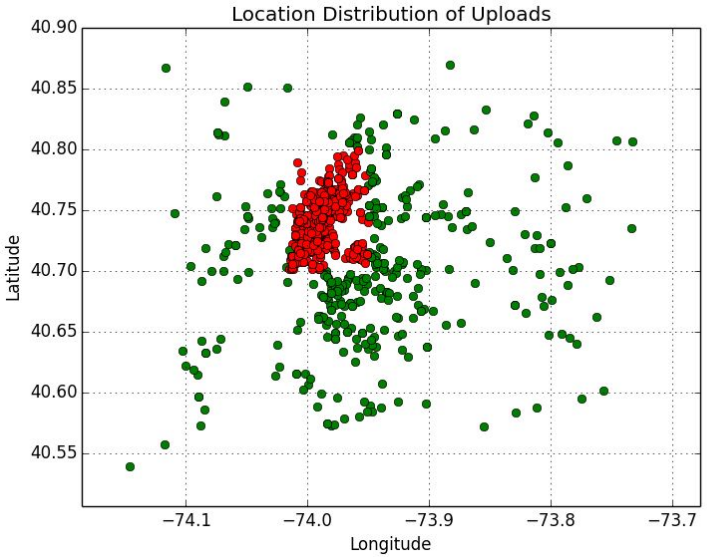
Instagram and Demographics (LA)

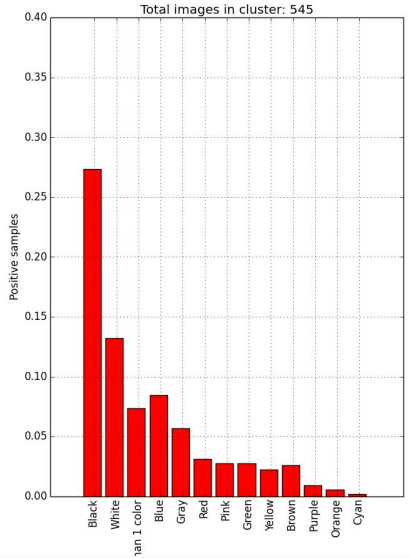


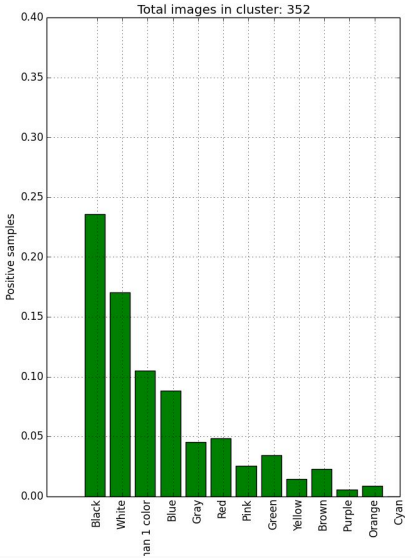


Instagram uploads and population density are positively correlated, but the median income is negatively correlated!

Intra-city Trends (NYC)

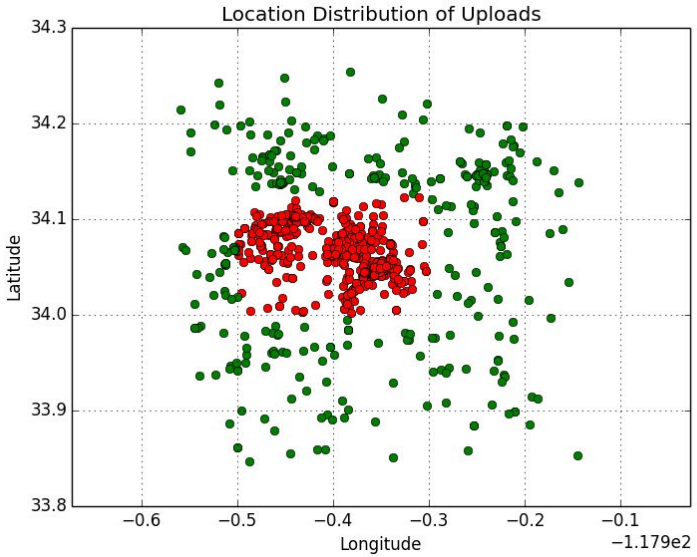


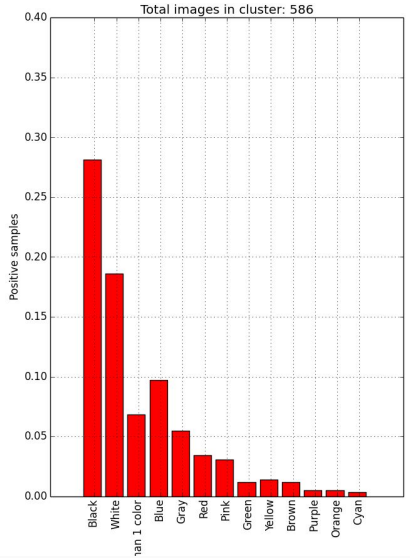


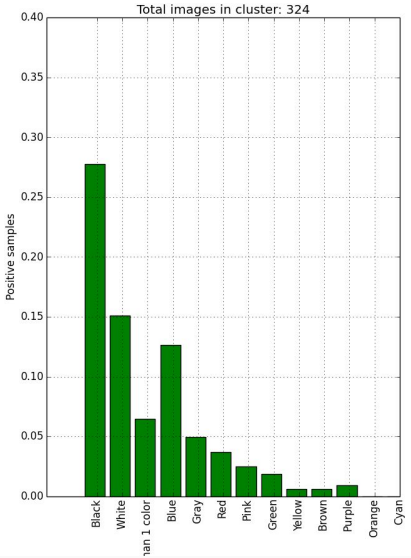


- Black dominates the Manhattan region
- White and multiple colors comparatively more trendy elsewhere

Intra-city Trends (LA)

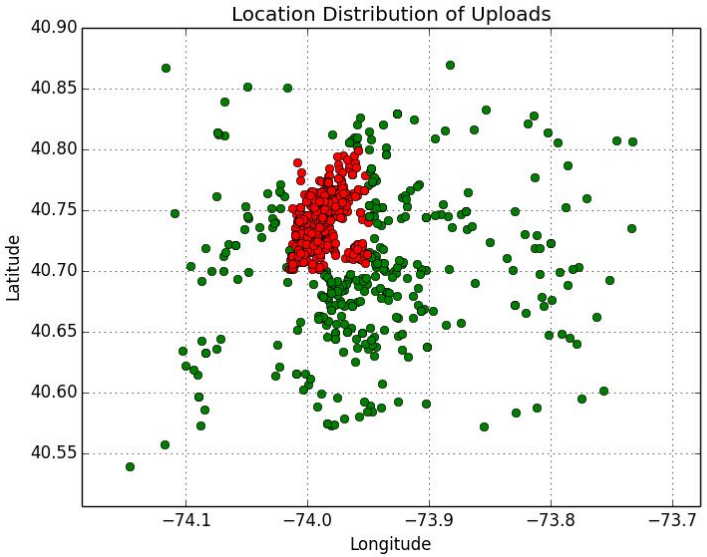


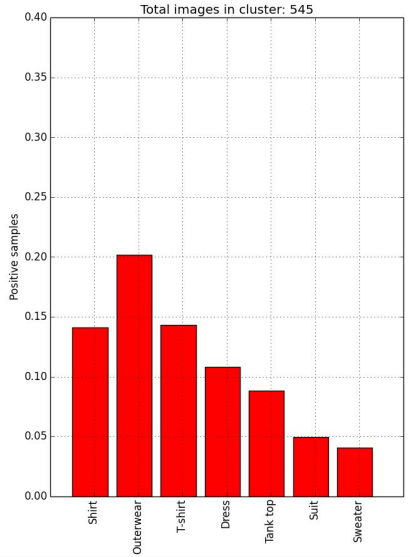


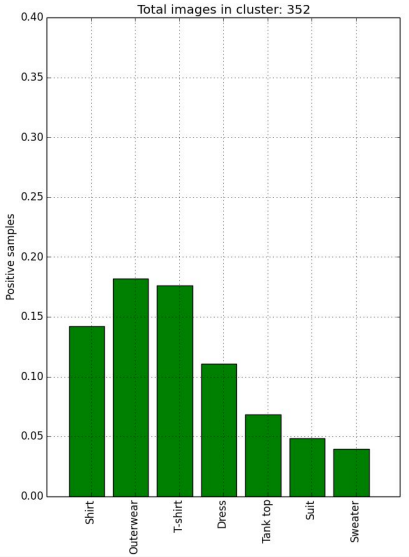


- Black and white predominant in downtown region
- Blue comparatively more popular elsewhere

Intra-city Trends (NYC)

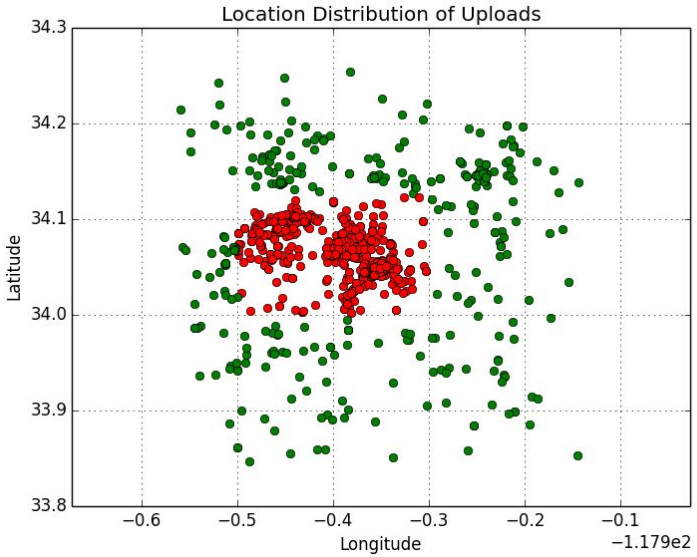


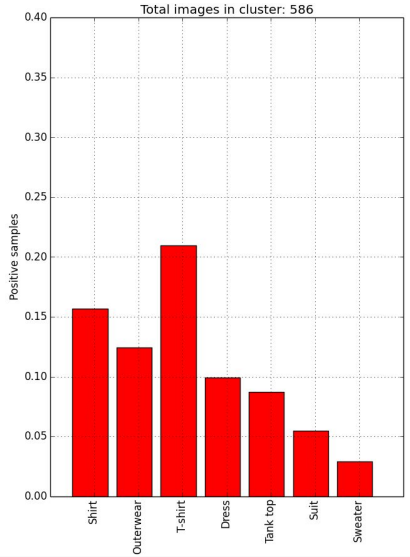


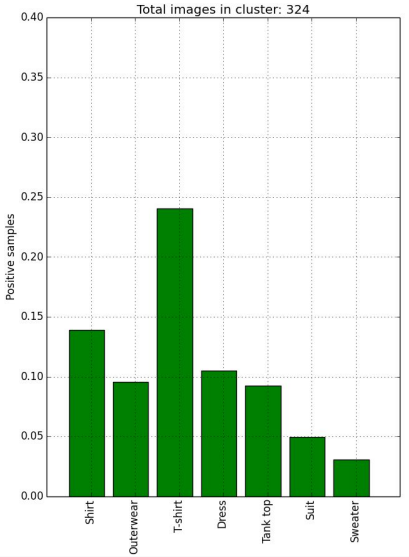


- T-shirts more popular elsewhere, as compared to Manhattan

Intra-city Trends (LA)







- T-shirts more popular elsewhere, as compared to downtown area

Thanks for listening!

Check out our code at:
https://github.com/angelaslin/street style-experiment