BUSTING BIAS IN THE FASHION INDUSTRY

BY

HEATHER HUYNH

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Adviser:

Professor Karrie Karahalios

ABSTRACT

Diversity in the media is a huge issue; lately there have been wins like more minority representation in films and TV, but one area that is known to be less diverse is the fashion industry. While many companies are trying to have more diversity, not just in race but in body type too, it's hard to tell who is doing a good job and how well they are doing it. We were interested in seeing if there was a way to easily compare these brands and their diversity through visualization which resulted in StyleRep.

StyleRep is an interface composed of four main parts: the "average", skin, shape, and clustering. These four sections each focus on a different way to represent diversity either through body shape or skin color. While not perfect, we were still able to see general patterns and trends in the fashion industry through our visualization. We hope StyleRep is the first of many steps towards trying to better represent the state of diversity within the fashion industry and beyond.

To my parents, for their love and support.
To James, for his support and proofreading.

ACKNOWLEDGMENTS

To Wendy and Tiffany for their help with this project and Silas for moral support.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Fashion is an industry that has had a long reputation of lacking diversity. In recent years, clothing companies have attempted to do better with diversity with actions like hiring Diversity and Inclusion (D&I) officers[1] to beginning to use plus size models in their advertisements[2]. However, it can be difficult to evaluate how effective these actions have been at actually improving diversity and inclusion in the fashion industry. One big indicator of the diversity of the fashion industry is the diversity of models and while a company can assure that they are trying to improve their diversity, without some sort of way to compare, it is hard to hold companies accountable for what they claim they will do.

For a long time, the media has presented an "idealized form of beauty" with thin, young, and Caucasian models in an attempt to encourage women to aspire to this ideal. It was thought that presenting an unattainable ideal could help companies sell more products as women tried to achieve this ideal.[3] This is the sort of behavior that fashion companies are now trying to fix through D&I initiatives due to public backlash. For example, there was consumer outrage at Abercrombie when their CEO stated he only wanted "All-American" people wearing his clothes.[3] A lot of this backlash also comes from the fact that consumers are becoming aware of the effects of media advertising. The fashion imagery that is shown in advertisements can affect peoples' behaviors and perceptions including their body image and consumption decisions.[3] Because of these effects, it is important for fashion companies to take increasing diversity, especially of their models, seriously.

While diversity often times refers to race of the models, it could also mean age or class diversity, not just in models used, but also the types of clothes that are sold at a particular store. Ben Barry found that variations in preference for the fashion industry's current beauty ideals are usually related to culture, age, and sexual orientation. For example, older women tended to favor inner beauty and health as opposed to physical beauty. [3] Since different brands and stores are targeted towards different audiences, this can have an impact on the sorts of models they use as well.

This work focuses on how to visualize photos of models from six different brands to see what trends occur in the models they are using. Our focus is on clothing worn by people who identify as female, so we only collected images from the women's section for each brand. To measure diversity of different brands, we focused on skin color and body shape for our visualizations. The goal of the visualizations is to present some ideas for how to represent or show this data as well as analyze some of the trends that might be seen when presenting the data in this way. Through these visualizations, we hope that consumers as well as brands are able to clearly see how diverse the models are or aren't.

CHAPTER 2: RELATED WORKS

2.1 COLOR TRENDS IN FASHION

Color trends in fashion have been studied extensively before. This color forecasting is part of a collective process known as fashion forecasting which tries to make predictions on future fashion trends. Through this, seasonal colors have been recognized as a powerful driving force of fashion-related or consumer products.[4] When looking at Turkish fashion, for example, Arik et al. found that in the summer months, people preferred light and pastel colors as opposed to darker hues in the winter. Additionally, some colors were found to be unrelated with seasons and there were differences between preferred colors of young people and older people for different seasons.[5] While this experiment looked at preferences of consumers, there has been little work done on what trends companies are putting out themselves.

2.2 IMPORTANCE OF DIVERSITY IN FASHION

The fashion modeling industry has long been criticized for using thin and exclusively anglo-looking models in advertising and runway shows.[6] This is especially true in fashion for those who identify as female. This lack of diversity can be very harmful long term both for society and consumers. In the US, people are exposed to an estimated 3000 advertising images per day and most young women read fashion magazines at least once a month.[3] These images have an influence on society's beauty standards which can unfortunately be idealized due to the disparity between models and the average person. Frequently, the National Organization for Women points out this large gap by claiming that the average weight of a model is 23% lower than that of an average woman which is large compared to an 8% differential 20 years prior.[6] Without some sort of intervention, this gap will only grow wider.

Why is an idealized standard so toxic? Fashion imagery influences the range of peoples' behaviors and perceptions, from consumption decisions to body image.[3] This means seeing certain images could affect ones own perceptions about their own body both positively and negatively. For example, researchers found that women with high levels of thin ideal internalization report positive feelings about their bodies, as well as the advertisements, after viewing average-sized models as opposed to thin models.[3] If most advertisements that people are viewing are of idealized or excessively thin models, then this could lead to widespread negative body image for many people which is bad for society overall.

Additionally, the point of creating this unattainable beauty standard for businesses is to sell more product. By ensuring that women cannot attain the ideal, the fashion industry intends for women to continually purchase products in an attempt to fulfill the ideal.[3] However, in recent times, this tactic has backfired as many companies have been called out in the media for a lack of diversity and promoting unhealthy body images. Research also has shown that consumers favorably evaluate models with appearances similar to their own. For example, older women preferred the attractiveness of mature models compared to younger models and viewers have a positive response when ethnicity matched between viewers and models.[3] To continue selling products, it might be beneficial for businesses to start shifting their model hiring practices towards a more diverse set of models and having some way to measure that diversity will be important.

While the fashion industry has been criticized over the years, it's clear that diversity isn't impossible to try and achieve. When looking at editorial fashion (like New York Fashion Week) versus commercial modeling, diversity in shape and color is more prevalent in commercial modeling because they are attempting to reach a certain buyer demographic.[6] Even with this diversity, though, it is clear they are not doing enough based on public perception. Images of models promote and disseminate ideas about how women should look, so businesses should be doing more to accomplish diversity in their models.[6] In order to get a holistic view of the industry overall, it will be important to sample images from a variety of clothing brands and see how they compare.

2.3 PRIOR VISUALIZATION WORKS

Fashion websites have lots of images that they show to consumers to entice them into buying clothes. To understand what images were important to use as well as good ways to visualize these images, I took inspiration from some prior visualization works.

2.3.1 Faces Engage Us

Faces Engage Us focused on what sorts of photos attracted the most engagement from users in the form of likes and comments. While Bakhshi's work dealt primarily with Instagram, the results show which photos attract the most user attention and can be extended to this work where we want to effectively visualize the photos that users will be interacting with the most on fashion websites. Bakhshi et al. found that photos with faces were 38% more likely to receive likes and 32% more likely to receive comments than other photos.[7] Since these are the photos users will likely engage with the most, this helps us figure out which photos are the most important to visualize since these have the most impact.

2.3.2 Zooming into an Instagram City

Hochman and Manovich were also focusing on Instagram and how Instagram photos for a particular location, Tel Aviv, differed over time. They focused on different aspects of the photos such as hue, upload time, and brightness.[8] The time aspect in particular for the visualizations was very stunning which is why I incorporated it into the average visualization for each brand. Having the time aspect for the average of the fashion images over time adds an extra dimension that users can explore to get a sense for changes in the sets of images gathered per day. Adding the time aspect was also similar to Jason Salavon's work, particularly his piece "Every Playboy Centerfold, The Decades (normalized)" (shown in Figure 2.1) [9]. In this piece, Salavon mean averages every Playboy centerfold foldout for four decades and displays them per decade. This gives the viewer a sense of what the photos overall look like without overwhelming them with the actual details. The prominence of skin color appears in Salavon's work which served as inspiration for applying a similar technique to fashion images in order to gain a sense of what sorts of models a brand is using. This work's average images are similar to Salavon's work, but groups each average of images by brand over time.

Figure 2.1: Every Playboy Centerfold, The Decades (normalized)

2.3.3 Clustering of Fashion Images

Clustering of fashion items is an area that has been explored before, especially by researchers in Computer Vision. One project that dealt with clustering fashion images and then visualizing the clusters is Streetstyle. Their goal was to use social media images to see fashion and style trends world wide.[10] In order to cluster well, the dataset they used was labeled with various style and fashion terms and they processed their images using a neural network. This is probably necessary because images from social media aren't standardized, so they need to undergo processing to be similar. They were able to cluster together similar styles in terms of clothing items worn and the visualizations looked really impressive.[10] This work took some inspiration from Streetstyle to do more basic clustering on images to see what sorts of results would appear. Because this clustering method is more simplistic, we had to manually separate out some of our images since there are no labels associated with the images. For example, any images that are clothing items, which Streetstyle does not have to worry about, was put in a separate category for clustering.

CHAPTER 3: DATASET AND COLLECTION

Originally, we had proposed for our project to explore the social media, mainly Instagram, of different brands. By using Instagram data we could 1) guarantee all the images would be the same size even across brands and 2) also find influencer images to get an idea of the diversity of influencers brands use to promote their products. Unfortunately, Instagram has strict rules about scraping their data, even through an API and using Instagram images would make it difficult to ensure the images would be in a format we could use for our processing. For example, a lot of the Instagram images we initially saw were either advertisement images for sales (which were all text) or had more than one person present in the image (which would make the body outline calculations difficult).

After discussing the idea, we considered using images from the websites for the brands. These images were more regular with there only being one person in the image (usually) and the images being similar except for the clothing worn/pose. However, to use these images we would need to gather them. Since the focus of this project was on fashion worn by people who identify as women, we only collected from the women's section of each of the six brands that we analyzed. The overall dataset contained 6661 unique images which ranged from full-body shots of models to photos of the clothing items themselves.

3.1 DATA COLLECTOR

This project used a Python-based web scraper for collecting images. Because it is Python-based, it is unable to handle websites where data is entirely loaded dynamically which limited which websites we could pull from. Additionally, it can only pull what is loaded on the page during the initial load, so if the website has an infinite scrolling feature to load more images/products (a common feature on many fashion websites), the collector was unable to get those images. As a result, there is a very uneven distribution of images from each website since some, like Fashion Nova, load a lot at once while others, like Topshop, do not. The statistics for each brand can be seen below in the Brand Selection section.

In a prior iteration of this project, we grabbed images from different clothing sections of different websites. To make this more standardized, we took from the main women's clothing page (like the "All" category) and to be more encompassing, if a brand had a "plus" category that was separate from the all women category, we also scraped the images from there. If a website provided many images of the clothing item or model on the same page (i.e. a rollover effect), we also grabbed those images which resulted in multiple images per clothing item per brand. The scraper was run once a day between February 19, 2020 to April 16, 2020 for each website in order to collect both image data and temporal data for each brand.

3.2 BRAND SELECTION

The six brands were chosen based on the type of clothing company they are and the ability to collect from those sites. The selected brands were Fashion Nova, Gucci, H&M, Ralph Lauren, Reformation, and Topshop. Three of these companies (Fashion Nova, H&M, and Topshop) have been known as fast fashion companies which cycle through new clothing more quickly than other clothing brands and tend to have cheaper prices. On the other end of the spectrum, Gucci and Ralph Lauren are more high-end or luxury brands and are expected to have less turnover in their clothing within the period we would be collecting data. This variety of types of brands was important since there could be trends seen in some brands, but not in others.

Another difference between these brands, besides their classification, is the amount that was able to be scraped at once from each brand. Some websites, like Fashion Nova, allowed our scraper to grab hundreds of images at once which means this page was loading hundreds of images initially for the user to look at. Other sites, like H&M and Topshop, would only load a handful of images at a time and lazy load the rest if the user scrolled. Since there is such a big discrepancy between the amount of images collected from each site, we calculated the average number of unique images gathered per day from each website as well as the average number of images overall gathered per day from each website. These results can be found in Table 3.1.

Table 3.1: Dataset Image Collection Statistics

Brand	Average Unique per Day	Average per Day
Fashion Nova	97.83	199.59
Gucci	5.93	136.31
H&M	8.16	35.98
Reformation	7.18	39.38
Ralph Lauren	3.61	35.43

Demographic information for each brand which could potentially be related to its labels (fast fashion, luxury, etc) would be desirable, but demographic information for brands can be difficult to get so the selected fashion brands are based on an initial classification, but within the visualization itself, the classification of the brands is left to the viewers. To help with classification, we wanted to provide some additional sliders for users to use that can help them determine their own labels for brands (for example, the cost of a white t-shirt from each of the brands could be related to whether a brand is more high-end or not), but due to time constraints were unable to add this to the visualizations themselves.

3.3 DATASET PROCESSING

As stated above, we had over 6000 images that were collected initially over the collection time period. However, it can be difficult to create visualizations and see trends in these images without some processing. Since we collected images day by day, there are bound to be duplicates between different days. We had one script that could process the dated image folders and pull out the unique photos for each brand using the MD5 hash for each image. While this method worked to get out most duplicates, there were still some cases where a duplicate could get missed due to the picture being slightly different in some way. To control for this, we also had a manual pass that occurred after separating the unique images since we needed to manually look through images to separate them into different categories.

Another issue with the initial dataset was that it included images depicting models and not models. For example, Gucci had images of all their items as well as their models wearing the items. However, when processing a visualization on skin color or body outline, we would not care much for the images showing handbags or the clothing without a model because we can't get that information from the picture. Thus, the dataset needed to be separated into different categories that we could use. The first classification I created was pictures that included faces of people (which are useful for skin color detection), pictures where people were shown, but their face was not clearly seen (often for pictures showing off pants), and pictures of clothes items. The big focus on faces was also influenced by the fact that users are likely to pay more attention to these photos not only because people are more attracted to photos of faces, but also because they probably want an idea of how the clothing will look on themselves before buying. Examples of images in each of these classes can be seen in Figure 3.1.

Since the pictures containing faces category has multiple types of images, I further separated out full-body images for each brand to help with finding the average outline for each brand. Comparing the outlines between full-body shots is a lot easier and clearer than comparing the outlines between full-body and varying half-body shots. This also reduced the amount of photos I would need to process for outlines to around half the original amount (2707 images). All of this classification was done by hand since a robust classifier would require manual validation and would have taken a longer time to create.

(a) People with faces (b) People without faces (c) Clothing Items

Figure 3.1: Sample Image Classification: These are two randomly selected images for each class that we initially created for the dataset. The full-body classification which was created later is a subset of images from (a).

After separating out the different categories manually, the data was further processed to get information about average skin color for different brands as well as the average outline for the brand. In a prior iteration of this project, we looked at the outlines for only 10 handpicked full-body images for each brand. This was faster to process and resulted in an average outline that we could see better. For this project, I tried both the 10 image outline (by selected the first 10 images in the full-body class for each brand) and the average outline of the full set of images for each brand.

For average skin color, all the images in the people with faces class were processed. This is because clothing only images resulted in “NaN” when being processed and if there was not enough of a person showing in the image, then a similar result would occur. While this removes photos from the dataset, it made the process of generating and averaging skin color better since NaN photos would not contribute meaningful information to the average. Since most websites have many images of the same model but with different views (full-body, upper body shot, etc), the removal of these images should not affect the average negatively since these models are still included in the average.

The technical information about these last two processing steps will be more extensively covered in the next section.

CHAPTER 4: VISUALIZATION

This project focuses on body shape and skin color as markers for diversity. Thus, all of the visualizations try to show the trends in these two attributes for the photos of each brand. Since this visualization overall is meant to be exploratory for users, the visualizations try to offer ways for users to view data between different brands as well as across the same brand. This section will cover each of the different visualizations created for this work and give examples as well as talk about some of the technical details for some of the visualizations. Extra technical details can be found in Appendix B.

This project originally had two iterations. The first was completed last year as a class project and this second iteration improves upon the first iteration. All of the visualizations mentioned were created for the second iteration, but both iterations will be referenced to help motivate certain design choices in each visualization.

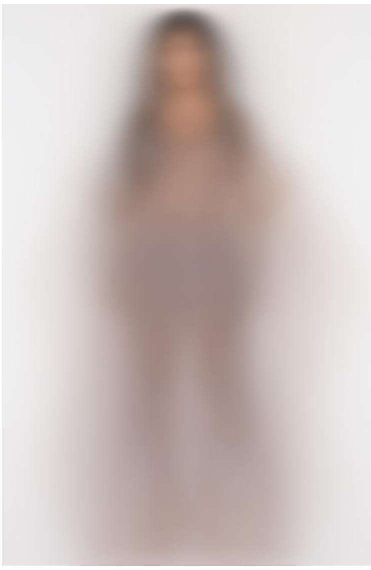
4.1 THE "AVERAGE"

Initially, one average was generated per brand in a prior iteration of this project. While some patterns could be seen across brands using the average image, especially for some of the brands, there were some issues with the dataset being used like different types of images taken from each site as well as images taken from each site at different times. Since this is a nice visualization to look at, I extended the original idea to include not just an average of the image, but an average over time. The comparisons between different brands might also be more effective for this set of averages because the images that were collected were standardized. Two visualizations were created using the average: the overall average using only the unique photos that were collected for each brand and the average per day of the images collected for one brand.

In order to get the average image, all images were averaged by pixel value. This required all images to be the same size. Within one brand, all images were usually the same size, but if some were not, the size that had the largest amount of images was used for the average image and all images not of that size were resized.

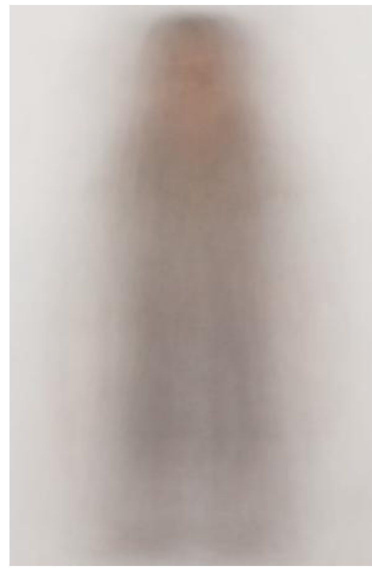
4.1.1 Average of Unique Images per Brand

As mentioned in the dataset section, I created a script that separated out the unique images from all the images that were collected per day from each brand. This is because each brand has at least some overlap between days of the images that are showing, so it adds a lot of extra processing time and also dilutes the average. An example of this visualization can be seen in Figure 4.1. Each visualization shows the brand and the number of images that were averaged for that brand. Since each brand is represented by one average image, viewers are able to easily compare between different brands.

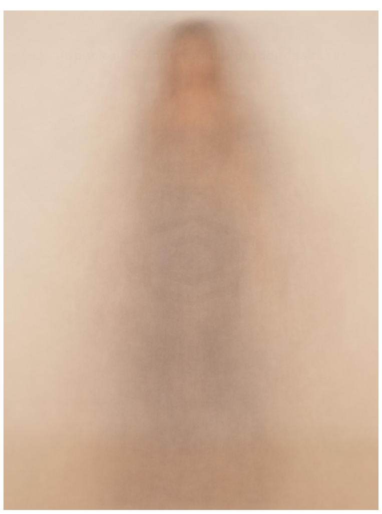


(a) Fashion Nova (5092 images)

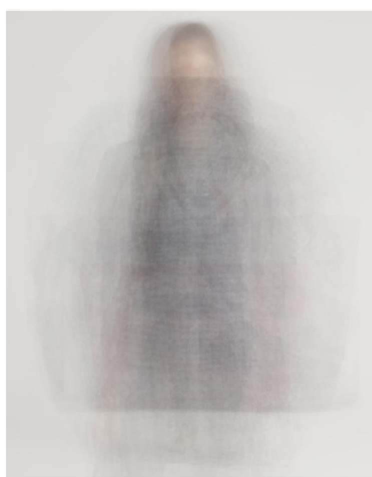
(b) Gucci (321 images)



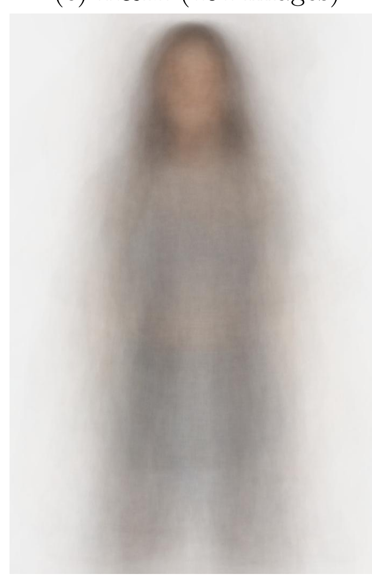
(c) H&M (404 images)



(d) Reformation (375 images)



(e) Ralph Lauren (196 images)



(f) Topshop (279 images)

Figure 4.1: Average of unique images for each brand. The brand and number of images in the overlay are shown below the image

4.1.2 Average of Images per Day per Brand

For this visualization, the images are generated the same way as the average of unique items per brand, but separated into averages of the set of images gathered per day for a brand. This visualization is a little harder to use for comparing between brands because each brand is represented by multiple average images and the trend that is visible from the averages. One example of this visualization is shown in Figure A.4. The images are arranged by time from left to right and up to down. The rest of the images can be seen in the Appendix A.

Figure 4.2: Average of Images per Day for Reformation

4.2 SKIN COLOR

The next few visualizations are focused on the skin color of the models that are being used in the images for each brand. For this part of the visualization, the goal was to create an easy way to see patterns in skin tone across different brands. While the average offers a glimpse of this, it could be hard to tell overall where each brand falls since some of the averages don't make it easy to see skin tone. We wanted to see if there was a more quantitative way to show skin tone which resulted in one of the skin tone graphs which can be seen in Figure 4.3.

To create these skin color visualizations, the skin color for each image needed to be calculated. We created a Python script that found the skin color for each image and averaged them together for each brand. This average skin color was defined as the average color of all skin pixels in all images of a brand. To find skin pixels, we used thresholding to isolate the skin. These explicitly defined regions were in both the HSV and YCbCr color space with the thresholds shown below [11]:

$$0 <= H <= 17 \text{ and } 15 <= S <= 170 \text{ and } 0 <= V <= 255$$

$$and$$

$$0 <= Y <= 255 \text{ and } 135 <= Cr <= 180 \text{ and } 85 <= Cb <= 135$$

Both the average skin color of the images and the overall average skin color of all images are stored in JSON and output from the script. For average skin color, we save both the HSV and the RGB values for each image and for the overall average of each brand. We also include the RGB and LAB values for the six reference colors that are used. These six reference colors were selected to represent a wide spectrum of skin colors and each overall average for each brand is compared to these reference colors to help with figuring out what the average color looks like.

4.2.1 Average Skin Color per Brand

To display the average skin color per brand that was calculated, we opted to use a scatterplot since it would be easy to see what colors the brands' averages were clustering around. On this graph, the x-axis represents the average skin tone of a brand based on its distance from one of six reference colors. The y-axis represents the range of skin tones found for a particular brand. The range was determined by taking the darkest skin tone found for a brand and the lightest skin tone for a brand and getting the difference between their values. The value comes from the third part of the HSV representation for a color. While skin tones do not scale just along the V axis, we found in experimenting that the V value tended to scale with darker skin tones, so we felt this was the best approximation of range for our graph.

Each dot on the graph represents a brand. When rolling over a dot, the brand's name appears above it. We have two versions of this graph. The first, as seen in Figure 4.3, displays range indicators with the dots to show the skin color range visually, which differs from our first iteration of this project. Since there are only six brands, there is a lot more space to show the range indicators so it should not add too much clutter. This differs from our first iteration where we removed the range indicators to improve clarity. Each dot is colored the color it represents based on its distance from one of the six reference skin colors. This entire graph was created using d3 which made it easy to update when data changed.

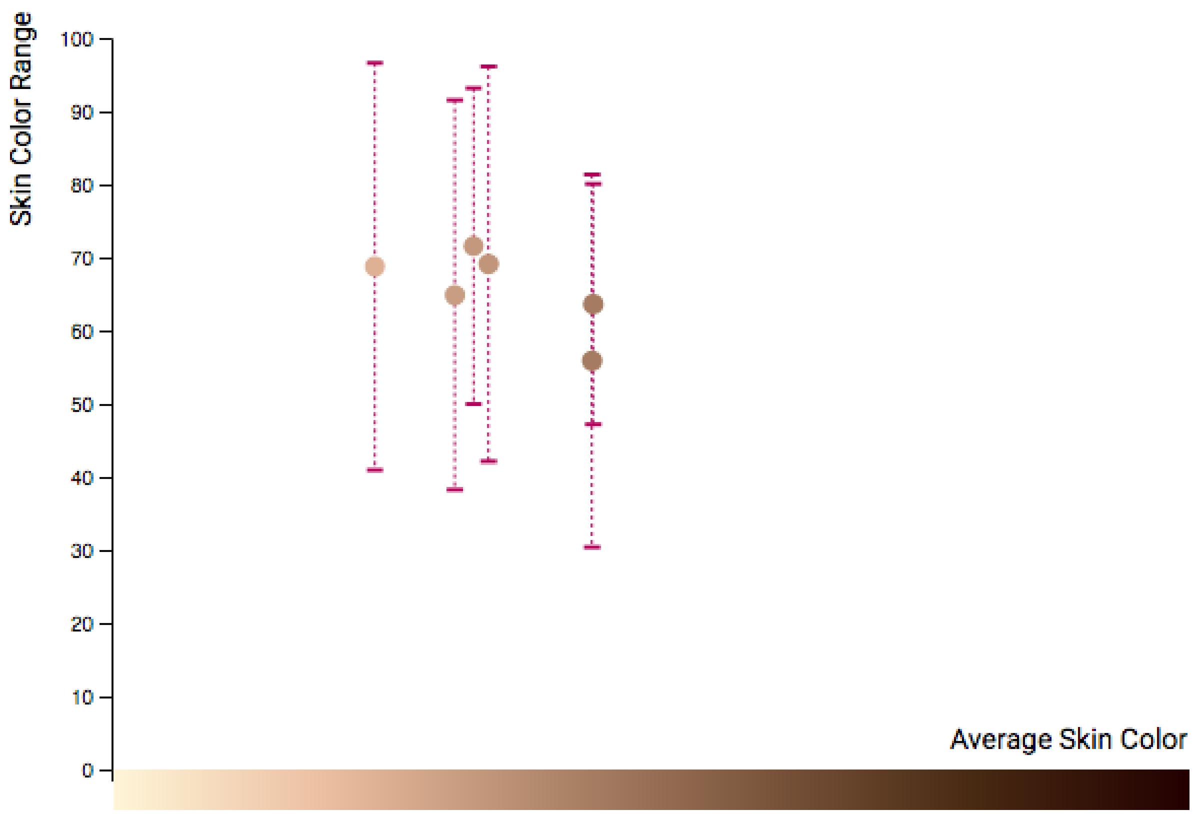


Figure 4.3: Average Skin Color per Brand with Range Indicators

The second version of this graph shows the variance of the skin colors for each brand as the size of the circle. Using the size of the circle creates a cleaner scatterplot when compared to the range indicators which is why we tried this method as well. The variance was calculated using the HSV values and the math.js variance function. This visualization can be seen in Figure 4.4.

This graph initially had a filter that was applied using a dropdown menu, but with the new data we removed the dimension of world region because it would require collecting from multiple brands from each region. Even with just six brands, the collection process was

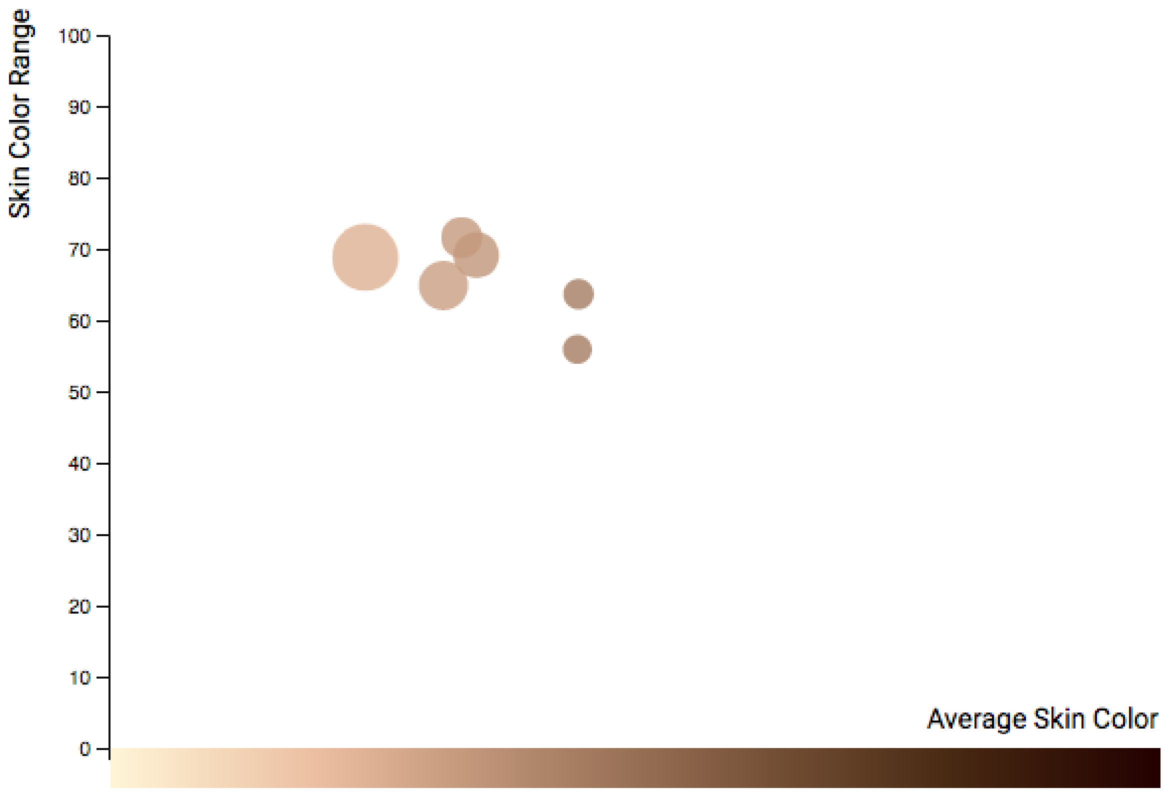


Figure 4.4: Average Skin Color per Brand with Variance as Dot Size

difficult, especially when we wanted to keep the process as standard as possible so it would be easier to compare the different brands to each other. Thus, we removed the region filter for this iteration because most of the brands we sampled from were international brands and would be classified together in one region.

Further implementation details are included in Appendix B.

4.2.2 Skin Color Distribution per Brand

When looking at the graph of the average skin colors for each brand, a natural question to ask is: "Why is the average that way?". For example, the average graph seems to imply that Fashion Nova has a fairly middling skin color average with a wide range. Showing the distribution of what average skin color each image has helps with getting an idea of why the average is so low.

Since the skin color data takes a long time to generate, I had to figure out a way to display skin color distribution using the data provided in my returned JSON. Luckily, the average skin color in both HSV and RGB for each image is provided in the JSON, so I was able to leverage this data for the skin color distribution chart. With the data I had, the best way to show the distribution was to use the six reference colors and basically classify each image for each brand into one of these six buckets. If a brand has many images classified in a lighter bucket, it would make sense that their average would be low as well. An example of this visualization can be seen in Figure 4.5 which shows Reformation, a brand with a low average skin color.

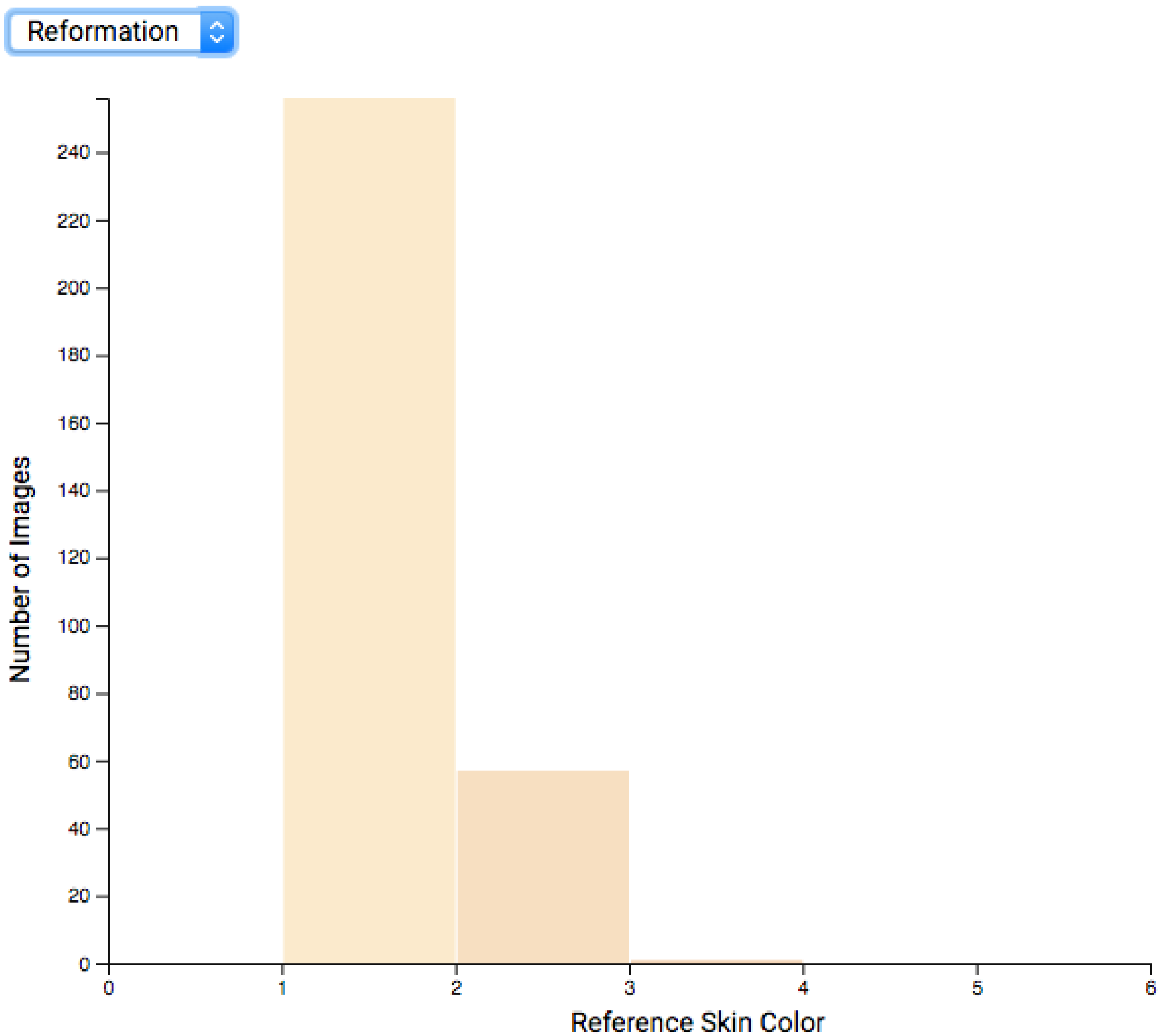


Figure 4.5: Skin Color Distribution for Reformation

We display these buckets using a histogram because it is easy to see at a glance which buckets have the most images. It is also easy to see a normal distribution if one exists which can be seen in Figure 4.6 for Fashion Nova. The user is able to switch between each brand to see what its distribution looks like, but since each brand has a different number of total images in the full-body dataset, we adjust the Y-axis as they change through different brands.

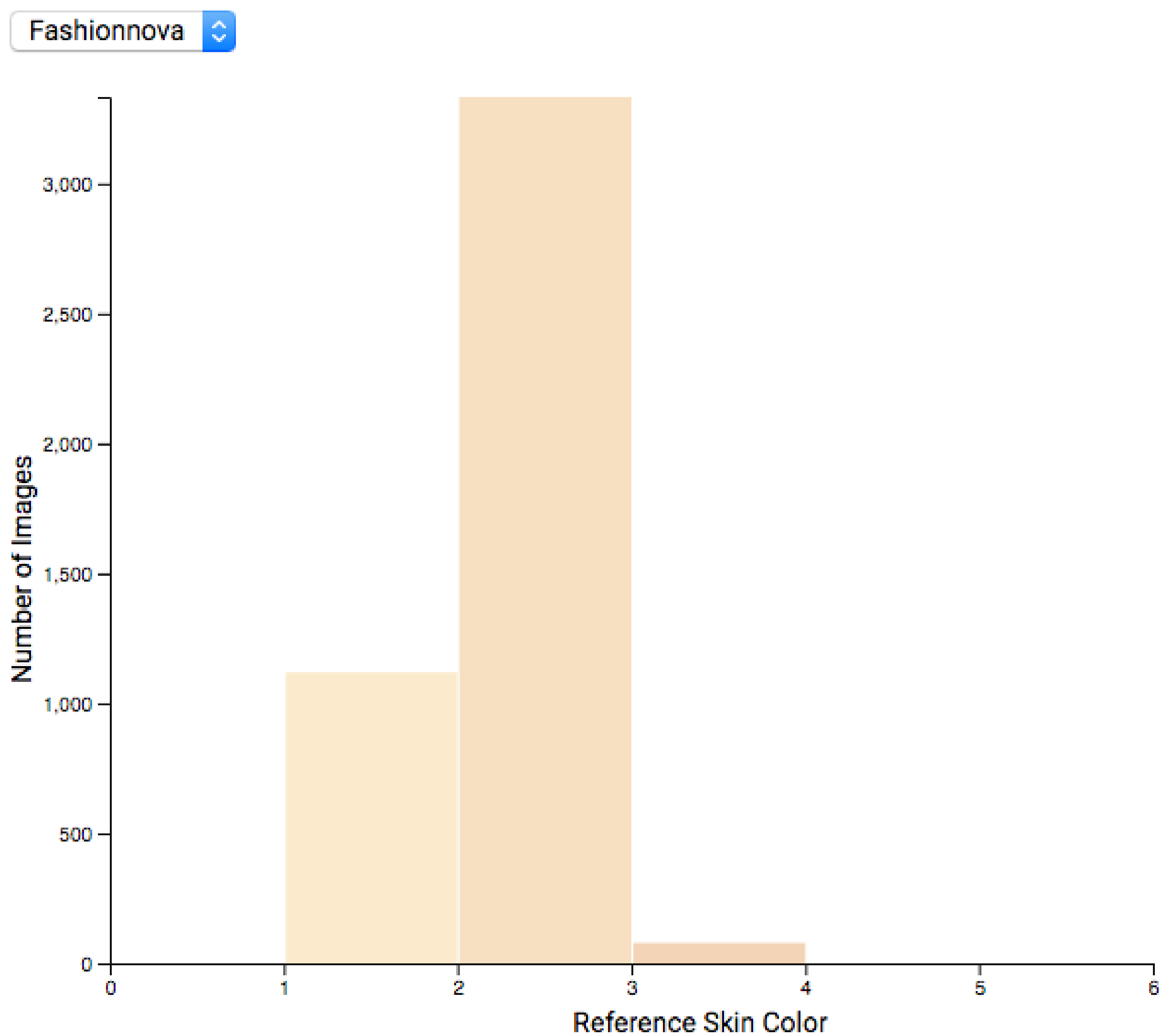


Figure 4.6: Skin Color Distribution for Fashion Nova

The JSON data only provides the average HSV and RGB values, so we utilized the math.js library to help perform some calculations on this data to generate the buckets. To sort the colors into the right bucket, we took the average of each image and found its euclidean distance to each of the reference colors in HSV space. Then we categorized each color by giving it the label of the reference color it was closest to using the minimum of these distances. All these labels were used together with d3's binning function to create a histogram.

Along the X-axis we have the different buckets (each of the six reference colors) while the

Y-axis shows the number of pictures within that bucket for that brand. Each bar is colored the color of the reference color for that bucket to help with visualizing whether a brand has photos with models that are mostly classified as lighter colored or darker colored skin colored models.

4.3 BODY SHAPE

As mentioned before, our measure of diversity was focused on both skin color and body shape for each brand. While we were able to find an average skin color, it is much harder to try and define an average body shape when images might have models in different poses, and we have no information about the actual camera being used. Because we have no camera information, it is difficult to create measurements for different parts of the body. Originally, we wanted to find some sort of average ratio of particular body regions (hips, bust, etc) and use that as our measurement of body shape, but it was difficult to figure out how to calculate this consistently even when we had a way to segment the different regions. Instead, we found a way to approximate the average visually by overlaying the outline of full-body model images. Since we would be overlaying the images for models within the same brand, knowing the camera parameters are not as important because one brand will probably shoot photos in a similar fashion.

The images used for the body shape visualization are a subset of the images used for the skin color visualization. For skin color, we were able to use all images that we had classified in the humans with faces category. A smaller category for full-body images was created from these images since full-body images showed the body shape of the models the best. Even though images with the model sitting down are technically full-body images, we tried to remove any image where the model was not standing or mostly standing from the full-body dataset because overlaying an outline of someone sitting would only add noise to the outline overlay.

To generate these outlines, we first tried applying regular edge detection methods; however, this would often also get edges within the model which was not useful for our purposes. Thus, using edge detection did not work for finding just the outermost outline. Instead, we then tried using background subtraction for all images. Most of the images in the full body dataset had a monotone background. Within one brand, though, different full body pictures might have different background (i.e. a blue background versus a white background). This meant we couldn't use more advanced techniques that were usually used for static backgrounds in videos, like the Gaussian Mixture Model. Our simpler algorithm identified background pixels by comparing their color to either the color of the corners of the image or the regional background color average of an image. Using this algorithm had a higher accuracy than using an advanced technique due to the nature of the images in our dataset. Unfortunately, this method was not perfect; while the accuracy was better, sometimes clothes that were similar in color to the background of the image would be mistakenly classified as background pixels which affected some outlines (as seen in Figure 4.7).

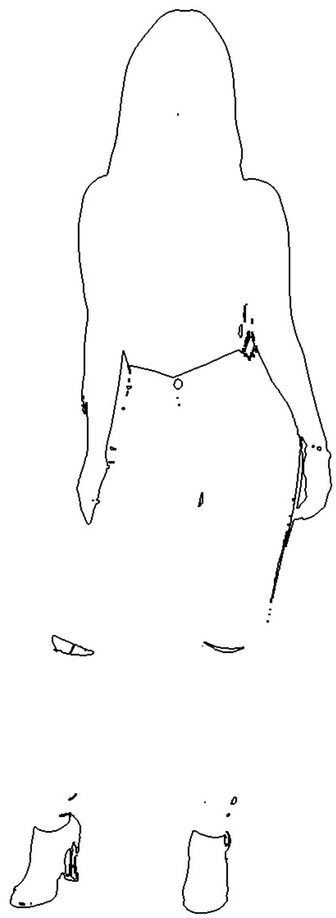


Figure 4.7: An outline affected by mistakenly classifying background pixels

Each pixel is tagged using this algorithm as a background or non-background pixel. After these pixels were classified, we applied an edge detection algorithm. We tried using some common edge detection algorithms, but ended up having to create one for our purposes. In our edge detection algorithm, we iterated through each pixel in a 3x3 window. If the proportion of background pixels in a window was between 0.2 and 0.8, we considered the central part of the 3x3 window as part of the outline. Using this algorithm, we colored the pixels we identified as the outline in black and the non-outline pixels in white.

With these outlines, we created two types of body shape overlays. The first uses all images within our full-body dataset for each brand while the second selects only the first ten images from each brand to create the overlay.

4.3.1 All Images

Because we had already created a full-body dataset, it made sense to try the outline generation algorithm on the full set and see what results we got. The generated outlines we created can be seen in Figure 4.10. For these outlines, we can see there is an opacity applied to them, but even with this opacity, certain regions get really dark. In particular, the Fashion Nova outline is very dark and messy which makes sense since it is composed of 2336 separate images compared to Gucci who only had 11 images in the full-body dataset. When looking at the dataset statistics, it makes sense that a brand we collected 100 unique images from per day resulted in such a large amount of full-body images.

In Figure 4.10, we see that each image has the corresponding brand underneath it with the number of images that were used for that brand's outline. Some of the component outlines, as seen in Figure 4.8, have some artifacts and are especially noticeable on the Reformation, Fashion Nova, and Ralph Lauren images. However, there were also many outlines that were actually very clean for each brand (Figure 4.9).

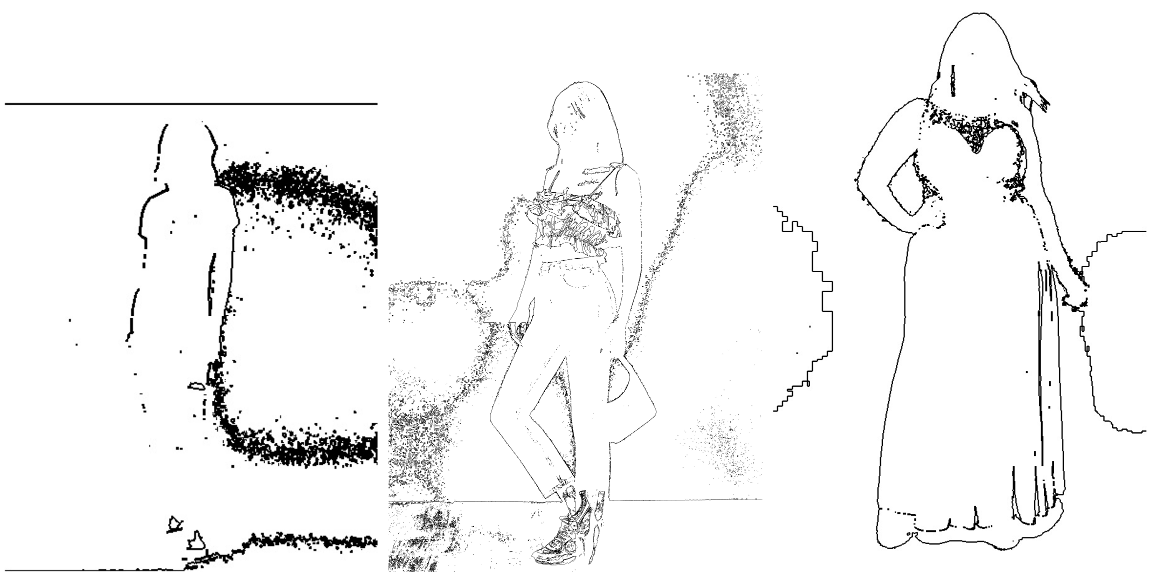


Figure 4.8: Messy Component Outlines

Due to the fact that each brand has a different number of images and that Fashion Nova has over 2000 images used for its outlines, it didn't make sense to also display all the component outlines with these total outlines like in the first iteration of the project. The difficulty of consistently seeing a clear average outline from these images inspired the idea

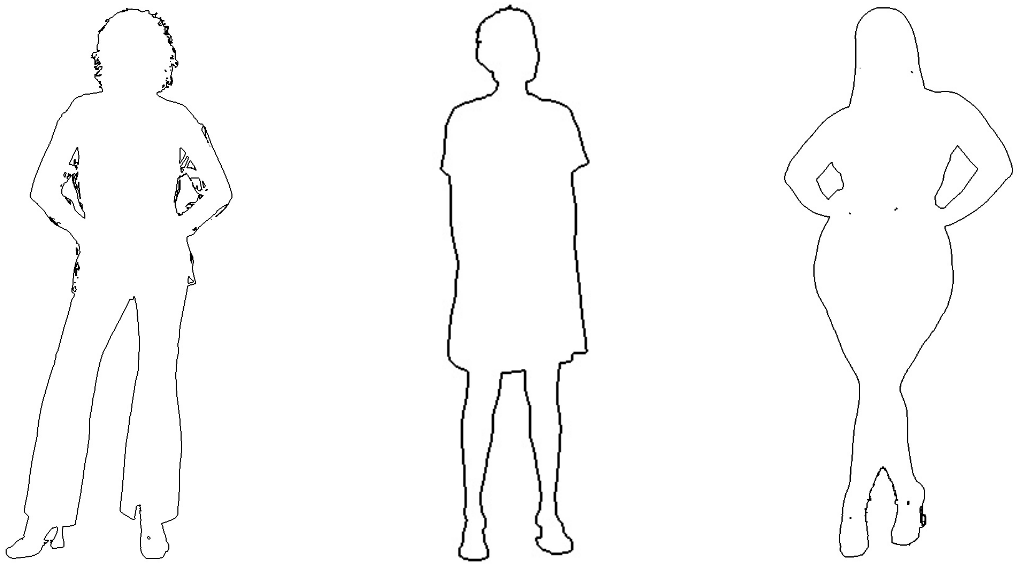


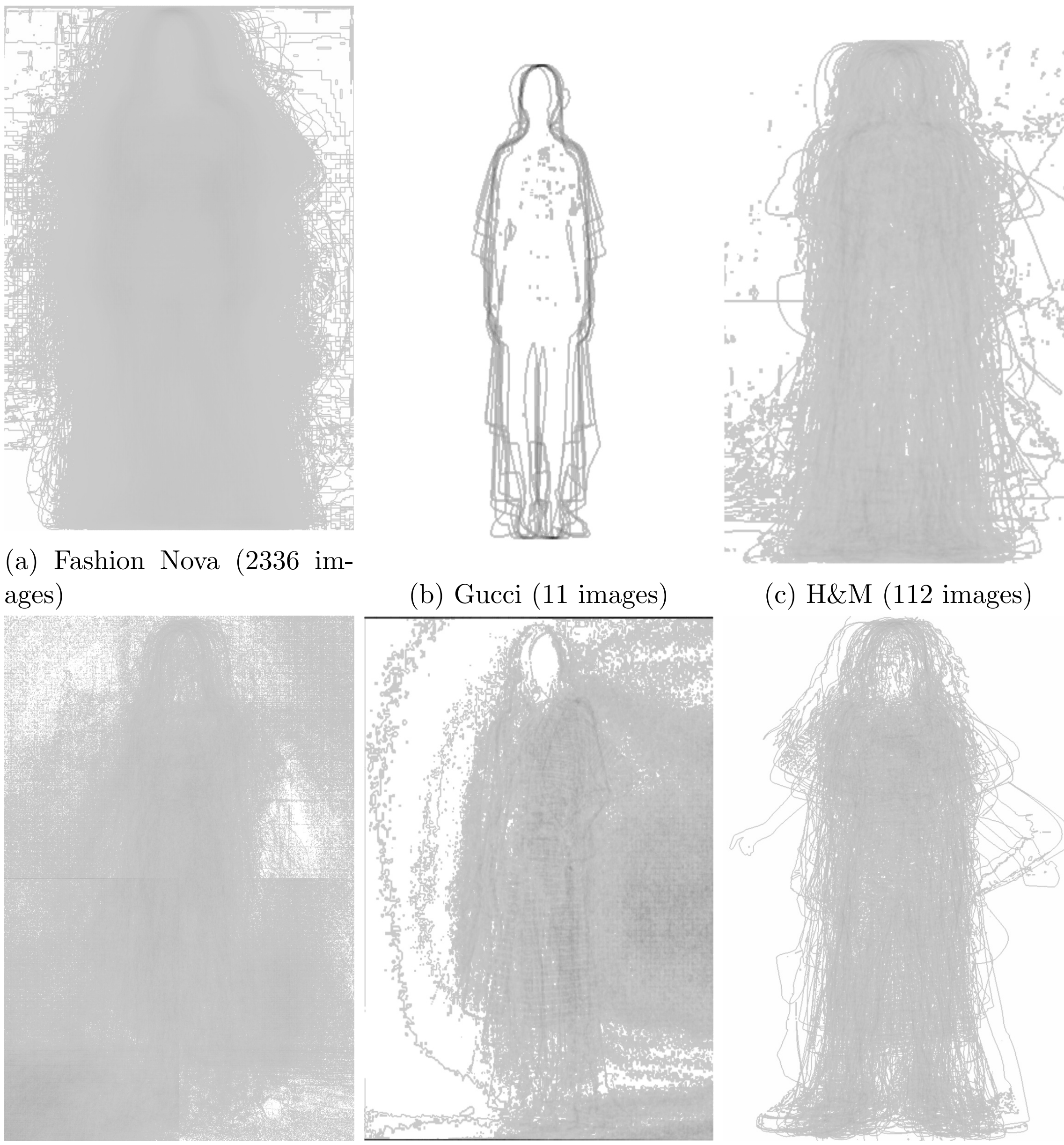
Figure 4.9: Clean Component Outlines

to take a smaller subset to create an outline overlay since each of the brands have some set of clean outline images.

4.3.2 Ten Image Subset

Looking at the overlays for all the full-body image outlines, the average body shape is a bit difficult to see. One of the big issues with generating these outlines was clarity. Displaying the individual outlines helps, but we also found that reducing the amount of outlines we laid on top of each other helped us better communicate what we wanted to show. Viewers could get a good idea of what a brand's average body shape looked like while not being overwhelmed by too many lines on the outline image. We arbitrarily selected ten outlines as our limit because when we tried stacking more than ten originally, the outline got too messy. There is a trade-off here since using more outlines could contribute to a better idea of the actual average body shape of a brand, but also be too cluttered to actually get any information. We also experimented with darker and lighter outlines before we settled on an in-between gray. A lighter outline was too hard to see while a darker outline, as shown in Figure 4.11, made it difficult to distinguish between outlines.

One interesting thing to note is that Gucci's outline for both all images and the ten image subset is the same since they only had ten full-body images in the subset. The nice thing is that the ten image generation takes a significantly less amount of computing power and time than the full dataset, so it would be possible to potentially try different combinations of ten images to see different outline overlays. The ten image outlines we created can be



(a) Fashion Nova (2336 images) (b) Gucci (11 images) (c) H&M (112 images)

(d) Reformation (129 images) (e) Ralph Lauren (57 images) (f) Topshop (62 images)

Figure 4.10: Outline overlays for all generated outlines for each brand. The brand and number of images in the overlay are shown below the image

seen in Figure 4.12 and compared to the all image outlines, it is definitely easier to see some sort of average body outline for each brand.

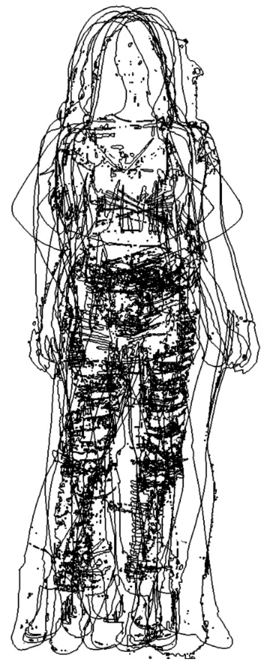


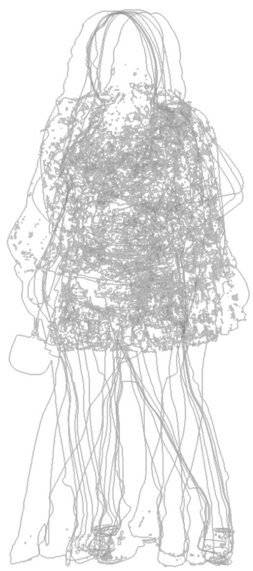
Figure 4.11: One of our original, dark outlines

4.4 CLUSTERING

The last visualization we created deals with clustering of images. We chose clustering because we thought it could potentially be useful for seeing both skin color and body shape trends for different brands. To vectorize the images, we first read them in as RGB images and flatten the image. These flattened images are then run through Principal Component Analysis (PCA) to reduce dimensionality and make them easier to process and work with. After, we cluster the images using K-Means which is one of the most popular clustering methods. While other work has done more complicated clustering using Convoluted Neural Networks and other Computer Vision techniques, we wanted to try a basic clustering first to see how it performed.

We ran clustering on the people with faces dataset since it didn't make sense to try and cluster on images that were obviously different (i.e. clothing item images with no people in them). Since we also wanted to look at poses, this was the subset of images that worked the best. Additionally, the entire clustering pipeline took a significant amount of time to run. Because Fashion Nova had a much larger amount of images to cluster compared to the other brands, we broke the Fashion Nova set into a smaller subset since we wanted to see how well the clustering was working and it would be difficult to see all 4000 images from Fashion Nova as clusters.

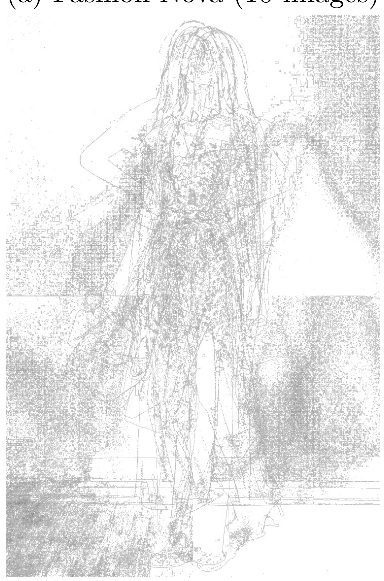
Figure 4.13 and Figure 4.14 show two examples of clusters our algorithm identified. The



(a) Fashion Nova (10 images)

(b) Gucci (10 images)

(c) H&M (10 images)



(d) Reformation (10 images)

(e) Ralph Lauren (10 images)

(f) Topshop (10 images)

Figure 4.12: Outline overlays for 10 outline images in each brand. The brand and number of images in the overlay are shown below the image

images in the resulting clusters are similar to each other, although the criteria for the similarity is not always clear. Additionally, there are some images clustered together that could be joined with another cluster, but some outliers are expected with this basic method. Additional cluster images for H&M can be found in Appendix A.

Figure 4.13: H&M Cluster 1: All of the images in this cluster have models in black clothing items except for two.

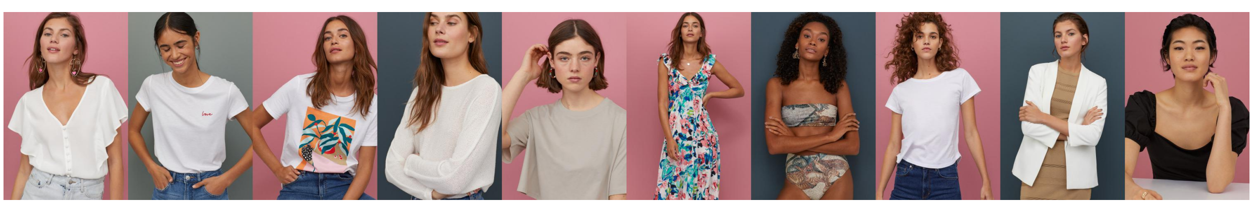


Figure 4.14: H&M Cluster 6: These images all have a colored background which differs from many of the other images in the dataset

CHAPTER 5: ANALYSIS AND DISCUSSION

In this section, we will cover the design iteration process that led to the final visualizations, analyze some of the trends found in the final visualizations, and discuss those results.

5.1 DESIGN ITERATIONS

This project was split into two iterations and there were a few different stages that our initial design went through before we settled on our current design. Our initial idea focused on being more interactive in order to engage the viewer, but also offer something unique compared to the traditional visualization. Our initial design sketches can be seen in Figure 5.1 and Figure 5.2.

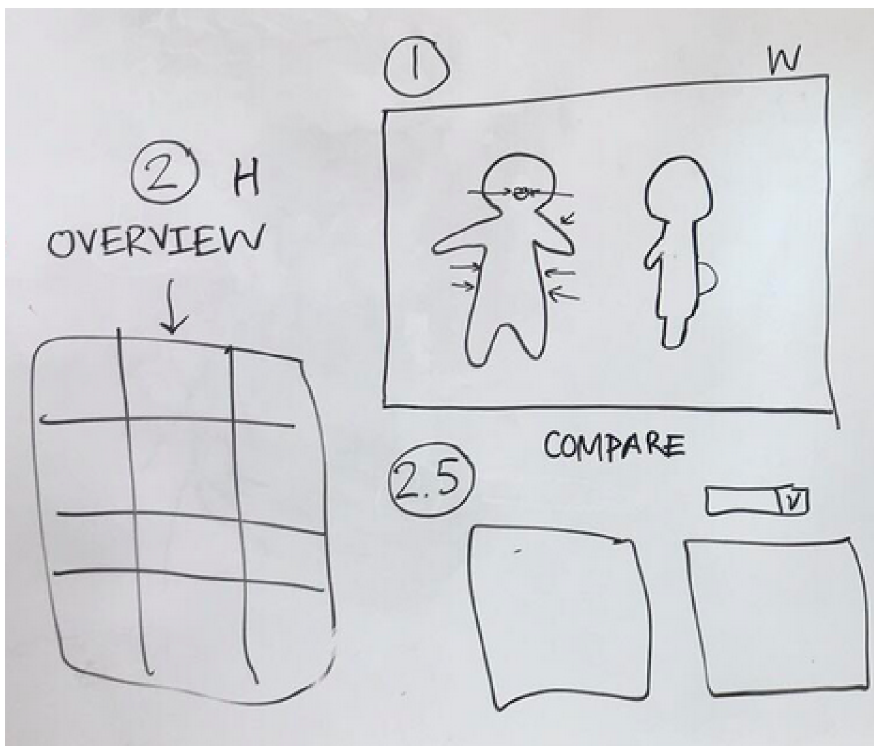


Figure 5.1: Design sketch for first two views

The first view in our sketch was an avatar view where the user could interactively find a brand where models were similar to their body shape. The idea was that a user could pull and push parts of the avatar which would then form a shape. We could match this shape to approximate measurements that we could then match to models from each brand. This is our first view because we felt it would draw users in. It also was the view that helped

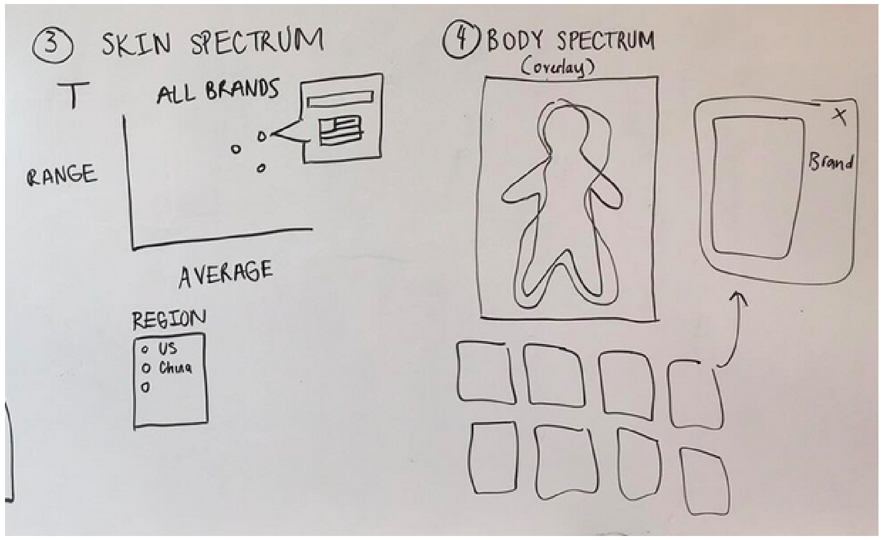


Figure 5.2: Design sketch for second two views

kickstart the rest of the visualization idea.

The second view in our sketch was the average view that was incorporated into our final design. From the design sketch and initial idea, the average images were supposed to form a mosaic of images that could be clicked. When clicked, the user would be taken to view 2.5 where they could select an average from another brand and compare these two images side by side. We decided to simplify this a bit in the final product since, at a glance, the grid helps to offer some of the comparison.

The third view in our sketch is the skin tone graph. On the sketch, we had planned for the x-axis to correspond to average skin tone, but no gradient was added yet. Additionally, we thought that it would be useful to have more information pop-up when a dot was clicked such as the region it was from. We also had the filters as checkboxes. Some of these features were carried over into our design proposal, but ultimately dropped in the final design.

The final view from our sketch is the body outline view. The original idea was to show an aggregated outline for all brands to get an idea of how the fashion industry was doing as a whole and then allow users to pick an average brand outline to view and compare with the overall average. With our original sketches, we were interested in figuring out how to display the information in such a way that users could easily compare across brands and draw conclusions.

5.1.1 First Design Iteration

Each of the sketched views were then turned into a more polished design. These are all shown in Figures 5.3, 5.4, 5.5, and 5.6. These were all created in Photoshop. These polished designs, for the most part, incorporated the aspects shown in the initial design sketches and added onto them.

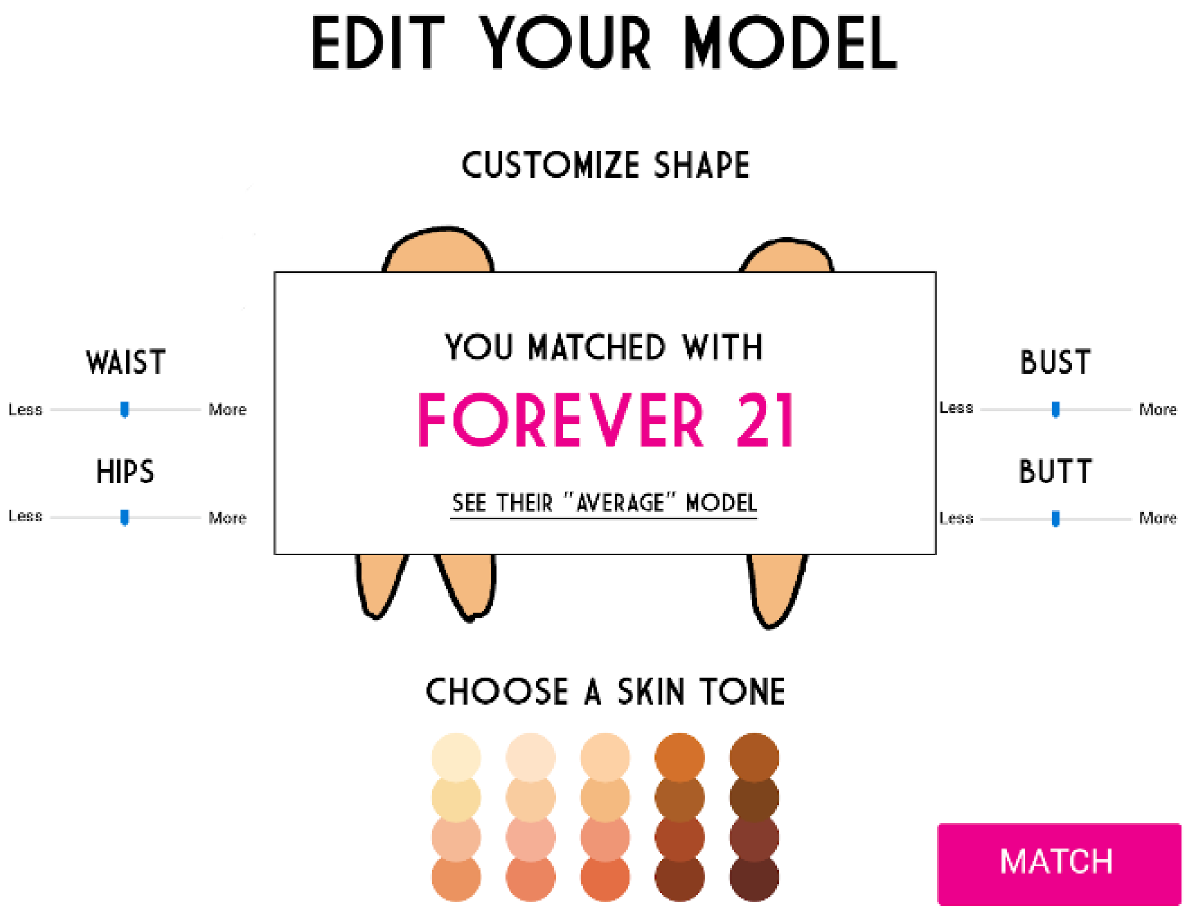


Figure 5.3: Proposed Avatar Design

For example, the polished avatar view offered sliders on either side of the avatars as well as a color palette to select skin tone. From our discussion during our presentation, we found that having a palette of colors was not good from a user experience side. This motivated us to focus on a six skin color system which was used for our skin tone graph later on.

In the polished skin view, we added two gradients: one for the bottom axis to represent average skin tone and one on the left axis to represent the range. The information box when clicking a dot was also changed to hold both representative images for a particular brand

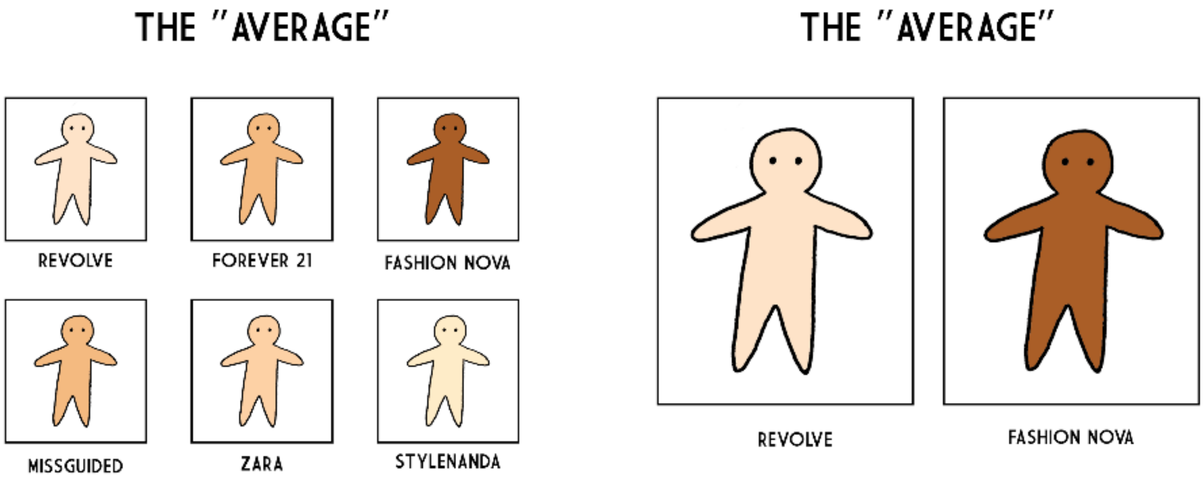


Figure 5.4: Proposed Average Design

and a bar showing the range and average skin tone on a gradient. The filters were still checkboxes and we also had a way to toggle the names of brands on and off.

The polished shape view kept the same layout as the sketched shape view, but filters were added on this page too. The idea was that a user could filter by a particular region to quickly see the average body type for that particular region and be able to compare across brands in that region. Additionally, the box showing the brand's outline would also display representative images used to calculate the outline.

In our final design for the first iteration of this project, we ended up only implementing three of the four views we initially proposed. This was mostly due to technical constraints on the avatar view that were difficult to resolve. Originally, we thought that we could use full body images of models to determine the measurements of different parts of their body like hips or waist. Using these, we could somehow match the user's avatar to the closest matching model body based on measurements. However, upon further research we realized it was impossible to try and get measurements from the photos of models because we knew nothing about the photo that we would need in order to measure other parts of the photo such as the camera length or even the known height of the model. Additionally, there's no guarantee that the model is always the same distance from the camera, so it would be difficult to reconcile different models that were different distances away.

We also thought that we could potentially use segmentation of the body to get measurements for each individual part we wanted to measure. One library we found to do this was a TensorFlow JS library called BodyPix. While this library makes it easy to segment a full

Figure 5.5: Proposed Skin Design

body photo, we found that the granularity of the segmentation that was offered was not enough for our purposes. Even if we wanted to calculate by hand approximate pixel values, we would have to perform heavy calculations in Javascript which is not an efficient way to perform calculations. An example segmentation we performed using BodyPix is shown in Figure 5.7.

The last issue with the avatar view was figuring out how to translate the values the user inputs on the sliders to pixel values in the photos. We didn't have a good way to correlate the pixels on the avatar to the pixels on the model. What does it mean for a user to move the slider one pixel to the right? One idea we had was to maybe use a ratio of pixels for each segment of the body and use these to correlate to models, but the segmentation wasn't fine-grained enough for us to use this measurement. The only way we could solve this was to write our own segmentation script, but because we had a lot of other processing work to do, like calculate average skin color and body shape outlines, we decided to scrap the avatar view in favor of finishing the other views first.

The average view, besides the simplification that was mentioned earlier, was similar to our proposed design for the first iteration. However, instead of implementing a mosaic view like in our initial sketch, we took the route proposed in our polished average view with the grid. This ended up working better anyway because we could not guarantee that all of the average images would be the same size which would make it difficult to properly display

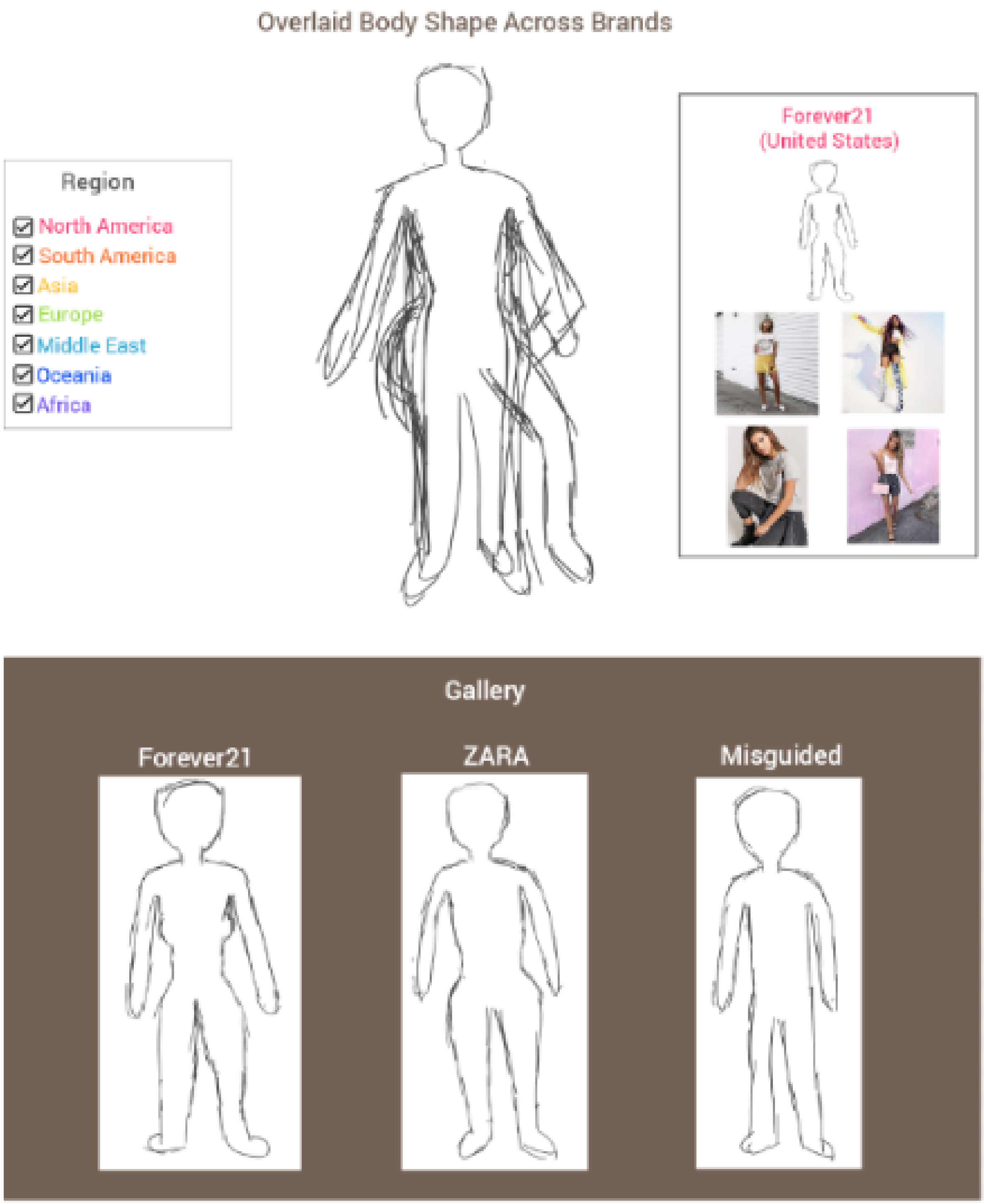


Figure 5.6: Proposed Shape Design

Figure 5.7: Example segmentation with BodyPix

all the images in a mosaic without it looking bad. For the implementation of the skin tone graph, it didn't make sense to have the left gradient, so we replaced these with a regular y-axis with numbers. Because adding range indicators made the graph too crowded, we also avoided displaying all the names of the brands at once and instead opted for brand name to show on mouse rollover. Finally, the click pop-up that was proposed for both the skin and shape views was scrapped because we decided to not show an overall average outline (due to lack of clarity of the overall outline) and the graph felt cluttered if we added more to that view. Instead, we replaced where the overall outline appeared with the brand's overall outline and displayed the individual outlines of the models used for a particular brand in the space where the representative images was supposed to be.

5.1.2 Second Design Iteration

Based off feedback on the first iteration, we made improvements to the second iteration. For the second iteration, we used entirely new data that was collected every day for about a month. Most notably, the average view went from a mosaic like view with one average for each brand to an image showing the average for each brand over time (day-by-day). This change was made to address some of the limitations of our scraping method which are covered in the Limitations section of this work.

Changes were also made to both the skin color visualization and the average body shape visualization. For example, the original skin color scatterplot included a region selector, but due to the number of brands in our second iteration dataset, we did not feel this filter would add additional information. The range indicators were also added back in since the smaller amount of brands meant we could have a clearer graph with range indicators. We also experimented with new visualizations like plotting the variance of the skin colors as the size of each point on the scatterplot. Overall, the second design tries to build off the visualizations from the first design iteration while adding new visualizations that can support the existing ones.

5.2 VISUALIZATION ANALYSIS AND DISCUSSION

In this section, we'll cover some of the trends we saw in the visualizations that were created and discuss some reasons for why those trends might have occurred. These are some of the more interesting trends, but the idea is that viewers will be able to draw their own conclusions when seeing the visualizations as well.

5.2.1 Average Images

Compared to the average of all unique images, the average per day seemed a lot clearer and less dull in color for most brands. This is probably due to just the amount of images used for each and the fact that averaging so many images could dilute the average. See Appendix A for the corresponding images. For Fashion Nova and H&M, we can see these two companies are changing colors a lot per day in their average images. This makes sense since these companies are typically known as fast fashion brands that change out their stock a lot. On the opposite side, we see Gucci seems have an average image that stays stagnant for a long period of time. There is also an interesting artifact of a line down the center of the image. Gucci changes their stock less often, so its average images per day reflects that. Additionally, this artifact is produced because of a lower body shot of black pants with a white background that is included in the data for Gucci. It looks like the average image is very sensitive to high contrast clothing images from these results (since the line is very prominent).

5.2.2 Variance and Average Skin Color

In Figure 4.4, we noticed that generally the brands with a darker skin color average have a lower variance which is indicated by the smaller circle radius. Since each brand's average is made up of several images' skin color values, the variance is saying most of these images are very close in skin color to the average skin color. This makes sense because if we look at the graph with range indicators, which shows the range of skin colors, it looks like most of the brands have very similar ranges. However, Reformation, a brand with a very light average skin color, has a skewed distribution towards more lighter skinned models, so the darker skinned models will contribute heavily to the variance.

5.2.3 Body Shape

In the outline images in Figure 4.12 and Figure 4.10, we noticed some interesting patterns. Even though the overlays aren't perfect, an idea of what the average outline looks like can be seen. For example, Gucci and Ralph Lauren seem to have very thin body outlines where as some of the other brands, especially Fashion Nova, seem to have much wider outlines. For brands with a messier outline for 10 images, like H&M, this could imply they have a variety of body shapes they are using when compared to a brand like Gucci who has a very defined outline. For Gucci, this makes some sense because it is known as a luxury brand for clothing, so the style of models they might go for could be different from a fast fashion or more consumer focused brand like Fashion Nova or H&M.

5.2.4 Interesting Dataset Observations

While we were only able to collect data for a month, there were still some interesting observations we could make about the dataset. One big thing we noticed was that a lot of the companies suddenly had pictures of masks as clothing items they were selling. This was evident when separating out clothing item pictures from the other types of images in the dataset. It can also be seen in some of the average images like in A.4 where the bottom row has a clear outline for the masks. This reflects current events that were happening at this time which is the COVID-19 outbreak.

We also noticed that the outline algorithm we used was working well for most brands, but worked really poorly on images from Reformation. Almost all of the outlines ended up with lots of artifacts that prevented a nice, clean outline. It could be because within our dataset, these images had solid colored background that were not white and had shadows on them.

These shadows likely affected some of the detection we used for the image outline.

Finally, we did see that different sites allowed us to scrape different amounts of photos each day. However, this led to an imbalance within our dataset where almost 70% of the images in the dataset is from one company. It is possible that sites like H&M had a similar rate of new clothing items as Fashion Nova, but due to how our scraper worked, we were not able to capture this.

CHAPTER 6: CONCLUSIONS AND FUTURE WORK

In this work, we focused on creating visualizations that showcased the skin color and body shape diversity of six different clothing brands.

6.1 LIMITATIONS

The limitations for this project cover limitations seen in both the first iteration of this project and the second iteration of this project. We tried to address some of the limitations from the first iteration, but some of them still apply to the second iteration.

One of the biggest limitations with our visualization is the accuracy of the skin detection script. While this script works really well for well-lit photos, not all photos from brands will be perfectly lit with a monochrome background. In the first iteration of this project, the brand that caused the most problems for us was StyleNanda, but these results apply to any brand. StyleNanda was unique among the brands we scraped during our first iteration because almost every single photo on their website included a background (some examples can be seen in Figure 6.1. While this makes the images less boring to look at, it also makes them more difficult to process. In our final set of images, we still had many images with backgrounds, especially in brands like Fashion Nova.

Figure 6.1: StyleNanda Images

Since our skin detection algorithm uses thresholding, this makes it susceptible to identifying colors in images as skin that are not skin. With the StyleNanda photos, the backgrounds sometimes get recognized as skin which can skew the average skin tone for the photo. The background is also an issue because any shadows cast by nearby trees or buildings onto skin will affect the values as well. These are issues to think about with using thresholding as a skin detection algorithm. While it works well for stereotypical clothes photos, it does not work well for more unique photos. Figure 6.2 shows parts of an image from StyleNanda where parts of the background are being categorized as skin.

Figure 6.2: Algorithm finding background as skin

This resulted in StyleNanda having a much darker average skin tone (Figure 6.3 than the rest of their counterparts even though their models, overall, do not have dark skin. This is surprising to see on the graph which is why it stands out as a limitation of our work. When compared to all brands in the first iteration of this project, the graph shows StyleNanda as having skin tones that are almost as dark as the African brands! One way to counteract this might be to normalize for shadows on faces. We could also attempt some sort of background subtraction to start with a cleaner image which might get us closer to the value we expect StyleNanda to have for their average skin tone.

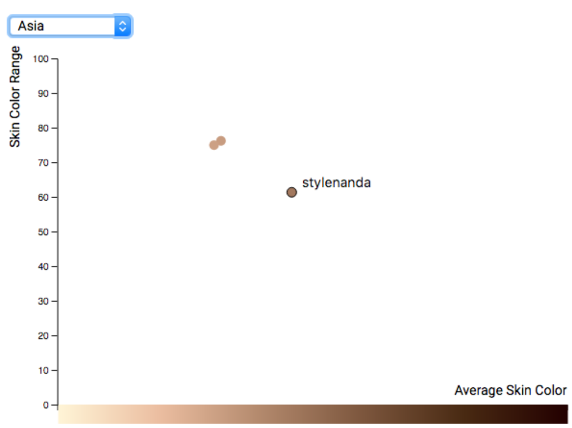


Figure 6.3: StyleNanda on Skin Graph

Another limitation is scraping. As mentioned earlier, scraping to collect data is unsustainable in the long run since a new script has to be written per website. Additionally, if the brand decides to update their website, the script will subsequently need to be updated. Between the period of time when we finished the first iteration of this project and the time when we started the second iteration, many of the brands we had originally scraped using our Python scraper were no longer working with our scraper. Also, the brands tend to update their websites fairly often. Just in the few weeks we were working on the first iteration, Forever 21 had updated their website at least twice. This resulted in us pulling different images with the same script depending on when we decided to scrape. This highlights another issue with scraping. By scraping our data, we are able to only represent the diversity of a brand as a snapshot in time. This means in a few weeks, this same snapshot might not be valid anymore. Because it represents a snapshot, this was a reason for us collecting data over time for the second iteration of this project.

This visualization is also very sensitive to the amount of images available in the dataset. With a smaller dataset, for example, the average images become much clearer. Adding even one image with a wildly different pose from the other images to this average will make that different pose stand out like a sore thumb. On the flip side, adding more images will hide some of these outliers, but might result in harder to interpret or see average images. We've created our average images using a different number of images for each brand, so better comparisons could be done between brands if all the averages were created using the same number of images.

Currently, for the shape view, the ten full body images per brand are picked manually. While we tried to select representative images in a standardized way, there is no good way to guarantee we will manually pick representative images. Thus, while the outlines can show us the general body shape of the models a brand uses, it is possible that these outlines are not fully representative. To make these more representative, I think finding a way to better display more than ten outlines together would be the first step. We tried a version with all the outlines together for the second iteration, but it isn't clear if this is very effective in communicating the average body shape. However, if the outline was more clear, it would definitely be a more representative image of the outlines for one brand.

Finally, because we are using background subtraction and then edge detection, sometimes artifacts get left by other objects in the picture. For example, the shadow can be seen in the outline for a lot of photos from FabIndia (Figure 6.4) which interferes with the overall outline shape. An improvement would be to eliminate this entirely, possibly by testing out more advanced background subtraction techniques. While the example in Figure 6.4 is not in the way of the outline, it is possible these shadow artifacts could mess up the outline, especially if the shadow is right behind the model. If the shadow can't be removed, finding a way to reduce the effect of the shadow on the outline might be an alternative.

Figure 6.4: Shadow Artifact

6.2 FUTURE WORK

As mentioned before, our scraper was a Python scraper which severely limited which sites we could pull images from. With an automated system and a Javscript-based web scraper, it might be possible to have a wider variety of brands to perform these visualizations on. Additionally, it could help with collecting more images from websites that don't show many images on the first page load like H&M. This might even out the dataset more so that it isn't dominated by one brand's images.

Another improvement that could be made is to the outline generation algorithm. As mentioned in limitations, our current methods leave some artifacts outside the outline, but also leave some artifacts inside the outline. There wasn't enough time to explore algorithms that allowed for removing all inside artifacts, but some of these could be applied to help with cleaning up the outlines more. Additionally, it is clear that stacking all outlines together makes for a very messy average outline, but with cleaner outlines, it is possible that this method can actually communicate a lot of information to the viewer.

Skin color detection was also a limitation that was mentioned previously. When looking at the data, some of the values seem weird at first, like Fashion Nova having a middling average skin color while sites like H&M and Ralph Lauren had pretty high average skin colors. However, it makes sense when looking through the data itself since these brands do indeed have models of both lighter and darker skin colors. However, I think there are still improvements that could be made to the current skin tone detection algorithm to make these calculations more accurate. Currently, the algorithm is really sensitive to backgrounds that aren't fully white, so it would be helpful to have a more accurate measurement.

Finally, something that we wanted to add but did not have a chance to was filters for the visualizations so that users could isolate brands based on different characteristics like the average price of a white t-shirt or possibly incorporating demographic information in some way. This way, users could not only see trends for different brands, but maybe also have an idea of how diverse different types of brands (luxury, fast fashion, sustainable, etc.) are overall.

6.3 CONCLUSION

In this work, we created visualizations that displayed two big aspects of clothing photos: model body shape and model skin color. These visualizations aim to help viewers see trends over time for these brands and decide for themselves whether brands are really adopting more diversity in their models. While not perfect, these visualizations are a great start for seeing trends in average skin color and body shape for a particular brand and could extend to seeing trends for many types of brands.

The goal of creating StyleRep was so we could further explore what diversity, if any, was present in the fashion industry. In particular, we were curious about temporal differences and if skin color and body shape could be used as a measure of diversity in models for a brand. This is reflected in our visualization through the use of time when displaying the average images and could be extended to some of the skin tone visualizations over a longer period of time. With StyleRep, we were able to focus on two main parts of diversity: body shape and skin tone. While there are other parts not captured in this, i.e. race, we see StyleRep as the beginning of finding ways to better quantify and represent different aspects of diversity. While there are areas that could still be improved upon, we hope that StyleRep is the first step towards finding better ways to represent diversity in both fashion and beyond.

CHAPTER 7: REFERENCES

[1] "Diversity at gucci." [Online]. Available: https://www.gucci.com/us/en/st/stories/gucci-equilibrium/article/actions-for-diversity-and-inclusion

[2] L. Prinzivalli, "Victoria's secret announced its first plus-size model — but there's a catch," 2019. [Online]. Available: https://www.allure.com/story/victorias-secret-plus-size-model

[3] B. Barry, "Selling whose dream? a taxonomy of aspiration in fashion imagery," *Fashion, Style, & Popular Culture*, vol. 1, no. 2, pp. 175–192, 2014.

[4] M. DeLong and B. Martinson, *Color and design.* A&C Black, 2013.

[5] B. Arik et al., "Seasonal and emotional associations of the colours and their effects on directing the turkish fashion," *Color Research & Application*, vol. 41, no. 5, pp. 523–529, 2016.

[6] A. Mears, "Size zero high-end ethnic: Cultural production and the reproduction of culture in fashion modeling," *Poetics*, vol. 38, no. 1, pp. 21–46, 2010.

[7] S. Bakhshi, D. Shamma, and E. Gilbert, "Faces engage us: Photos with faces attract more likes and comments on instagram," *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 965–974, Apr. 2014.

[8] N. Hochman and L. Manovich, "Zooming into an instagram city: Reading the local through social media," *First Monday*, vol. 18, no. 7, 2013.

[9] J. Salavon, "Every playboy centerfold, the decades (normalized)," 2002. [Online]. Available: http://salavon.com/work/EveryPlayboyCenterfoldDecades/

[10] K. Matzen, K. Bala, and N. Snavely, "Streetstyle: Exploring world-wide clothing styles from millions of photos," *arXiv preprint arXiv*, vol. 1706, no. 01869, 2017.

[11] CHEREF-Medhi, "Skindetection," 2018. [Online]. Available: https://github.com/CHEREF-Medhi/SkinDetection

APPENDIX A: EXTRA VISUALIZATIONS

This appendix is for the visualizations that were not shown in the main portion of the paper due to space issues.

A.1 AVERAGE IMAGES

Figure A.1: Average of Images per Day for Fashion Nova

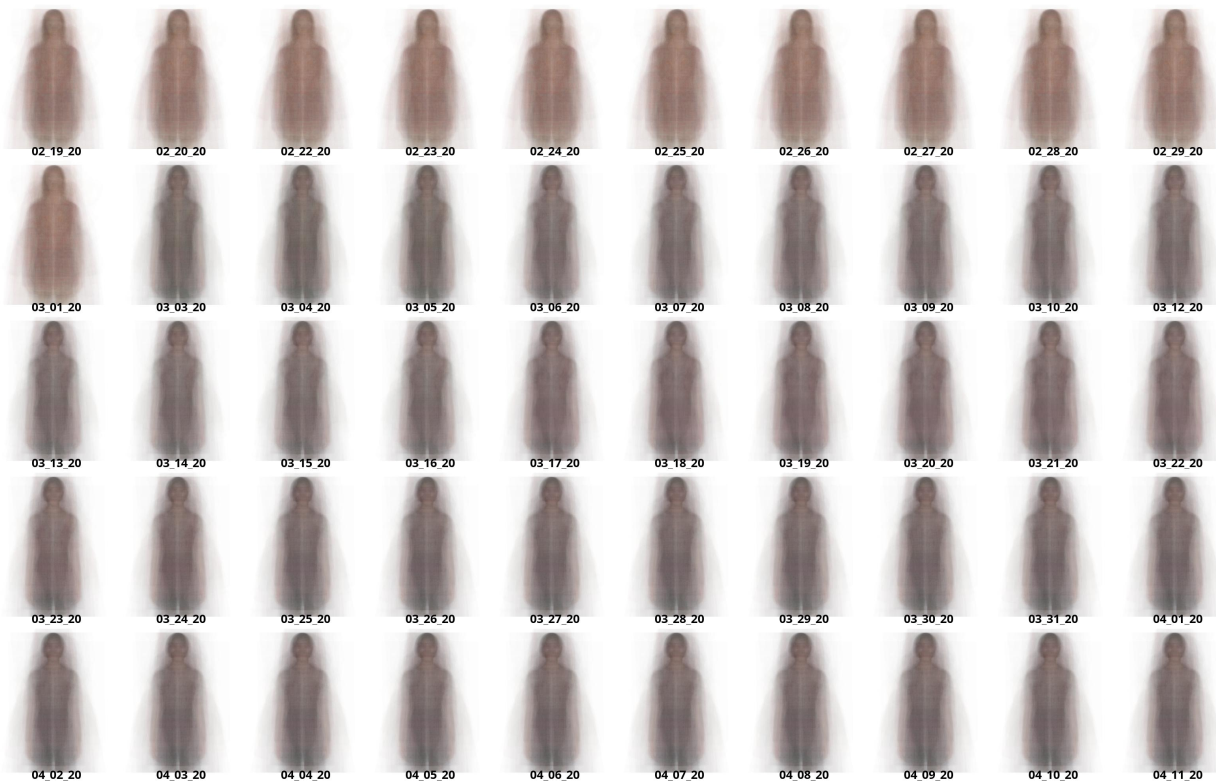


Figure A.2: Average of Images per Day for Gucci

Figure A.3: Average of Images per Day for Topshop

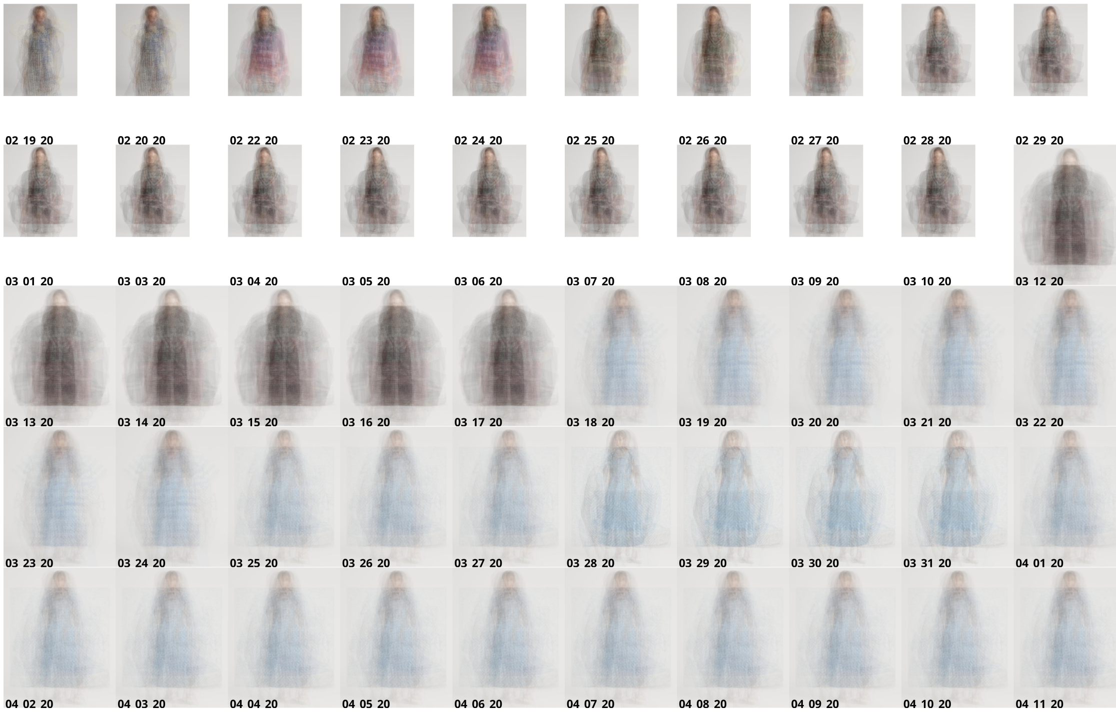


Figure A.4: Average of Images per Day for Ralph Lauren

Figure A.5: Average of Images per Day for H&M

A.2 SKIN COLOR GRAPHS

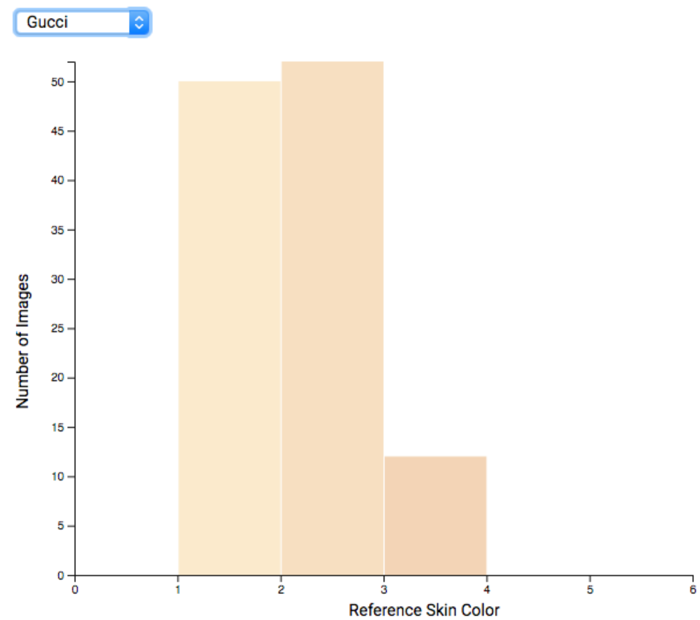


Figure A.6: Skin Color Distribution for Gucci

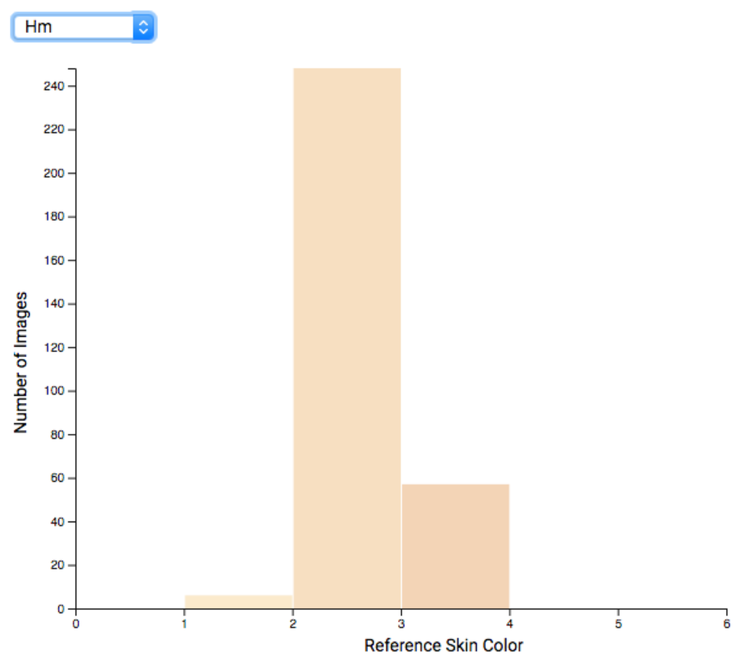


Figure A.7: Skin Color Distribution for H&M

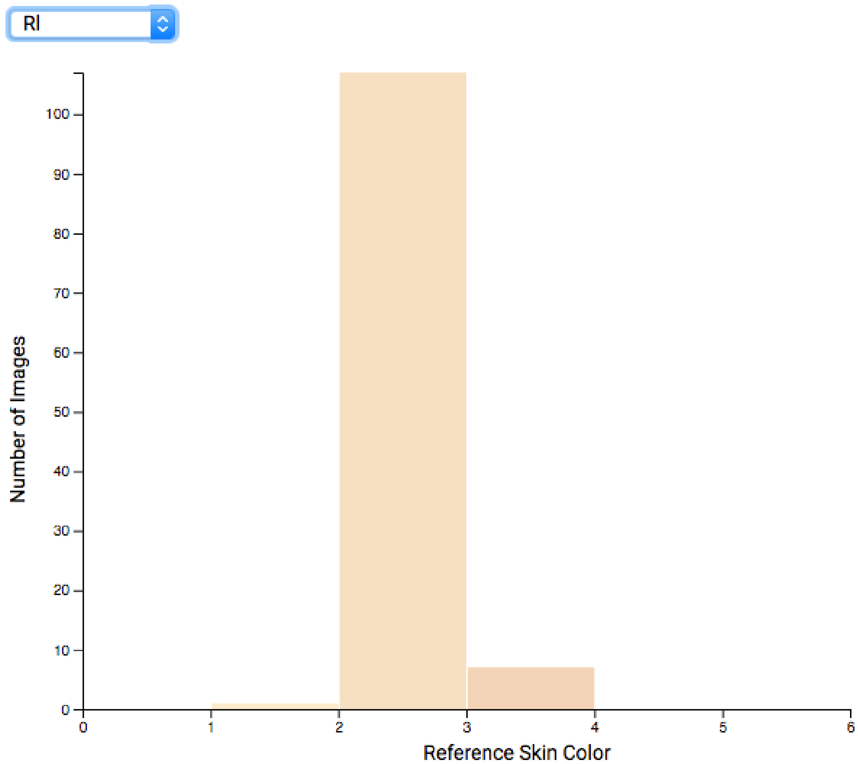


Figure A.8: Skin Color Distribution for Ralph Lauren

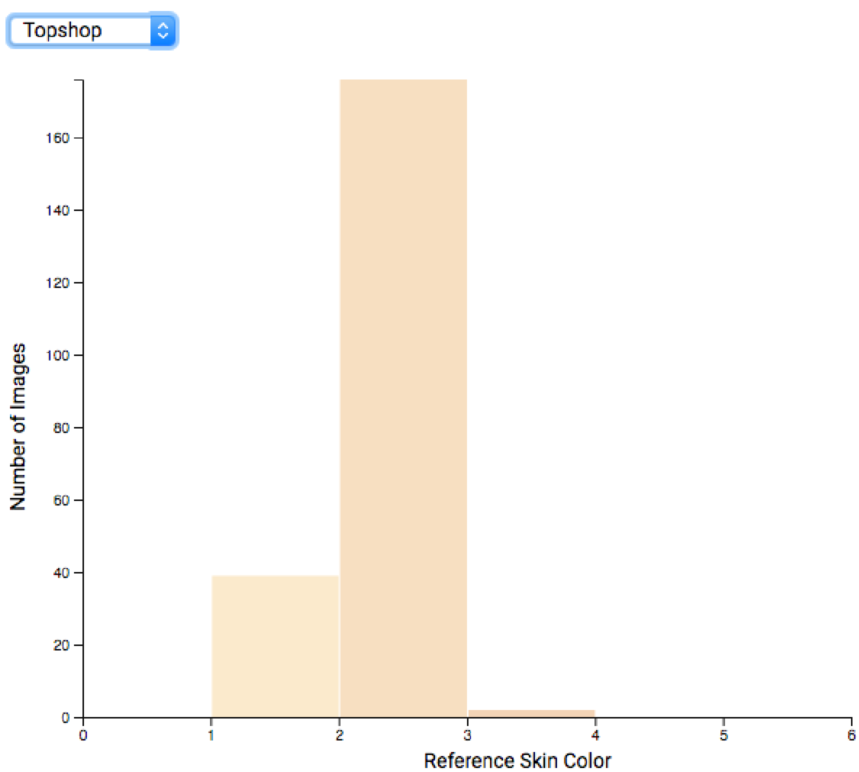


Figure A.9: Skin Color Distribution for Topshop

A.3 CLUSTERING IMAGES

Figure A.10: H&M Cluster 2

Figure A.11: H&M Cluster 3

Figure A.12: H&M Cluster 4

Figure A.13: H&M Cluster 5

Figure A.14: H&M Cluster 7

Figure A.15: H&M Cluster 8

Figure A.16: H&M Cluster 9

APPENDIX B: IMPLEMENTATION DETAILS

B.1 SKIN COLOR GRAPHS

Generating the graph in Figure 4.3 itself posed some challenges; first, the x-axis of a scatterplot graph in d3 cannnot be a gradient. To get this gradient, we had to figure out how to create a color scale in d3 that interpolated between the six different colors. We were able to do this using d3-interpolate which allows for many different types of interpolation, including color interpolation. Once this scale was created, it could be used for both the gradient and the colors for the dots.

The original JSON data also needed to be transformed into a format better suited for d3 visualization, but this was done in Javascript after reading in JSON data using d3.json(). One of the most difficult parts in transforming the data was figuring out the best way to iterate through it; since I didn't necessarily know the keys ahead of time and the data that was read in was a JSON object, it took trial and error before I was able to use Object.keys() to get a list of keys I could use to iterate.