

# The Shades of Racism in the Beauty Industry\*

An exploration of inclusivity and tokenism within fashion and makeup

Huda Sahaf

18 April 2023

This paper aims to investigate the history of racism in the beauty industry and how this impacts the inclusivity in the market. The complexions of models are assigned a lightness value and analyzed across 8 years. These results are compared to the foundation shade ranges across all major makeup brands across the U.S revealing a continuation of compounding racism within the industry as a whole.

```
fashion_tones <-  
  read.csv(  
    file = "https://raw.githubusercontent.com/the-pudding/data/master/vogue/faces.csv",  
    skip = 0  
  )  
  
fashion_frequency <-  
  read.csv(  
    file = "https://raw.githubusercontent.com/the-pudding/data/master/vogue/models.csv",  
    skip = 0  
  )  
  
makeup_shades <-  
  read.csv(  
    file = "https://raw.githubusercontent.com/the-pudding/data/master/makeup-shades/shades",  
    skip = 0  
  )  
  
#### Save data ####  
  
write_csv(fashion_tones, here::here("inputs/data/fashion_tones.csv"))
```

---

\*Code and data are available at: [https://github.com/hsahaf/Racism\\_Beauty\\_Industry.git](https://github.com/hsahaf/Racism_Beauty_Industry.git).

```
write_csv(fashion_frequency, here::here("inputs/data/fashion_frequency.csv"))
```

```
write_csv(makeup_shades, here::here("inputs/data/makeup_shades.csv"))
```

```
### Clean Fashion Tones Dataset:
```

```
cleaned_fashion_tones <-  
  clean_names(fashion_tones)
```

```
cleaned_fashion_tones <-  
  cleaned_fashion_tones |>  
  rename(  
    skintone = tone,  
    lightness_value = l  
  )  
head(cleaned_fashion_tones)
```

	date	model	skintone	lightness_value
1	1/1/2018	Lupita Nyong'o	#402C2C	0.2127451
2	10/1/2015	Lupita Nyong'o	#58342B	0.2574230
3	3/1/2009	Michelle Obama	#844930	0.3541176
4	4/1/2015	Serena Williams	#875038	0.3750000
5	7/1/2014	Lupita Nyong'o	#85524F	0.4166667
6	10/1/2016	Lupita Nyong'o	#8F5347	0.4205882

```
### Modeling Full Spread of Complexion Values on Vogue Covers from 2000-2018
```

```
cleaned_fashion_tones$clean_date <- ymd(cleaned_fashion_tones$clean_date)
```

```
entire_spread <- cleaned_fashion_tones |>  
  ggplot(aes(x = lightness_value, y = clean_date)) +  
  geom_point(aes(color = lightness_value < 0.4)) +  
  labs(  
    x = "Lightness Value of Vogue Models",  
    y = "Date of Cover",  
    title = "Complexions of Models on Covers of Vogue: 2000-2018"  
  ) +  
  theme_classic()
```

```
datebreaks <- seq(as.Date("2001-01-01"), as.Date("2018-09-01"), by = "1 year")
```

```
print(entire_spread)
```

Table 1: Models With More than 5 Appearances on Vogue

Model	Number of Covers	Lightness Value
Angelina Jolie	5	0.67
Cate Blanchett	5	0.68
Charlize Theron	6	0.68
Gisele Bundchen	9	0.67
Gwyneth Paltrow	5	0.69
Kate Hudson	5	0.67
Keira Knightley	5	0.71
Nicole Kidman	7	0.70
Penelope Cruz	5	0.62
Reese Witherspoon	5	0.63
Rihanna	5	0.62
Sarah Jessica Parker	6	0.62

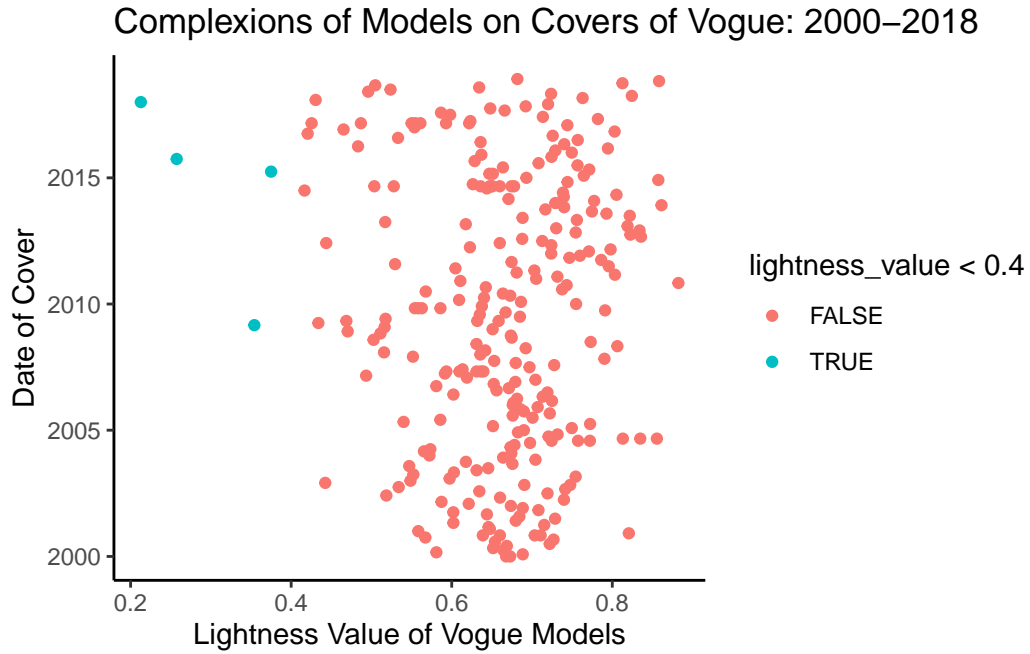


Figure 1: Complexion Values of All Models on Vogue Covers from 2000-2018

Figure 1 is a graph of the lightness value of every model that has been on the cover of Vogue from 2000 to 2018. Lightness value is

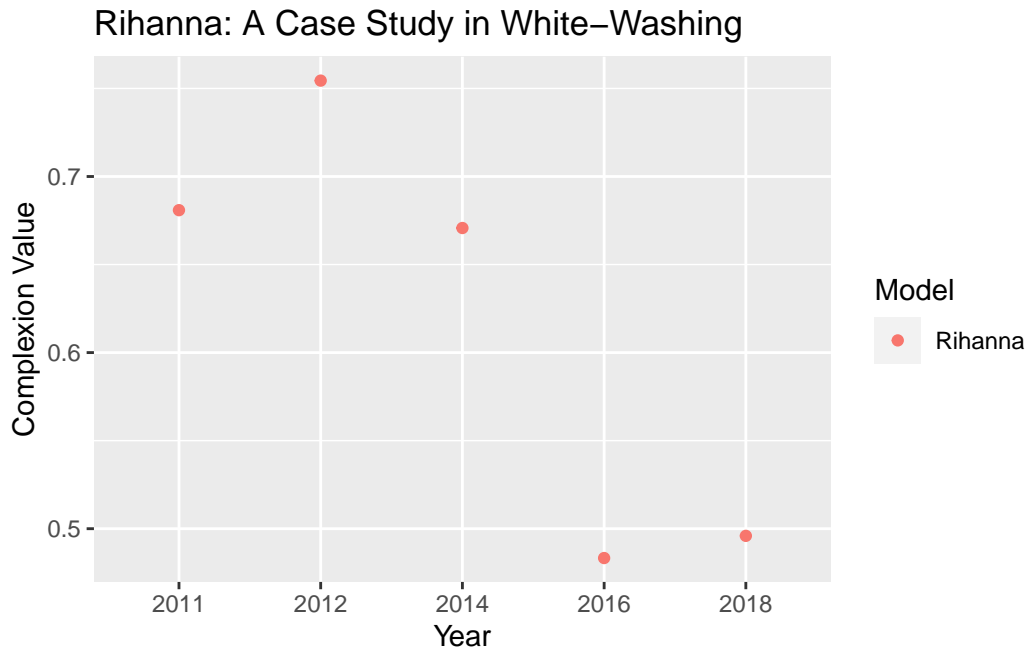


Figure 2: Rihanna's Appearances on Vogue

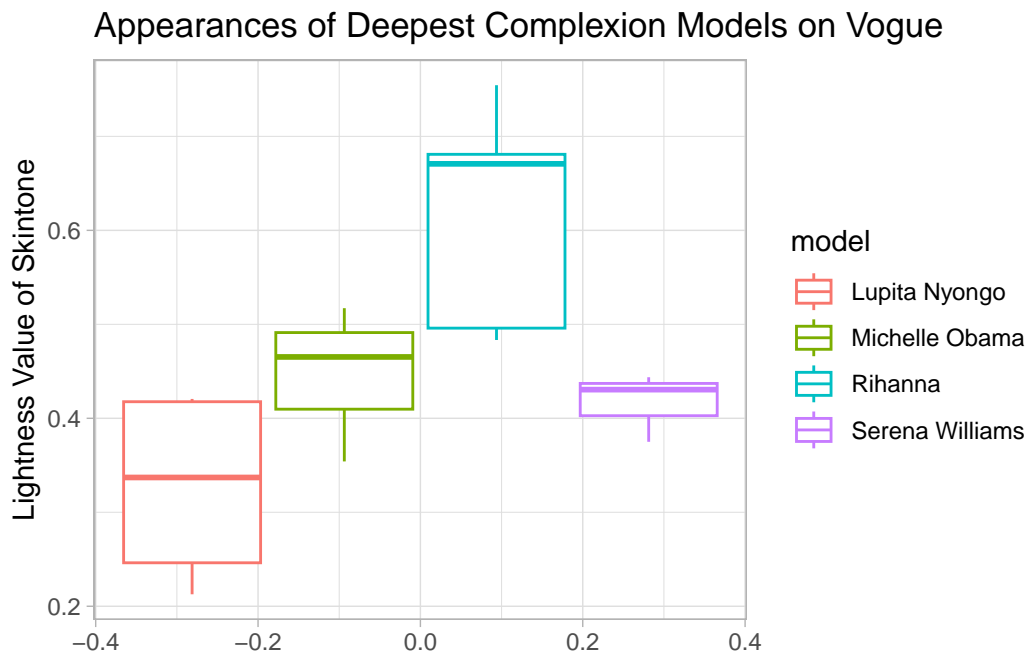


Figure 3: Top 4 Deepest Complexion Models on Vogue

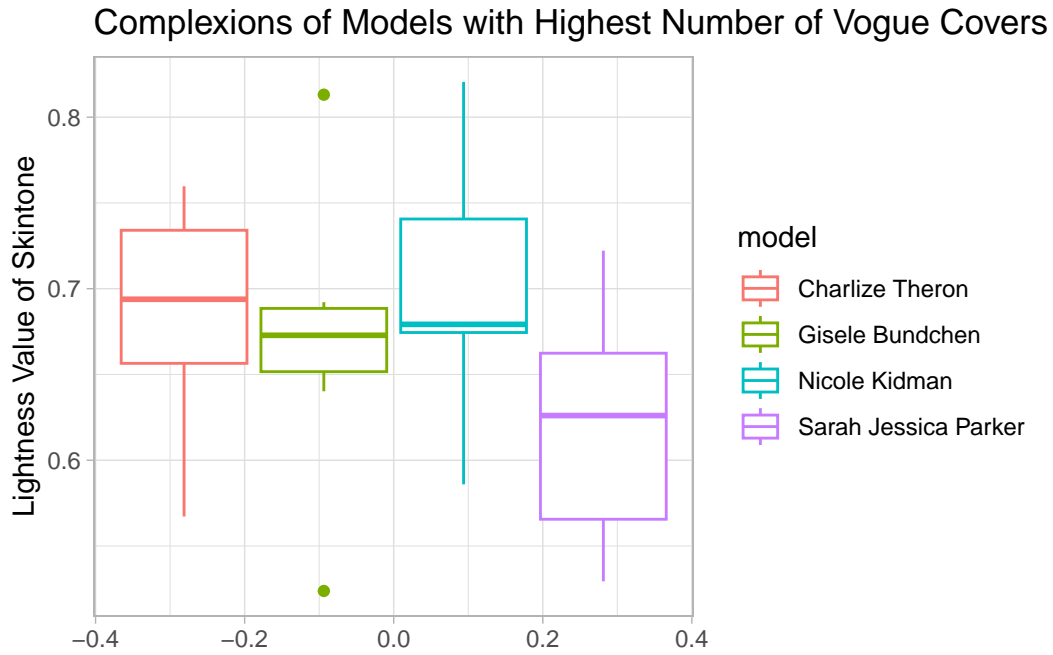


Figure 4: Models with Highest Number of Appearances

Table 2: US Makeup Brands: Foundation Shade Range

Makeup Brand	Number of Foundation Shades
Beauty Bakerie	30
Black Opal	12
Black Up	18
Bobbi Brown	30
Covergirl + Olay	12
Estée Lauder	42
Fenty	40
Iman	8
L'Oréal	22
Lancôme	40
Laws of Nature	17
MAC	42
Make Up For Ever	40
Maybelline	40
Revlon	22
bareMinerals	29

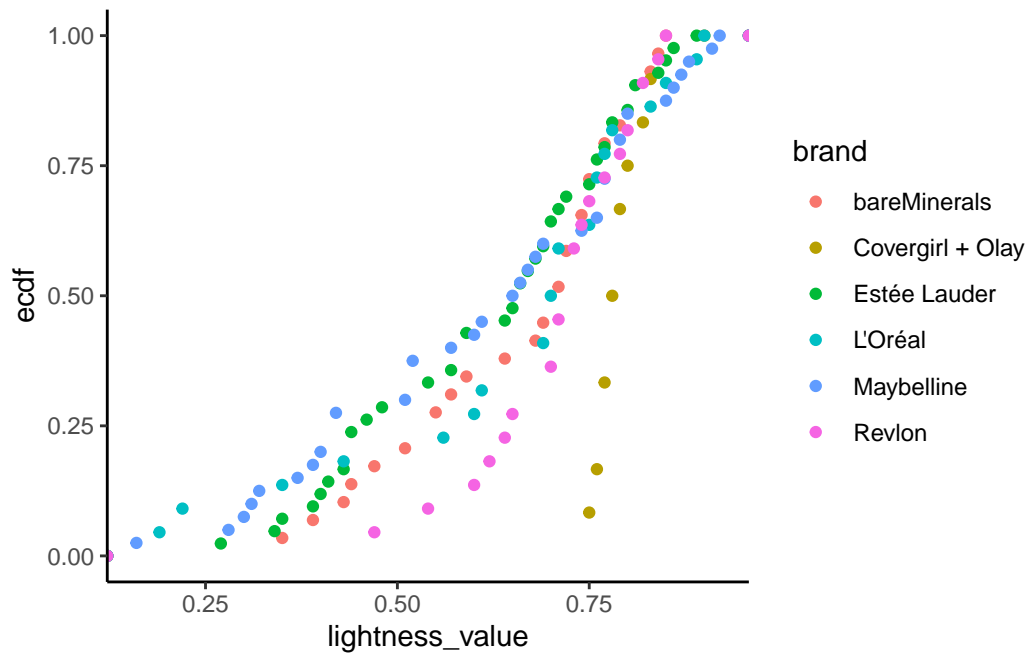


Figure 5: The Shade Distribution of Best-Selling Foundations in the U.S

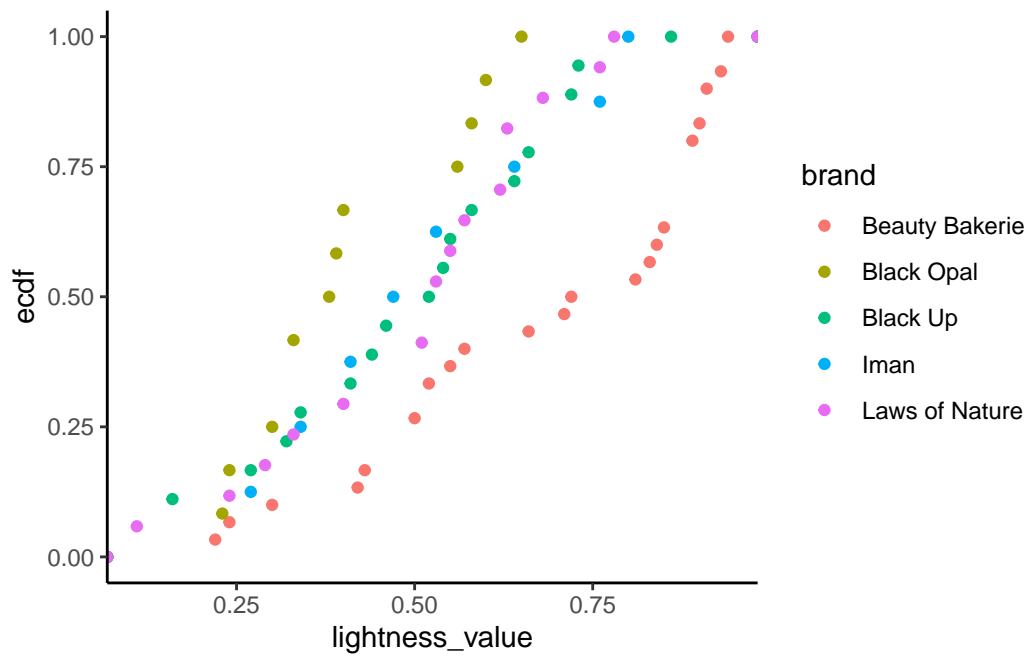


Figure 6: The Shade Distribution of BiPOC-owned Makeup Brands

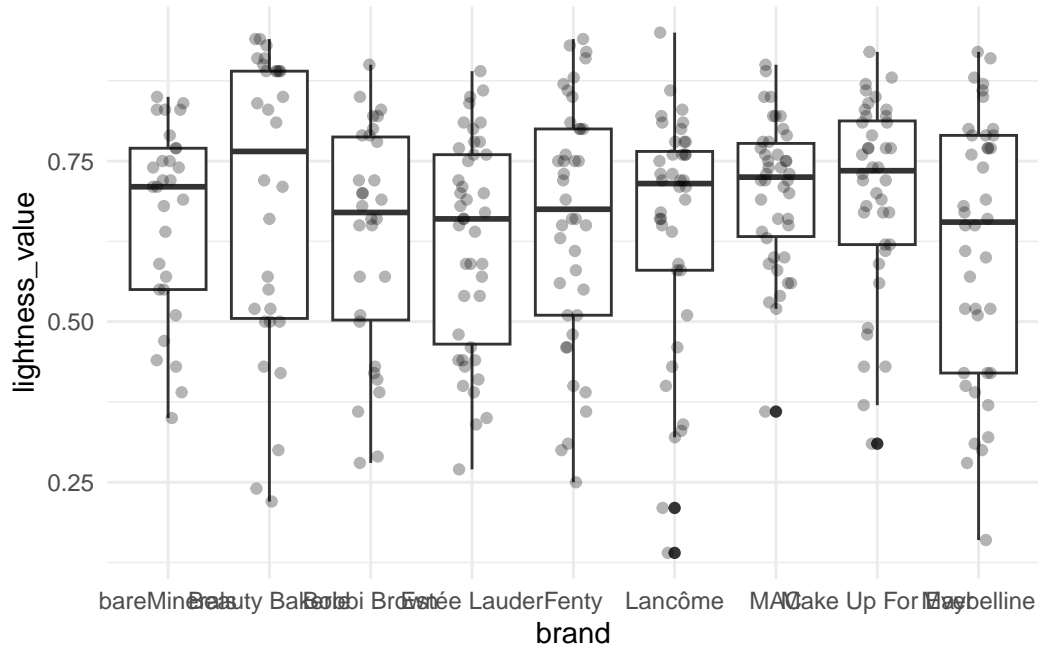


Figure 7: The Shade Distribution of Largest Shade Ranges Offered in the U.S

## 1 Introduction

You can and should cross-reference sections and sub-sections. For instance, Section [2](#) and Section [4.1](#).

## 2 Data

### 2.1 Data Management

All analysis on the data sets were done using R (R Core Team 2020). The data was read in, cleaned, and manipulated using `dplyr` (Wickham et al. 2023), `tidyverse` (Wickham et al. 2019), and `janitor` (Firke 2023). The graphs and tables in this paper were coded using the following packages: `ggplot2` (Wickham 2016), `kableExtra` (Zhu 2021), `knitr` (Xie 2023), `readr` (Wickham, Hester, and Bryan 2023), and `RColorBrewer` (Neuwirth 2022).

## 2.2 Source and Sampling

In order to conduct an analysis on the various parts of the beauty industry, three separate data sets were used. The Pudding is a digital publication that makes its data sets open to the public *The Pudding Repository* (2023). A closer look is taken at Vogue using the first two data sets used for the article *Colorism in High Fashion* (Handa, Thomas, and Diehm 2019). All covers of Vogue from the year 2000 to 2018 are analyzed and categorized in these data sets. The female models on each cover are identified, and then their skin tones are filtered out of the cover, and an average color value is assigned based on the all skin tone pixels. These color values are then drained of any hue or saturation and therefore assigned a lightness value based on how light or deep the complexion of the model is, which allows for the comparison of all models on the covers of Vogue over the span of 8 years. The first data set contains the name of the model, date of their appearance on Vogue, the hex code of their skin tone, and their lightness value once the hue and saturation has been removed. This includes multiple appearances of the same model. The second data set contains the names of the models, the hex code of their skin tone, the number of covers they have made appearances in, and their average lightness value across these appearances.

The data set behind The Pudding article, *Beauty Brawl* is the third data set used in this paper (Li, Thomas, and Manian 2018). It contains data about the foundation shade ranges through lightness values of multiple beauty brands around the world. For the sake of this paper, we are focusing on beauty retailers based in the US, especially since Vogue is also based in the United States.

Our data is of penguins (`?@fig-bills`).

Talk more about it.

Also bills and their average (`?@fig-billssssss`). (You can change the height and width, but don't worry about doing that until you have finished every other aspect of the paper - Quarto will try to make it look nice and the defaults usually work.)

Talk way more about it.

Here's a dumb example of how to use some references: In paper we run our analysis in R (R Core Team 2020). We also use the `tidyverse` which was written by (`thereferencecanbewhatever?`) If we were interested in baseball data then (`citeLahman?`) could be useful.

We can use maths by including latex between dollar signs, for instance  $\theta$ .



### **3 Results**

## **4 Discussion**

### **4.1 First discussion point**

If my paper were 10 pages, then should be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

### **4.2 Second discussion point**

### **4.3 Third discussion point**

### **4.4 Weaknesses and next steps**

Weaknesses and next steps should also be included.

## **Appendix**

### **A Additional details**

## References

- Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Handa, Malaika, Amber Thomas, and Jan Diehm. 2019. *Colorism in High Fashion*. <https://pudding.cool/2019/04/vogue/>.
- Li, Jason, Amber Thomas, and Divya Manian. 2018. *The Diversity of Makeup Shades Beauty Brawl*. <https://pudding.cool/2019/04/vogue/>.
- Neuwirth, Erich. 2022. *RColorBrewer: ColorBrewer Palettes*. <https://CRAN.R-project.org/package=RColorBrewer>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- The Pudding Repository*. 2023. The Pudding. <https://github.com/the-pudding/data>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data*. <https://CRAN.R-project.org/package=readr>.
- Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.