Tidy Tuesday - Week 4

Grace Hu

01/04/2021

Task at hand

For this week, I will be looking at the Makeup data taken from here. As someone who has an interest in beauty and makeup, I know that race representation in the cosmetics industry has been a problem. In the recent years, more and more companies are aware that shade range in foundation products is very important, but I wanted to explore that further with the data.

Challenge 01

For the first task, I want to mimic the plots found in the related article. However, I did not know how to colour every individual data point with the given HEX codes, so I looked online for some references. I found some helpful code from Florence V. Dubois' GitHub repository and learned some new features of `ggplot`!

Data Wrangling

```
shades_cat <- shades_raw %>%
  mutate(l_cat = floor(lightness * 20), UID = as.factor(row_number())) %>%
  group_by(l_cat) %>%
  mutate(num = row_number()) %>%
  ungroup()


shades_cat_colours <- shades_cat$hex
```

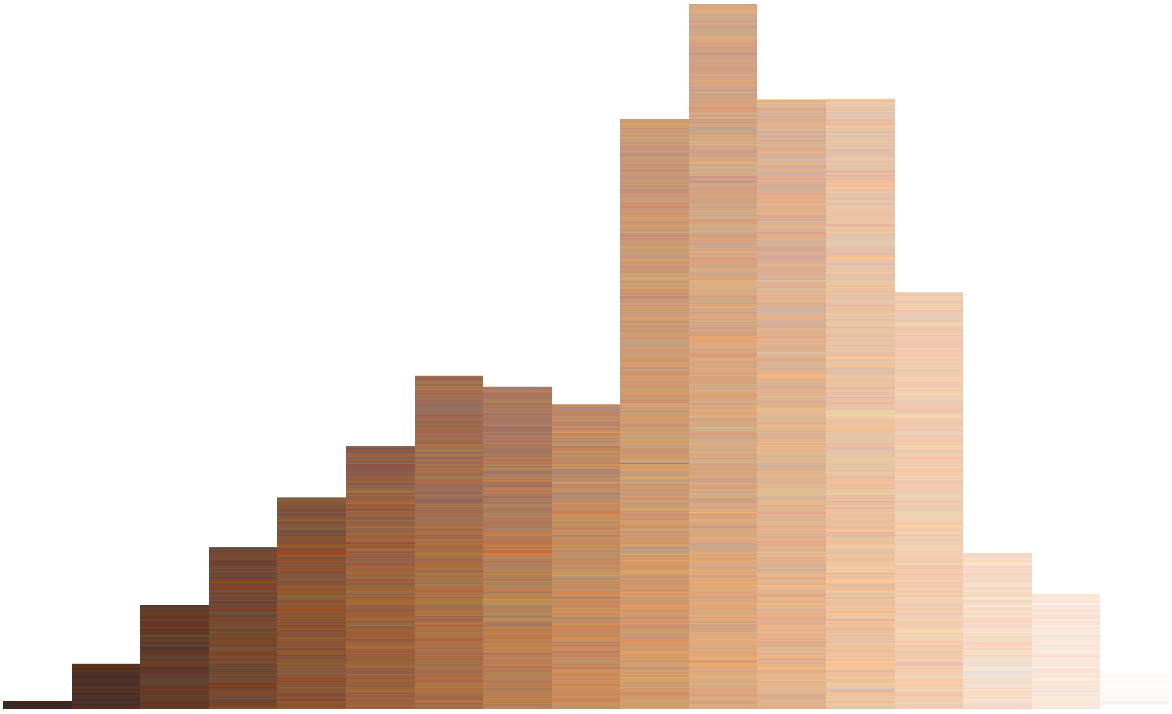
Data Visualization

```
ggplot(data = shades_cat, aes(x=l_cat, y = num)) +
  geom_raster(aes(fill = UID)) +
  scale_fill_manual(values = shades_cat_colours) +
  scale_x_discrete() +
  labs(
    title = "Foundation shade ranges",
    subtitle="from all makeup companies at Sephora and Ulta"
  ) +
  theme_void() +
  theme(legend.position="",
        axis.title.x=element_blank(),
```

```
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank(),
        axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank())
```

Foundation shade ranges

from all makeup companies at Sephora and Ulta



Challenge 02

Now, I am interested in seeing what the foundation range is like for any collections with 20 or less shades.

Data Wrangling

```
shades_bot_20 <- shades_cat %>%
  group_by(brand, product) %>%
  mutate(count = n()) %>%
  ungroup() %>%
  filter(count <= 20) %>%
  group_by(l_cat) %>%
  mutate(num = row_number()) %>%
  ungroup()


shades_bot_20_colours <- shades_bot_20$hex
```

Data Visualization

```
ggplot(data = shades_bot_20, aes(x=l_cat, y = num)) +
  geom_raster(aes(fill = UID)) +
  scale_fill_manual(values = shades_bot_20_colours) +
  scale_x_discrete() +
  labs(
    title = "Foundation shade ranges",
    subtitle="from foundation collections with 20 or less shades at Sephora and Ulta"
  ) +
  theme_void() +
  theme(legend.position="",
        axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank(),
        axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank())
```