Extra credit

INFO 2950 - Spring 2023

Ming DeMers

5/9/23

Setup

Load packages and data:

```
library(tidyverse)
library(tidytuesdayR)
library(skimr)
library(grid)
library(directlabels)
```

Extra credit

Question: As "The Joy of Painting" progressed, how does Bob Ross's use of colors change?

Does he use fewer colors? Are there particular colors that remain consistent? Certain colors that see increased/decreased use?

The Data

This data set chooses 18 distinct colors and tracks them over the 30 seasons of "The Joy of Painting." Each observation is one painting, its season and episode, and what color the painting uses, by name and hexcode. We are primarily interested in the colors and their hexcodes, so we do some wrangling to recreate a dataframe for our purposes, below.

```
# Get the Data (as suggested by the R4DS project)
bob_ross <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday
```

```r
# Thomas Mock (2022). Tidy Tuesday: A weekly data project
# aimed at the R ecosystem.
# https://github.com/rfordatascience/tidytuesday.

head(bob_ross)
```

```
# A tibble: 6 x 27
  painting_index img_src    painting_title season episode num_colors youtube_src
           <dbl> <chr>      <chr>           <dbl>   <dbl>      <dbl> <chr>
1            282 https://w~ A Walk in the~      1       1          8 https://ww~
2            283 https://w~ Mt. McKinley        1       2          8 https://ww~
3            284 https://w~ Ebony Sunset        1       3          9 https://ww~
4            285 https://w~ Winter Mist         1       4          3 https://ww~
5            286 https://w~ Quiet Stream        1       5          8 https://ww~
6            287 https://w~ Winter Moon         1       6          4 https://ww~
# i 20 more variables: colors <chr>, color_hex <chr>, Black_Gesso <lgl>,
#   Bright_Red <lgl>, Burnt_Umber <lgl>, Cadmium_Yellow <lgl>,
#   Dark_Sienna <lgl>, Indian_Red <lgl>, Indian_Yellow <lgl>,
#   Liquid_Black <lgl>, Liquid_Clear <lgl>, Midnight_Black <lgl>,
#   Phthalo_Blue <lgl>, Phthalo_Green <lgl>, Prussian_Blue <lgl>,
#   Sap_Green <lgl>, Titanium_White <lgl>, Van_Dyke_Brown <lgl>,
#   Yellow_Ochre <lgl>, Alizarin_Crimson <lgl>
```

```r
# remove the first row, and fix color naming scheme
bob_ross <-
  select(bob_ross, -1) |>
  mutate(
    across(Black_Gesso:Alizarin_Crimson, as.logical)
  )
```

```r
#| label: color-vectors

# a rainbow 7-color scheme
roygbiv <- c('#FA3D1E', '#FAF749', '#3CFA46', '#488EFA',
                      '#8B4513', 'black','white')

# the actual hex codes of Ross's color palette
pal_colors <-  c("#4E1500", #Alizarin_Crimson
                 "#000000", #Black_Gesso
                 "#DB0000", #Bright_Red
```

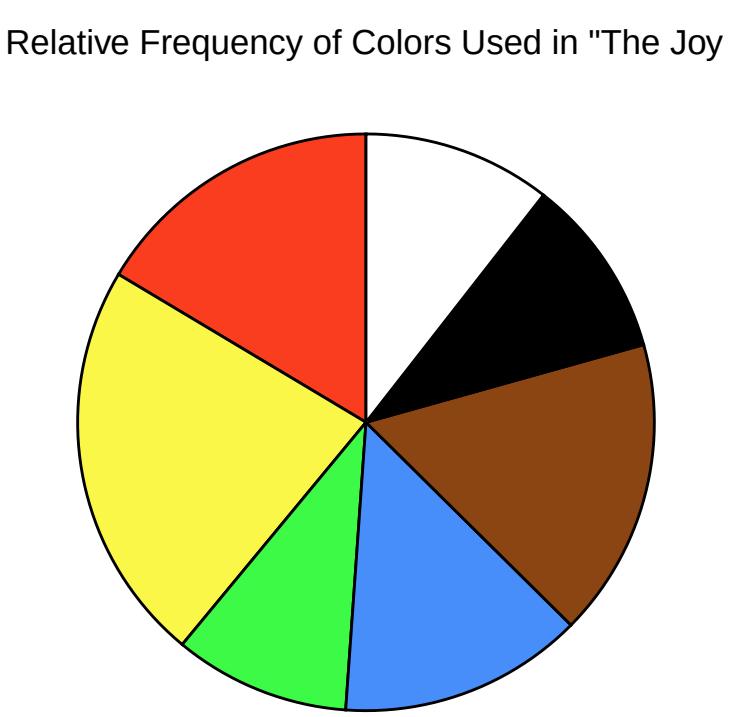
```
                  "#8A3324", #Burnt_Umber
                  "#FFEC00", #Cadmium_Yellow
                  "#5F2E1F", #Dark_Sienna
                  "#CD5C5C", #Indian_Red
                  "#FFB800", #Indian_Yellow
                  "#000000", #Liquid_Black
                  "#FFFFFF", #Liquid_Clear
                  "#000000", #Midnight_Black
                  "#0C0040", #Phthalo_Blue
                  "#102E3C", #Phthalo_Green
                  "#021E44", #Prussian_Blue
                  "#0A3410", #Sap_Green
                  "#FFFFFF", #Titanium_White
                  "#221B15", #Van_Dyke_Brown
                  "#C79B00" #Yellow_Ochre
)
```

```
#| label: group-pivot-mutate-bob-ross

#  create a df that shows the frequency of each color use, by season
color_freq <-
  bob_ross |>
  group_by(season) |>
  summarise(across(Black_Gesso:Alizarin_Crimson, sum)) |>
  pivot_longer(cols = !season,
               names_to = 'color',
               values_to = 'count') |>
  mutate(color = case_when(
    color == "Alizarin_Crimson" ~ "Alizarin Crimson",
    color == "Black_Gesso" ~ "Black Gesso",
    color == "Bright_Red" ~ "Bright Red",
    color == "Burnt_Umber" ~ "Burnt Umber",
    color == "Cadmium_Yellow" ~ "Cadmium Yellow",
    color == "Dark_Sienna" ~ "Dark Sienna",
    color == "Indian_Red" ~ "Indian Red",
    color == "Indian_Yellow" ~ "Indian Yellow",
    color == "Liquid_Black" ~ "Liquid Black",
    color == "Liquid_Clear" ~ "Liquid Clear",
    color == "Midnight_Black" ~ "Midnight Black",
    color == "Phthalo_Blue" ~ "Phthalo Blue",
    color == "Phthalo_Green" ~ "Phthalo Green",
```

```
    color == "Prussian_Blue" ~ "Prussian Blue",
    color == "Sap_Green" ~ "Sap Green",
    color == "Titanium_White" ~ "Titanium White",
    color == "Van_Dyke_Brown" ~ "Van Dyke Brown",
    color == "Yellow_Ochre" ~ "Yellow Ochre"))
```

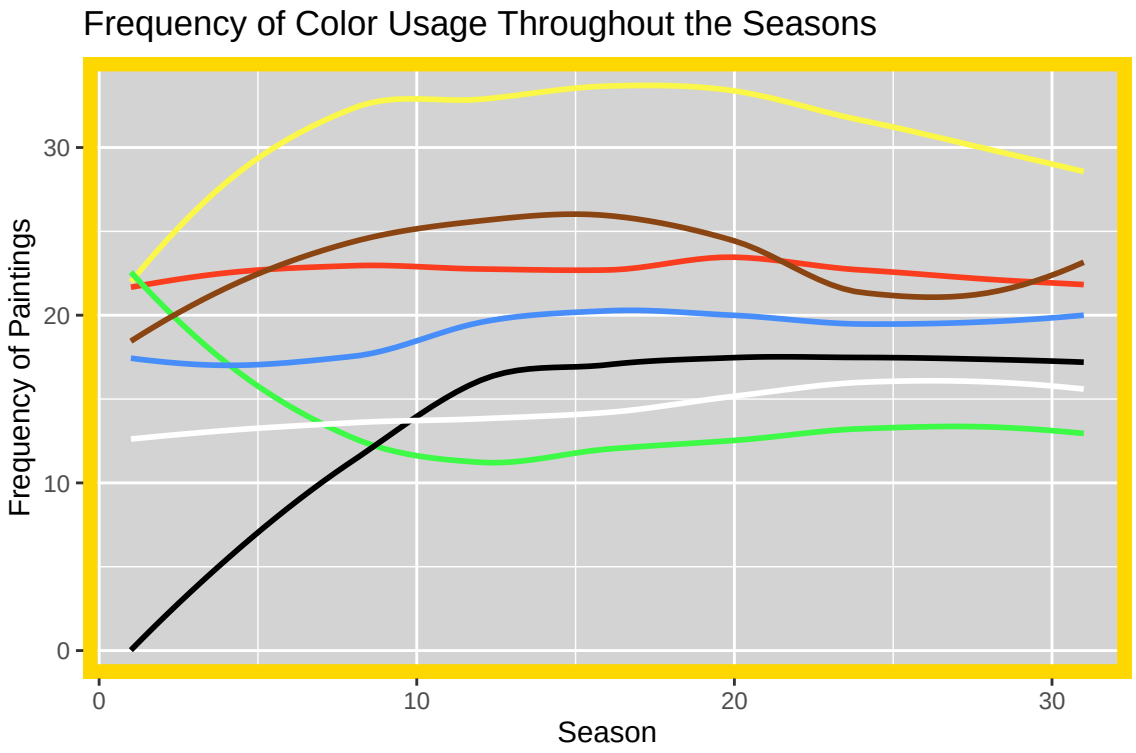
```
# a df that further flattens the colors into a 7-color scheme
color_freq_simpl <- bob_ross |>
  group_by(season) |>
  summarise(across(Black_Gesso:Alizarin_Crimson, sum)) |>
  mutate(
    Red = Alizarin_Crimson + Bright_Red + Indian_Red,
    Yellow = Cadmium_Yellow + Indian_Yellow + Yellow_Ochre,
    Green = Phthalo_Green + Sap_Green,
    Blue = Phthalo_Blue + Prussian_Blue,
    Black = Black_Gesso + Liquid_Black + Midnight_Black,
    Brown = Burnt_Umber + Dark_Sienna + Van_Dyke_Brown,
    White = Liquid_Clear + Titanium_White
  ) |>
  select(season, Red:White) |>
  pivot_longer(cols = !season,
               names_to = "color",
               values_to = "count"
  ) |>
  mutate(color = fct_relevel(color,
            "Red", "Yellow", "Green", "Blue", "Brown", "Black", "White"))

# a df that has the total sums of each color
color_sum_simpl <-
color_freq_simpl |>
  group_by(color) |>
  summarize(
    sum = sum(count))

# simple pie chart
color_sum_simpl |>
  ggplot(aes(x = '',y = sum, color = color, fill = color)) +
  scale_fill_manual(values = roygbiv) +
  geom_bar(stat="identity", color = 'black', show.legend = F) +
  coord_polar("y", start=0) +
  theme_void() +
```

```
  labs(
    title = "Relative Frequency of Colors Used in \"The Joy of Painting\"",
  )
```




```
# simple line graph
color_freq_simpl |>
  ggplot(aes(x = season, y = count, color = color)) +
  geom_smooth(se=F, show.legend = F) + scale_colour_manual(values = roygbiv) +
  labs(
    title = 'Frequency of Color Usage Throughout the Seasons',
    x = 'Season',
    y = 'Frequency of Paintings',
    color = 'Color'
  ) +
  theme(panel.border = element_rect(colour = "gold",
                                    fill=NA,
                                    linewidth=5),
        panel.background = element_rect(fill = "lightgrey"),
```

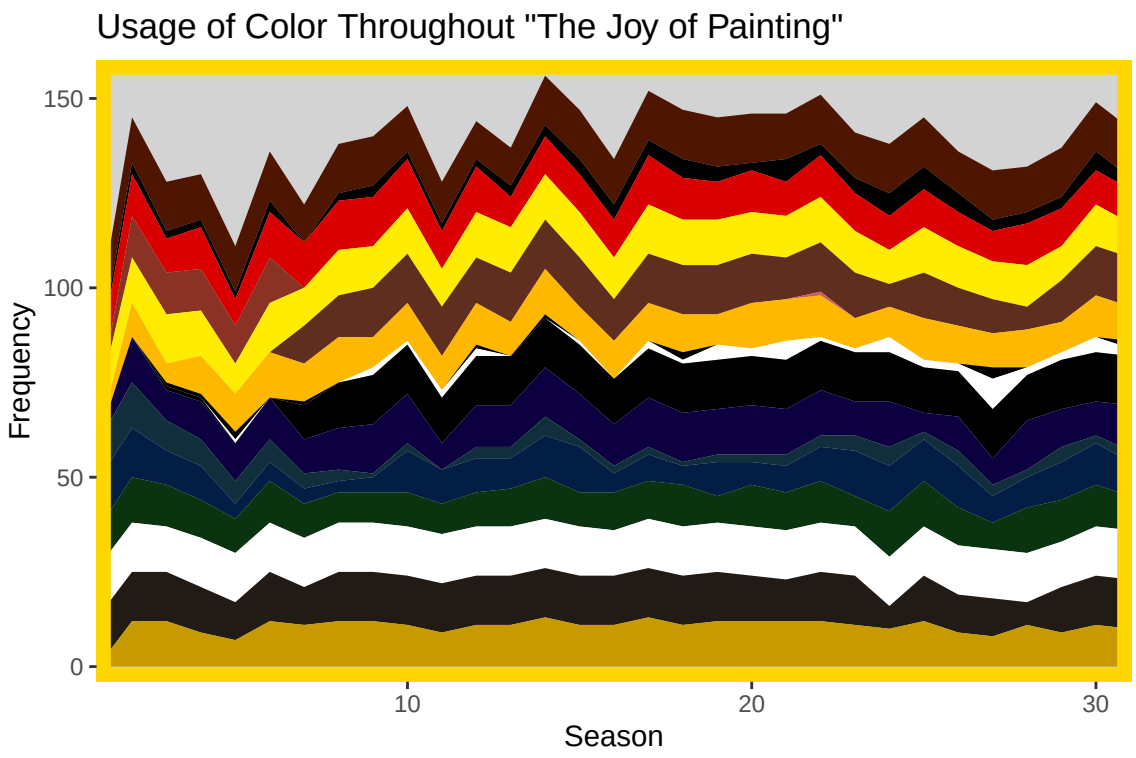
```
)
```



How does Bob Ross use seven colors differently over his 30 seasons of "The Joy of Painting?"

We first categorize his colors into a 7-color palette that loosely follows the rainbow (ROYGBIV baby!). The relative frequency chart indicates that yellow is his most used color, followed by brown and red. In a way, this pie chart looks like a painters palette! (I couldn't figure out how to add a background image). We next examine the line chart to see how colors changed over time. It seems that yellow has remained largely dominant in his works throughout the 30 season, whereas black started with a very low frequency, under 10 paintings for the first 8 seasons, but progressed to around 19 paintings a season. White, brown, blue, and red have stayed relatively the same, hovering around 12, 22, 20, and 22, respectively, with brown having the most variance of around 3 paintings. Yellow and green seem to mirror each other, where green was initially used more than Yellow in Season 1, they soon saw a steep increase and decrease such that Yellow is used more than 30 times each season, and Green around 11 times.

```
#| label: color-freq-plot-area

# plot a stack line chart of each actual color used
color_freq |>
  ggplot(aes(x = season, y = count, fill = color)) +
  geom_area(show.legend = F) + scale_fill_manual(values = pal_colors) +
  theme(panel.border = element_rect(colour = "gold",
                                    fill=NA,
                                    linewidth=5),
        panel.background = element_rect(fill = "lightgrey"),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank()) +
  scale_x_continuous(expand = c(0,0)) +
  scale_y_continuous(expand = c(0,4)) +
  labs(
    title = "Usage of Color Throughout \"The Joy of Painting\"",
    x = 'Season',
    y = 'Frequency'
  )
```



The Mountain Mesa

We have a more granular look with this stacked line chart, where each of his actual colors are displayed. What results is a plot that looks almost like a painting itself! Perhaps it is a landscape of a mountainous mesa, complimented with a lake, forest, and some sand in the foreground. We can choose to show the legend or not, but either way the graph is a bit hard to interpret. We can understand that many light-yellow to gold colors are used in his paintings, as well as dark blues. This makes sense as Bob Ross's subject was often nature, thus his palette would be colors that would likely be resembling those found in nature.

Moreover, we can see that every season, Ross uses anywhere from 110 to 153 unique colors in his paintings. Most colors are used pretty consistently, too. Notably, "Liquid clear" sees little usage until half way through the show's lifespan, as well as "Black Gesso" and "Midnight Black." This shows he already has choices of Black and White, and only choose to expand his palette in later seasons. As such, perhaps his paintings became more detailed and nuanced in color.

Nevertheless, these visualizations are more for the visual appetite rather than the inquiring mind. The lack of a legend, scale ticks, simple and not-entirely-accessible-theme create for plots that ought to be admired, not analyzed - much like the works of Bob Ross.