Assignment 3 - Drafting Visualizations

AUTHOR
Sam Lance

PUBLISHED
February 20, 2025

```
library(ggplot2)
library(tidyverse)
library(here)
library(janitor)
library(patchwork)
library(lubridate)
library(forcats)
library(gghighlight)
library(stringr)
library(maps)
library(ggtext)
library(showtext)
library(scales)
library(rnaturalearth)
library(rnaturalearthdata)
```

1 🧘 Project Intentions

1.1 Chosen Visualization Option

I have chosen Option 1 where I answer one overarching question and address three smaller sub-questions.

1.2 Questions to Answer

Overarching Question:

How are volcanoes spatially distributed/ clustered, and why is this the case?

Sub-Questions:

1. Are volcanoes clustered in specific regions?
2. What is the primary rock type that makes up volcanoes?
3. What geologic setting do volcanoes primarily form in?

Changing My Question:

After looking at my data, I realized there was significant clustering around the Ring of Fire, with not a lot of interesting temporal trends. This led me to pivot my infographic from being more of a volcano explanation document to one helping people to see why most volcanoes form around the Pacific Ocean.

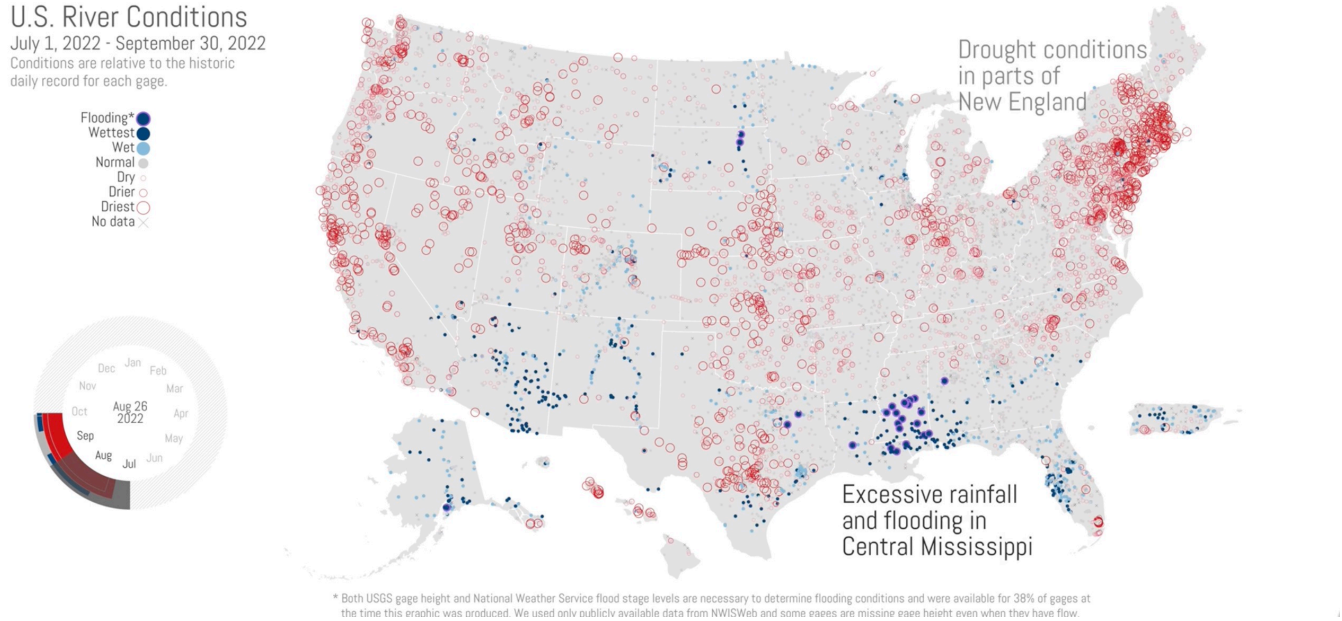
1.3 Variables Used to Answer Questions

My dataset, downloaded from the Smithsonian Global Volcanism Program, is primarily categorical data, with every global volcano described with an array of traits. To understand the spatial distribution of this data, I utilized the latitude and longitude of each volcano, as well as the volcano type (volcanic landform) to show that the majority of volcanoes were of a specific type. To understand more about that specific volcano type (composite volcanoes) I looked at the rock type and geologic setting to see if they align with the characteristics of composite volcanoes.

1.4 Inspiration Visualizations

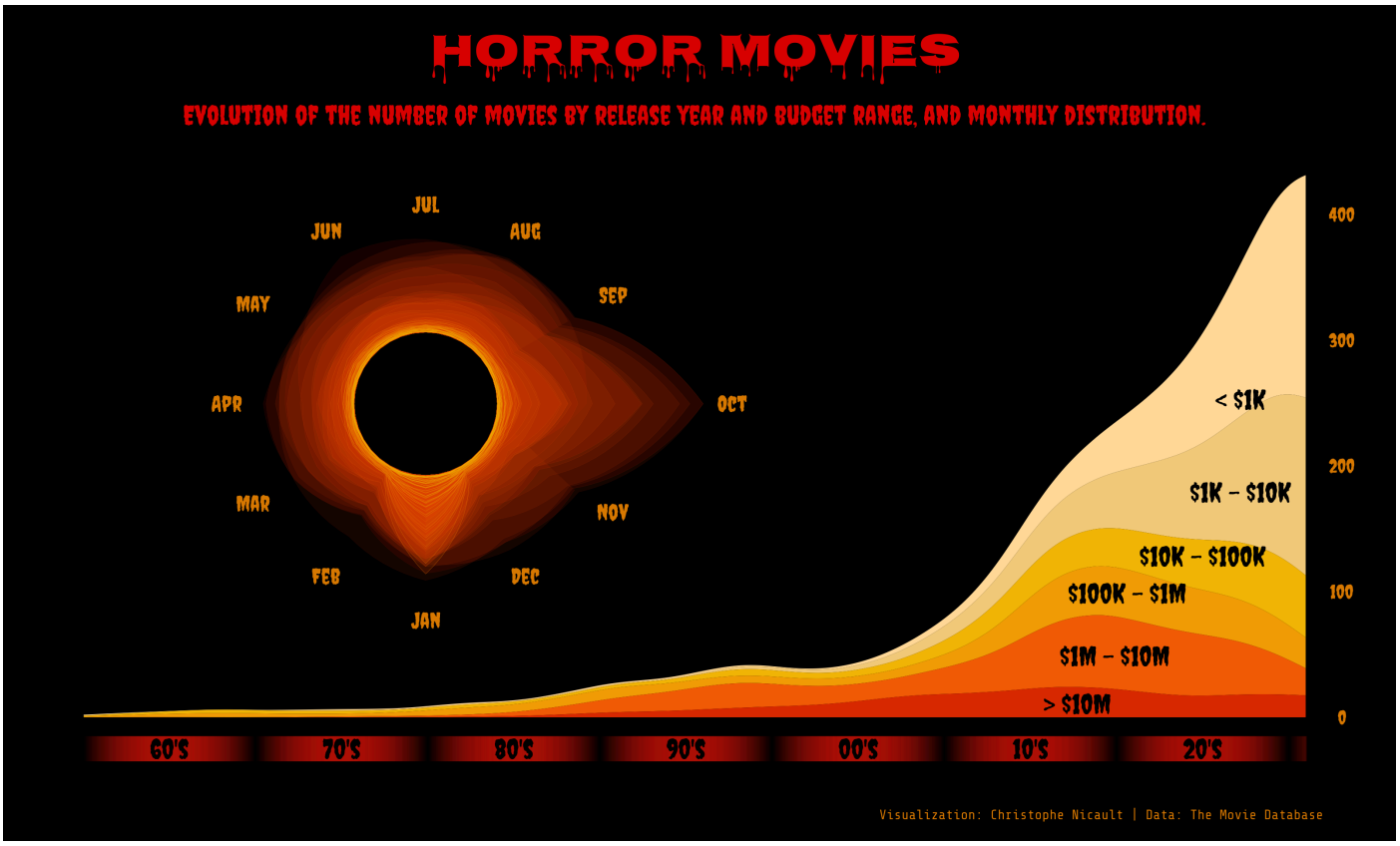
Inspiration Visualization 1:

As I was making my map, this visualization of US River Conditions from Elmera Azadpour inspired my layout. I enjoyed the relatively simple grey basemap and legend, with annotations throughout to provide the majority of the content. While I have chosen a darker theme overall, the simplicity of this plot was my guiding point.



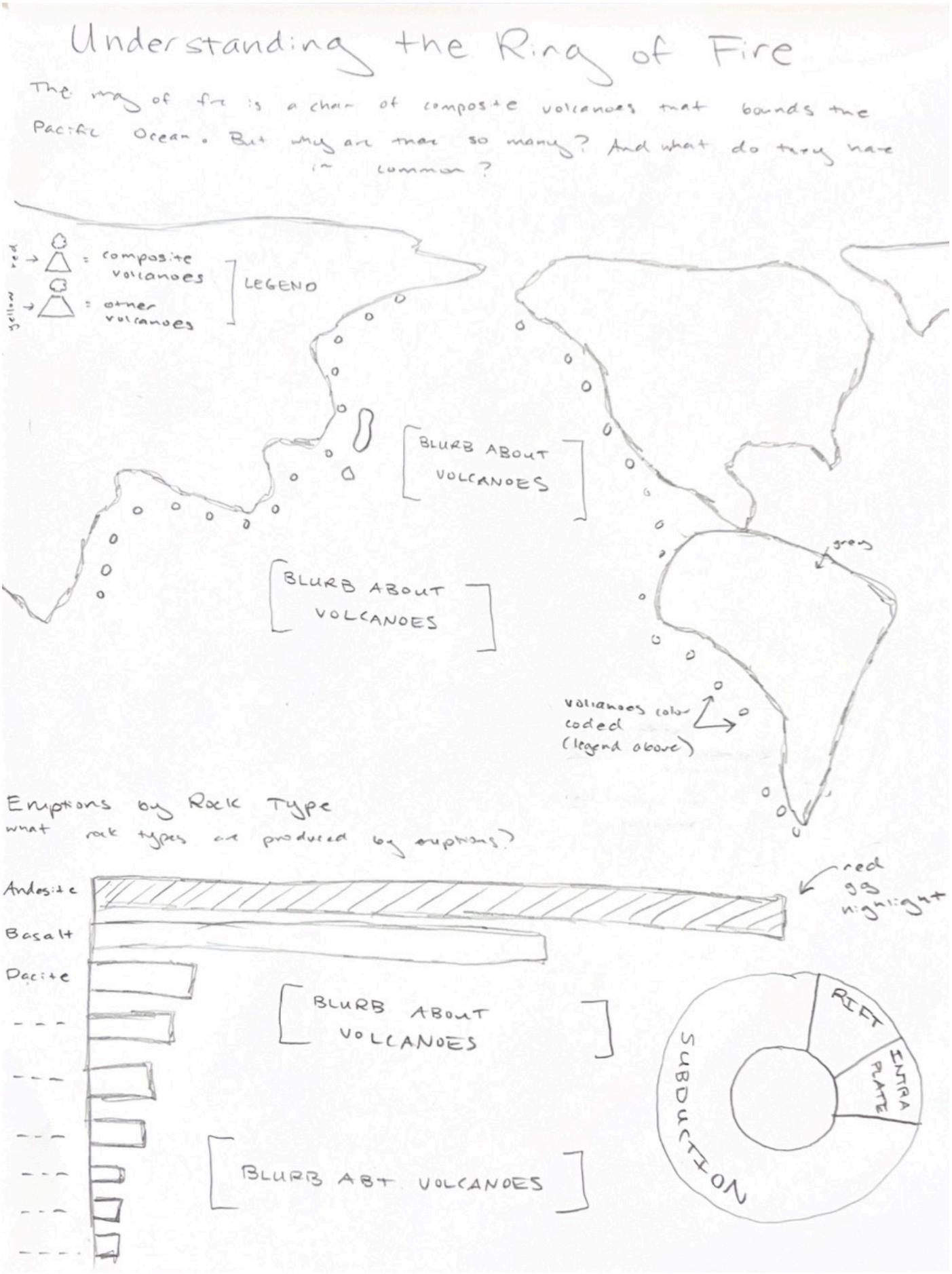
Inspiration Visualization 2:

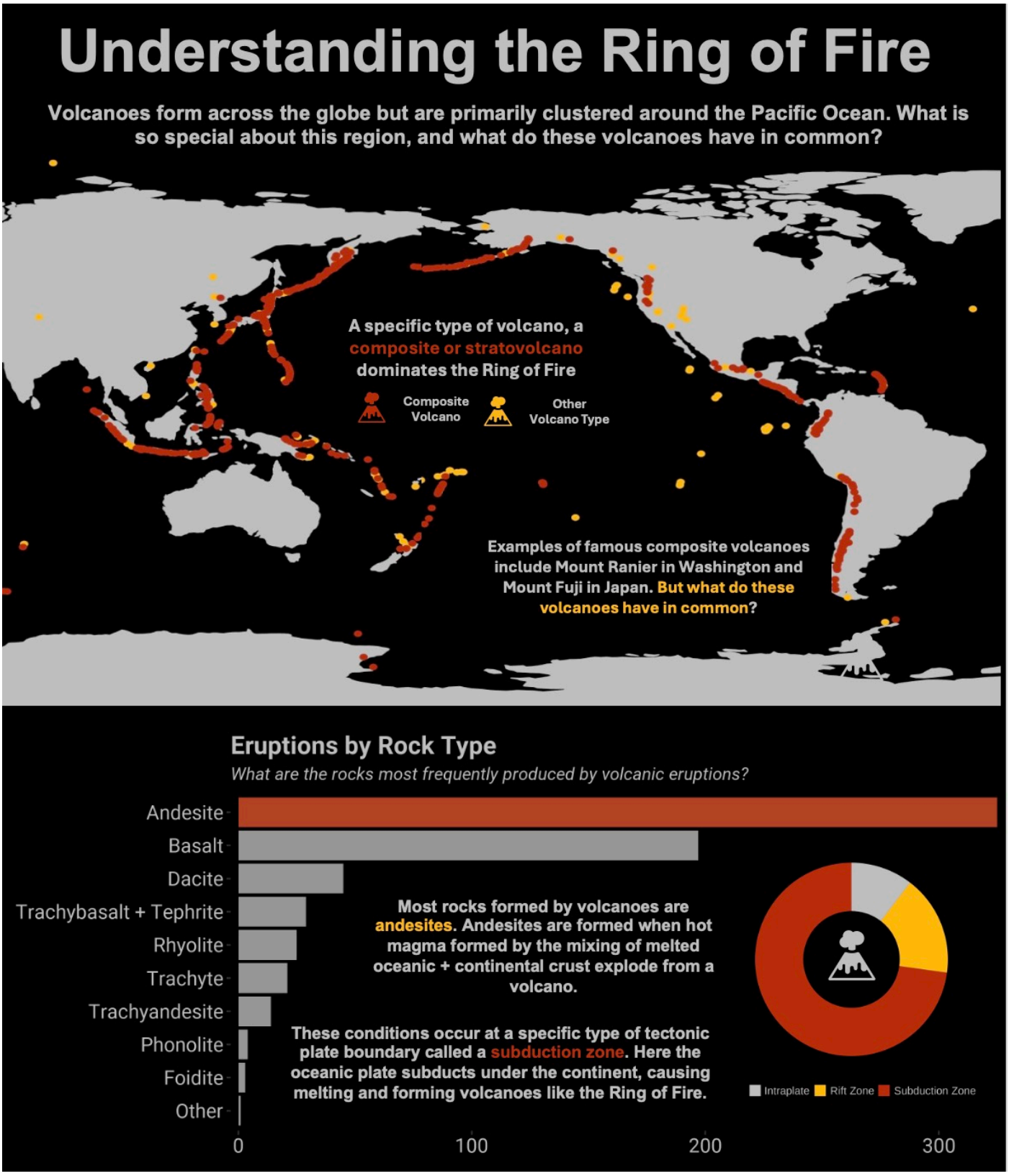
This visualization from Christophe Nicault from the text and themes class inspired the color scheme of my plot. I thought for the topic of volcanism, a darker theme with red accents would be perfect, and I enjoyed the use of different hues of a similar theme (red, orange, yellow).



1.5 Hand-Drawn + Powerpoint Visualization Drafts

Note: I know my mockup version is blurry… I'm working on it :)


Understanding the Ring of Fire
The ring of fire is a chain of composite volcanoes that bounds the Pacific Ocean. But why are there so many? And what do they have in common?
red
= composite volcanoes
yellow
= other volcanoes
LEGEND
BLURB ABOUT VOLCANOES
BLURB ABOUT VOLCANOES
grey
volcanoes color coded (legend above)
Eruptions by Rock Type
what rock types are produced by eruptions?
Andesite
Basalt
Dacite
red bg highlight
BLURB ABOUT VOLCANOES
BLURB ABT VOLCANOES
RIFT
INTRA PLATE
SUBDUCTION



2 ✍️ Initial Visualization Drafts

2.1 Note Before Grading

All graphics here are the version that would render best in Quarto, I have all my actual graphics saving to the right proportions in ggsave in another file. If you could, could you primarily review the mock graphic above when providing feedback so you can see everything as it should appear? Thank you! :)

2.2 Loading Data

```
#LOADING INITAL DATA
volcano_og <- read.csv(here("data", "volcano_list.csv"))
volcano_pop_og <- read.csv(here("data", "volcano_pop.csv"))

#CREATING MODIFIED VOLCANO DATA
volcano_mod <- volcano_og |>
  clean_names() |>
  filter(last_known_eruption != "Unknown") |>
  mutate(
    eruption_year = case_when(
      grepl("BCE", last_known_eruption) ~ -as.numeric(gsub(" BCE", "", last_known_erupti
      grepl("CE", last_known_eruption)  ~ as.numeric(gsub(" CE", "", last_known_eruption
      TRUE ~ as.numeric(last_known_eruption))) |>
  arrange(eruption_year)

volcano_mod_1 <- volcano_mod |>
    filter(eruption_year > 0)

#CREATING ROCK TYPE DF
#levels(as.factor(volcano_mod$dominant_rock_type))

rock_type <- volcano_mod |>
  filter(last_known_eruption != "Unknown") |>
  filter(!dominant_rock_type %in% c("No Data (checked)", "")) |>
   mutate(rock_type = recode(dominant_rock_type,
                             "Andesite / Basaltic Andesite" = "Andesite",
                             "Basalt / Picro-Basalt" = "Basalt",
                             "Trachybasalt / Tephrite Basanite" = "Trachybasalt + Tephrite
                             "Trachyandesite / Basaltic Trachyandesite" = "Trachyandesite"
                             "Phono-tephrite /  Tephri-phonolite" = "Other",
                             "Trachyte / Trachydacite" = "Trachyte"))

#FILTERING FOR COMPOSITE + NON-COMPOSITE VOLCANOES

composite_volcanoes <- volcano_mod_1 |>
  filter(volcano_landform == "Composite")

other_volcanoes <- volcano_mod_1 |>
  filter(volcano_landform != "Composite")
```

2.3 Loading Design Elements

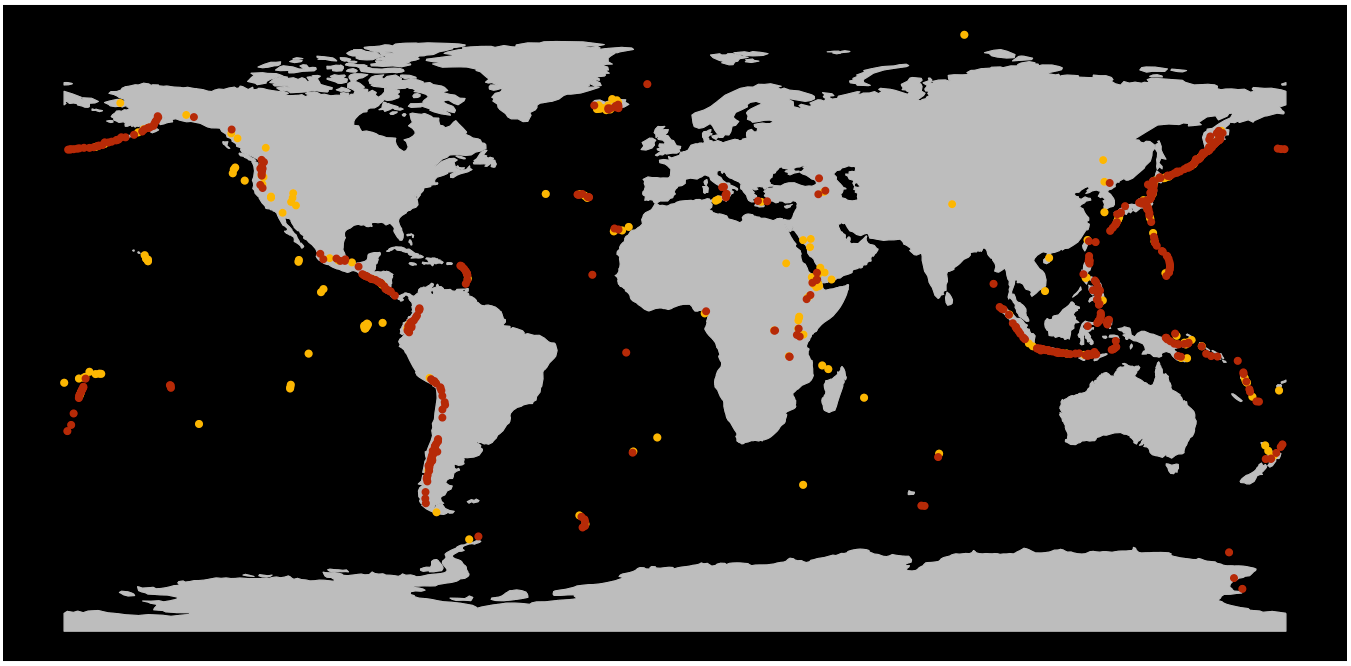
```
#importing colors
dorange <- c("#BA2D0B")
porange <- c("#ffba08")
v_grey <- c("#c0c0c0")
v_black <- c("#000000")

#fonts
font_add_google(name = "Roboto", family = "roboto")
showtext_auto()
```

2.4 Map of Composite + Other Volcanoes

```
#| echo: true
#| message: false

world <- ne_download(scale = 110, type = "countries", category = "cultural")
```

```
Reading layer `ne_110m_admin_0_countries' from data source

`/private/var/folders/7r/2ngr651x2rd0q43kgmpyvwvw0000gn/T/Rtmpw0Pkg9/ne_110m_admin_0_cou
ntries.shp'
  using driver `ESRI Shapefile'
Simple feature collection with 177 features and 168 fields
Geometry type: MULTIPOLYGON
Dimension:     XY
Bounding box:  xmin: -180 ymin: -90 xmax: 180 ymax: 83.64513
Geodetic CRS:  WGS 84
```

```
volcano_map <- ggplot(data = world) +
  geom_sf(fill = v_grey, color = v_grey) +
  geom_point(data = other_volcanoes,
             aes(x = longitude, y = latitude, color = porange),
             size = 0.75) +
  geom_point(data = composite_volcanoes,
             aes(x = longitude, y = latitude, color = dorange),
             size = 0.75) +
  scale_color_identity() +  # Use scale_color_identity to preserve the colors specified
  theme_void() +
  labs(title = NULL,
       x = NULL, y = NULL) +
  theme(legend.position = NULL,
        legend.title = element_blank(),
        plot.background = element_rect(fill = v_black))

volcano_map
```

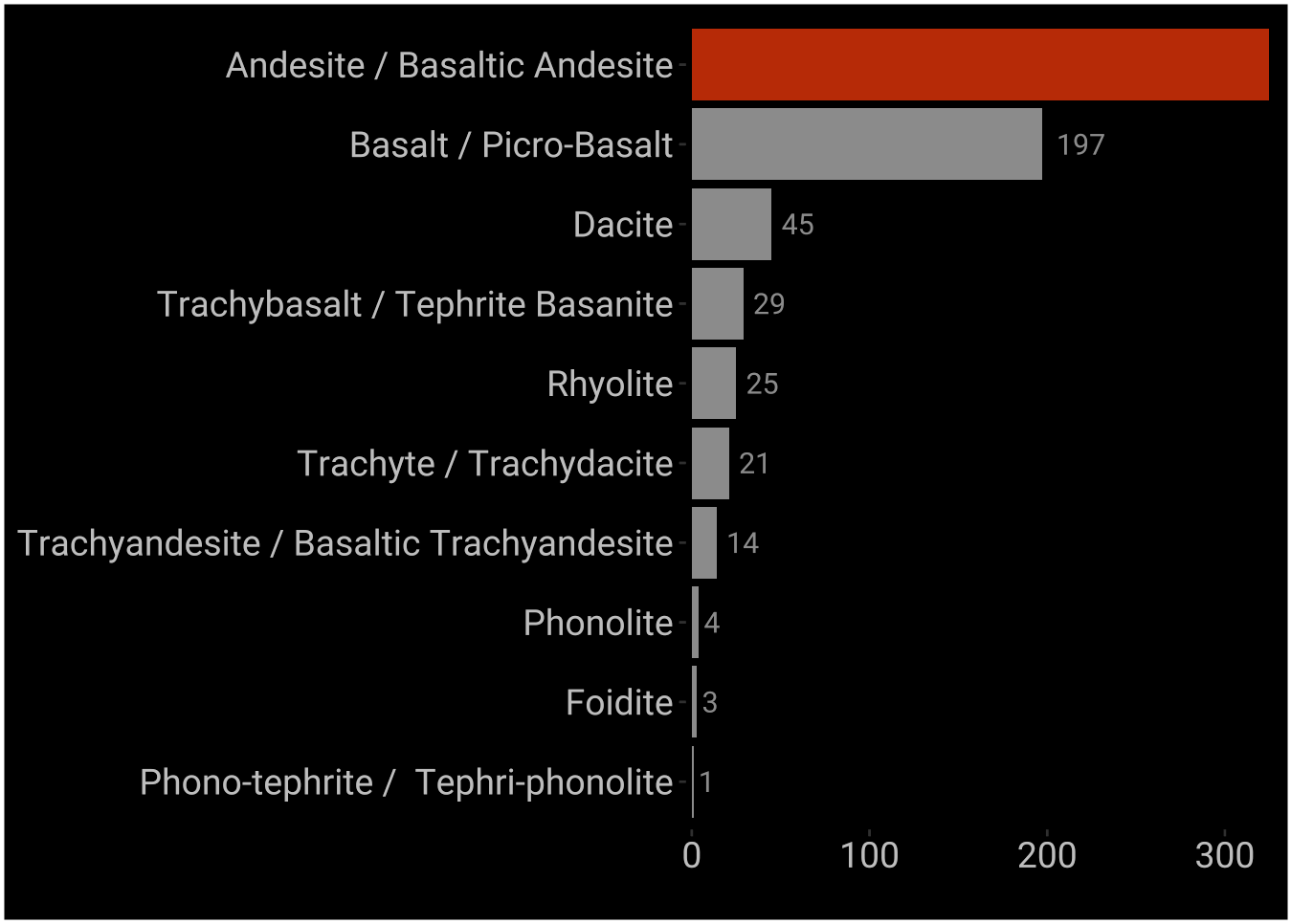


```
ggsave(
  filename = here::here("images", "volcano_map.png"),
  plot = volcano_map,
  device = "png",
  width = 8,
  height = 7,
  unit = "in"
)
```

2.5 Majority Rock Type of Volcanoes

```
#| echo: true
#| message: false

rock_type <- rock_type |>
  filter(eruption_year > 0) |>
  mutate(dominant_rock_type = fct_rev(fct_infreq(dominant_rock_type)))

rock_plot <- rock_type |>
  ggplot(aes(x = dominant_rock_type)) +  # Reverse the order of the factor levels
  geom_bar(fill = dorange) +
  geom_text(stat = "count", aes(label = ..count..), hjust = -0.3,
```

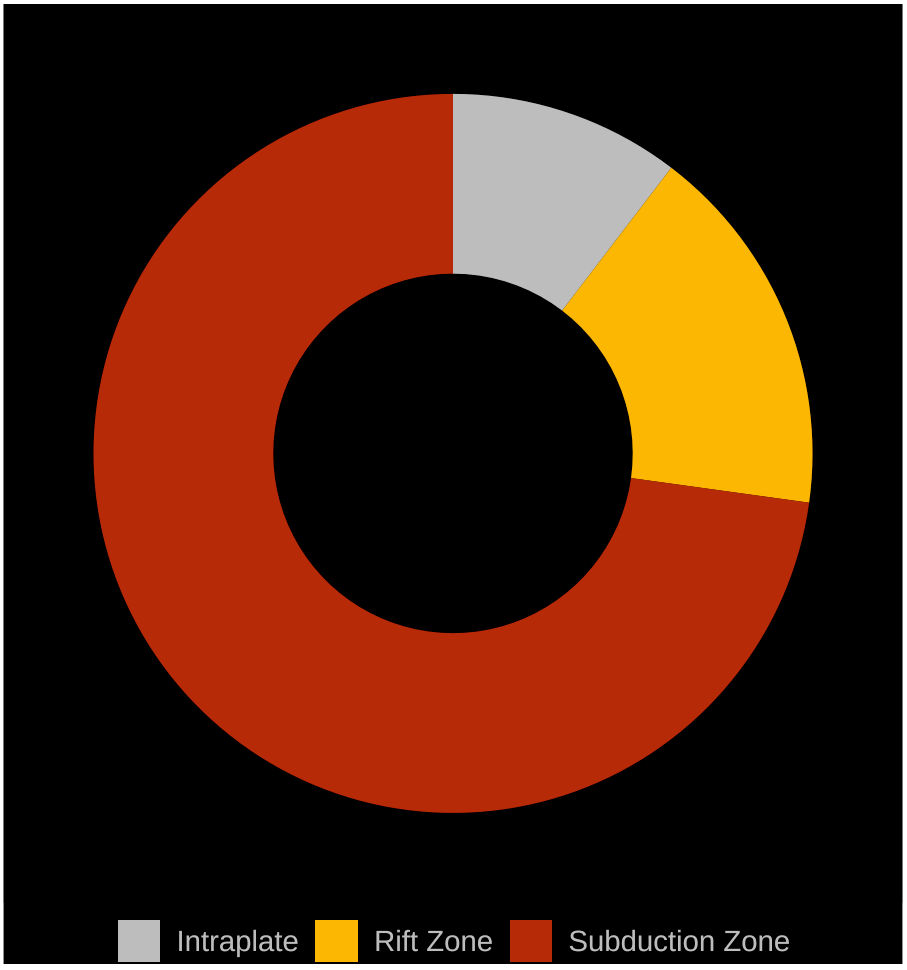
```
            color = v_grey, family = "roboto", size = 4) +
  coord_flip() +
  scale_y_continuous(expand = expansion(mult = 0.01)) +
  #scale_x_discrete(labels = function(x) strwrap(x, width = 15)) +  # Wrap text after 15
  labs(x = NULL,
       y = "Total Number of Eruptions",
       title = NULL) +
  gghighlight(dominant_rock_type == "Andesite / Basaltic Andesite",
              use_direct_label = FALSE) +
  theme(axis.text.x = element_text(color= v_grey, size = 14, family = "roboto"),
        axis.text.y = element_text(color= v_grey, size = 14, family = "roboto"),
        axis.title.y = element_text(color= v_grey, size = 14, family = "roboto"),
        plot.background = element_rect(fill = v_black),
        panel.background = element_rect(fill = v_black),
        panel.grid = element_line(color = v_black))

ggsave(
  filename = here::here("images", "rock_type.png"),
  plot = rock_plot,
  device = "png",
  width =9,
  height =4,
  unit = "in"
)
rock_plot
```



2.6 Geologic Setting of Volcano Formation

```
#| echo: true
#| message: false

volcano_loc <- read_csv(here("data", "volcano_count.csv"))

# Compute percentages
volcano_loc$fraction = volcano_loc$eruption_count / sum(volcano_loc$eruption_count)

# Compute the cumulative percentages (top of each rectangle)
volcano_loc$ymax = cumsum(volcano_loc$fraction)

# Compute the bottom of each rectangle
volcano_loc$ymin = c(0, head(volcano_loc$ymax, n=-1))

# Make the plot
zone <- ggplot(volcano_loc, aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=tectonic_sett
  geom_rect() +
  coord_polar(theta="y") +
  xlim(c(2, 4)) +
  scale_fill_manual(values = c("#c0c0c0",  "#ffba08", "#BA2D0B" )) +
  theme_void() +
```

```
  theme(plot.background = element_rect(fill = v_black),
        panel.background = element_rect(fill = v_black),
        legend.text = element_text(color = v_grey, size = 11),
        legend.title = element_blank(),
        legend.position = "bottom")

zone
```




```
ggsave(
  filename = here::here("images", "form_zone.png"),
  plot = zone,
  device = "png",
  width =5,
  height =5,
  unit = "in"
)
```

3 🤔 Reflecting on Process

3.1 Challenges Faced

My primary struggle that I am continuing to work through is my map. It has been incredibly difficult for R to use a projection centered around the Pacific Ocean where I can fully see the trend of the Ring of Fire. Currently, my solution is creating two clipped versions of the map, and pasting them together in my final visualization, which, while crude, has been very effective.

Additionally, I am still working on fine tuning my graphs. I think I need to slightly change some of my colors, work on my text wrapping in my donut plot, and think more about the colors within my bar chart.

3.2 GGPlot Extensions/ Tools Used

I am currently using RNaturalEarth in order to make my maps, but no other extensions.

3.3 Feedback Needed from Instructors

I would love feedback on my annotations and if the general message of the plot is clear for non Earth Scientists. The overall message is meant to first show the ring of fire, then explain how it forms based on the data from actual volcanoes, but am unsure if I am accomplishing that now.