

Homework 5

Ethan Chang - ehc586

This homework is due on Feb. 28, 2023 at 11:00pm. Please submit as a pdf file on Canvas.

Problem 1: (4 pts) We will work with the `ufo_sightings` dataset.

Since 2000 (inclusive), what are the top 10 cities that have reported the most UFO sightings? Create a new dataframe to answer the question. No plots are necessary.

```
ufo_sightings %>%  
  filter(year >= 2000) %>%  
  count(city) %>%  
  arrange(-n) %>%  
  head(10)
```

```
## # A tibble: 10 x 2  
##   city      n  
##   <chr>   <int>  
## 1 seattle  406  
## 2 phoenix  360  
## 3 portland 314  
## 4 las vegas 303  
## 5 los angeles 294  
## 6 san diego 279  
## 7 houston  224  
## 8 chicago  218  
## 9 miami    198  
## 10 tucson   196
```

Problem 2: (4 pts)

Next, how has the number of UFO sightings changed for five states since 1940? Please follow these steps:

1. Filter the dataset to keep the following five states: AZ, IL, NM, OR, WA
2. Keep only the records from 1940 and onwards.
3. Find the number of records for each year and state.
4. Output the new table below your code block.

Your final table should be in long format and have three columns, `year`, `state`, and `count`. You will plot this table in Problem 3.

```
new_df <- ufo_sightings %>%  
  filter(state %in% c("AZ", "IL", "NM", "OR", "WA")) %>%  
  filter(year >= 1940) %>%  
  count(year, state) %>%  
  rename(count = n)
```

```
new_df
```

```
## # A tibble: 288 x 3  
##   year state count
```

```
##      <dbl> <chr> <int>
## 1  1944 NM      2
## 2  1945 NM      2
## 3  1946 AZ      1
## 4  1946 IL      1
## 5  1946 WA      1
## 6  1947 IL      2
## 7  1947 NM      2
## 8  1947 WA      2
## 9  1949 WA      2
## 10 1950 OR      1
## # ... with 278 more rows
```

Problem 3: (2 pts)

Use the new dataframe you made in Problem 2 and add an appropriate color scale from the `colorspace` package to the plot below.

```
new_df %>%
  ggplot(aes(x = year, y = count, color = state)) +
  geom_line() +
  xlab("Year") +
  ylab("UFO Sightings (Count)") +
  theme_bw() +
  scale_color_discrete_divergingx(palette = "Zissou 1")
```

