Data Visualization II: Multivariate Plots - Homework (Answer Key)

Ben Graham and Gaea Morales

Summer 2022

Creating more complex figures

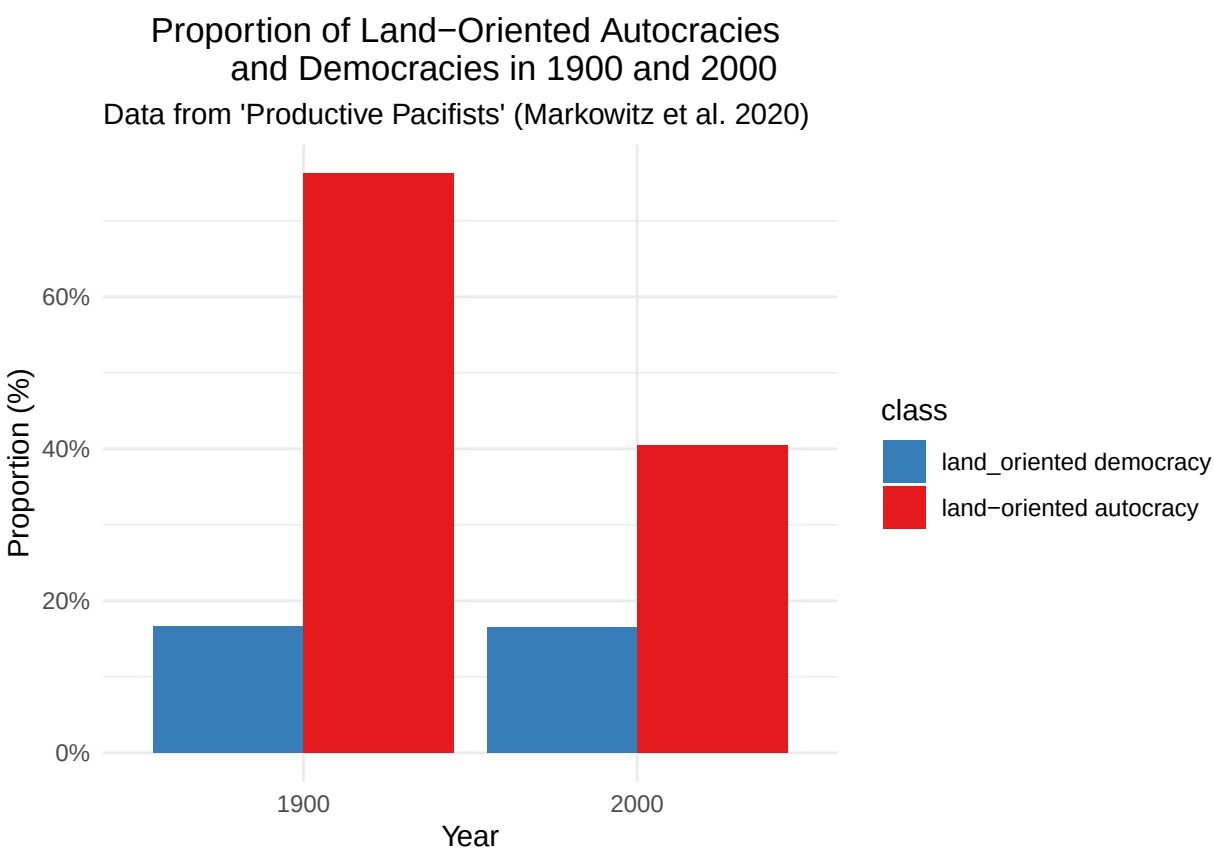
Complete the following assignment and then save your R script and figures into the associated google drive folder. The R script should be titled HW_DV2_[YOUR INTIALS].

We will be working with different data from those used in the walkthrough and group work. Load in `ProductivePacifists_Data.RData`, which is located in the "Training Data" folder, as well as the relevant packages.

Exercise 1: Create a barplot that compares the proportion of democracies and autocracies that are land oriented in 1900 and in 2000 (i.e., two pairs of two bars). Use the `land_oriented_medium_binary` and `autocracy_BX`. Add an appropriate title and axis labels.

```
library(tidyverse)
library(ggplot2)

load("ProductivePacifists_Data.RData")

# Preprocessing using the tidyverse

# Small dataset of the number of countries for 1900 and 2000
countries <- dat %>%
  filter(year %in% c(1900, 2000)) %>%
  group_by(year) %>%
  summarise(country_count = n())

# Create a small dataset that counts the number of
  # autocracies and democracies that are land oriented
  # to merge with country-count object and ultimately generate
  # the proportions for each

regime_prop <- dat %>%
  group_by(year, autocracy_BX, land_oriented_medium_binary) %>%
  summarise(count = n()) %>% # the number of countries in each year,
    # that are either autocratic or not, and land-oriented or not
  filter(year %in% c(1900, 2000)) %>%
    # filter only for the years of interest
  left_join(countries, by = "year") %>%
  mutate(prop = count/country_count) %>%
    # get the proportion of countries for each combination using
      # the merged country count variable
```

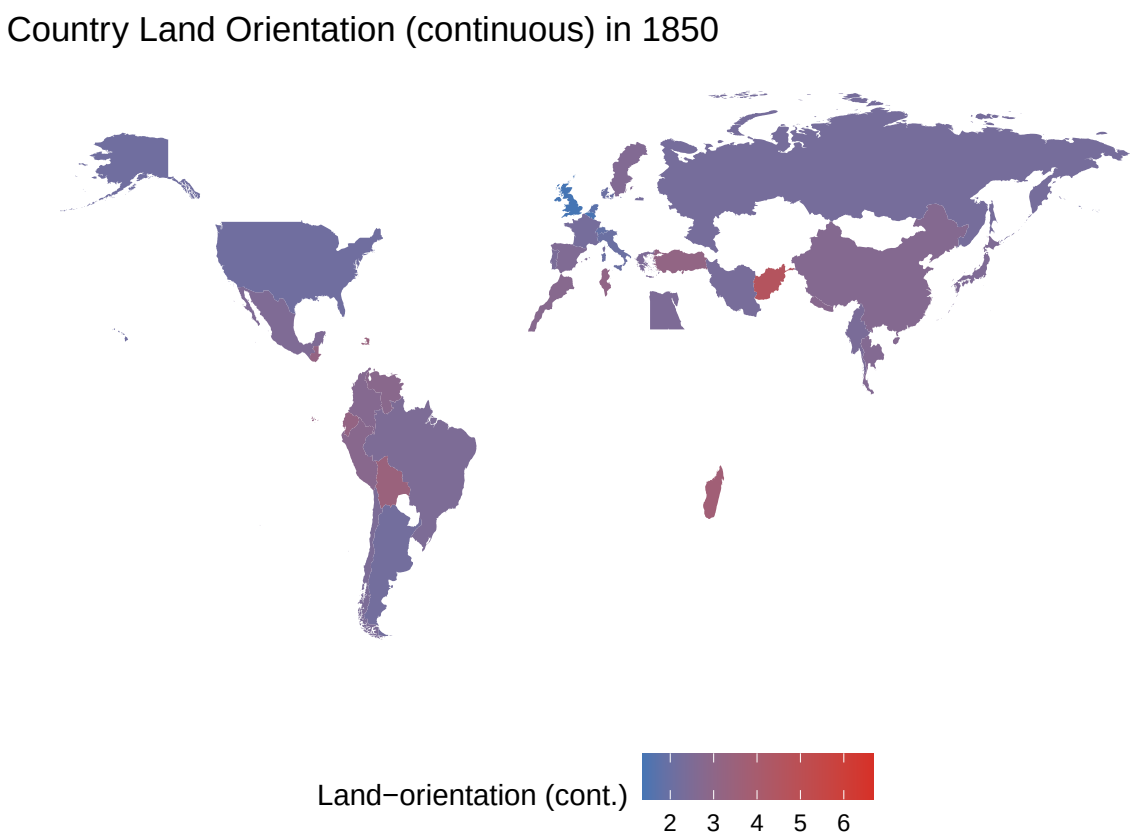
```r
  filter(!(land_oriented_medium_binary == 0),
         !is.na(autocracy_BX)) %>%
      # filter only for the land-oriented cases
  mutate(class = ifelse(autocracy_BX == 1, "land-oriented autocracy",
                        "land_oriented democracy"))

ggplot(regime_prop, aes(x = factor(year), y = prop, group = class, fill = class)) +
    # we want to set year as factor so that we don't have extraneous years plotted
  geom_bar(stat = "identity", position = "dodge") + # use own y-axis and unstack barplots
  scale_fill_manual(values = c("#377eb8", "#e41a1c")) +
  theme_minimal() + # remove gray area and minor grid lines
  scale_y_continuous(labels = scales::percent) + # change scale to percent
  labs(x = "Year", y = "Proportion (%)", title = "Proportion of Land-Oriented Autocracies
       and Democracies in 1900 and 2000",
       subtitle = "Data from 'Productive Pacifists' (Markowitz et al. 2020)") +
    # changing labels
  theme(plot.title = element_text(hjust = 0.5))
```



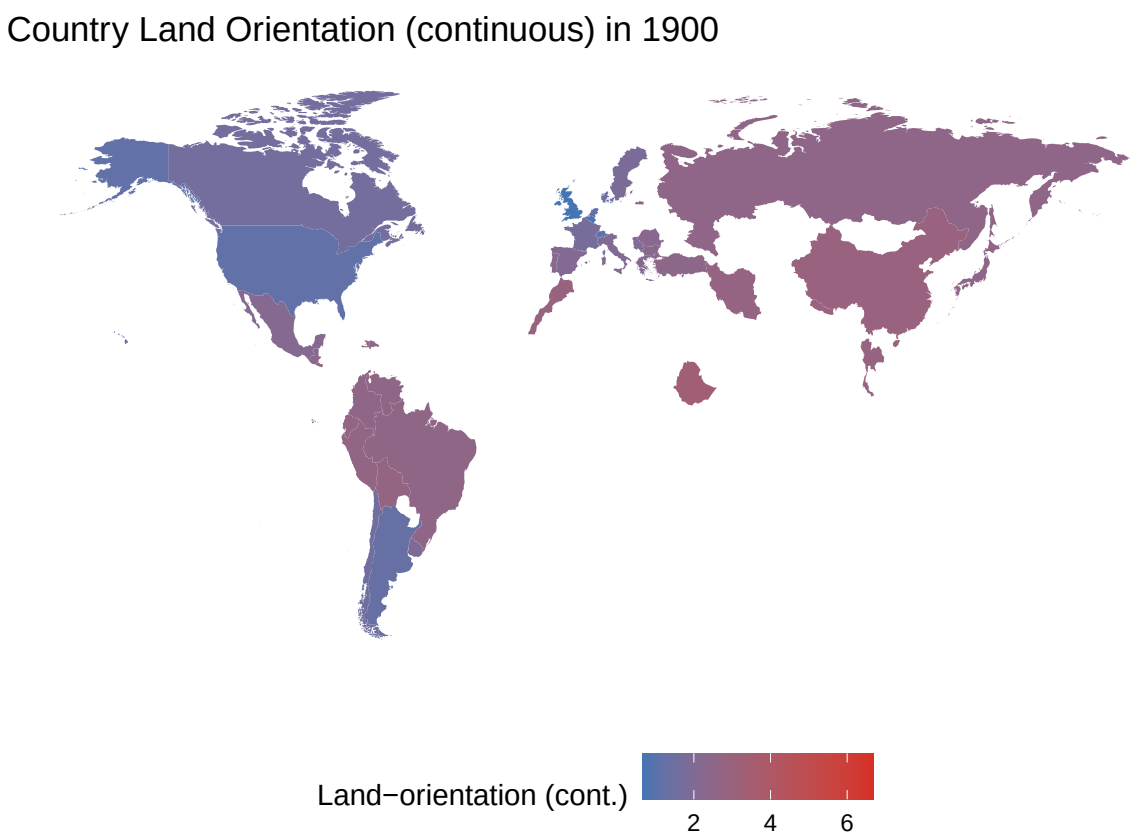
Exercise 2: Create four map plots, one with data from 1850, one from 1900, one from 1950, one from 2000. In each plot, show a heatmap of how land-oriented each country is. Use the variable `land_oriented_medium_continuous_trim`.Make sure to load in the necessary packages, and add the necessary information for each plot (title and labels, legend).

```r
library(maps)
library(ggplot2)
library(dplyr)
library(countrycode)
```

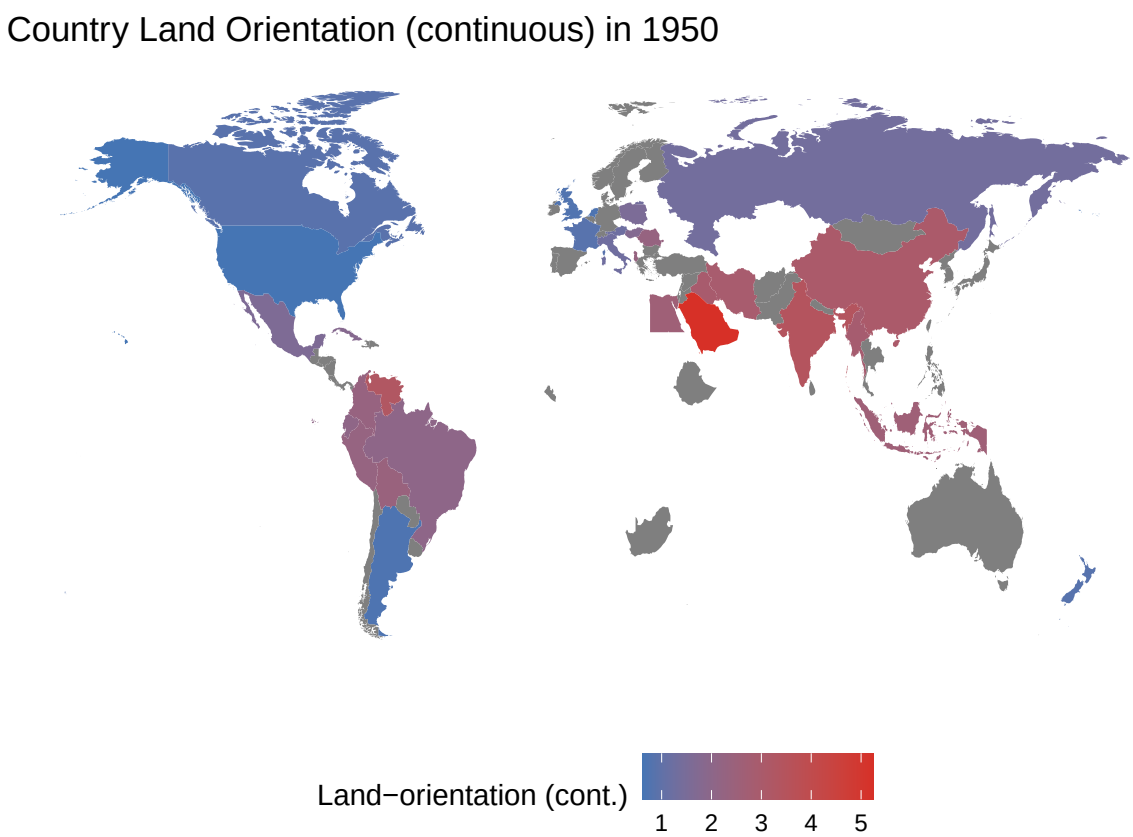
```r
# Extract map data
dat_map <- map_data("world")

# Pre-processing (adding country variables to merge on to map data)
dat_map$gwno <- countrycode(dat_map$region,
                            origin = "country.name",
                            destination = "gwn") # using the Gleditsch & Ward
                               # countrycode

  # Note: we already have the country code (gwno) in the productive pacifists data

# Merge map data and productive pacifists data
merged <- full_join(dat_map, dat, by = "gwno")

# 1850
p1850 <- ggplot(subset(merged, year == 1850),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "Country Land Orientation (continuous) in 1850",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +

# BONUS: in maps, you can also remove the whole background using "element_blank()"
    # as we don't technically need the grid lines
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
        axis.ticks.x = element_blank(), axis.ticks.y = element_blank(),
        axis.text.x = element_blank(), axis.text.y = element_blank(),
        legend.position = "bottom")
p1850
```



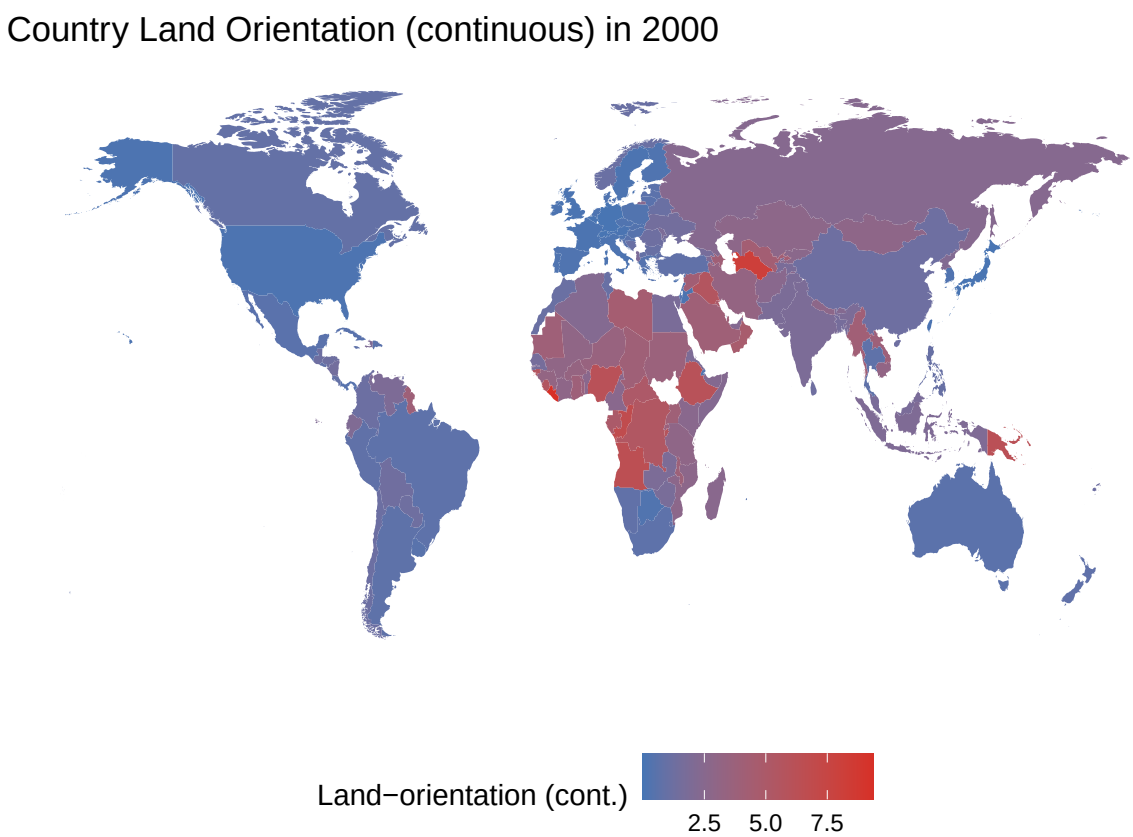

```
# 1900

p1900 <- ggplot(subset(merged, year == 1900),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "Country Land Orientation (continuous) in 1900",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +

# BONUS: in maps, you can also remove the whole background using "element_blank()"
    # as we don't technically need the grid lines
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
        axis.ticks.x = element_blank(), axis.ticks.y = element_blank(),
        axis.text.x = element_blank(), axis.text.y = element_blank(),
        legend.position = "bottom")
p1900
```




```r
# 1950

p1950 <- ggplot(subset(merged, year == 1950),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "Country Land Orientation (continuous) in 1950",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +

# BONUS: in maps, you can also remove the whole background using "element_blank()"
    # as we don't technically need the grid lines
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
        axis.ticks.x = element_blank(), axis.ticks.y = element_blank(),
        axis.text.x = element_blank(), axis.text.y = element_blank(),
        legend.position = "bottom")
p1950
```



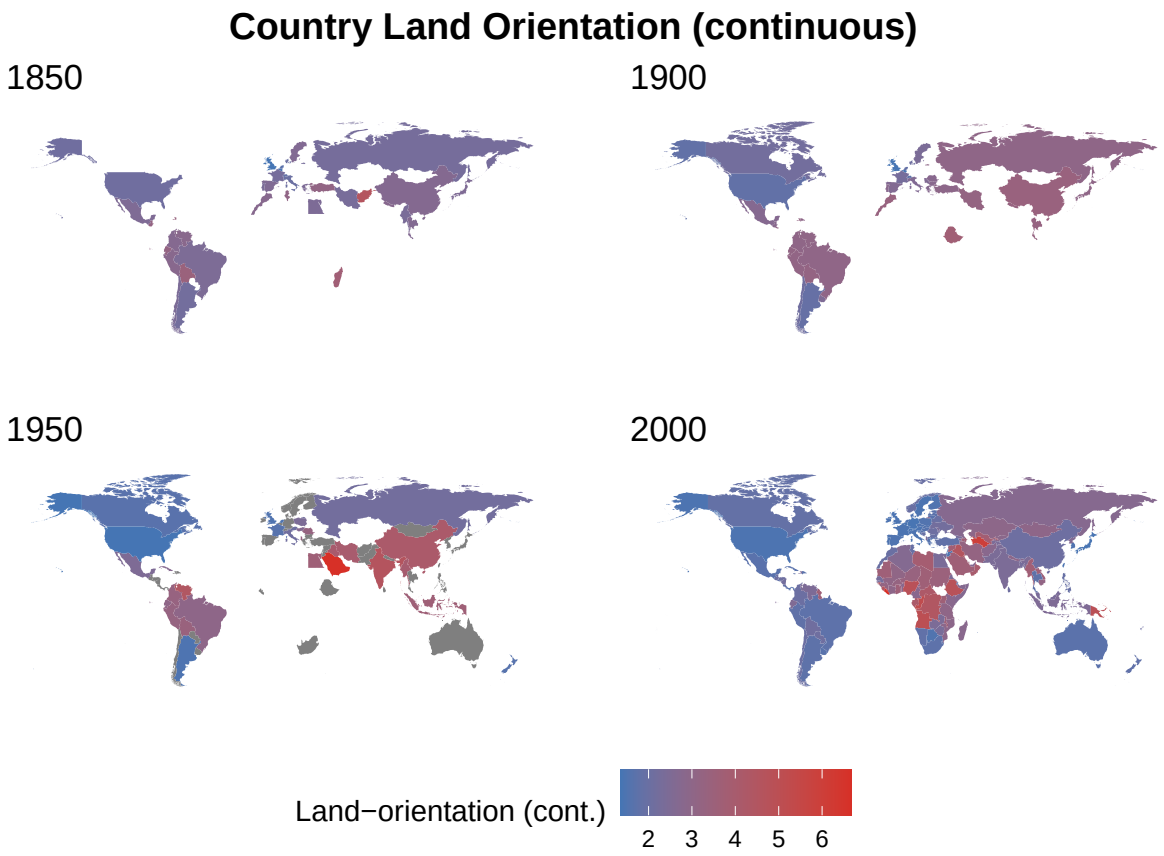

```
# 2000

p2000 <- ggplot(subset(merged, year == 2000),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "Country Land Orientation (continuous) in 2000",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +

# BONUS: in maps, you can also remove the whole background using "element_blank()"
    # as we don't technically need the grid lines
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
        axis.ticks.x = element_blank(), axis.ticks.y = element_blank(),
        axis.text.x = element_blank(), axis.text.y = element_blank(),
        legend.position = "bottom")
p2000
```



Bonus: Combine all four plots into a single figure. Label each subplot with the year. Add a legend.

```
library(ggpubr)
# Resaving plots with an adjustated title (year only)

p1850 <- ggplot(subset(merged, year == 1850),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "1850",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
        axis.ticks.x = element_blank(), axis.ticks.y = element_blank(),
        axis.text.x = element_blank(), axis.text.y = element_blank())

p1900 <- ggplot(subset(merged, year == 1900),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "1900",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
```

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

p1950 <- ggplot(subset(merged, year == 1950),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "1950",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
        axis.ticks.x = element_blank(), axis.ticks.y = element_blank(),
        axis.text.x = element_blank(), axis.text.y = element_blank())

p2000 <- ggplot(subset(merged, year == 2000),
  aes(x = long, y = lat, group = group, fill = land_oriented_medium_continuous_trim)) +
  geom_polygon() +
  scale_fill_gradient(low = "#4575b4", high = "#d73027") +
  labs(title = "2000",
       x = "",
       y = "",
       fill = "Land-orientation (cont.)") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_blank(),
        axis.ticks.x = element_blank(), axis.ticks.y = element_blank(),
        axis.text.x = element_blank(), axis.text.y = element_blank())


p_all <- ggarrange(p1850, p1900, p1950, p2000, ncol =2, nrow =2,
                   common.legend = T, legend = "bottom")
  # combines the plots, and specifies two columns and two rows for the figure,
    # that the plots share one legend, and to place the legend at the bottom

annotate_figure(p_all, top = text_grob("Country Land Orientation (continuous)",
               color = "black", face = "bold", size = 14)) # creates one common title
```


Country Land Orientation (continuous)
1850
1900
1950
2000
Land−orientation (cont.)
2
3
4
5
6