Problem set 1

Your name here

Due 10/4/2021 at 5pm

NOTE: *Start with the file* `ps1_2021.Rmd` *(available from the github repository at https://github.com/aeggers/IntroQSS-F21/tree/main/assignments). Modify that file to include your answers. Make sure you can "knit" the file (e.g. in RStudio by clicking on the* `Knit` *button). Submit both the Rmd file and the knitted PDF via XXXX.*

Question 1: US presidential election results

Load the data:

```
data_path <- "https://raw.github.com/aeggers/IntroQSS-F21/main/data/"
df <- read_csv(str_c(data_path, "yearly_county_pres_results_wide.csv"))
```

This data includes `dem_vote_share` (the share of the two-party vote won by the Democratic candidate) and `dem_county_share` (the share of counties won by the Democratic candidate) in each U.S. presidential election since 1960.

1a) Make a plot showing the Democratic vote share (vertical axis) in each year (horizontal axis). Draw a point for each year and connect them with a line.

```
## replace this with your 1a plot code
```

1b) Make a plot showing the Democratic county share (vertical axis) in each year (horizontal axis).

```
## replace this with your 1b plot code
```

1c) Now make a plot showing Democratic county share (vertical axis) and Democratic vote share (horizontal axis), again connecting the points with a line. (Hint: use `geom_path()`.) Label each point with the corresponding year. (Hint: use `geom_text()`.)

```
## replace this with your 1c plot code
```

Now load this dataset, which is the same data organized differently:

```
df2 <- read_csv(str_c(data_path, "yearly_county_pres_results_long.csv"))
```

1d) Make a plot showing both the Democratic vote share and Democratic county share (vertical axis) in each year (horizontal axis), with a different color for each series. Your figure should include a legend.

```
# replace this with your 1d plot code
```

1e) Which of these two plots do you prefer, and why?

Question 2: democracy and GDP

Load the data. This is an extract from the V-Dem dataset (https://www.v-dem.net/). Variables include `country_name` and `continent_name` (self-explanatory), `polyarchy` (V-Dem's measure of democracy), `pop` (World Bank measure of population), and `gdppc` (GDP per capita), all from 2010. The full dataset (available in the `vdemdata` `R` package) contains many more variables and years.

```
vd <- read_csv(str_c(data_path, "vdem_2010_extract.csv"))
```

2a) Make a scatterplot of the V-Dem polyarchy score (vertical axis) against GDP per capita (horizontal axis). Make the color of the dots reflect the continent, and the size reflect the population. Show the horizontal axis on the log scale.

```
# replace this with your 2a plot code
```

2b) Now make the same figure faceted by population. Add a linear regression line (use `geom_line(method = lm)`). How does the relationship between GDP per capita and democracy differ across continents?

```
# replace this with your 2b plot code
```