Skill Check 1.2 - Group 4
Data on maps - geom_map()

By Sarah, Spencer, and David

Characteristics of geom_map() type data sets:

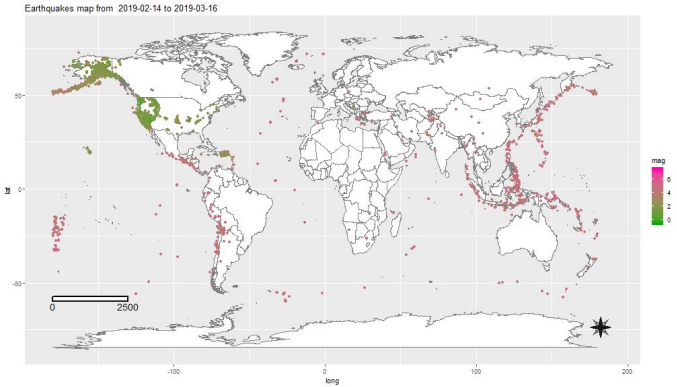
Characteristics:

- typically found on maps
- Represents:
  - singular data
  - categorical data
- Typically relating to people, animals, outbreaks, money

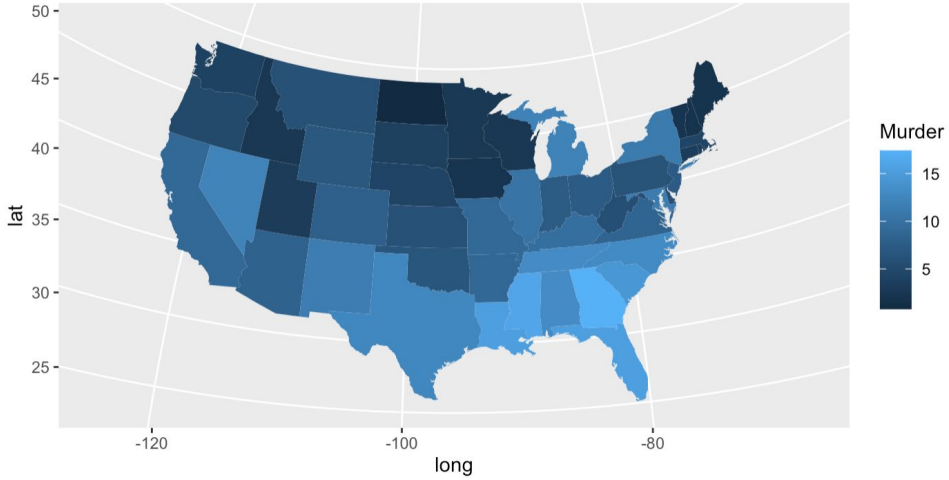
Examples:

- Fortune 500 company data set indexes.
- Earthquake magnitude ratings.
- Geographically important places relating to their relative intensity on a map.

A good example of this type of data set would be the amount of Covid-19 cases in a certain area.



Product Graph




```r
```{r setup, include=TRUE}
library(ggplot2)
# install.packages("maps")
library(maps)  #required for state map data
# install.packages("mapproj")
library(mapproj)
library(dplyr) # required for arrange()
library(viridis)
```
```

```r
```{r, inlcude=TRUE}
states_map <- map_data("state") # extracts data from the states map

# Make a data set of US crime data using the USArrests data set

data("USArrests")
USArrests$region <- row.names(USArrests)

crimes <- data.frame(state = tolower(rownames(USArrests)),
                     USArrests)
crime_map <- merge(states_map, crimes, by.x = "region", by.y = "state")
crime_map <- arrange(crime_map, group, order)
head(crime_map)

#Plot murder rate on the map

basemap <- ggplot(crime_map, aes(x=long, y=lat, group=group, fill=Murder)) +
  geom_polygon() + coord_map("polyconic")

basemap
```
```

Different Options/Arguments

- Mapping: Set of aesthetic mappings created by aes() or aes_().
  - If specified and inherit.aes = TRUE (the default), combined with default mapping which is at the top level of the plot.
  - An important aspect is that you must supply mapping if there is no plot mapping.
- Data: The data to be displayed in this layer. There are three options: If NULL, the default, the data is inherited from the plot data as specified in the call to ggplot().
  - data.frame will override the plot data. All objects will be fortified to produce a data frame.
  - Seefortify() for which variables will be created.
  - function will be called with a single argument, the plot data.
  - return value must be a data.frame, and will be used as the layer data.

Cont.

- stat: statistical transformation to use on the data for this layer, as a string.
- Other arguments passed on to layer().
  - often aesthetics, used to set an aesthetic to a fixed value
  - Like colour = "red" or size = 3.
  - may also be parameters to the paired geom/stat.

Cont.

- Map: Data frame that contains the map coordinates.
  - typically be created using fortify() on a spatial object
  - It must contain columns x or long, y or lat, and region or id.
- Na.rm: If FALSE, the default, missing values are removed with a warning.
  - If TRUE, missing values are silently removed.

Cont.

- **map_id**
- alpha
- colour
- fill
- group
- linetype
- size
- subgroup

Cont.

- Show.legend: NA, the default, includes if any aesthetics are mapped.
  - FALSE never includes
  - TRUE always includes.
  - can be a named logical vector to finely select the aesthetics to display.
- Inherit.aes: If FALSE, overrides the default aesthetics, rather than combining with them.
  - most useful for helper functions that define both data and aesthetics
  - shouldn't inherit behaviour from the default plot specification, e.g. borders().

Resources

- https://remiller1450.github.io/s230s19/Intro_maps.html
- Dr. Waldrop!!!
- https://jrnold.github.io/r4ds-exercise-solutions/data-visualisation.html