cta-crimes

Rei Bertoldi

3/3/2020

Reading in railstation shape file, maintaining geometry column for mapping later.

```
railstations <- st_read("CTA_RailStations/CTA_RailStations.shp") %>%
  set_names(to_snake_case(colnames(.))) %>%
  st_set_crs(102671) %>%
  st_transform(2163) %>%
  st_buffer(100) %>%
  select(longname, lines, gtfs, geometry)
```

```
## Reading layer `CTA_RailStations' from data source `C:\Users\User02\Desktop\Harris\cta-crime\CTA_Rails
## Simple feature collection with 144 features and 9 fields
## geometry type:  POINT
## dimension:      XY
## bbox:           xmin: 1101572 ymin: 1842222 xmax: 1182667 ymax: 1969903
## epsg (SRID):    NA
## proj4string:    +proj=tmerc +lat_0=36.66666666666666 +lon_0=-88.33333333333333 +k=0.9999749999999999
```

Reading in crime data from 2001 to 2018.

```
crimes <- read_csv("crimes_2001.csv") %>%
  set_names(to_snake_case(colnames(.)))
```

Filtering crime incidences labeled CTA, standardizing the geometery column between the crime and railstation data and converting data column to DateTime.

```
crimes_fil <- crimes %>%
  filter(
    str_detect(location_description, "CTA") &
      !(str_detect(location_description, "BUS"))
  ) %>%
  filter(!is.na(longitude)) %>%
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326, remove = FALSE) %>%
  st_transform(2163) %>%
  mutate(
    date = as_date(
      date,
      format = "%m/%d/%Y %H:%M:%S %p",
      tz = "America/Chicago"),
    month = month(date),
    year = year(date)
  ) %>%
  select(id, date, primary_type, month, year, latitude, longitude)
```

Reading in railstation shape file again but this time converting geometry column to lat/long coordiate points to join with crime observations later.

```
stations <- st_read("CTA_RailStations/CTA_RailStations.shp") %>%
  set_names(to_snake_case(colnames(.))) %>%
  st_set_crs(102671) %>%
  st_transform(2163) %>%
  sfc_as_cols(names = c("longitude", "latitude")) %>%
  select(gtfs, longitude, latitude) %>%
  st_set_geometry(NULL)
```

```
## Reading layer `CTA_RailStations' from data source `C:\Users\User02\Desktop\Harris\cta-crime\CTA_RailS
## Simple feature collection with 144 features and 9 fields
## geometry type:  POINT
## dimension:      XY
## bbox:           xmin: 1101572 ymin: 1842222 xmax: 1182667 ymax: 1969903
## epsg (SRID):    NA
## proj4string:    +proj=tmerc +lat_0=36.66666666666666 +lon_0=-88.33333333333333 +k=0.9999749999999999
```

Joining stations and crime observations, summarizing count of crime observations by station and month.

```
crimes_per_stop <- st_join(crimes_fil, railstations, join = st_within) %>%
  filter(!is.na(longname)) %>%
  st_set_geometry(NULL) %>%
  mutate(new_date = as_date(paste0(month(date),"/01/",year(date)), format = "%m/%d/%Y", tz = "UTC")) %>%
  group_by(gtfs, new_date) %>%
  summarise(
    crime_count = n()
  ) %>%
  left_join(stations)
```

Reading in rail line shape data and creating an object of unique lines and colors.

```
line_color <- railstations %>%
  select(lines, gtfs) %>%
  st_set_geometry(NULL)

rail_lines <- st_read("CTA_RailLines/CTA_RailLines.shp") %>%
  set_names(to_snake_case(colnames(.))) %>%
  st_transform(2163)
```

```
## Reading layer `CTA_RailLines' from data source `C:\Users\User02\Desktop\Harris\cta-crime\CTA_RailLine
## Simple feature collection with 154 features and 9 fields
## geometry type:  LINESTRING
## dimension:      XY
## bbox:           xmin: 1101572 ymin: 1842222 xmax: 1182667 ymax: 1969903
## epsg (SRID):    NA
## proj4string:    +proj=tmerc +lat_0=36.66666666666666 +lon_0=-88.33333333333333 +k=0.9999749999999999
```

Plotting crime counts at Chicago Transit Authority (CTA) rail stations over time. Each point is a CTA rail station stop with the size of the point determined by number of crime observations for that month - the larger the point, the greater the observation count. Rail station stops are overlayed atop rail lines, which have been colored by line color.

```
p <- crimes_per_stop %>%
  left_join(line_color, by = "gtfs") %>%
  ggplot() +
  geom_sf(data = rail_lines, aes(color = lines), alpha = .5) +
  geom_point(aes(x = longitude, y = latitude, size = crime_count, group = gtfs, color = lines)) +
  scale_color_manual(values = c("Green Line" = "green",
```

```
                                "Red Line" = "red",
                                "Yellow Line" = "yellow",
                                "Brown Line" = "brown",
                                "Green Line (Lake)" = "green",
                                "Blue Line (O'Hare)" = "blue",
                                "Orange Line" = "orange",
                                "Purple Line, Evanston Express" = "purple",
                                "Red, Brown, Purple (Express)" = "red",
                                "Green Line (Englewood)" = "green",
                                "Brown, Orange, Pink, Putple (Express), Green, Blue" = "brown",
                                "Brown, Purple (Express)" = "brown",
                                "Blue Line" = "blue",
                                "Brown, Orange, Pink, Purple (Express)" = "brown",
                                "Green (Lake), Pink" = "green",
                                "Green (Lake), Pink" = "green",
                                "Blue Line (Congress)" = "blue",
                                "Pink" = 'pink',
                                "Brown, Orange, Pink, Purple (Express), Green" = "brown",
                                "Red, Yellow, Purple, Evanston Express" = "red",
                                "Orange & Green Lines" = "orange",
                                "Blue Line (Forest Park)" = "blue",
                                "Brown, Purple (Express), Red" = "red",
                                "Brown, Green, Orange, Pink, Purple (Exp)" = "brown",
                                "Brown, Orange, Pink, Purple (Express)" = "brown",
                                "Green, Orange" = "green",
                                "Green, Pink" = "green",
                                "Purple Line" = "purple",
                                "Brown, Purple" = "brown",
                                "Pink Line" = "pink",
                                "Red, Purple Line" = "red")) +
transition_components(new_date) +
labs(title = 'CTA Stop Crime Counts (2001 - 2018)',
     subtitle = '{frame_time}') +
theme(
  legend.position="none",
  line = element_blank(),
  rect = element_blank(),
  axis.text = element_blank(),
  axis.title = element_blank(),
  panel.grid.major = element_line(colour = "transparent")
)
```