My title*

My subtitle if needed

Cristina Burca

September 18, 2024

Abstract

First sentence. Second sentence. Third sentence. Fourth sentence.

```
data <- read.csv("Motor Vehicle Collisions with KSI Data - 4326.csv")
```

```
clean_data <- clean_data %>%
  mutate(
    # Extract the first pair of numbers inside the brackets
    geometry_clean = gsub(".*\\[\\[(-?[0-9.]+),\\s*(-?[0-9.]+)\\].*", "\\1,\\2", geometry)
  )

# Split into longitude and latitude
clean_data <- clean_data %>%
  separate(geometry_clean, into = c("longitude", "latitude"), sep = ",", convert = TRUE)
```

```
# Group by longitude and latitude, and count the number of accidents at each unique location
accident_density <- clean_data %>%
  group_by(longitude, latitude) %>%
  summarise(accident_count = n())
```

```
`summarise()` has grouped output by 'longitude'. You can override using the
`.groups` argument.
```

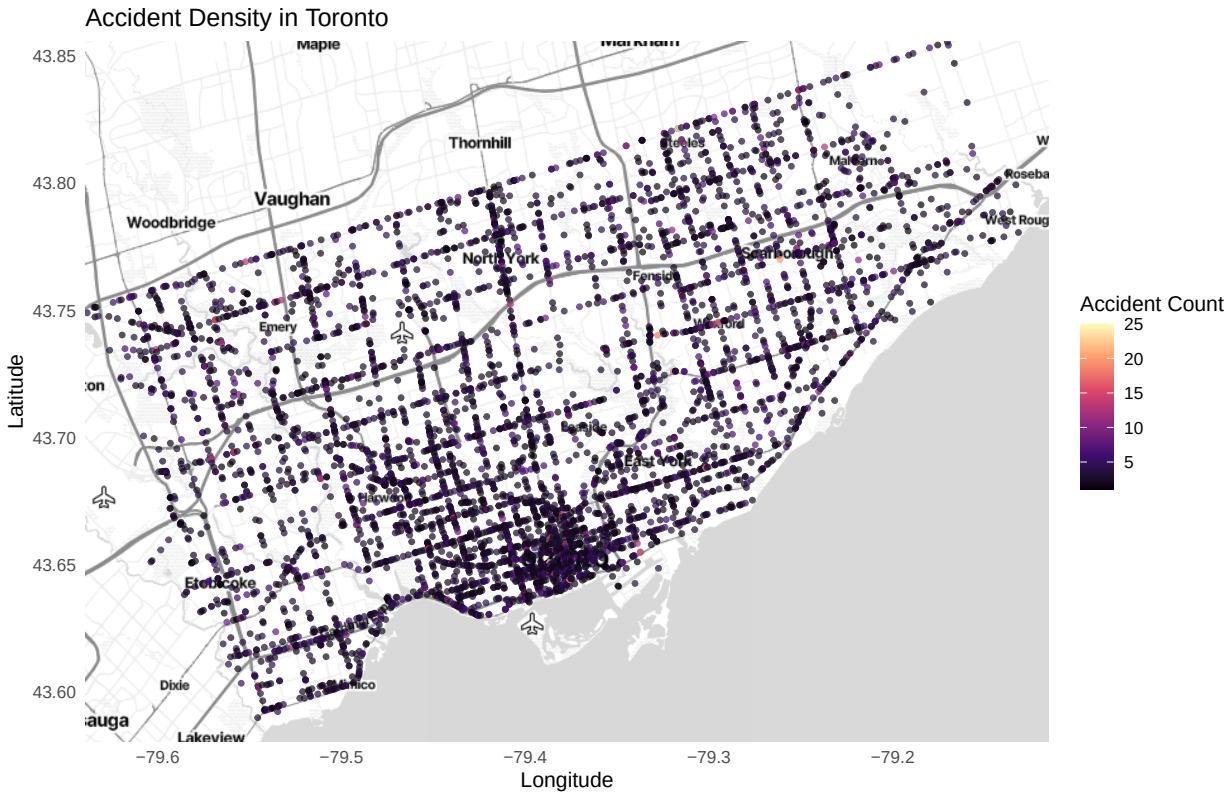
```
# Convert to sf object (spatial data frame) for mapping
sf_accident_density <- st_as_sf(accident_density, coords = c("longitude", "latitude"), crs = 4326)
```

```
# Define the bounding box for Toronto (approximate values)
bbox <- c(left = -79.6393, bottom = 43.5804, right = -79.1152, top = 43.8555)
# Fetch the map tiles from Stadia Maps
toronto_stadia_map <- get_stadiamap(bbox = bbox, zoom = 11, maptype = "stamen_toner_lite")
```

```
i © Stadia Maps © Stamen Design © OpenMapTiles © OpenStreetMap contributors.
```

*Code and data are available at: LINK.

```
# Plot the map using ggmap and overlay accident points
ggmap(toronto_stadia_map) +
  geom_point(
    data = sf_accident_density,
    aes(x = st_coordinates(geometry)[, 1], y = st_coordinates(geometry)[, 2], color = accident_count),
    alpha = 0.7, size = 1
  ) +
scale_color_viridis_c(option = "magma")  +
  labs(x = "Longitude", y = "Latitude", color = "Accident Count", title = "Accident Density in Toronto")
  theme_minimal()
```




```
sf_high_density <- sf_accident_density %>%
  filter(accident_count > 5)
```

```r
# Plot the map with fewer points but with varying sizes
ggmap(toronto_stadia_map) +
  geom_point(
    data = sf_high_density,
    aes(x = st_coordinates(geometry)[, 1], y = st_coordinates(geometry)[, 2], size = accident_count, col
    alpha = 0.7
  ) +
  scale_size_continuous(range = c(2, 8)) +  # Control the size range based on accident count
  scale_color_viridis_c(option = "plasma", direction = -1) +  # Optional: change or invert the color sca
  labs(x = "Longitude", y = "Latitude", color = "Accident Count", size = "Accident Count", title = "High
  theme_minimal()
```