

SDS 192 Project 3

Maps

Sabzara Ali

Tabitha Githinji

2024-12-06

Question: How does the distribution of Chipotle stores across the United States align with state populations? Additionally, when analyzing the proportion of stores relative to the total number of Chipotle locations and adjusting for population size, what insights can we uncover about factors influencing store placement, such as population density, urbanization, or regional economic activity

```
# Include all libraries you may need for your project here.
```

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.5.1      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr      1.0.2
```

```
-- Conflicts ----- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
```

```
x dplyr::lag()     masks stats::lag()
```

```
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(maps)
```

Attaching package: 'maps'

The following object is masked from 'package:purrr':

map

```
library(ggplot2)
library(sf)
```

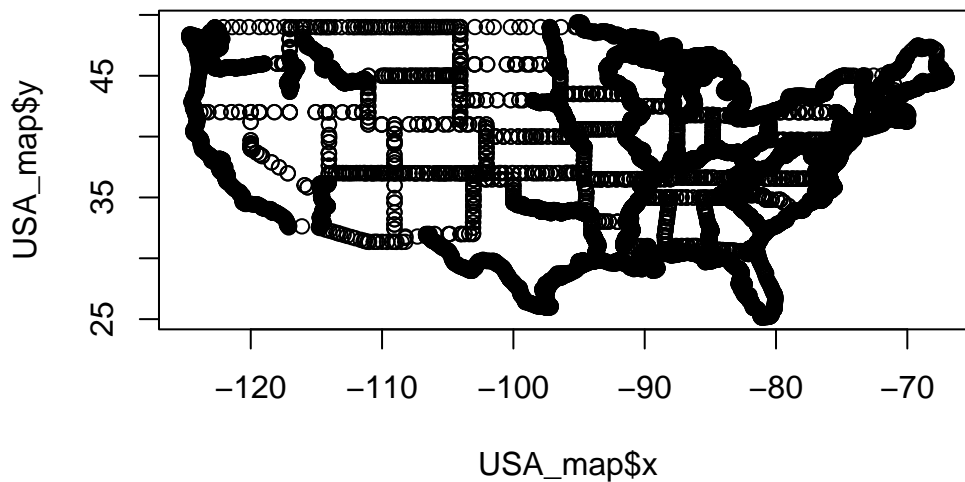
Linking to GEOS 3.11.0, GDAL 3.5.3, PROJ 9.1.0; sf_use_s2() is TRUE

```
# If you need to read in csv data do it here.

chipotle_stores <- read.csv("chipotle_stores.csv")
```

This map displays the distribution of Chipotle stores across the United States

```
# Group member 1 work here!
# converting the chipotle data to shape file
chipotle_sf <- st_as_sf(chipotle_stores,
                        coords = c("longitude", "latitude"),
                        crs = 4326)
# creating the USA map
USA_map <- maps::map("state", plot = FALSE, fill = TRUE)
plot(USA_map)
```



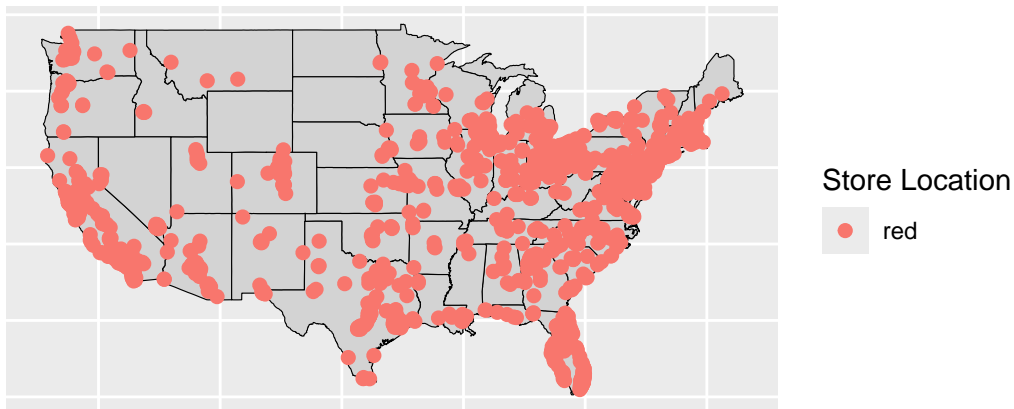
```

# Convert maps package map object to sf object using st_as_sf() again:
USA_sf <- USA_map |>
  st_as_sf(coords = c("x", "y"), crs = 4326)

# Plot both layers:
ggplot() +
  # Plot the USA map (USA_sf should already be an sf object)
  geom_sf(data = USA_sf, fill = "lightgray", color = "black") +
  # Plot the Chipotle store locations on top of the USA map
  geom_sf(data = chipotle_sf, aes(color = "red"), size = 2) +
  # remove axis ticks/label
  theme(
    axis.ticks.x = element_blank(),
    axis.ticks.y = element_blank(),
    axis.text.x = element_blank(),
    axis.text.y = element_blank()
  ) +
  labs(
    title = "Chipotle Store Locations in the USA",
    color = "Store Location",
    caption = "Source: Chipotle store locations dataset from Kaggle"
  )

```

Chipotle Store Locations in the USA



Source: Chipotle store locations dataset from Kaggle

In this section, we aim to visualize the proportion of Chipotle stores in each state across the United States. To achieve this, we first create a population map to provide context and to ensure we don't create the population map again.

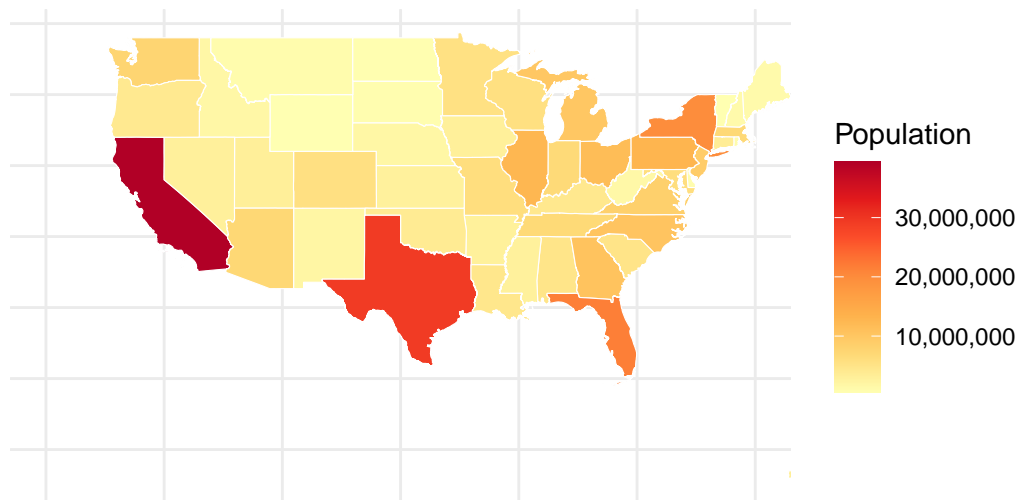
We first create the population map of the United States.

```
# Group member 2 work here!
#Creating the population map
US_pop <-
  tidycensus::get_acs(
    geography = "state",
    variables = c(population = "B01003_001"),
    geometry = TRUE
  ) |>
  mutate(NAME = tolower(NAME))

US_pop|>
  filter(NAME != "alaska")|>
  ggplot() +
    geom_sf(aes(fill = estimate), color = "white", size = 0.2) +
    scale_fill_distiller(palette = "YlOrRd",
                        direction = 1,
                        name = "Population",
                        labels = scales::comma
                      ) +
    theme_minimal() +
    labs(
      title = "Population by State in the United States",
      subtitle = "Choropleth map of U.S. population",
      fill = "Population") +
    theme(
      axis.ticks.x = element_blank(),
      axis.ticks.y = element_blank(),
      axis.text.x = element_blank(),
      axis.text.y = element_blank()
    ) +
    xlim(-130,-70)
```

Population by State in the United States

Choropleth map of U.S. population



Next, we create a map displaying the proportion of Chipotle stores in each state relative to the total number of stores across the United States.

```
# Creating showing the proportion of chipotle stores for each state

#Custom function to calculate store proportions by state

calculate_proportion <- function(df, total_stores) {
  df |>
    group_by(state) |>
    summarise(
      total_stores_per_state = n(),
      proportion_each_state = (n() / total_stores) * 100
    )
}

# Apply the custom function to calculate store proportions
total_stores <- nrow(chipotle_stores) # Total number of stores in the dataset
chipotle_stores_prop <- calculate_proportion(chipotle_stores, total_stores)

# Print the resulting data frame to check the results
print(chipotle_stores_prop)
```

```
# A tibble: 48 x 3
  state      total_stores_per_state proportion_each_state
  <chr>          <int>          <dbl>
1 Alabama           15           0.571
2 Arizona           85           3.23
3 Arkansas           6           0.228
4 California        421          16.0
5 Colorado          79           3.00
6 Connecticut       24           0.913
7 Delaware          9           0.342
8 Florida          177           6.73
9 Georgia           61           2.32
10 Idaho            4           0.152
# i 38 more rows
```

```
# Load the USA map for plotting
USA_map <- maps::map("state", plot = FALSE, fill = TRUE)

# States data to join with the chipotle data
states <- as.data.frame(cbind(abbrev = state.abb, lower_state = tolower(state.name), upper_state = state.name))

# Join chipotle data with states to have lowercase state names for merging
chipotle_stores_prop_lower <- states |>
  inner_join(chipotle_stores_prop, by = c("upper_state" = "state"))

# Convert the map object to an sf object for plotting
USA_sf <- USA_map |>
  st_as_sf(coords = c("x", "y"), crs = 4326)

# Join the map with chipotle store proportions
USA_map_joined <- USA_sf |>
  left_join(chipotle_stores_prop_lower, by = c("ID" = "lower_state"))
)
```

creating the choropleth map

```
# Create choropleth map
ggplot() +
  geom_sf(data = USA_map_joined, aes(fill = proportion_each_state)) +
  scale_fill_distiller(palette = "YlOrRd", direction = 1, name = "Proportion of Stores (%)")
  labs(
    title = "Proportion of Chipotle Stores by State in the United States",
```

```

    subtitle = "Choropleth map showing the percentage of Chipotle stores per state relative to 1
    caption = "Data sourced from Chipotle store locations(kraggle) and U.S. Census population
theme_minimal() +
theme(
  panel.background = element_rect(fill = "azure2",
                                   colour = "black",
                                   size = 0.5),

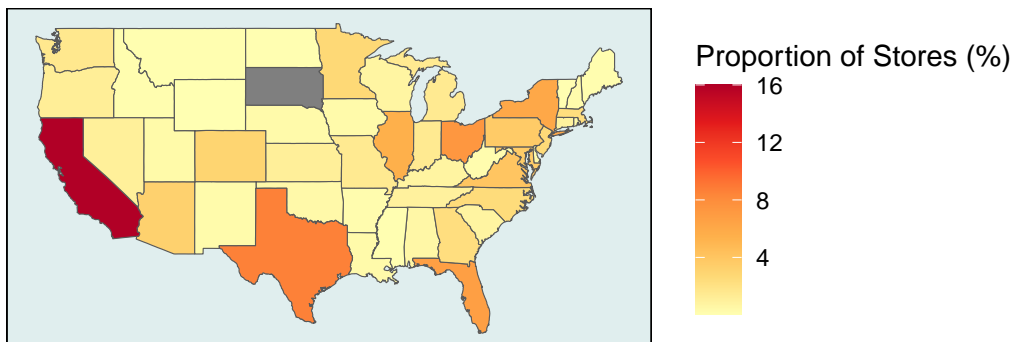
  panel.grid.major = element_blank(),
  panel.grid.minor = element_blank(),
  axis.ticks.x = element_blank(),
  axis.ticks.y = element_blank(),
  axis.text.x = element_blank(),
  axis.text.y = element_blank()
)

```

Warning: The `size` argument of `element_rect()` is deprecated as of ggplot2 3.4.0.
 i Please use the `linewidth` argument instead.

Proportion of Chipotle Stores by State in the United States

Choropleth map showing the percentage of Chipotle stores per state relative to 1



store locations(kraggle) and U.S. Census population estimates.

From the plot created, it is evident that the map closely resembles a population map, as it shows a similar distribution pattern across the states.

To create a map that accounts for population differences, we calculate the number of Chipotle stores per million residents in each state.

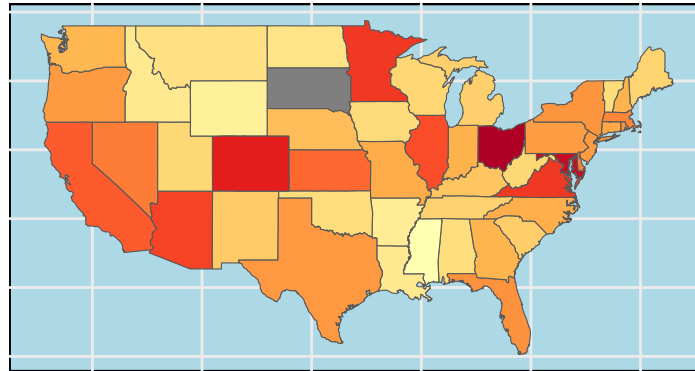
```

# Calculate stores per million people for each state
chipotle_stores_with_population <- chipotle_stores_prop_lower |>
  inner_join(US_pop |> select(NAME, estimate), by = c("lower_state" = "NAME")) |>
  mutate(stores_per_million = (total_stores_per_state / estimate) * 1e6)
# Merge population-adjusted data with USA map
USA_map_with_population <- USA_sf |>
  left_join(chipotle_stores_with_population, by = c("ID" = "lower_state"))
# Create a map based on stores per million residents
ggplot() +
  geom_sf(data = USA_map_with_population, aes(fill = stores_per_million)) +
  scale_fill_distiller(palette = "YlOrRd", direction = 1, name = "Stores per Million") +
  labs(
    title = "Chipotle Stores Adjusted for State Population",
    subtitle = "Stores per Million Residents",
    caption = "Data: Chipotle and US Census"
  ) +
  theme_minimal() +
  theme(
    axis.ticks.x = element_blank(),
    axis.ticks.y = element_blank(),
    axis.text.x = element_blank(),
    axis.text.y = element_blank(),
    panel.background = element_rect(fill = "lightblue", color = "black"),
    legend.position = "bottom"
  )

```


Chipotle Stores Adjusted for State Population

Stores per Million Residents



Data: Chipotle and US Census

```
# Group member 3 work here!
```

Blog post

In this project, we explored the distribution of Chipotle restaurants across the United States using geospatial analysis. Leveraging a dataset from Kaggle, which provides the location of every Chipotle restaurant in the country, we sought to understand the company's location strategy. The dataset was uploaded by Jeffrey Braun, a UChicago graduate, who claimed to have scraped the data from the Chipotle website. (Braun 2020). Specifically, we aimed to analyze how Chipotle store locations are distributed in relation to state populations and to investigate whether store concentration correlates with population density.

We began by visualizing the locations of Chipotle restaurants on a U.S. map. This involved converting the dataset into a spatial format, enabling us to plot the coordinates of each store. The visualization revealed that Chipotle locations are primarily concentrated in urban centers, with states like California, New York, and Texas hosting the most stores. These states, with their high population densities and large urban hubs like Los Angeles, New York City, and Houston, naturally align with a higher number of Chipotle locations. To assess how population influences store distribution, we created a population map of the United States. This step was crucial as it provided a baseline for comparing the distribution of Chipotle locations to population data. From there, we developed a custom function to calculate each state's proportion of Chipotle stores relative to the national total. Using this data, we created a

choropleth map that graphically represented the store proportions, offering a clear depiction of states with higher or lower concentrations of Chipotle locations. The resulting map closely mirrored the population map, suggesting that states with larger populations generally host more Chipotle stores. However, we wanted to investigate further to determine if less-populated states had a comparable number of stores per million residents to more populous states.

We then took our analysis a step further by considering population-adjusted metrics. Recognizing that absolute store counts alone do not account for state population differences, we integrated U.S. Census data to calculate the number of Chipotle stores per million residents in each state. This approach allowed us to explore whether smaller states with lower populations had comparable store density to larger, more populous states. Using this new proportion, we created another choropleth map to visualize the distribution of Chipotle stores on a per-capita basis, offering a more balanced view of store density relative to population size. The population-adjusted analysis highlights interesting patterns. States with, such as Minnesota, Ohio, and Pennsylvania, have a proportionally higher number of Chipotle stores per capita, even if their total store counts are relatively modest. On the other hand, some highly populated states have a relatively lower store density per capita, despite hosting a large number of locations overall. This suggests that Chipotle's placement strategy is not solely driven by population size but also considers other factors such as high-traffic urban centers, economic activity, and potentially regional preferences.

Our findings align with insights from Jordan Bean's study, which highlights Chipotle's strategic focus on urban centers in high-density areas while also noting a recent shift toward suburban areas (Bean 2023). This strategic adaptation may reflect evolving trends, such as population migration and hybrid work patterns, as Chipotle seeks to expand its reach into new markets.

In conclusion, this project highlights the value of integrating geospatial and demographic analysis to understand business strategies. By creating maps that account for both absolute store counts and population-adjusted metrics, we uncovered patterns that reveal the brand's strategic focus on urban areas and densely populated states. These findings can help inform business decisions of Chipotle and other brands, from identifying growth opportunities to refining location strategies. Lastly, this project highlights how geospatial tools and demographic data in R can be applied to real-world problems, offering practical insights for optimizing retail operations and enhancing strategic planning.

Bean, Jordan. 2023. "Beyond the Burrito: The Analytics Behind Chipotle's Location Strategy." *Medium*, May. <https://jordanbean.medium.com/the-analytics-behind-chipotles-location-strategy-2d91a23fd66>.

Braun, Jeffrey. 2020. "Chipotle Locations: A CSV File for All Chipotle Locations in the US." <https://www.kaggle.com/jeffreybraun/chipotle-locations>.