Michelin Wrapped - The Globally Recognized Restaurants of San Francisco: 2023

Esha Shah

Introduction

As the year comes to a close in the next few weeks, we do what we do every December - reminisce. We chat about our Spotify Wrapped results, which movie releases were the best, and major events of the year. However, there is one that isn't talked about enough each year: the release of the Michelin guide.

The word "Michelin" is an iconic one that represents luxury, grandeur, and the "best-of-the-best." But people don't know about its roots. In a small town in France in 1889, brothers Andre and Edouard Michelin founded a tire company that is still around today. In order to boost tire sales, they started a little red book guide of travel information. Since then, it has evolved to include hotels, and most notably, restaurants. A Michelin star can be awarded to an establishment as recognition of the quality of their cooking - one for good, two for exceptional, and three for exquisite. It is among the highest honors a restaurant can strive for and has the power to bring in immense amounts of business.

Each year, Michelin hosts a live award ceremony for different regions that awards Michelin stars to new restaurants. California's ceremony occurred earlier this year on July 18th. California is significant because in 2007, one of its major cities - San Francisco - was the second US city to be added to the guide. Each ceremony represents how far restaurants have grown and evolved since the last, and 2023 was no different.

While restaurants are simply places for people to go eat, they represent so much more. The available cuisines can say a lot about a location. Moreover, restaurants are places of gathering - where people can come together with the sole intention of eating and talking - and make lasting memories.

This data analysis intends to provide more insight into the 26 Michelin star restaurants currently active in the city of San Francisco.

Method + Findings

There were 4 main data tables used.

1. [Registered Business Locations - SF government's database page](): This lists every business registered with the City and County of San Francisco.
2. [San Francisco Analysis Neighborhoods - SF government's database page:]() This shows where each neighborhood in the city is located.
3. Michelin - created by Esha Shah: This was a hand-created table of the 26 Michelin star restaurants in the City of San Francisco. The information came from the [official Michelin Guide website]() and includes the name, number of stars, cuisine, etc of each restaurant.
4. [Total Population by Race]() - Census Reporter (from the Census Bureau)

First, I downloaded and cleaned the Registered Business Location database, which was a huge table of 317,364. I filtered it to only include San Francisco city-specific businesses and selected it to only include relevant columns. To analyze the dataset a bit more and get a sense of trends, I converted the "business start date" column to a date type using lubridate and plotted the trend of businesses registered over the last 10 years using ggplot2.

Businesses Registered in SF Over the Last 10 years
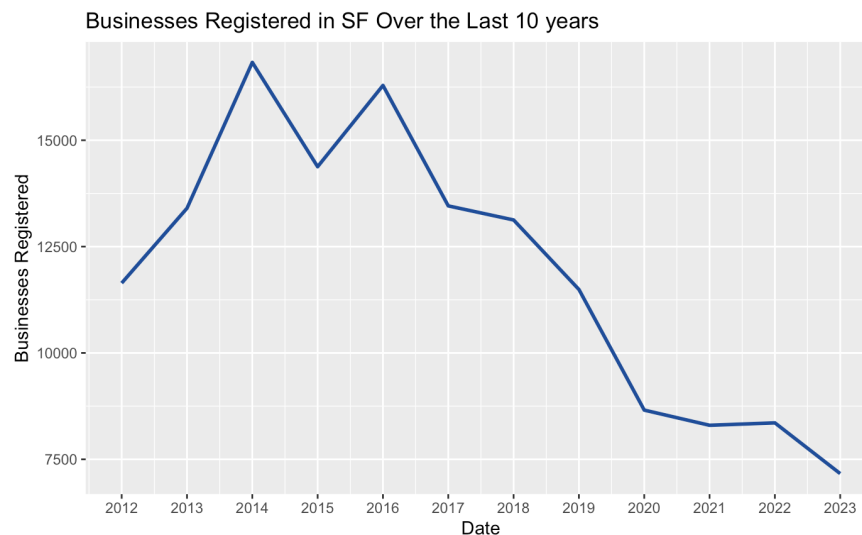
From this, we can see that there is a downward trend of businesses being registered since 2016. There is a dip near 2020, which could be due to the COVID-19 pandemic. As someone who lives in the Bay Area, I know that there also are increasing concerns about the safety of the city. In addition, after COVID, the city has become much quieter, as the primary industry - software - has allowed a mostly work-from-home lifestyle. These could all be reasons for the decrease. 2023 data is still being collected, as the dataset is updated daily, so the drop-off at the end is not significant right now.

The Registered Business Location database also contains a column for "neighborhood." San Francisco is organized into 41 different main neighborhoods. I grouped the dataset by neighborhood and counted the number of businesses in each. I found that the Financial District/South Beach neighborhood has the most number of active businesses at 40,323. The second most is the Mission neighborhood at 20,430. I then combined the analysis neighborhood data with my grouped table to include the geom MULTIPOLYGON column, which allowed me to create an interactive map of polygons that outlined each neighborhood.
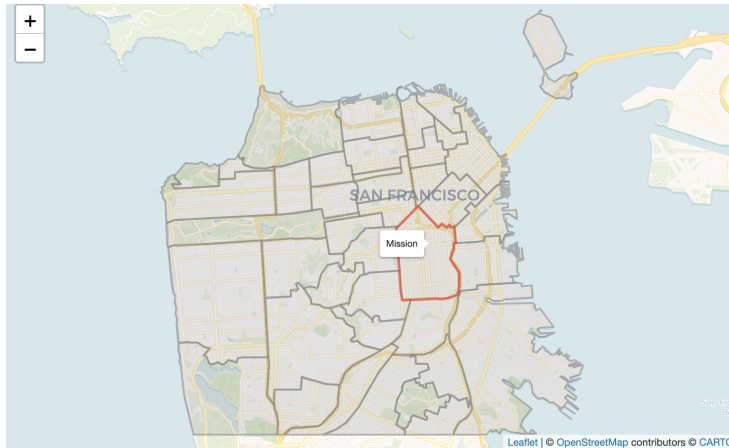

Now, to isolate just the Michelin star restaurants from the Registered Businesses database, I joined my Michelin data table with the registered businesses one. I was left with a table of 26 rows (one for each restaurant) and 12 columns of information. Then, I used a loop to combine the Street Add and State columns, which allowed me to add columns with the longitude and latitude of each address using geocode from ggmap. I had to google a lot of this methodology, but it worked in the end.

Then, using leaflet, I plotted the locations of each of the restaurants on top of the previous map. This resulted in an interactive map that outlines each neighborhood and has clickable markers for each restaurant.
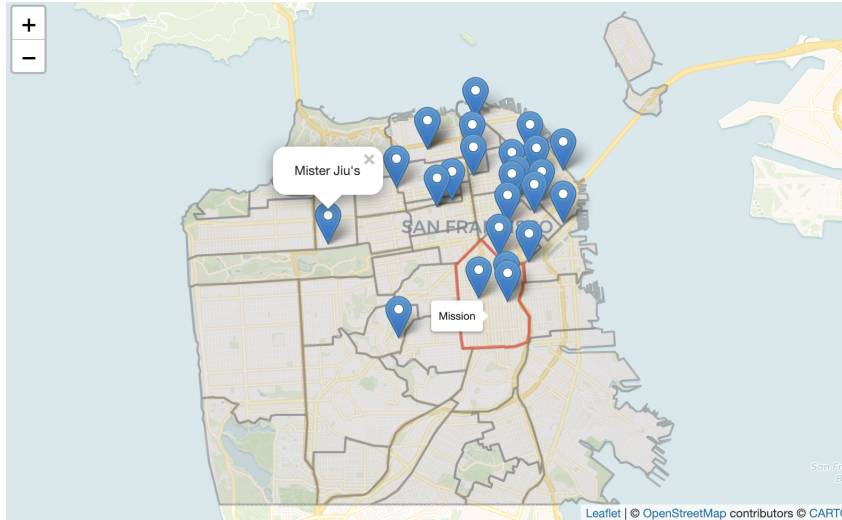
From the map, you can see that some neighborhoods don't have any Michelin star restaurants. I investigated which ones by grouping my Michelin table by district and summarizing the count of restaurants in each in another column. Only 14 of the 41 districts in SF city have at least one Michelin star restaurant.
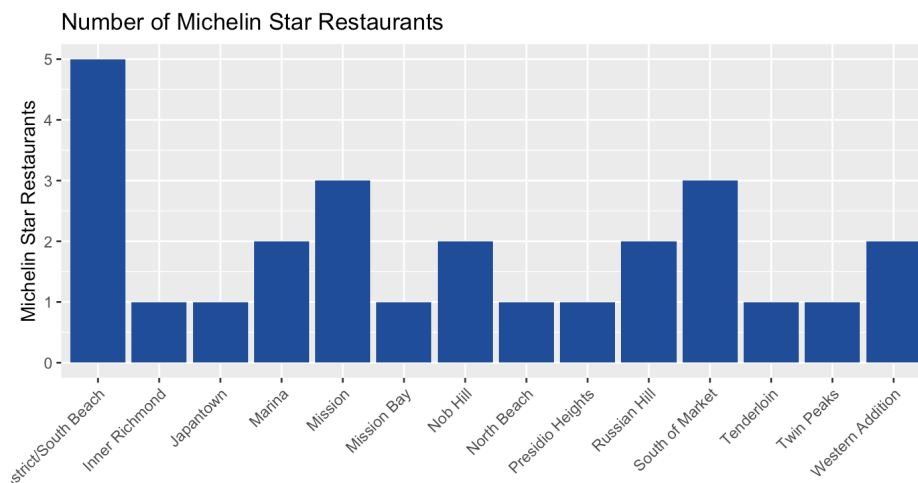
The Financial District/South Beach neighborhood has the most Michelin Star Restaurants - 5 of them! This is consistent with the fact that they have the most number of active businesses in general, which I found earlier.

I then decided to see which cuisine type was the most popular. I grouped by cuisine type, and then merged the rows with Japanese, Asian, Korean, Chinese, and Thai food into one umbrella branch named "Asian." As someone who grew up in San Francisco, these groups are usually categorized as Asian food on a broad scale, so to truly see the cuisine distribution, I grouped these together. The most popular two cuisines for Michelin star restaurants are Contemporary - which represents a mix of flavors, and Asian.

The type of food demanded can say something about the demographics of a place. Food is a representation of people's identities and is one of the core ways to share that with others. I used information from the census ACS to get a table of the population by race totals in San Francisco. I converted these values to proportions and mapped them into a pie chart with some aid from Google. White and Asian are the most populous groups. The higher Asian population may explain the large number of Asian Michelin star restaurants in the area.

Population Distribution by Race

Source: Census Bureau Data
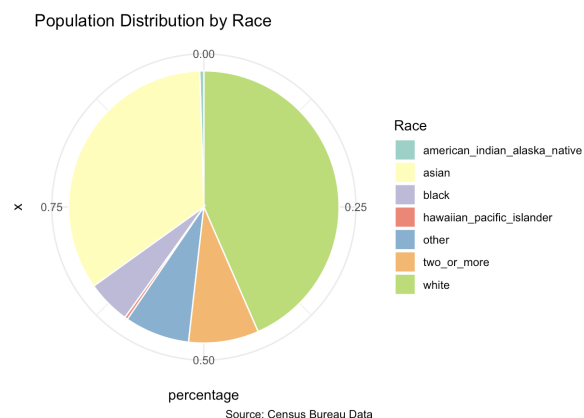
Conclusion

Food is a unifying thing, and the Michelin star system represents those restaurants that do an especially exquisite job of delivering an unforgettable dining experience. As we head into the new year and wait for 2024's Michelin guide release, let's be sure to share a few bites to eat with those that we love.

Assignment Details

Elements of the data interview Rmd file:

Up to 100 points for the written summary of your findings. Spelling and grammar is accurate. Omits needless words. Emphasizes news values of timeliness, prominence, proximity, oddity, emotional appeal, conflict to make a general news audience care about the analysis. Is narrow and deep rather than broad and shallow. Is free of all fact errors.

For the data interview (you do not need to do all of these) :

10 points for loading the data correctly
10 points for loading the packages into your library correctly
10 points for using filter() effectively

10 points for using arrange() effectively

12 points for using mutate() effectively

12 points for using count() effectively

18 points for using group_by() and summarize() effectively

18 points for binding data effectively

20 points for joining data effectively

20 points for using case_when() effectively

20 points for using st_read() effectively

20 points for using tabular data from the tidycensus package (effectively)

20 points for using geo-data from the tigris or tidycensus packages (effectively)

25 points for using any lubridate function effectively

25 points for using conditionals or loops effectively

Either 40 or 50 points for your charts

Up to 100 points for your map: 70 points for making a map of points or polygons. Additional 15 points for mapping both points and polygons. Additional 15 points for making the map interactive.

50 points for GitHub repository* (not required)

50 points for HTML page published to GitHub pages* (not required)