Highlighting the Impact of COVID-19 on the US Restaurant Industry

Lucas Fernandes, Paul Mara, Kenny Rhodes, Rachael Sallie

Table of Contents

Overview and Motivation

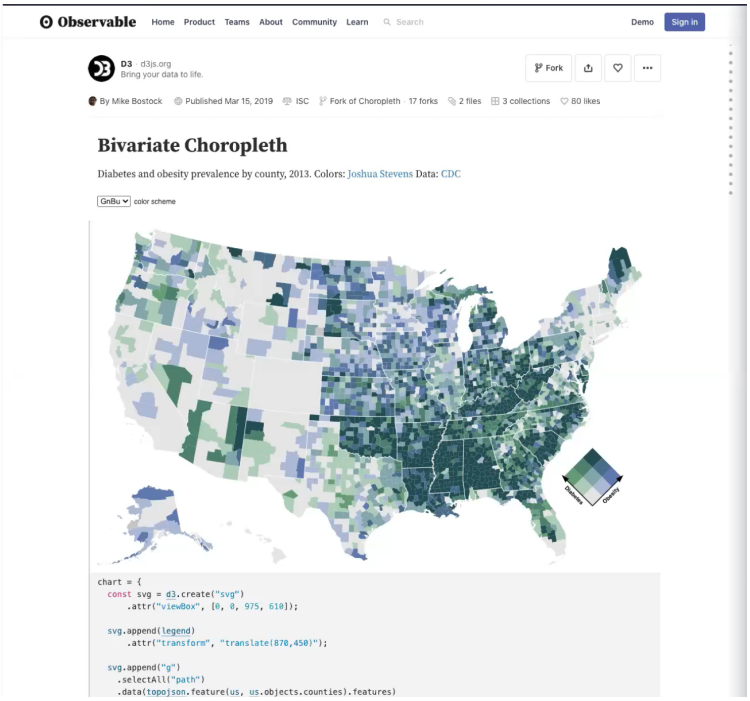
Our initial motivation was to create a project on the impact of COVID-19. After exploring data online, we eventually settled on narrowing our focus to the restaurant industry. The team was interested in working with maps, so two of the factors that interested us were time (effects of landmark events) and location (state to state differences).

As we developed our plans, we considered a variety of angles to tackle this topic from. Should we narrow our focus to a couple states with different COVID-19 responses or should we look at the impact of the pandemic country-wide? What parts of the restaurant industry should we investigate? Our goal was to tell the story of the restaurant industry, so we decided to develop a series of visualizations that tell the story in terms of employees (Is there more unemployment?), the restaurants themselves(How have sales been impacted?), and the customers (Are they avoiding restaurants? Dining in vs taking out?). We consider how specific changes over the course of the year have affected the industry, whether they are concrete differences, like state mandates, or just the indirect effects of fear of the virus.

We look to investigate these topics with the data we have available to us, and ultimately connect them into one cohesive story.

Related Work and Impact on Project Direction

Our initial inspiration came from Mike Bostock's bivariate choropleth map that was demonstrated by Professor Harrison during class. We saw a lot of potential in implementing a similar map and wanted to investigate interactivity and data presentation through maps.



Our topic, however, was inspired by a kaggle dataset named "Restaurant Business Rankings 2020". We had previously been investigating Covid related datasets, but found that initial topic to be too broad. This dataset narrowed our focus to one area affected by the pandemic: the restaurant industry. This dataset includes information about the top 100 independently owned businesses as well as the top 250 restaurants(including chains) in the United States. Though it lacked some of the information we wanted to explore, it did set us on the right path to further investigate other datasets that would help us dive deeper into the topic.

Another big inspiration was the concept of scrollytelling, as seen in Cuthbert Chow's visualization on the best college major. As our concept grew, we realized that we wanted to approach the subject from several angles, and put them together to tell the story of restaurants during the pandemic. Scrollytelling stood out as a compelling format to tell this story.

Lastly, we thought that our interactive maps should have an intuitive zooming feature after being inspired by Mike Bostock's Click to Zoom map example.

Research Questions

In a broad sense our project hopes to answer the question of how COVID-19 has affected the restaurant and dining industry in the USA. Specifically, we look at a few questions:

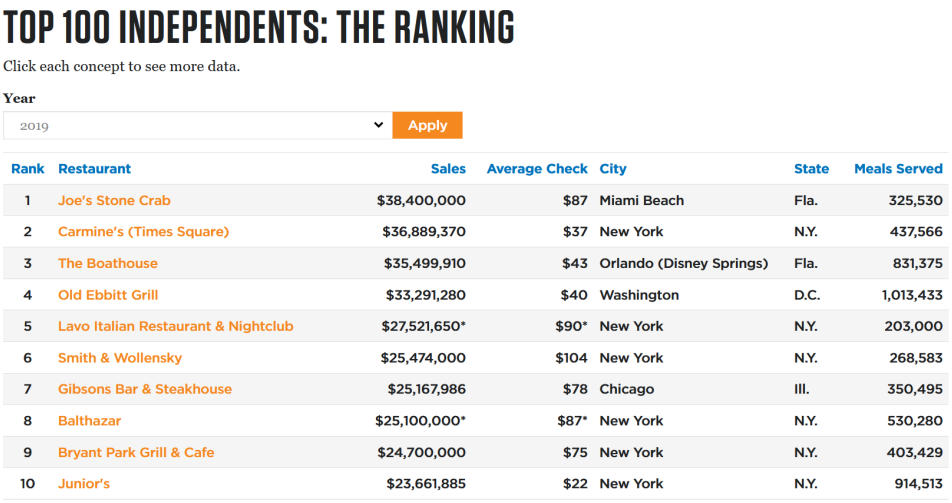
- How has employment in the restaurant industry changed?
- How successful are restaurants?
- How many people are dining out?
- How have specific lockdown measures impacted the industry?
- Are some states performing better than others due to their responses?

While investigating restaurant success, new questions developed. One such question was “What factors are affecting restaurant success?”. Sales increases and decreases alone were not enough to infer the reasons behind these changes. This led us to also look at meals served as a measure of whether more or fewer people were getting food from restaurants. Interestingly enough, in some cases, sales increased while meals served decreased. This might suggest that restaurants have increased prices to compensate for reduced business. Many restaurants also surprisingly increased both sales and meals served despite the pandemic. This led to the question of “What proportion of meals sold are take-out rather than dine-in?”

Data and Exploratory Analysis of Data

Data Source: Restaurant Business Magazine

Though the previously mentioned Restaurant Business Rankings dataset was our starting point and a good overview of the 2020 data, we ultimately went back to its source for developing one of our own visualizations. As we wanted to learn the impact of Covid-19, we first needed to develop a baseline of the state of the industry from prior years. [Restaurant Business Magazine], the source of our original dataset, fortunately contained data from prior years.



TOP 100 INDEPENDENTS: THE RANKING

Click each concept to see more data.

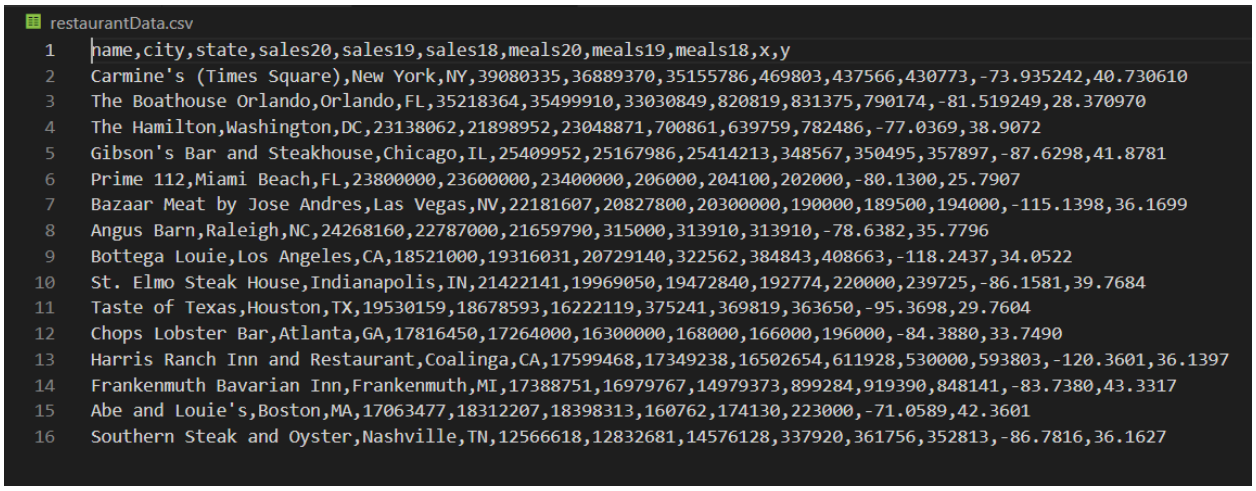
Year

2019 Apply

Rank	Restaurant	Sales	Average Check	City	State	Meals Served
1	Joe's Stone Crab	$38,400,000	$87	Miami Beach	Fla.	325,530
2	Carmine's (Times Square)	$36,889,370	$37	New York	N.Y.	437,566
3	The Boathouse	$35,499,910	$43	Orlando (Disney Springs)	Fla.	831,375
4	Old Ebbitt Grill	$33,291,280	$40	Washington	D.C.	1,013,433
5	Lavo Italian Restaurant & Nightclub	$27,521,650*	$90*	New York	N.Y.	203,000
6	Smith & Wollensky	$25,474,000	$104	New York	N.Y.	268,583
7	Gibsons Bar & Steakhouse	$25,167,986	$78	Chicago	Ill.	350,495
8	Balthazar	$25,100,000*	$87*	New York	N.Y.	530,280
9	Bryant Park Grill & Cafe	$24,700,000	$75	New York	N.Y.	403,429
10	Junior's	$23,661,885	$22	New York	N.Y.	914,513

100 is a relatively small number of data points, especially when considering the number of restaurants that exist in the country, so we decided to frame this part of our story as a case study on how the country's most successful restaurants fared during the pandemic. First, our data was narrowed down to restaurants that appeared in the data throughout the past few years. After deciding on a map as a medium for the user to interact with our visualization, we further reduced the data to a set of restaurants in different cities, so it would be easier for users to navigate between cities.

Ultimately we wrote our own CSV file containing data from this website for two main reasons. The first being that we could not directly download this data from Restaurant Business Magazine, and the second being that the data lacked coordinates to place points onto a map. Manually writing our CSV allowed us to overcome these challenges, and additionally format the remaining data into the form most useful for our visualization. Based on all of the restrictions we set for our data, we ultimately choose 15 restaurants to look at in our case study.


restaurantData.csv

```
name,city,state,sales20,sales19,sales18,meals20,meals19,meals18,x,y
Carmine's (Times Square),New York,NY,39080335,36889370,35155786,469803,437566,430773,-73.935242,40.730610
The Boathouse Orlando,Orlando,FL,35218364,35499910,33030849,820819,831375,790174,-81.519249,28.370970
The Hamilton,Washington,DC,23138062,21898952,23048871,700861,639759,782486,-77.0369,38.9072
Gibson's Bar and Steakhouse,Chicago,IL,25409952,25167986,25414213,348567,350495,357897,-87.6298,41.8781
Prime 112,Miami Beach,FL,23800000,23600000,23400000,206000,204100,202000,-80.1300,25.7907
Bazaar Meat by Jose Andres,Las Vegas,NV,22181607,20827800,20300000,190000,189500,194000,-115.1398,36.1699
Angus Barn,Raleigh,NC,24268160,22787000,21659790,315000,313910,313910,-78.6382,35.7796
Bottega Louie,Los Angeles,CA,18521000,19316031,20729140,322562,384843,408663,-118.2437,34.0522
St. Elmo Steak House,Indianapolis,IN,21422141,19969050,19472840,192774,220000,239725,-86.1581,39.7684
Taste of Texas,Houston,TX,19530159,18678593,16222119,375241,369819,363650,-95.3698,29.7604
Chops Lobster Bar,Atlanta,GA,17816450,17264000,16300000,168000,166000,196000,-84.3880,33.7490
Harris Ranch Inn and Restaurant,Coalinga,CA,17599468,17349238,16502654,611928,530000,593803,-120.3601,36.1397
Frankenmuth Bavarian Inn,Frankenmuth,MI,17388751,16979767,14979373,899284,919390,848141,-83.7380,43.3317
Abe and Louie's,Boston,MA,17063477,18312207,18398313,160762,174130,223000,-71.0589,42.3601
Southern Steak and Oyster,Nashville,TN,12566618,12832681,14576128,337920,361756,352813,-86.7816,36.1627
```

For such a small set of data, rather than make a new visualization to analyze it, we explored the tables already existing on Restaurant Business Magazine's website. We specifically paid attention to changes in rank, differences in sales, and potential causes, such as serving fewer meals (in the case of price decreases) or charging more per meal (in cases of sales increasing while meals served decreased). It was shocking to discover that many restaurants had increased sales in 2020 over previous years. We had gone into this expecting to see restaurants taking hits to their sales, so discovering the opposite seems especially interesting to visualize, as others may also not realize how successfully some restaurants have adapted.

Data Source: Open Table

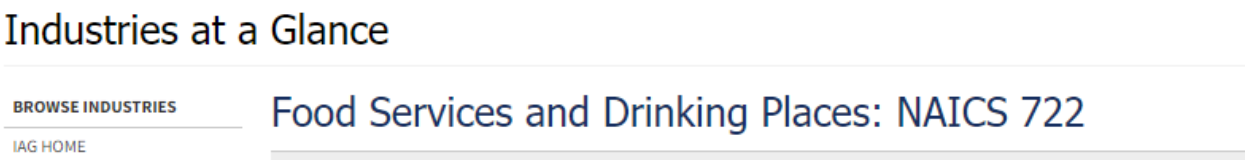
https://www.opentable.com/state-of-industry

As a part of our early exploration phase of the project we had stumbled upon Open Table's state of the industry page which hosts daily updated data sets showcasing the state of the restaurant industry across the United States. Included near the end of the page is the "Seated diners at restaurants open for reservations" data set which really stood out to the team as it captures data for many largely populated states in various parts of the country since May 1st, 2020 and is

updated daily. After pulling the data from the website some cleaning needed to be done as the set included other countries and unfortunately had an error with the state of New York. This meant we were left visualizing 7 states which were California, Florida, Illinois, Massachusetts, New Jersey, Pennsylvania, and Texas. Additionally we trimmed the data from daily to the first day of each month captured except for the month of March 2021 as this was the same month as when this project was conducted, in its case it remained daily. All date columns were renamed to follow a new name scheme so that we could process it far easier into our visualization. A snippet of our csv can be seen below.

Type	Name	MAY1202(	JUN12020	JUL12020	AUG12020	SEP12020	OCT12020	NOV1202(	DEC12020	JAN12021
state	California		39.95	46.78	60.48	90	74.54	78.21	54.44	
state	Florida		50.97	48.7	50.09	100	77.07	80.69	66.09	88.26
state	Illinois		23.37	56.54		70.5	55.91	21.43	21.63	41.35
state	Massachusetts			51.91	73.94	91.45	75.94	55.17	47.48	53.55
state	New Jersey			55.91	76.99	100	100	67.85	50.38	71.54
state	Pennsylvania			56.28	63.52	96.34	80.71	52.82	32.94	26.13
state	Texas	37.03	50.77	40.95	64.63	83.48	71.37	80.68	51.42	75.02

Data Source: United States Department of Labor



The US Department of Labor collects statistics per industry. It usually collects the statistics up on a large scale and for a large number of people. For this data, we narrowed the focus to Food Services and Drinking Places, as this seemed like the closest match to the categorical “restaurant” for the United States. Rolled up into this industry included the following statistical categories:

- Workplace Statistics
- Earnings and Hours
- Fatalities, Injuries, and Illnesses (this one would have been fascinating, but wasn’t updated with 2020 data)
- Workplace Trends (also not updated)

Eventually, we picked out three statistics that all relate to the relative state of the workers in the industry: Unemployment rate, total employees, and the average number of hours/week each worker actually logged work. The total employees shows how many people work in the industry. Discrepancies between this number and the unemployment rate mean people may have changed jobs into another industry. Lastly, the number of hours per week was thought to be an interesting measure because it looks at who is left working during the pandemic.

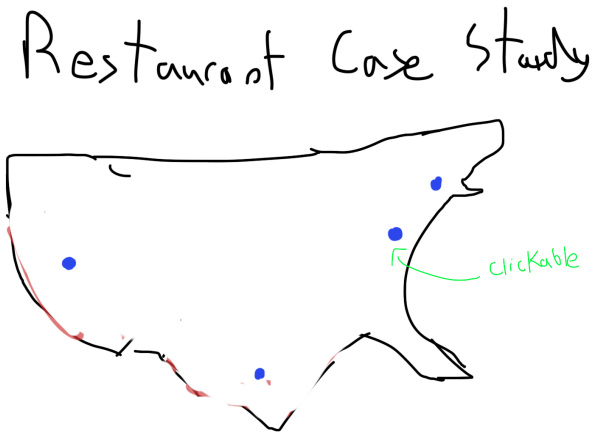
Lastly, there were already primitive graphs built for this data, so there wasn't much prototyping required to analyze what type of graph to use for the data. However, they were built in a single, straight scatter plot rather than being organized by months in the year

Design Evolution and Implementation

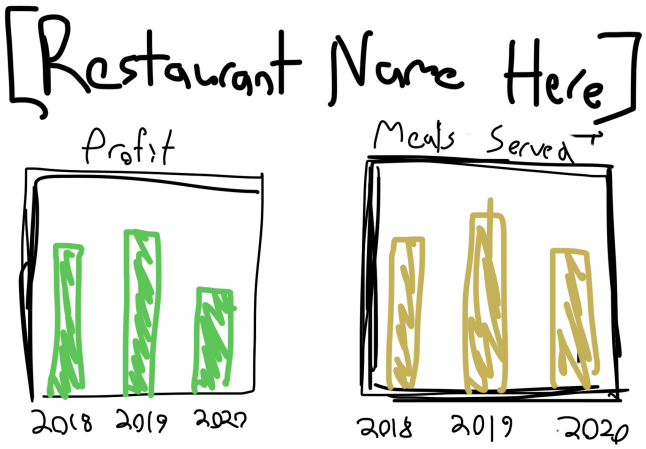
After brainstorming ideas and determining what was possible based on the data available, we created concept sketches for our set of visualizations.

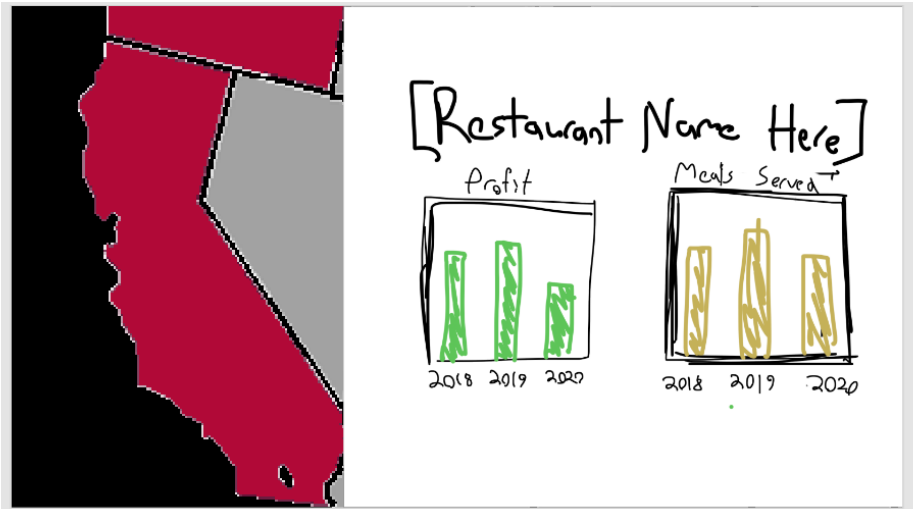
Restaurant case studies

One visualization idea was to create an interactive map as a case study on a small number of specific restaurants. The plan was to have dots representing restaurants located on a map of the USA in their corresponding spatial positions. Each dot would be clickable to overlay a new visualization containing a set of bar graphs.

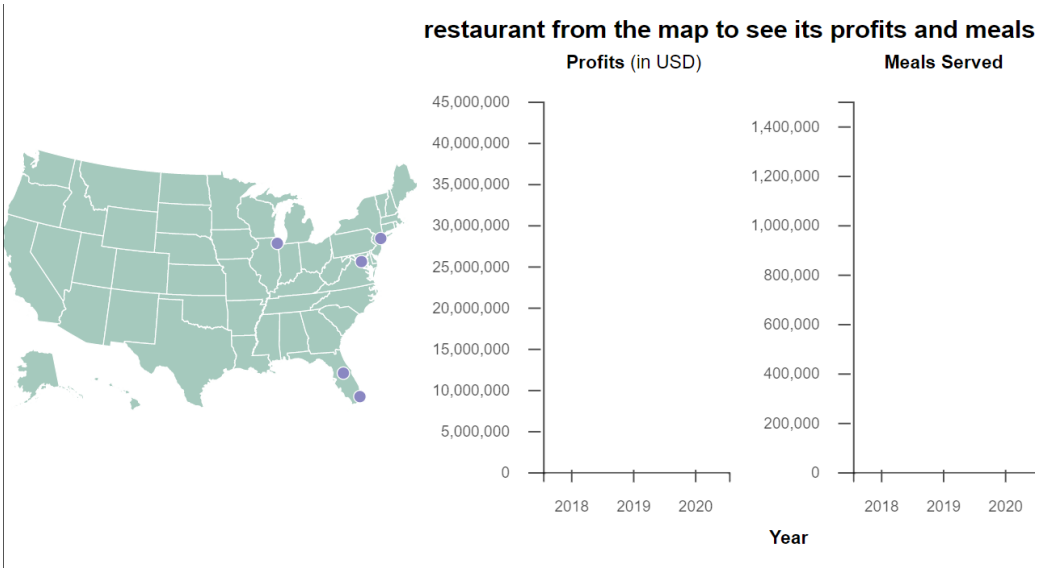


Bar graphs were chosen as they are an effective tool for comparing different amounts. The decision to place the bars into two separate graphs was made for two main reasons. Firstly, the two sets of data used very different scales along the y-axis, and placing them into the same graph may cause confusion. Secondly, the intention of the graph is not to compare meals to sales in any given year, and users may intuitively try to do that should they be placed within the same graph. Instead this format allows the user to analyze trends over the years for each topic with fewer distractions. One bar graph represents the sales of the selected restaurant from 2018 to 2020, and the second represents the meals served. The aim of the inclusion of both is to contextualize the change in profits as either being due to more/fewer meals sold, vs a change in price or other factor.

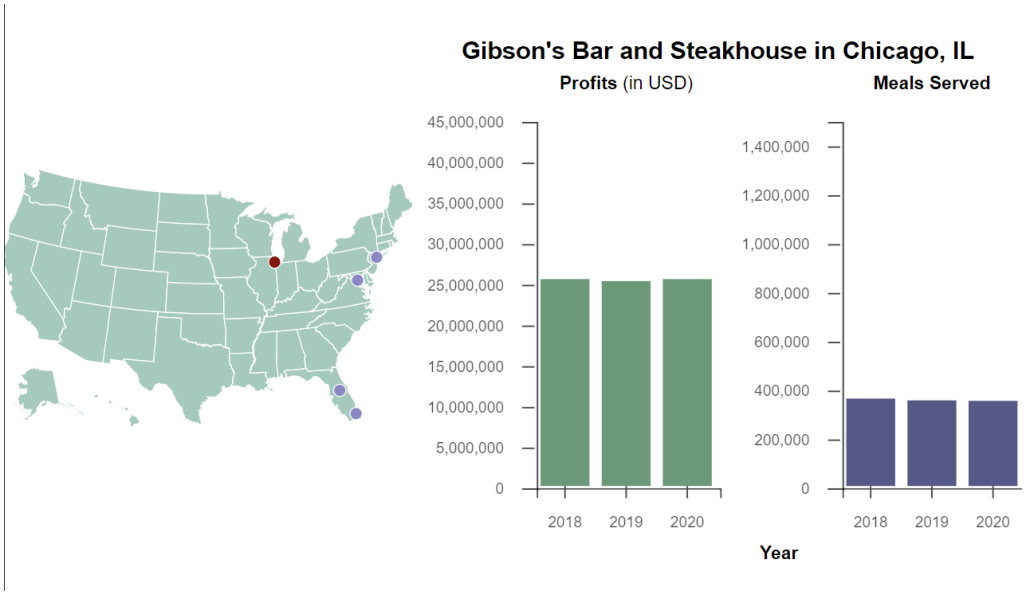




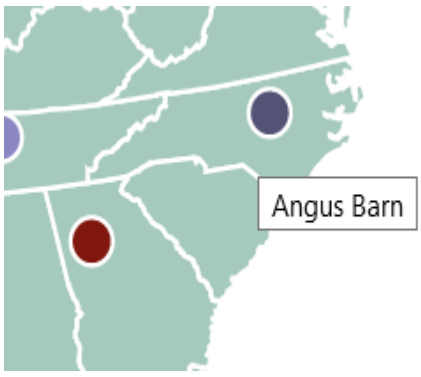
After more discussion, we decided to adjust the concept to have both the map and bar graphs visible at the same time. This adjustment would allow the user to explore the data more easily without having to navigate back and forth between two pages. Between this final sketch and the first d3 implementation, we decided to focus on restaurants country-wide, as there were few states that had a significant number of restaurants in our dataset. The data used for this visualization comes from Restaurant Business Magazine's top 100 restaurants.



v0: visualization on start

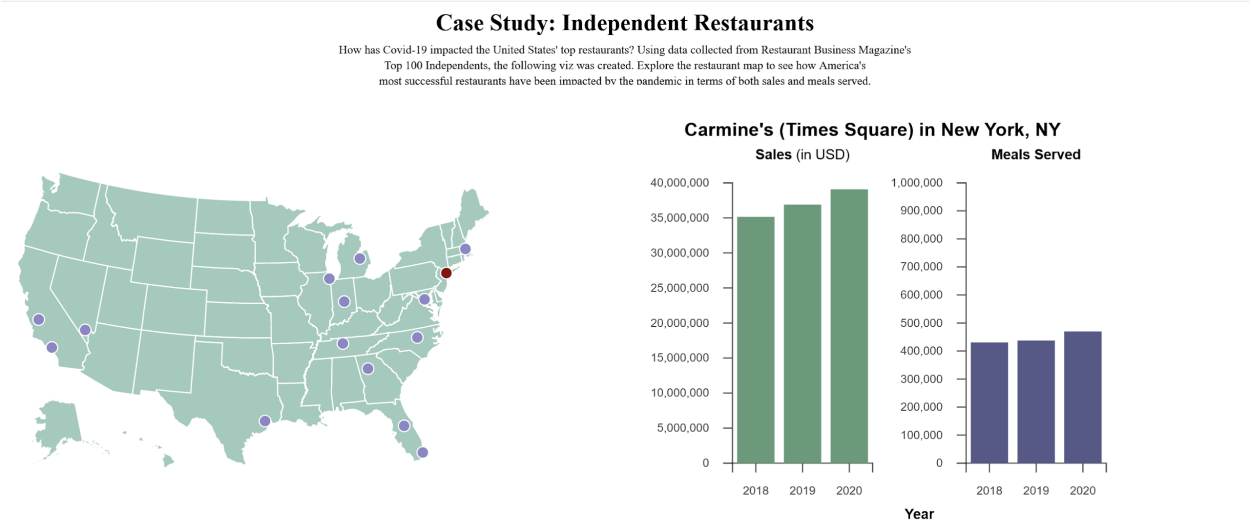


v0: after selecting a restaurant on the map



Like in the concept sketch, each dot representing a restaurant is located in its corresponding spatial position. Upon hovering over a dot, the dot becomes a darker shade to imply that it is clickable. Upon clicking a dot, it becomes red (until a different restaurant is selected) to make it clear which restaurant's data the user is looking at. Hovering over any dot displays a tooltip with the name of the restaurant.

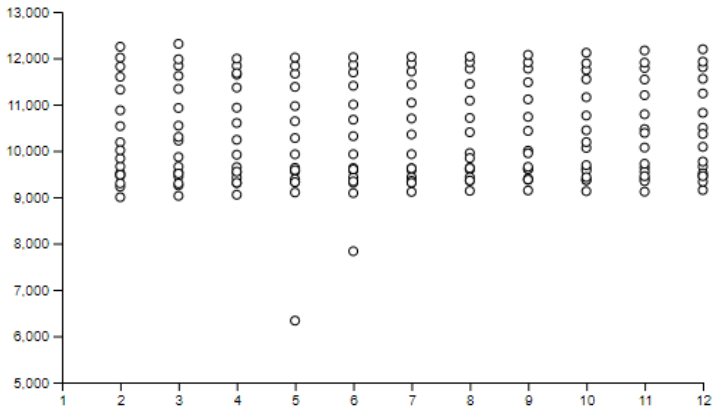
The bar graphs use two different values to represent that they measure two different types of data. Additionally, the bars animate to their new values when a different restaurant is selected to draw the eyes of the user to the change, as an attempt to avoid the issue of “change blindness”.



v1: The map fully populated with data and accompanying text added

The final version of this visualization before it is integrated into the React project can found here: https://resallie.github.io/RestaurantCaseStudy/

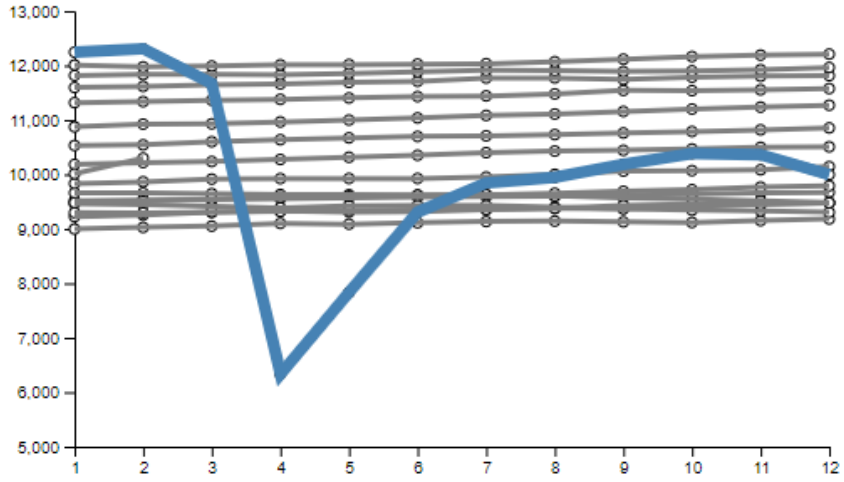
Labor Statistics Design Evolution



v0: Simple Scatter plot of Total Employees By Month of the year

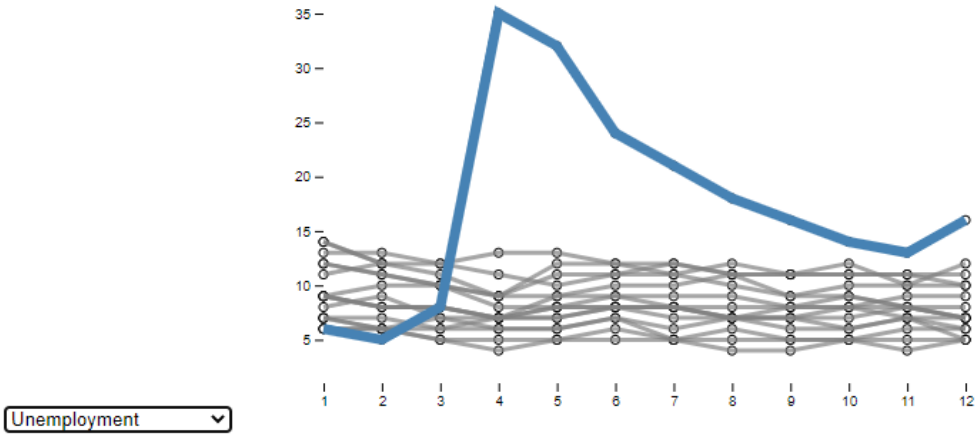
The first iteration of this design was a scatter plot showing the points. We wanted it to show the yearly trends especially as it comes to later implementations, but it obscures which

point is associated with what year. So we needed to connect each year together, and make it obvious which year was which.



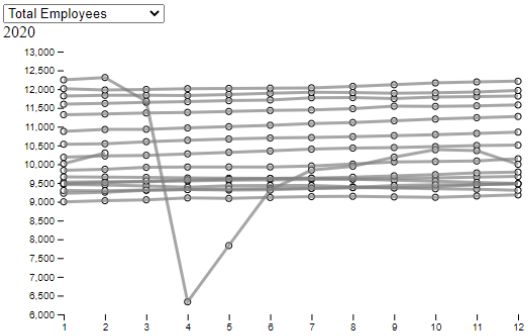
v1: Added a line and mouseover listener to each year line

The second iteration of this design showed the lines connecting each of the points. It does a good job of showing the stability of the data and then how crazy the 2020 data is comparatively. However, no one would have known that this dip is 2020! Labeling the chart is currently lacking, so I needed it to look better in that regard.



v2: Choose the data from a dropdown

Before integration could begin, we needed the ability to select the data from the possible charts. After this happened, we got to merging the vis together. This involved adding it to a react component to add to the overall scrolling visualization features.



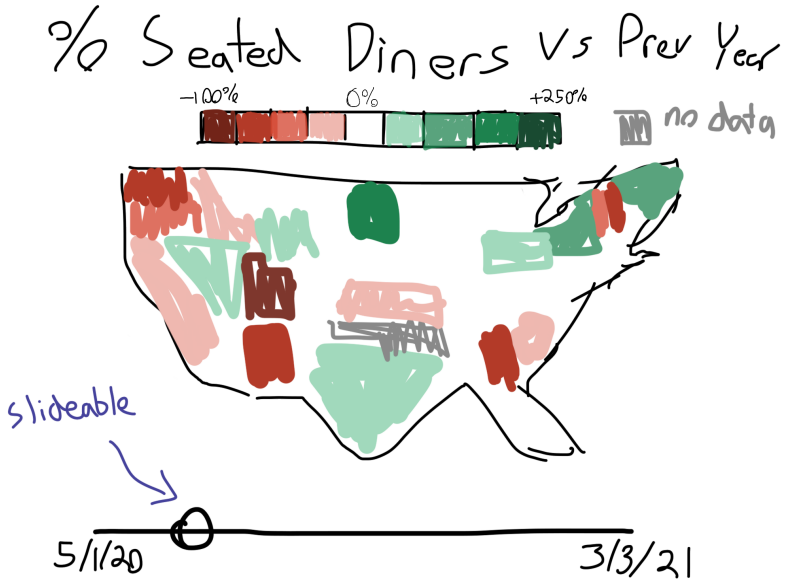
The drop in employees in 2020 was dramatic in the restaurant service industry: the number of workers was cut in half from 1/20-4/20. The total employees axis is in thousands of employees.

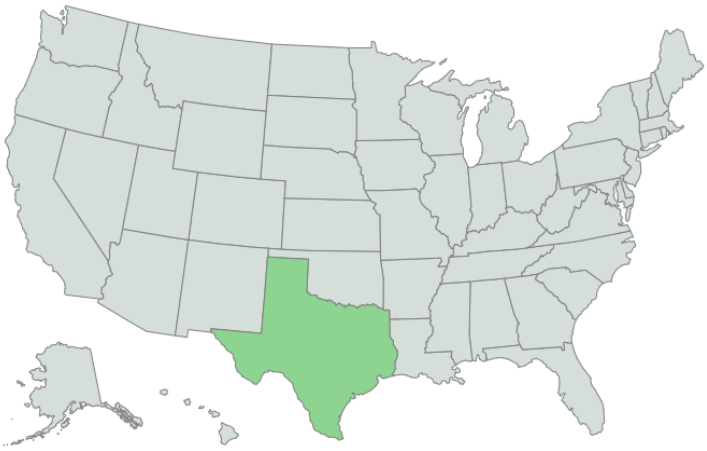
v3: added annotations

The last important step was adding a year text when the user highlighted one of the path lines. Also, adding a description to draw the users attention to elements of the chart was key to improving the interaction with the vis. Done!

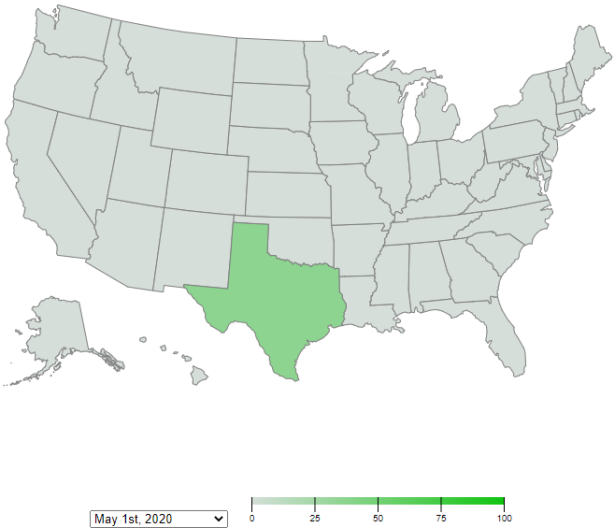
Seated Diners Design Evolution

For this visualization a simple static map using a similar development process to the restaurant case study visualization was implemented. This concept uses a choropleth map to visualize how many seated diners there were each day of the pandemic compared to the same day the previous year. For the color, it seemed a clear choice to use white as a baseline for no change, while increasing the red value to represent a decrease in diners, and increasing the green value to represent an increase in diners (Additionally using a gray to fill in states on days we had no data for). For this initial design, we choose to use a slider to allow users to explore the data for different dates.

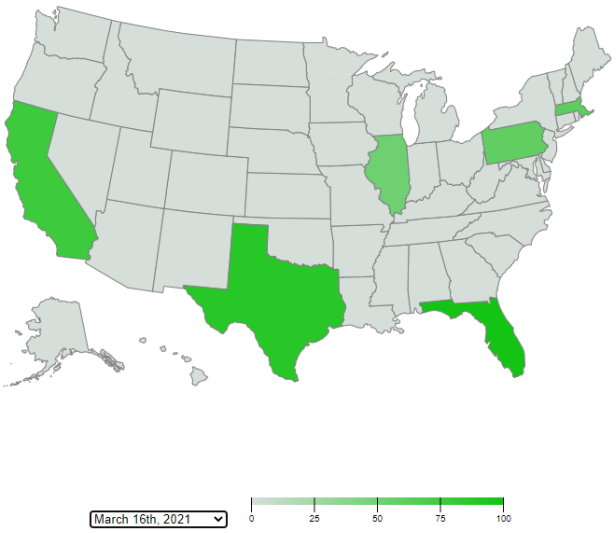




With this base map visualization we wanted to visualize the percent increases in time of seated diners in restaurants as the data set from Open Table began on May 1st, 2020 where many of the states surveyed were still relatively closed for dine in access. Using a choropleth design with this map we were able to visualize the percent of restaurants captured by Open Table to see how customers are responding to states reopening plans. As restaurants accept more reservations there may begin to be a relationship with sentiment of customers willingness to return to these establishments. This is visualized below where the first figure captures May 1st, 2020 and the second figure does the same for March 16, 2021.



May 1st, 2020



March 16th, 2021

Using these changes in color emphasize the level of return restaurants are seeing in the states captured. Building this visualization as a react component was quite interesting as it gave more freedom to scale the methods I ran to help visualize the changes through the react button. Given more time and data this visualization could be used to visualize the rest of the United States if given the opportunity.

Final Implementation

Our implementation of visualization was built using d3 and react. React was used for its great development server, ease of deployment, and component based organization. The development server allows for live compilation of code in development and the component based organization allows for independent work on visuals that can then be formatted together.

React's components allow for props, which can be passed into a component and displayed within the component like a normal HTML element. This was used for the scrolling animation which wraps the visuals in a component that handles evaluating whether or not the vis is onscreen.

Evaluation

Restaurant Case Study

This visualization led to some surprising discoveries. Though some expectedly had decreased sales in 2020, despite the pandemic, many restaurants actually had an increase in both sales and meals sold! With more time, this would be fascinating to explore further. Perhaps this is due to restaurants adapting quickly to offering take-out and delivery, or perhaps it is

biased due to the data being the top 100 restaurants, which have the funds and fame to continue succeeding despite the pandemic. It would be interesting to investigate the success of average restaurants to see if they fared just as well.

One area this visualization is lacking in is the fact that due to using the same scales for every restaurant, it becomes difficult to see the differences in the bars for restaurants that sell relatively fewer meals or make much less money. A possible solution would be to use custom scales for each restaurant in order to highlight the yearly differences. This however results in the trade-off of reducing the user's ability to compare data across restaurants for the benefit of more easily being able to compare a restaurant to itself over the years. Another alternate solution would be to only use restaurant data that falls within a specific range, making it easy to compare both year to year data and restaurant to restaurant data.

In terms of allowing the user to explore the data, it does its job well. Color changes and animations make the effects of user interaction clear. The bar charts clearly show whether sales and meals served in 2020 have increased, decreased, or stayed the same.

Labor Statistics Chart

This visualization went in an unexpected direction for me as well. When first examining the statistics, I assumed that they would all trend as dramatically as the unemployment and total employees charts. However, some just didn't fit this trend. Wages, for one, didn't really take much of a hit during the pandemic. Hours worked also stayed about average but went slightly down.

One overall takeaway from the labor statistics charts is that many people lost their jobs during the pandemic, but the ones who kept working made almost as much as they normally did if no pandemic had occurred. This statistic in particular brought an idea to light that I had not considered: a loss in a business may likely be distributed across a subset of individuals employed by the business and the rest leave relatively unscathed. This is compared to an approach like the normal hours taking a big hit and the number of unemployed people not increasing much.

Another overall takeaway is the scale of the disturbance to the industry. Unemployment to 35 percent?? So 1 in 3 people were out of work, that is very significant. The total employees in thousands tells an even darker story, where 50% of people stopped working for restaurants between 1/20 and 4/20.

Clearly, the industry was rocked by COVID-19, but not everyone was affected in the same way. Given more time to create a visualization, additional categories and statistics could be included, and more annotations added depending on what the user was looking at.

Seated Diners at Restaurants Open for Reservations Across the United States

The data used in this visualization was a part of an early reference we found during our data exploration process from Open Table. Seeing the dramatic changes that have occurred over the year as well as states independently altered their reopening plans due to the ongoing pandemic as well as even seeing the months states allowed restaurants to reopen. Creating this choropleth has showcased the differing sentiments by state about returning to restaurants and this visual has shown that states with weak to no restrictions such as Texas.

I do think though there's a lot of room for improvement with this visualization. I believe allowing users to see the current percentage of "seats covered" by restaurants when compared to 2019 would show how a state is either performing just as well or less to the baseline. Taking it a step further I would have made individual line charts to simulate time lines for states in order to juxtapose this data to the announcements of changes in a state's reopening plan. I think this is an incredibly important attribute to include with this dataset to identify trends based on the stage of the plan.