**How to help restaurants survive COVID-19**

#### Overview of industry, business, or problem

Congress passed a $25 billion COVID-19 bailout for the airline industry but not one tailored to the restaurant industry, which is four times bigger in terms of sales and 18 times bigger in number of jobs (restaurant industry is the nation’s second-largest private-sector employer with an employee base of 15.6 million). Restaurant industry losses are on track to top $240 billion by the end of 2020 — more than any other industry.

#### Define the specific problem that should be solved

How to help the restaurant industry survive?

1. Identify who needs a delivery partner, technology immersion, switch to curbside pick-up or keep their strategies.
2. Determine if restaurants must sell groceries and build/fortify partnerships with farmers.

How to answer those questions:

* Analysis of consumer preferences (availability to coming back to restaurants or not, concerns about the virus, likely to buy food through delivery services, curbside pick-up or in-store, likely to buy prepared meals or ingredients to cook at home, likely to use a mobile app to buy food, contactless payment methods or cash instead).
* Analysis of foot traffic and mobility patterns: detecting businesses in areas with less foot traffic, businesses far away from recreational areas, transit stations, in areas where people are less likely to come back to restaurants and are more concerned about getting the virus.

1. According to the Blueprint for Restaurant Revival (a plan elaborated by the National Restaurant Association that includes short and long-term relief to restart the restaurant industry), the industry asks for:

* Federal funding to programs accessible by states and localities to partner with restaurants and nonprofits to prepare meals for vulnerable populations including seniors and underprivileged children and healthcare workers. These programs would allow nonprofits, which have been collaborating with restaurants since the beginning of the pandemic, to continue partnering with restaurants to prepare food for these at-risk communities and hospital workers.
* Increase access to Restaurant Meals Program for low-income americans.

How to contribute to this effort:

* Identify vulnerable populations in the city selected: affordability of food, groups under the Restaurant Meals Program and excluded people that potentially could be included in the next few months. This information is valuable for non-profit organizations trying to build partnerships with restaurants to keep their operations.

#### Why does this problem matter?

Restaurants need to know what investments are likely to help them survive. With consumer fears over human-to-human contact at an all-time high, tools that allow restaurants to conduct business while eliminating touchpoints have risen in popularity. **This trend is likely to continue even once dining rooms reopen**, so restaurants will be investing in systems that support contactless dinings, such as mobile payment and ordering.

Regardless of when restaurants decide to reopen, **the takeout orders** that allow them to hang on during the roughest weeks **will continue to be essential to the slow rebuilding of their business during the next year**.

#### Potential Audience. Who cares about the results?

1. Tech companies should be one of the most interested audiences, in the sense that restaurants are going to need to invest in more low contact technologies to provide contactless dinings (mobile payment, ordering). That translates into apps and websites usable by all age groups, investment in network security, tools to make delivery and pick-up smoother, apps to improve the pick-up experience and also some kind of share apps where diners can place a single order and pick up multiple items at different brands.
2. Non-profit organizations doing campaigns to help local restaurants (ex. The Meal Bridge in Athens (GA), DC Metro Area, Seattle, Minneapolis, Los Angeles, ChefsForAmerica (nationwide emergency food relief program). In the Blueprint for Restaurant Revival, the Restaurant Association is explicitly asking for:
   1. federal funding to programs connecting restaurant owners and chains with vulnerable groups and healthcare workers, to keep the operation in restaurants and,
   2. expand the population eligible for benefits under the Restaurant Meals Program. For the Household Pulse Survey 2020 it’s possible to extract information related to affordability of food and people receiving free-meals. We can analyze and identify people who are not part of the initiative but they should.
3. Community Supported Agriculture, Community Farmers Markets: Their partnerships with restaurants are crucial to keep their own operations.

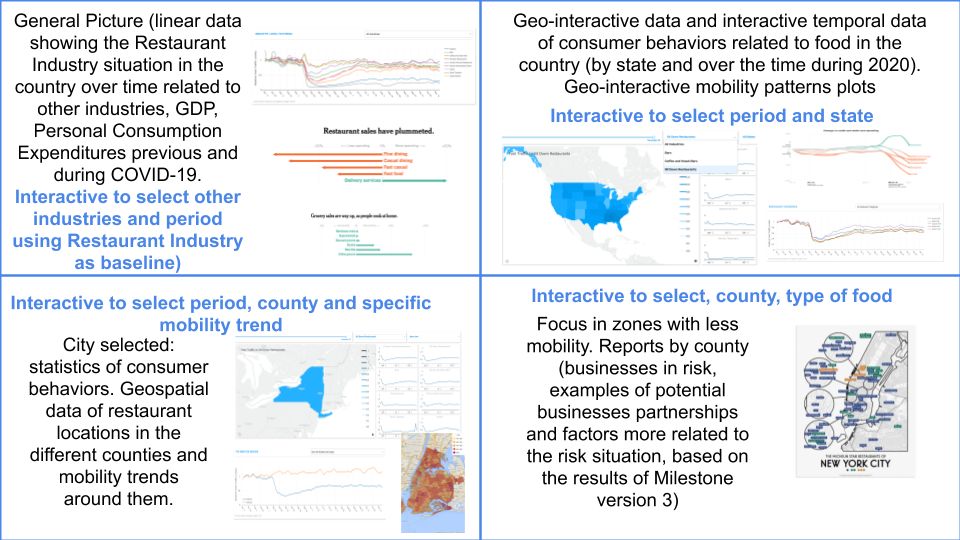
#### What are the datasets that you will consider to solve this problem?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Category** | **Datasets** | **Source** | **Features** | **Topics** | **Related topics** |
| Bureau Economic Analysis Data | 1. Domestic Product and Income by Industry and Expanded Detail  2. Personal Consumption Expenditures by Major Type of Product  3. Income and Employment by Industry (until 2019) | <https://www.bea.gov/data/by-place-us> | Measured per qtr; will be updated Nov 25. | Trends of GDP in food service, employment, incomes(Table 1.5.3. Real Gross Domestic Product, Expanded Detail, Quantity Indexes) | Determine GDP trends in food services vs food purchased during the pandemic.  Compare performance with other categories of durable-nondurable goods and services. |
| Analysis of expenditures in food purchased and food services (Table 1.5.1. Percent Change From Preceding Period in Real Gross Domestic Product, Expanded Detail and Table 1.5.2. Contributions to Percent Change in Real Gross Domestic Product, Expanded Detail) | Determine the change in personal consumption expenditures in food purchased for consumption vs food services per qrt 2018-2020. |
| Monthly Retail and Food Services Sales and Inventories | 1. Estimates of Monthly Retail and Food Services Sales  2. Estimates of Monthly Retail Inventories/Sales Ratios | <https://drive.google.com/drive/folders/1HjBamjglzwTK_PZ7VxyPBZDtPe7nWiTk> | Data 1992 to 2020. | Compare performance of food/beverage stores vs food services. |  |
| Mobility Patterns | 1. Apple Mobility Reports | <https://drive.google.com/drive/folders/1NB1oEsFE33XczpebAINhuFH828U2fCbk> | Relative amount of route requests from every region/date.. Information broken into states, counties, date from Jan 2020-now | Transit | Can look at transportation type preferences by county: walking, driving, transi. Identify zones with more foot traffic. |
| 2. Descartes Lab Mobility Change | <https://drive.google.com/drive/folders/1oDDY1Vhpaxa29Iy1_Mj6zU1aZw6pEkGr> | The distance a typical member of a given population moves in a day (kms). Information broken into states, counties, dates. | Transit | Can look for radius by county to know where people are more likely to move for getting food. |
| 3. Google Community Mobility Data | [https://www.google.com/covid19/mobility](https://www.google.com/covid19/mobility/)/ | Global information is broken down into counties.    Keeps track of mobility changes as a percent.  Industries tracked: retail, grocery, parks, transit, workplaces, residential | Transit | Can look at public transport changes, mobility for different purposes.  The trends of mobility to groceries, parks, workplaces and residential give us indirect information of the potential flow of people to restaurants close to those areas. |
|  | 4. Foursquare Community Mobility Data | <https://drive.google.com/drive/folders/1_lJWAae0MtYr7U_k3RUqBcaDhqa3siRv> | Visits, average duration in minutes and median visit length in minutes to different categories of places. | Visits, duration of visits to Food stores and Fast Food Restaurant. | Compare visits to food stores and fast food restaurants with other categories. |
| Household Pulse Survey |  | <https://www.census.gov/data/experimental-data-products/household-pulse-survey.html> | Phase 1 (April/August) asked about employment status, food security, housing, physical and mental health, access to health care, and educational disruption.  Phase 2 (August - current) asks about benefits, spending patterns, and availability of financial resources, post-secondary education disruptions, the capacity to telework, and travel practices. | Affordability of food, free meals and spending use of the Economic Impact Payment | Recognize groups eligible for the social food programs but not included (insights for the National Association of Restaurants). |
| Shopping and purchase preferences. | Shopping modalities, payment modalities, resumed/avoided eating at restaurants. Use of credit cards, apps to buy online. Consumer preferences (prepared food vs ingredients to cook at home) |
|
| Trips and teleworking variables | Fewer transit trips, planned trips, trips to stores (give us information about likely to leave the home to buy meals vs use of delivery) |
| COVID-19 | The COVID Tracking Project | <https://covidtracking.com/data/national/> | COVID Data sets per location | States  New tests  Cases (confirmed plus probable)  Negative PCR tests (people)  Cumulative hospitalized/Ever hospitalized  Currently hospitalized/Now hospitalized  Deaths (confirmed and probable)  Recovered  Total test results | Connecting to Household expenditures and mobility patterns to understand the impact accordingly. |
| Restaurants | Yelp dataset | <https://drive.google.com/drive/folders/1mp2texeym4VJbnPQFFFxnYyNFMjNRInu> | Name, location, status (open, closed), attributes (take-out, outdoor dining, parking), categories (type of food), hours, stars, reviews. | Restaurant current status |  |

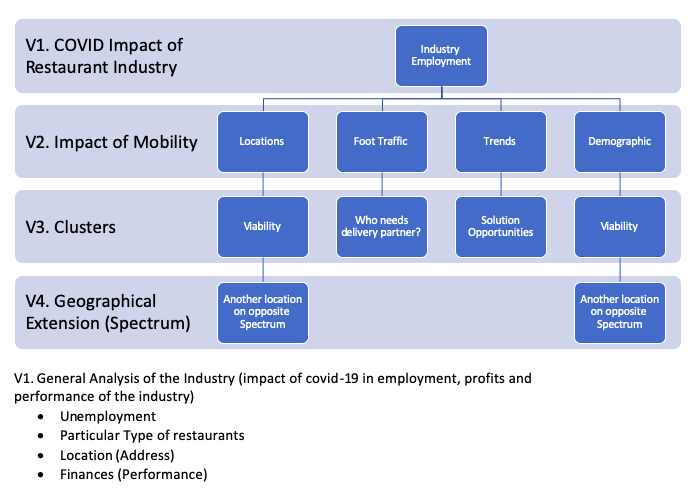
#### Methods

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Milestone** | **Topic** | **Data** | **Questions** | **Visualizations and Methods** |
| COVID-19 Impact on Restaurant Industry | General Picture of the Restaurant Industry related to the rest of industries. | Bureau Economic Analysis Data  The COVID Tracking Project | GDP trends in food services vs food purchased during the pandemic.  Performance with other categories of durable-nondurable goods and services.  Contrast industries performances with COVID-19 tracking along the time. | Linear plots, over time (univariate) analysis. Check if the Restaurant sector is one of the most/less affected. |
|
| General Picture of consumption expenditures and Consumer Behaviors (country/cities) | Bureau Economic Analysis Data | Determine the change in personal consumption expenditures in food purchased for consumption vs food services (2018-2020) | Linear plot over time of change in personal consumption expenditures in food purchased for consumption vs food services.  Linear plot of food categories related to others in the same period. |
| Household Pulse Survey | Spending Patterns (general and by city) | Spending use of the Economic Impact Payment per category (barplots) and over the time. Contingency tables. |
| Shopping and purchase preferences and relation with COVID concerns (general and by city). | Analysis of shopping modalities, payment modalities, resumed/avoided eating at restaurants, use of credit cards, apps to buy online, consumer preferences (prepared food vs ingredients to cook at home).  Correlation between COVID-19 concerns and changes in shopping behaviors (heatmaps, Pearson coefficients, interactive plots over time to see changes during the year and contingency tables).  Trips analysis: Transit trips, planned trips, trips to stores over time, correlation with COVID-19 concerns using heatmaps and contingency tables.  Based on that, do hypothesis and statistical tests ( ANOVA, chi-squared tests). |
| Impact of COVID-19 in mobility and restaurant situation in one particular city. | Mobility patterns | Apple Mobility Reports | Transportation type preferences by county: walking, driving, transi. Identify zones with more foot traffic. | Heatmaps, barplots, line-plots (univariate analysis over time and per location) |
| Descartes Lab Mobility Change | Look for radius by county to know where people are more likely to move for getting food. | Heatmaps, scatterplots, line-plots (univariate analysis over time and per location) |
| Foursquare Community Mobility Data | Visits, average duration in minutes and median visit length in minutes to different categories of places. | Compare visits to food stores and fast food restaurants with other categories. Hypothesis, barplots, scatterplots, heatmaps (univariate analysis over time and per location) |
| Google Community Mobility Data | The trends of mobility to groceries, parks, workplaces and residential give us indirect information of the potential flow of people to restaurants close to those areas | Compare mobility to different places. Hypothesis, barplots, scatterplots, heatmaps (univariate analysis over time and per location) |
| Operational situation of restaurants | Restaurants Yelp dataset | Identify location, status (open, closed), attributes (take-out, outdoor dining, parking), categories (type of food), hours, stars, reviews. | Barplots, group by restaurants county. Contingency tables. Hypothesis counties more/less affected and why. |
| Insights for restaurants in the city selected | Consumers situation in city selected | Household Pulse Survey dataset | Based on the city selected, in-deep analysis of consumer preferences to identify types of consumers more likely or not to use technology to buy food, to choose delivery, to resume eating at restaurants, to do more trips, etc. | Grouping and recognizing different types of consumers. Chi-squared tests, barplots, contingency tables. Clustering. |
| Recognize groups eligible for the social food programs but not included (insights for the National Association of Restaurants). | Grouping and recognizing vulnerable groups out of free meal programs. Clustering. |
|  | Variability mobility and restaurant situations by county and demographics. Recognize groups in risk. | Restaurants Yelp data  Mobility data | Merge mobility and restaurant situation by counties, locations.  Identify how to help restaurants in risk (who needs delivery, partnership with other restaurants and local businesses). | Hypothesis, contingency tables. Clustering.  Identify predictors of risk (using the current status of the restaurants). Logistic Regression, Decision Trees (or other Ensemble Tree Method to relate mobility variables as features to determine current status of restaurants) |
| Geographical Extension | Another location on the opposite spectrum (weather, spreading of the virus, consumer preferences, mobility trends, such as more/less use of cars, use of technology). |  |  |  |

#### Interface



#### Milestones



#### Timeline

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Date | Timing | Milestone version | Detailed Tasks | Team member | Deliverable Date |
| Week 1 |  |  |  | Team | Team Formation |
| Week 2 |  |  |  | Team | Work on idea formation |
| Week 3 |  |  |  | Team | Project Description (11-12-2020) |
| Week 4 |  |  |  | Team | Project Scoping (11-19-2020) |
| Week 5 | from 11-21-2020 to 12-03-2020 | Final collection of datasets for V1 and V2 | Collection of dataset related to employment in the Restaurant Industry during 2020  Collection of Restaurant Industry by types of food and other useful categories | Alejandro | Dataset Details (12-03-2020) |
| Starting Data Cleaning V1-V2 | Household Pulse Survey | Daniela |
| Domestic Product and Income by Industry, Personal Consumption Expenditures by Major Type of Product, Income and Employment by Industry (until 2019) | Pablo |
| Monthly Retail and Food Services Sales and Inventories | Troy |
| Mobility datasets | Team |
| Week 6 | from 12-04-2020 to 12-10-2020 | Finish Data Cleaning V1-V2 | Integrate clean datasets in GitHub | All members | Data Cleaning Methodology (12-10-2020) |
| Clean scripts and notebooks with the pipeline approach used. | All members |
| Report with Data Cleaning Section. Diagrams showing the cleaning process. | Daniela |
| Starting EDA | EDA first approach general picture Restaurant Industry | Team |
| EDA first approach mobility trends | Team |
| EDA first approach consumer preferences, purchase preferences, relation with COVID concerns, spending patterns visualizations | Daniela |
| Week 7 | from 12-11-2020 to 12-17-2020 | V1 progress: 100%  V2 progress: >60% | Consolidate and connect results, extra plots and analysis required and include the EDA Section in the Report. | All members | Basic EDA (12-17-2020) |
| Week 9 | from 12-18-2020 to 01-14-2021 | V1 progress: 100%  V2 progress: 100%  V3 progress: >80% |  |  | In-depth Analysis + Front End Mock-up (01-14-2021) |
| Week 10 | from 01-15-2021 to 01-25-2021 | V1 progress: 100%  V2 progress: 100%  V3 progress: 100% |  |  | Frontend Design (01-25-2021) |
| Week 11 | from 01-26-2020 to 02-01-2021 | V1 progress: 100%  V2 progress: 100%  V3 progress: 100%  V4 progress: 100% (optional) |  |  | Analysis Complete (02-01-2021) |
| Week 12 | from 02-01-2021 to 02-08-2021 | V1 progress: 100%  V2 progress: 100%  V3 progress: 100%  V4 progress: 100% |  |  | Presentation Draft, Datafolio and updated report (02-08-2021) |
| Week 13 | from 02-09-202 to 02-13-2021 | V1 progress: 100%  V2 progress: 100%  V3 progress: 100%  V4 progress: 100% |  |  | Final Report (02-13-2021) |

#### Concerns

* Impact of Covid
  + Waves of new COVID cases will likely affect the next series of data. Since the environment is always changing, levels of confidence in predictions from data will not be too high. Trends can still be found regardless
  + Vaccine trials. Open questions about effectiveness and subsequent actions of consumers and industry. Promises of a vaccine could keep restaurants open who would have otherwise closed
* Datasets
  + Mobility.
    - Choosing appropriate locations to do representative case studies, and then finding the data to support it.
      * Some of our current data may miss a key element to understanding the impact of mobility in certain regions on the restaurant industry
* Team
  + Work distribution
    - With the loss of one of our members, we need to replan our distribution and be aware of the impact of feasible output with one person less.
* Strategy
  + We are aware that our project covers a lot of areas that are really important and impactful in a number of industries. We have been careful to keep our scope close to our goals, but there is definitely the possibility to explore widely as we jump into the data in a way that may lead to introducing open-ended questions

#### References

Context

<https://www.restaurant.org/home>

<https://www.nysra.org/uploads/1/2/1/3/121352550/state_restaurant_association_survey_pressrelease_81620.pdf> [New York State Restaurant Association Survey]

<https://www.vox.com/21445198/bar-and-restaurant-bailout>

<https://www.safegraph.com/data-examples/covid19-commerce-patterns>

Other Datasets

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PFLAH4>

[Foursquare Community Mobility Data with Basemap (US)

<https://aws.amazon.com/marketplace/pp/prodview-cjhkgxpn6vcce?ref_=srh_res_product_title#overview> [This free data set contains indexed foot traffic to 19 categories of venues including Casual Dining and Fast Food. The indexed data is broken out geographically, with included data for National, SF, NYC, LA, and Seattle. The data is normalized against U.S. Census data to remove age, gender and geographical bias. Data is provided daily from 02/19/2020. Access pending]

<https://aws.amazon.com/marketplace/pp/prodview-lwgwq45xwf2nu?ref_=srh_res_product_title#overview> [This free sample data set contains the highest-accuracy visit data AND venue data for ~6M commercial venues in the US for June 2019 - June 2020. Access pending too]

<https://data.cityofnewyork.us/Health/Restaurants-rolled-up-/59dk-tdhz> [NYC restaurants]

Here is the dataset talking about types of restaurants and foot traffic in different metro areas: <https://www.safegraph.com/data-examples/covid19-commerce-patterns>