

A decorative network diagram in the top-left corner of the slide. It features a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are solid blue, some are solid grey, and some are hollow blue. The lines connecting them are thin and grey. The overall shape of the network is roughly triangular, pointing towards the top-left corner.

# Project 1

CSR Consulting:  
Christian R., Rich K., Stefany L

A decorative network diagram in the bottom-right corner of the slide. It features a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are solid blue, some are solid grey, and some are hollow blue. The lines connecting them are thin and grey. The overall shape of the network is roughly triangular, pointing towards the bottom-right corner.

The background of the slide is a light gray network pattern. It consists of numerous small circles, some of which are solid gray and others are hollow with a gray outline. These circles are interconnected by a web of thin, light gray lines, creating a complex, organic-looking structure that resembles a molecular or neural network.

# Summary & Motivation

**Hypothesis: Vernon Hills, IL has the potential to host a new high-end restaurant, using data insights to understand competition and underserved market segments.**

## Summary

We are targeting our presentation on opening a restaurant in Vernon Hills. Using different datasets from Yelp and Census data to decide where to open, type of restaurant, hours of operation, and the competitive landscape.

Our project is to uncover patterns and competitive landscaping in order to determine critical early decisions such as where to establish their business, the competitive landscape to identify underserved market segments, and other data driven insights.

## Initial Questions

- **Where will restaurant be located?**
- **What type of restaurant?**
- **What expense level to target?**
- **What are best hours of operation of our restaurant?**

# How to answer our Questions?

Where will the restaurant be located

- Vernon Hills, IL

Type of restaurant

- Filter by price range, food category

Expense level to target

- Identify household income

What are best hours of operation

- Identify opportunity, pickup/delivery capable

The background of the slide is a light gray network diagram. It consists of numerous small circles, some of which are solid gray and others are hollow with a gray outline. These circles are interconnected by a web of thin, light gray lines, creating a complex, organic pattern that resembles a molecular structure or a data network. The overall aesthetic is clean and technical.

# Data Cleanup & Exploration

## Data Cleaning Process

- Massive amounts of data available but not relevant
- Dropped unnecessary columns in CSV



```
# dropping unnecessary columns
census_data_df = census_data_df.drop(["Fact Note", "Value Note for Lake Forest city, Illinois",
                                     "Value Note for Mundelein village, Illinois", "Value Note for shire village, Illinois",
                                     "Value Note for Libertyville village, Illinois", "Value Note on Hills village, Illinois",
                                     "Value Note for Lake County, Illinois", "Lake County, Illinois"
                                     =1)
```



The background of the slide is a light gray network diagram. It consists of numerous small circles, some of which are double-outlined, connected by thin, light gray lines. The connections form a complex, interconnected web across the entire slide.

# Data Analysis

# Data Sources

## Yelp API

- Search by location radius (Vernon Hills City Hall)
- Filter for name, categories, price range (\$), hours, location (latitude/longitude), # of reviews, and ratings

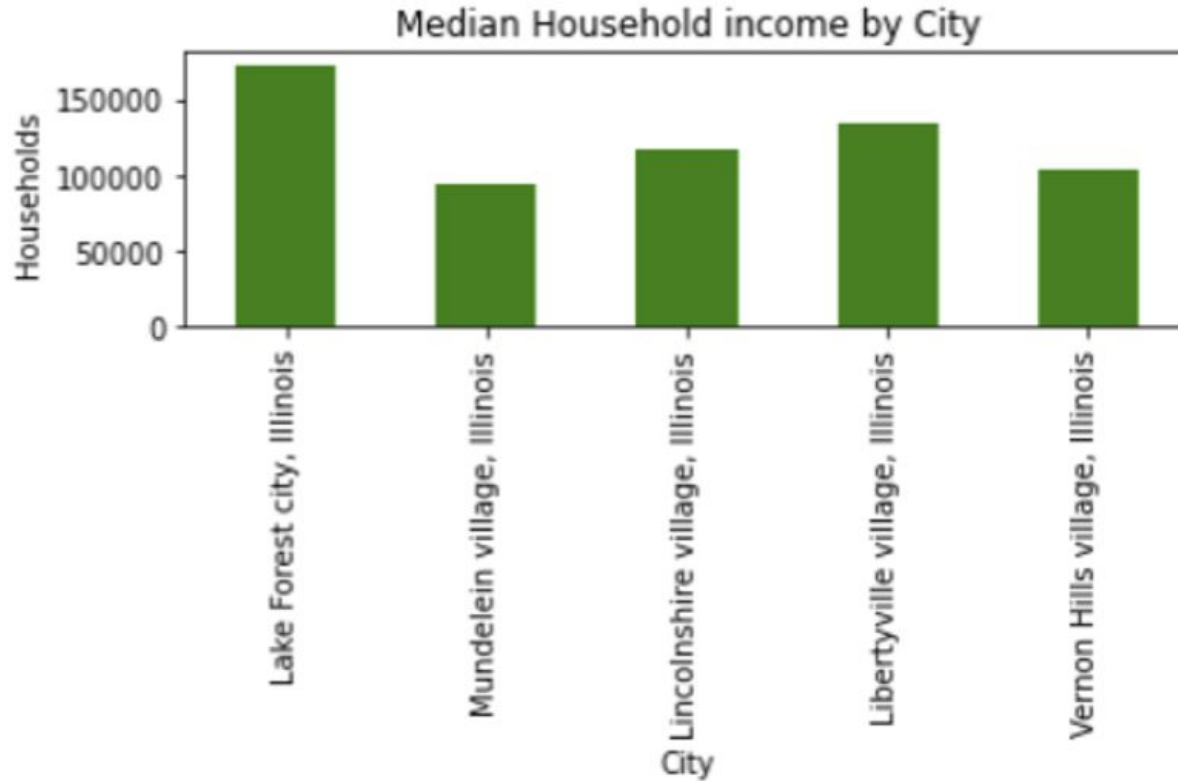
## US Census Bureau

- Find household income, amount of households, broadband



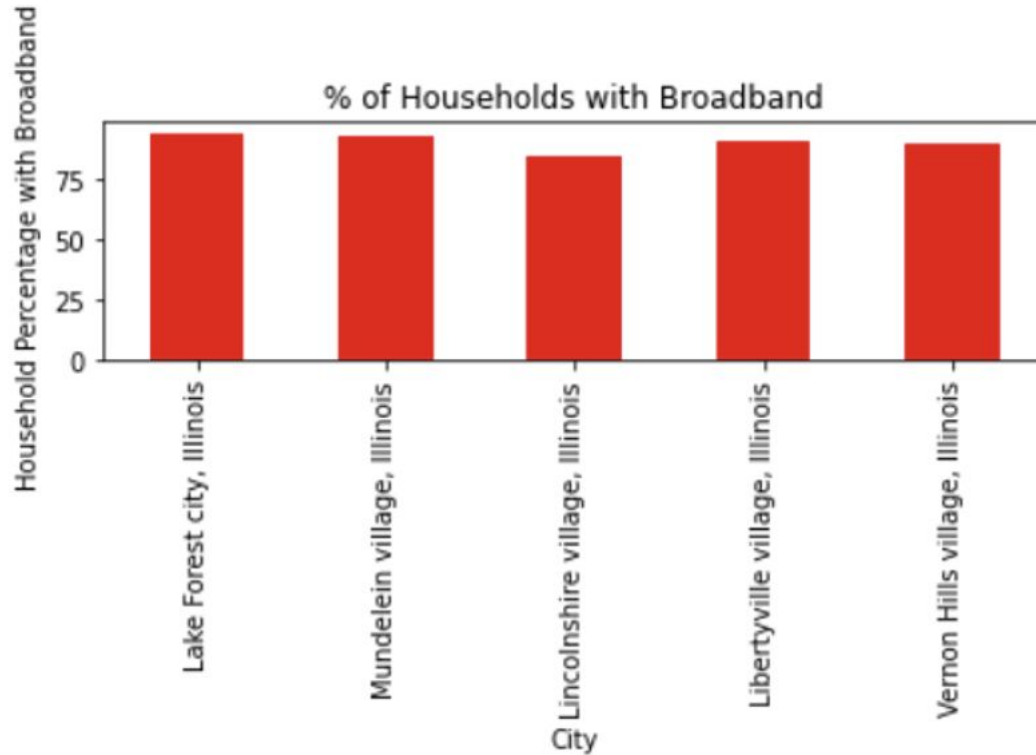


## Vernon Hills Median Income



Source: US Census

## Percentage of Households with Broadband



Source: US Census

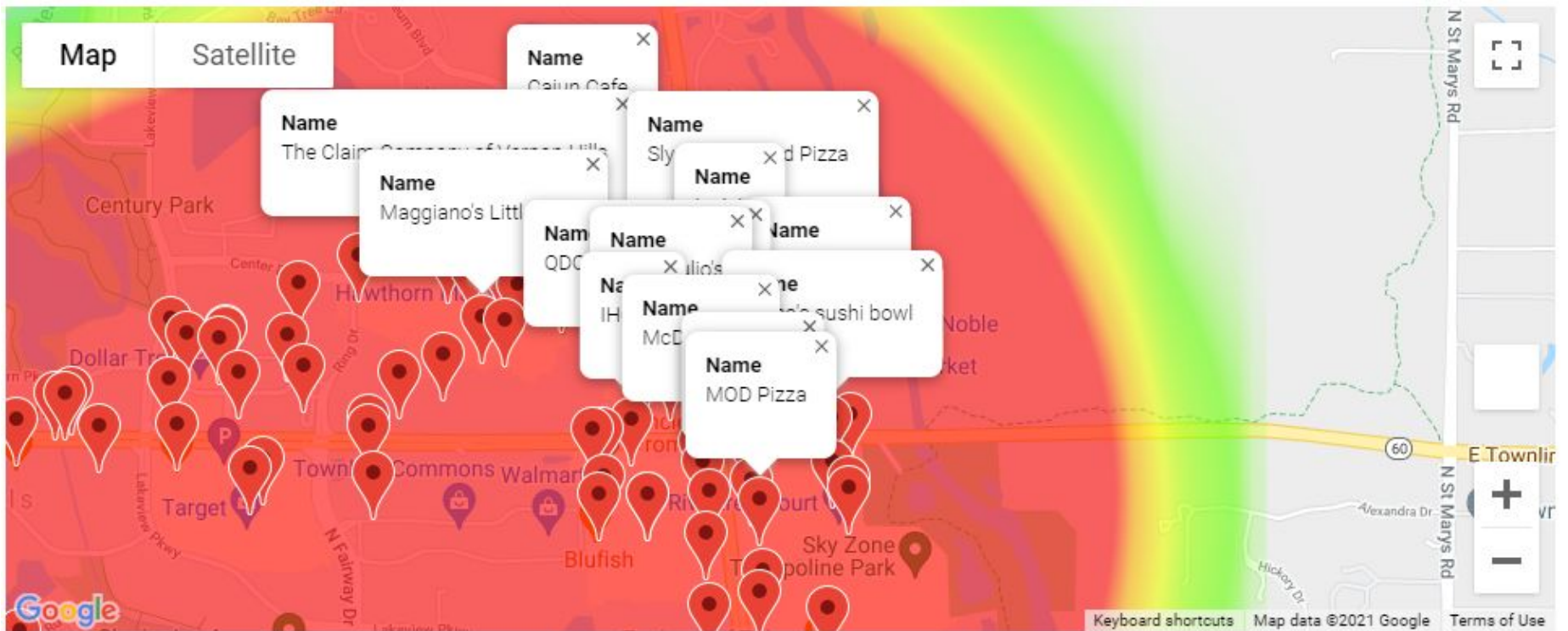
# Data Analysis

- ⦿ Pulled Yelp business IDs within radius to CSV file
- ⦿ Looped through CSV business IDs
- ⦿ Business IDs used to extract info
- ⦿ Census data used to support location

In [5]: `business_df`

Out[5]:

	index	name	coordinates_lat	coordinates_long	categories	review_count	rating	Services
0	0	bb.q Chicken - Vernon Hills	42.240060	-87.941100	chickenshop	20	4.5	[delivery, pickup]
1	1	Bonta	42.199130	-87.933020	cafes	499	4.5	[pickup, delivery]
2	2	Ramen House Shincan	42.238635	-87.941136	ramen	73	4.0	[pickup, delivery]
3	3	Lazy Dog Restaurant & Bar	42.244520	-87.945130	burgers	727	4.0	[pickup, delivery]
4	4	Zocalo Mexican Bar and Grill	42.230035	-87.976905	mexican	213	4.5	[delivery, pickup]
...	...	...	...	...	...	...	...	...
212	212	ALDI	42.252458	-88.019747	grocery	8	3.5	[]
213	213	Jani's Newstand	42.243160	-87.954115	convenience	3	1.5	[]
214	214	Starbucks	42.255691	-87.948152	coffee	25	2.0	[delivery]
215	215	Dunkin'	42.197669	-88.019592	donuts	9	2.5	[delivery]
216	216	Baskin-Robbins	42.253096	-88.020134	icecream	6	1.0	[delivery]



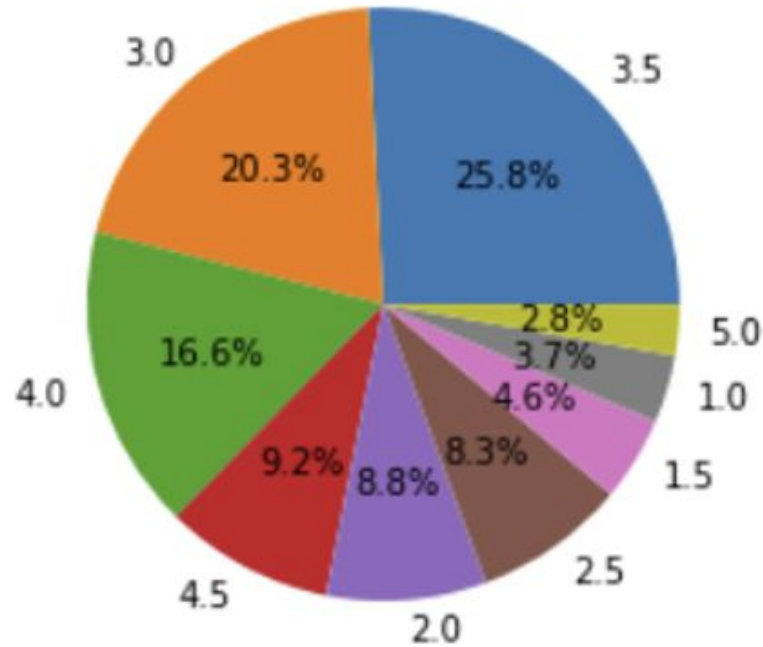
Heatmap based on average ratings and pins with restaurant names  
Average rating in response to the Vernon Hills City Hall.

The background of the slide is a complex network diagram. It consists of numerous small circular nodes, some of which are highlighted with a darker blue dot in the center. These nodes are interconnected by a web of thin, light gray lines. Some lines are solid, while others are dashed. A larger, faint dashed circle is also visible, encompassing a portion of the network on the right side of the slide.

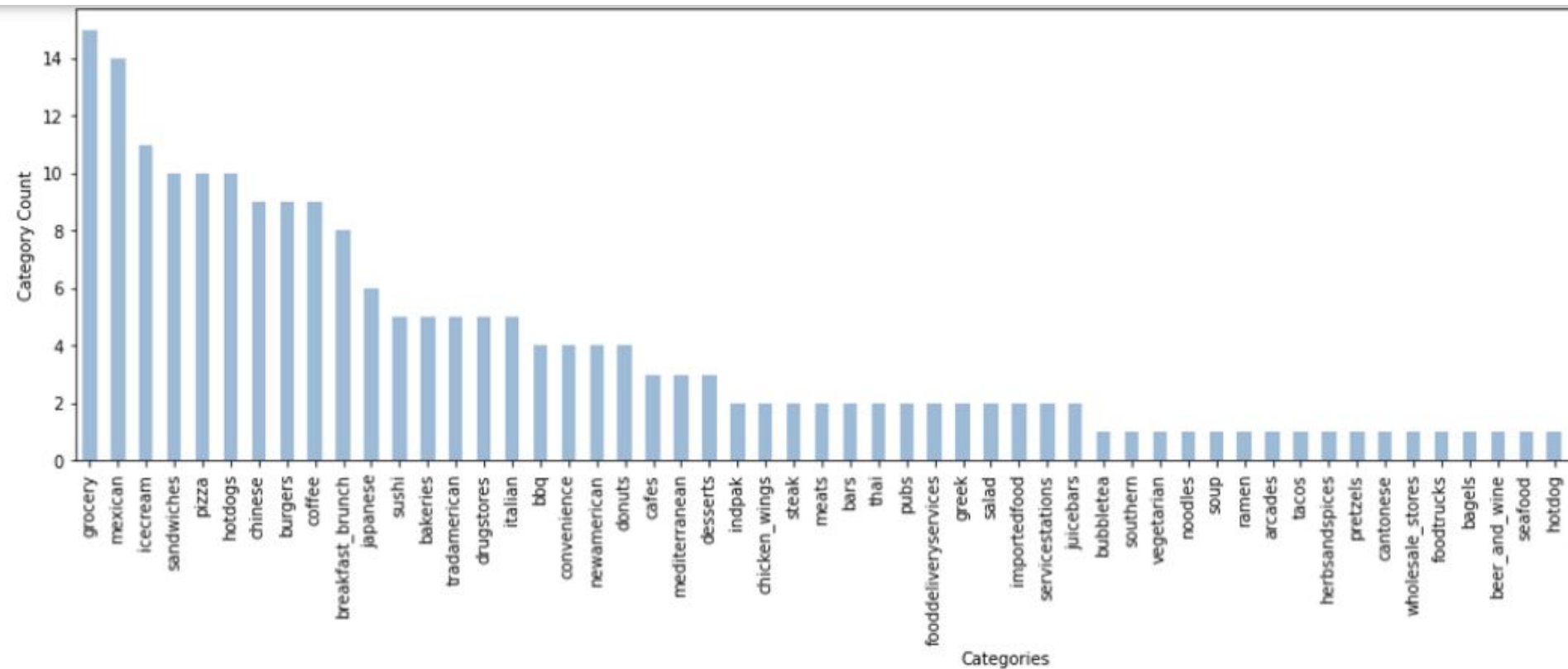
# Discussion

# Percentage of Restaurants with Rating

Restaurant Rating Breakdown







Amount of Restaurants with Categories

The background of the slide is a light gray network pattern. It consists of numerous small circles, some of which are solid gray and others are hollow with a gray outline. These circles are interconnected by a web of thin, light gray lines, creating a complex, organic structure that resembles a molecular or neural network.

# **Issues & Limitations**

## Limitations

- Yelp categories are redundant and may have some overlap
- Businesses may not have a Yelp profile
- # of reviews received vs rating
- Price (\$\$\$) did not pull for many
- Did not have time to dig deeper into age groups or race
- Would like to develop business tool for investors



Categories	review_count	rating	Service
bakery	20	4.5	[pickup,
bars	499	4.5	[delivery]
ben	73	4.0	[pickup,

```
Out[15]: grocery      15
mexican      14
icecream     11
pizza        10
sandwiches   10
hotdogs      10
burgers       9
coffee       9
chinese       9
breakfast_brunch 8
japanese      6
sushi         5
drugstores    5
tradamerican  5
italian        5
bakeries      5
convenience   4
donuts        4
newamerican   4
bbq           4
cafes         3
desserts      3
mediterranean 3
steak         2
greek         2
indpak        2
chicken_wings 2
bars          2
thai          2
importedfood  2
servicestations 2
beer_and_wine 2
```

# Thanks!

## Any Questions?



# Works Cited

- ◎ United States Census Bureau. (2019). *United States QuickFacts*.  
<https://www.census.gov/quickfacts/fact/table/US/PST045219>
- ◎ Yelp Fusion API. (2020). Yelp Fusion. <https://www.yelp.com/developers/documentation/v3/business>
- ◎ Reed, Alex [Sigma Coding. (2018, December 30). *Working With The Yelp API In Python | Part One* [Video].  
YouTube. <https://www.youtube.com/watch?v=GJf7ccRIK4U>