# Average Restaurant Reviews by Income

GEN BUS 760

Team Silver

August 7, 2022

# Summary

- We are looking to open a new full service fine dining restaurant in an underserved market in the United States.

- We leveraged Yelp data from 2005 to 2022 where category_name = 'Food' or 'Restaurants'

- We included Average Income by Postal Code data from the US Census Bureau

- Limits:
    - Avg Income >$80,000 Avg Income
    - Total Pop >10,000 people
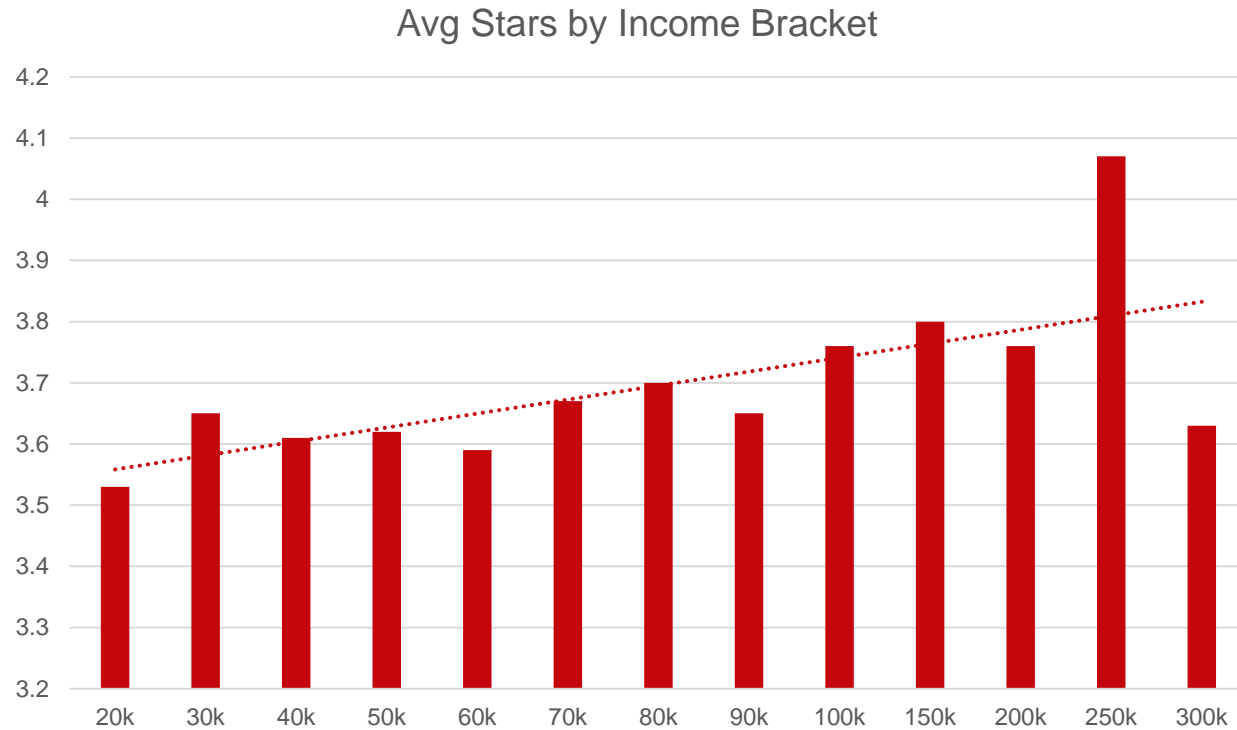
# There appears to be <u>no</u> relationship between Avg Stars and Avg Income

### AVERAGE_INCOME
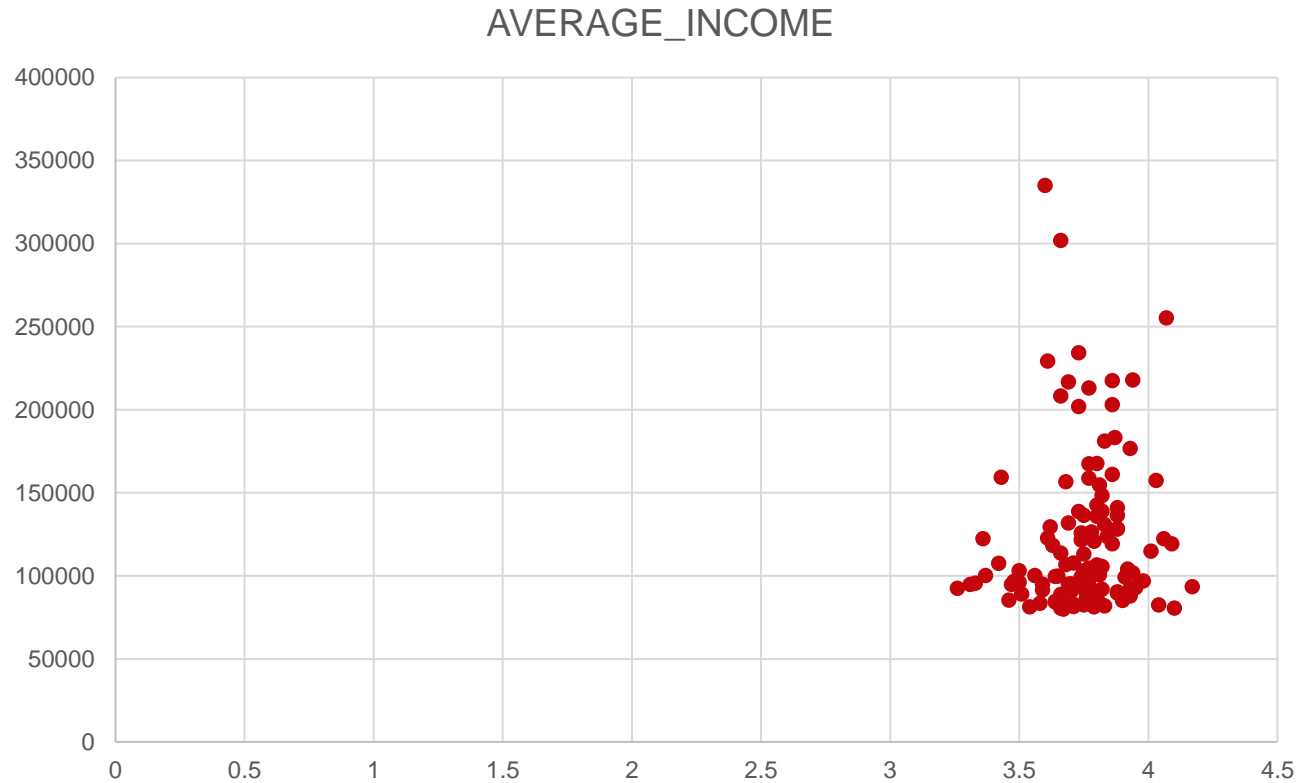


### Limits
- Income >$80,000
- Population >10,000

# Analyzing Income Bracket by Avg Stars reveals a positive correlation

**Avg Stars by Income Bracket**

Limits
- Population >10,000

# There is a slight positive relationship between Avg Stars and Review Index
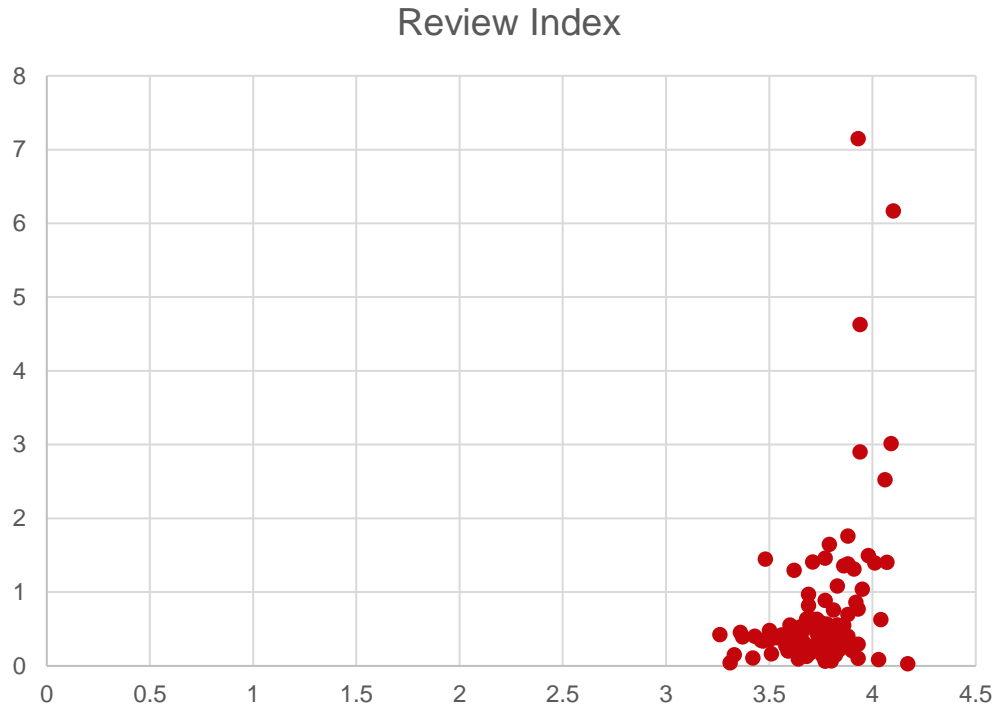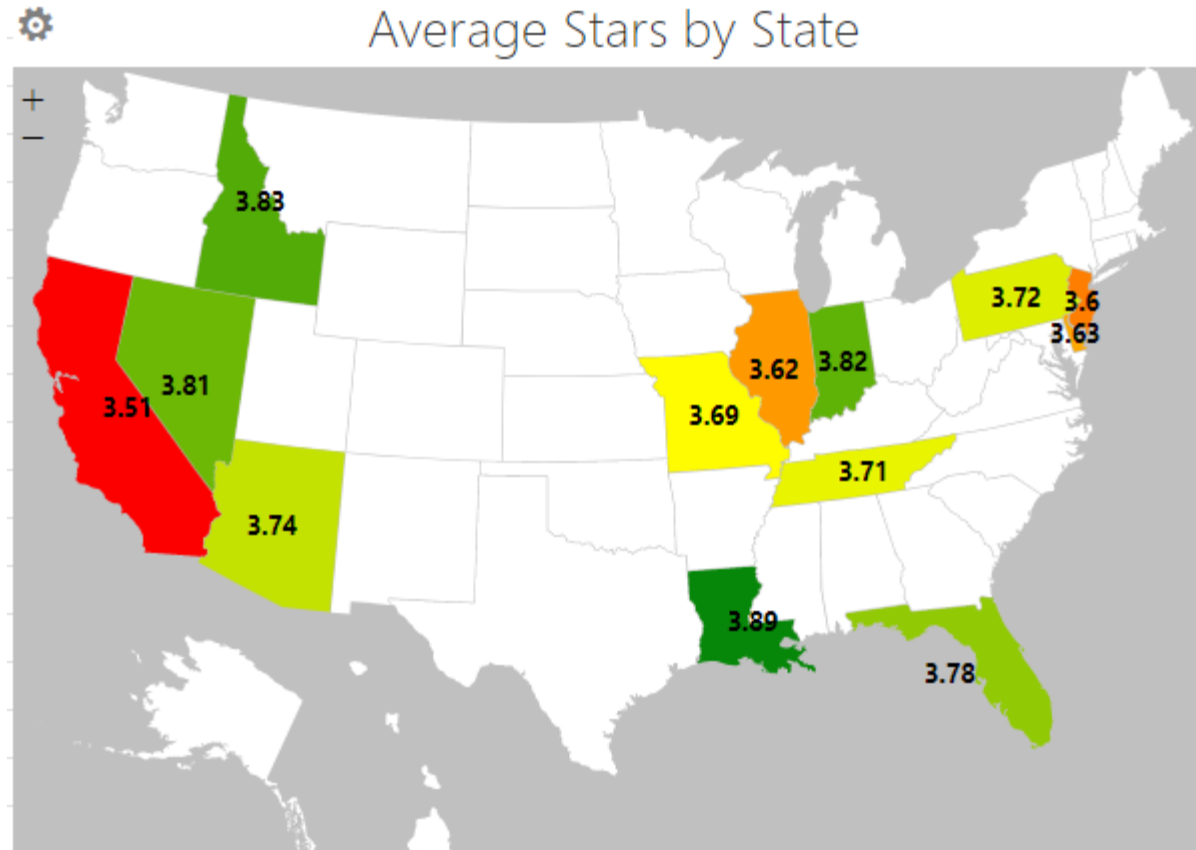
**Review Index**

Limits
- Income >$80,000
- Population >10,000

Review Index = Num Reviews / Total Pop
- Review Index > 1 =
    Total Pop > Num Reviews
- Review Index < 1 =
    Num Reviews > Total Pop

Zip codes with a higher Review Index tend to have higher Avg Stars

# Average Stars by State



Average Stars by State

| State | Average Review |
|-------|---------------|
| AZ | 3.74 |
| CA | 3.51 |
| DE | 3.63 |
| FL | 3.78 |
| ID | 3.83 |
| IL | 3.62 |
| IN | 3.82 |
| LA | 3.89 |
| MO | 3.69 |
| NJ | 3.6 |
| NV | 3.81 |
| PA | 3.72 |
| TN | 3.71 |

# Conclusion and Next Steps

- We would have liked to see a stronger relationship between income and average stars

- A clear opportunity would be represented by a high-earning zip code with a low average stars ratings, but we did not identify such a zip code

- As a next step, we would like to gather more Yelp data from states that were not present in the dataset

# Appendix – Code

```sql
25    SELECT ROUND(AVG(r.stars),2) AS AVG_Stars, PCI.zipcode,
26    ROUND(AVG(pci.avg_income)) AS Average_Income,
27    AVG(total_pop) AS TOT_POPULATION, COUNT(r.stars)AS Num_Reviews
28    FROM business b
29    INNER JOIN POSTCODE_LEVEL_AVERAGES PCI
30        ON PCI.ZIPCODE = b.postal_code
31    INNER JOIN review r
32        USING (business_id)
33
34    WHERE b.business_id IN
35        (SELECT business_id
36        FROM category
37            WHERE category_name = ('Restaurants')
38                OR (category_name ='Food'))
39            AND YEAR(r.review_date) >= 2014
40            AND pci.avg_income > 80000
41            AND TOTAL_POP > 10000
42
43    GROUP BY PCI.zipcode
44    ORDER BY Average_Income DESC
45    ;
46
47    SELECT *
48    FROM STAR_RATING2
49    WHERE NUM_REVIEWS > 100
50    ;
51
```