



Predicting Affluence


Danielle Mizrachi
Jessica Ertel
Leland Roberts
Vanessa Alvarado



Problem Statement

New Light Technologies was contacted by The Wildflower Cafe, a high end restaurant looking to strategically expand their business to new areas of Connecticut.

Questions

1. Can Yelp price estimates for restaurants in Connecticut accurately predict median household income, mean household income, per capita income or adjusted gross income?
 2. Can we confidently advise The Wildflower Cafe which city to expand their business?
- 



Process

01

Acquisition

02

Cleaning

03

Feature
Engineering

04

Modeling

Data Acquisition & Cleaning



IRS

The IRS produces state income tax statistics by zip code. We used 2017 data from Connecticut.



CT Income

The CT.gov produces data on per capita income, mean and median household income for towns in CT. We used data available from 2015.



Yelp

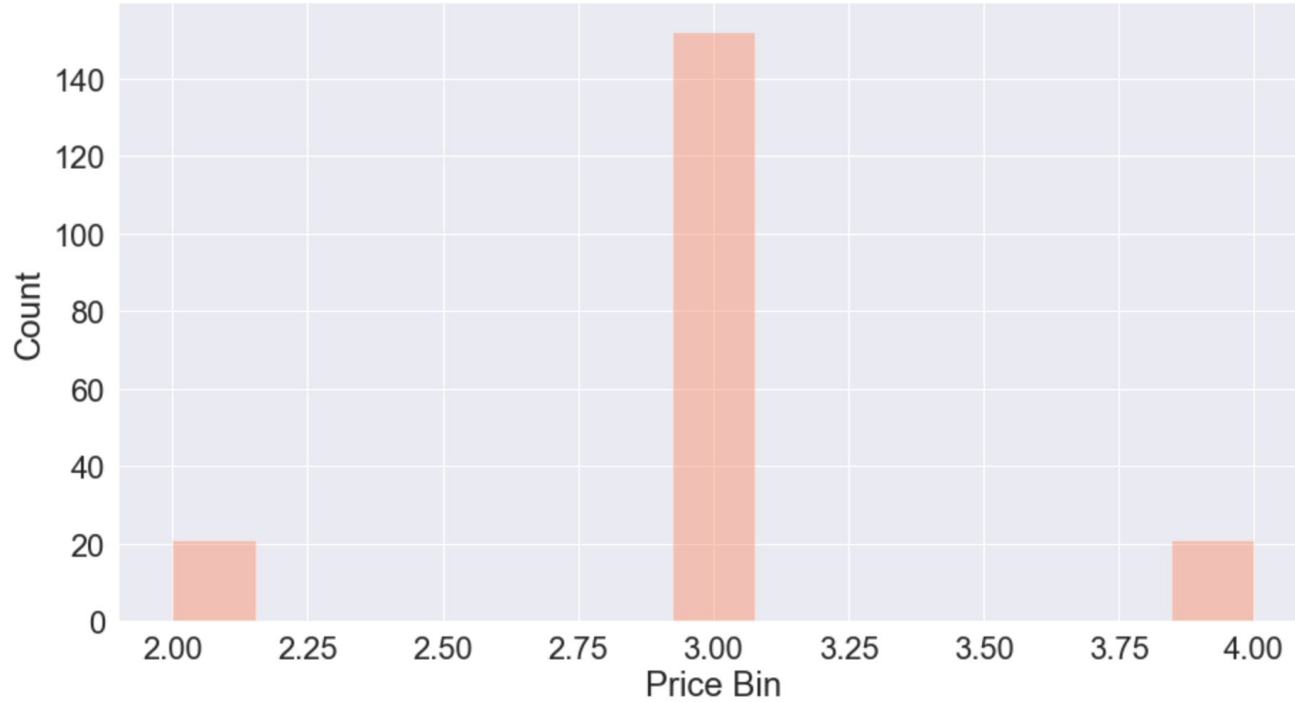
We used Yelp's API to gather restaurant data from the state of CT

The background of the slide is a collage of various images of building windows and facades. The images are arranged in a grid-like pattern, with some showing close-ups of window frames and others showing wider views of building exteriors. The colors are mostly warm, with a lot of orange and yellow tones, suggesting a sunset or sunrise setting. The central text is overlaid on a solid orange rectangular area.

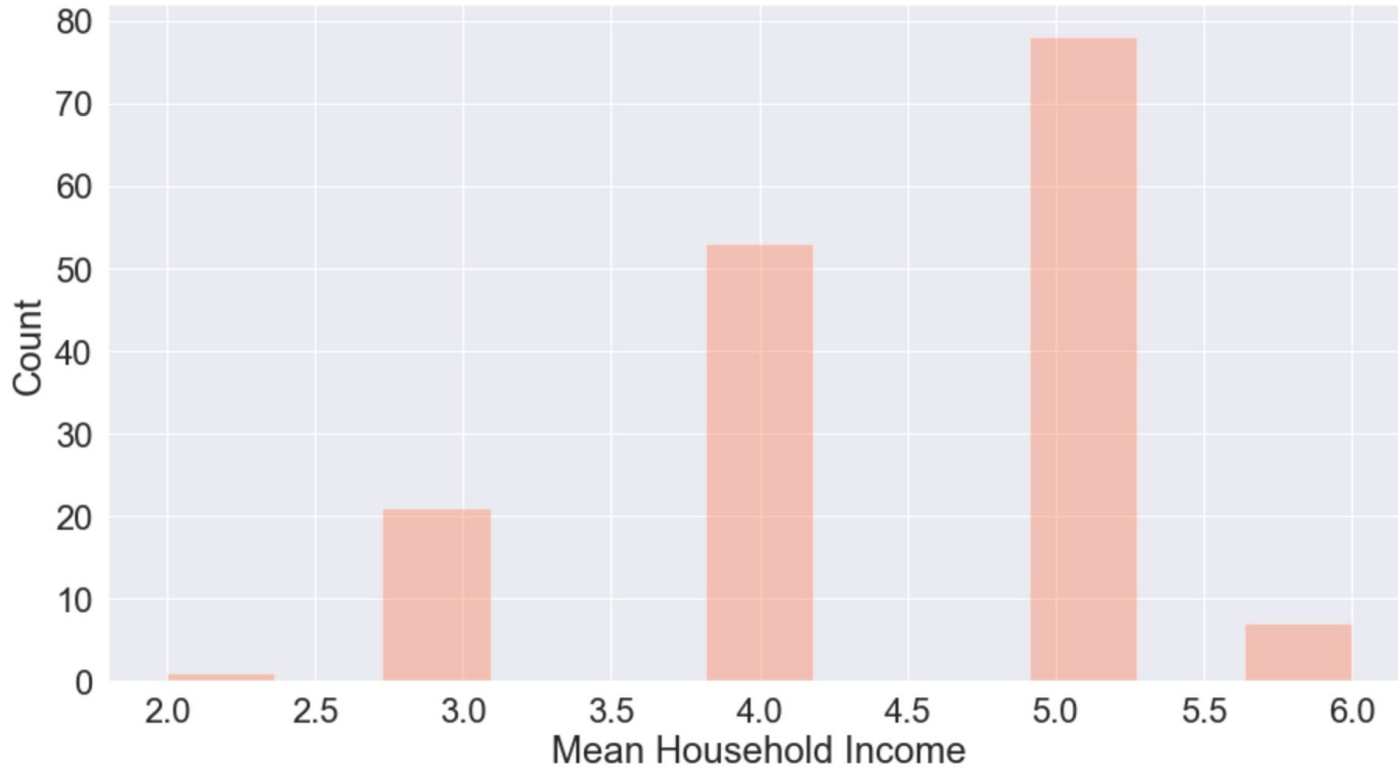
Feature Engineering

1. Average Restaurant Price
2. Price Categories
3. Average Income Bin

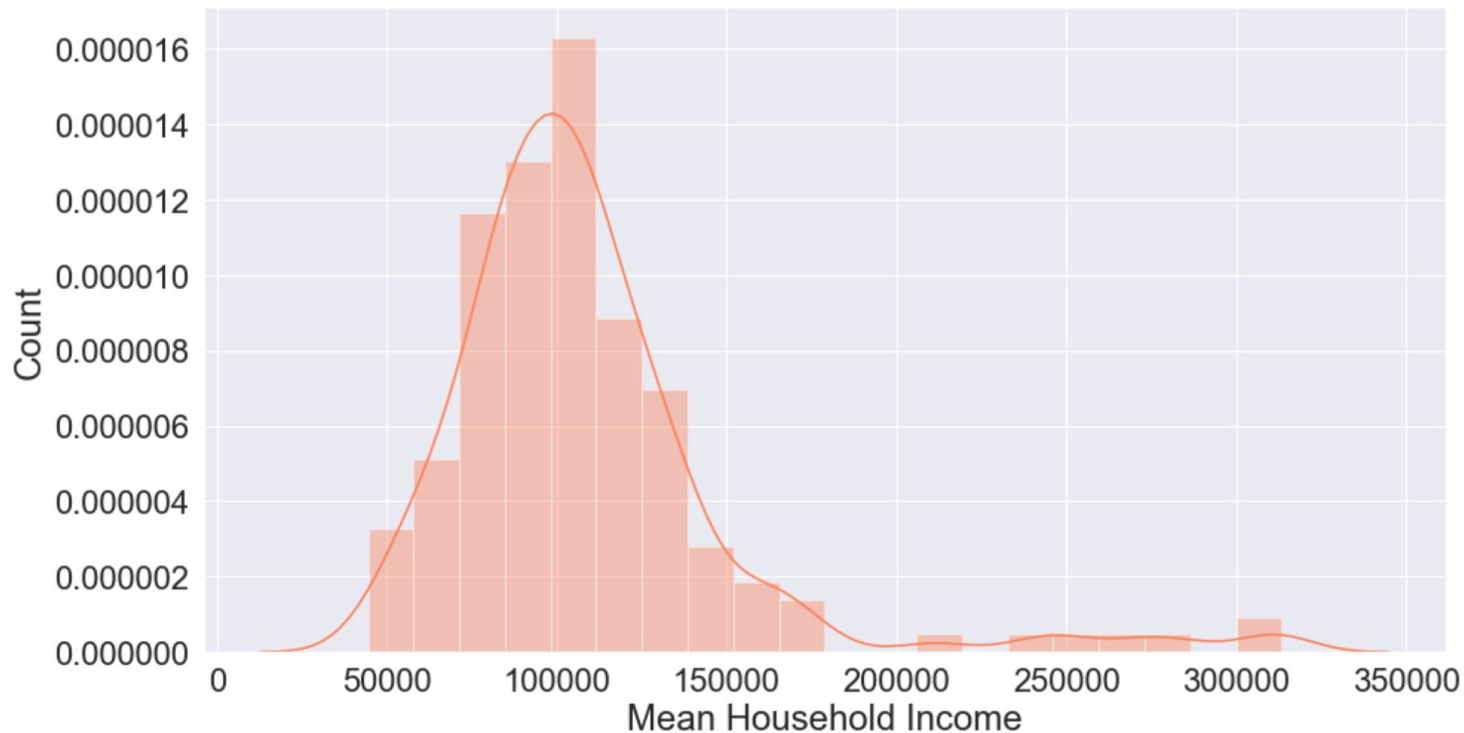
Distribution of IRS Income Bin



Distribution of CT Income Bin



Distribution of CT Mean Household Income



Modeling

Used both classification
and regression to model

Baseline score: 78%
Logistic Regression
accuracy score: 79.5%
(test)
Linear Regression, r^2 : .16

Limitations

Data and imbalanced
classes

CONCLUSIONS AND FUTURE WORK



Conclusions

- Using Yelp pricing scales might be a better indication of how residents of a certain city rate the cost of product and services.
- Yelp is not an accurate predictor of affluence
- We can't confidently advise Wildflower on where they should open their next restaurant



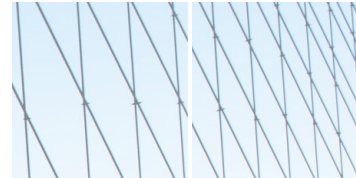
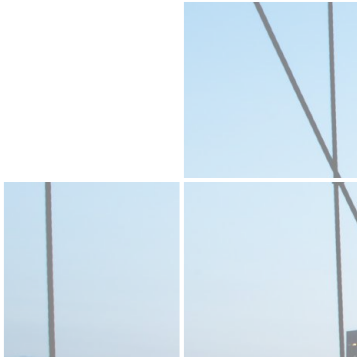
Future Work

- Gather more income data
 - Census API
- Expand search radius
 - Yelp



THANKS

Does anyone have any
questions?



CREDITS

- ◀ Presentation template by [Slidesgo](#)
- ◀ Icons by [Flaticon](#)
- ◀ Infographics by [Freepik](#)
- ◀ Images created by [Freepik](#)
- ◀ Author introduction slide photo created by Freepik
- ◀ Text & Image slide photo created by Freepik.com

