Neighborhoods of Oahu

Trends and Insights for Incoming Residents

Brock Schmalzel,
August 2020

Where to live on Oahu?

- Thousands of people move to the island of Oahu each year from the U.S. mainland and from abroad
- When deciding where to live, there are other factors beyond cost of rent and household amenities. Some incoming residents may intend to open a business while others come with their families and care about nearby schools and crime rates.
- Is it possible to provide insights about the different neighborhoods of Oahu to incoming residents?

Data

- **Oahu Zip Codes:** 56 zip codes on the island of Oahu
- **Foursquare City Guide:** This mobile app provides information on places and venues in a given area
- **Honolulu Police Department's 2018 Annual Report**: This report provides crime data for each of Oahu's 8 crime districts
- **2018 Strive HI Performance Results:** Hawaii State Department of Education data on public schools

Processed Data

- Foursquare was used to query venues in a 1000 meter radius around the latitude/longitude coordinates for each zip code. Top 10 most common venues in each zip code were calculated
- Normalized public high school test data (sum proficiency = total math, language, and science scores)
- Normalized crime data uses total crimes for each crime district

Zip Code	Sum Proficiency	School(s)	Total Crimes	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
96701	0.068594	Aiea High	0.2453	Mobile Phone Shop	Burger Joint	Golf Driving Range	Bank	Jewelry Store
96706	0.351234	Campbell High	0.4075	Home Service	Park	Clothing Store	Golf Course	Restaurant
96707	-1.124775	Kapolei High	0.4075	Video Store	Coffee Shop	Supermarket	Business Service	Doctor's Office
96731	-0.528091	Kahuku H&I	-0.9182	Airport	Trail	Yoga Studio	Dive Bar	Flower Shop
96734	0.304127	Kailua High, Kalaheo High	-0.9182	Beach	Construction & Landscaping	Food Truck	Tennis Court	Yoga Studio

Methodology

K-Means Clustering

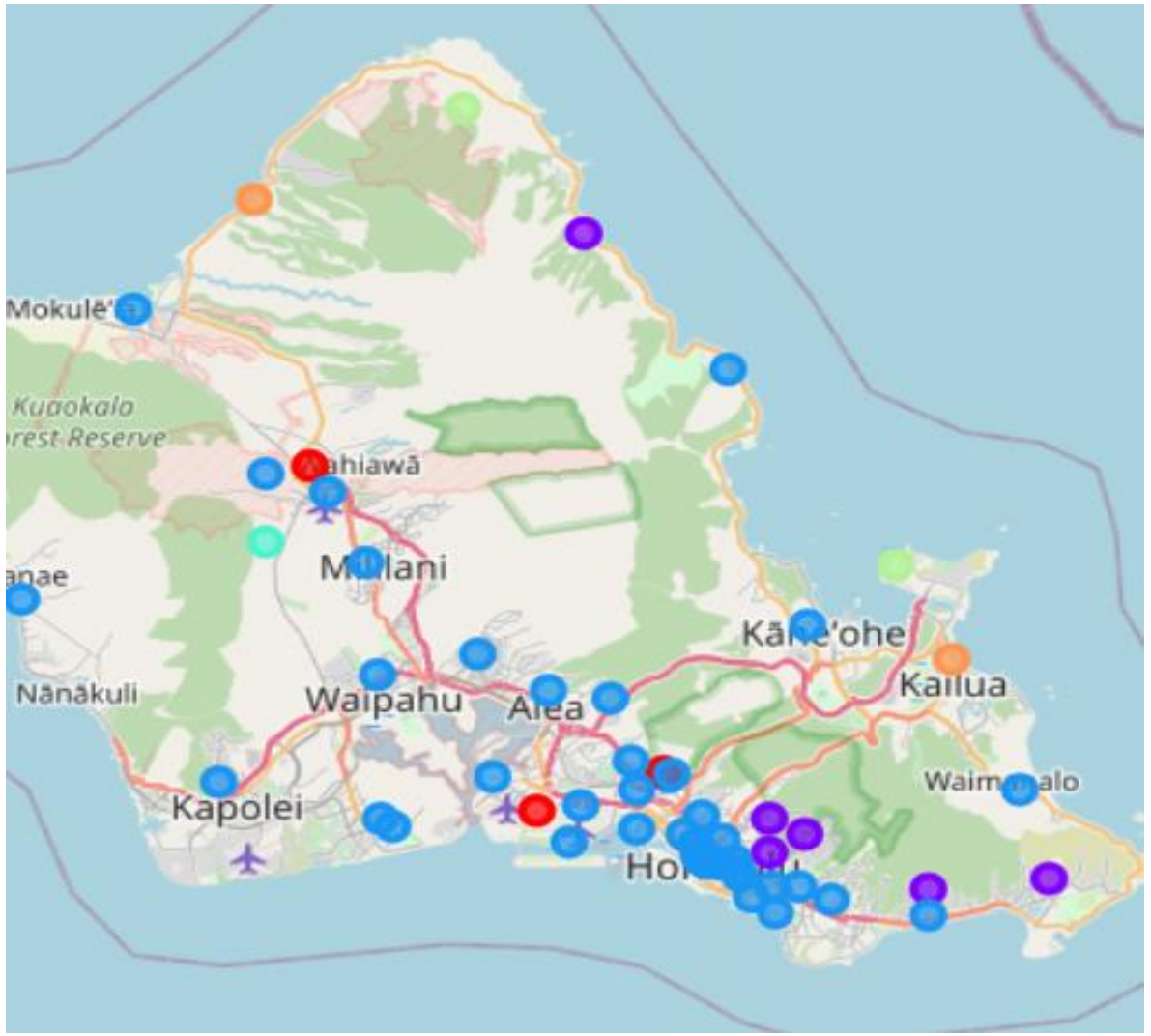
- K-means clustering was applied to the venue data to identify trends among the various zip codes.
- Once clustered, the most common venues and—more importantly—the venues not found in other clusters would be identified.
- Three versions of this algorithm was completed using:
  - Venue data for all Oahu Zip Codes
  - Venue, school, and crime data for Oahu Zip Codes with at least one public high school
  - Crime per population and wealth index data for Oahu Zip codes with at least one public high school

Regression

- Pearson's Correlation Coefficient calculated to analyze correlation between crime and school performance.
- Ordinary least squares method was used to perform multiple linear regression to see how well crimes-per-population and wealth index could predict school performance

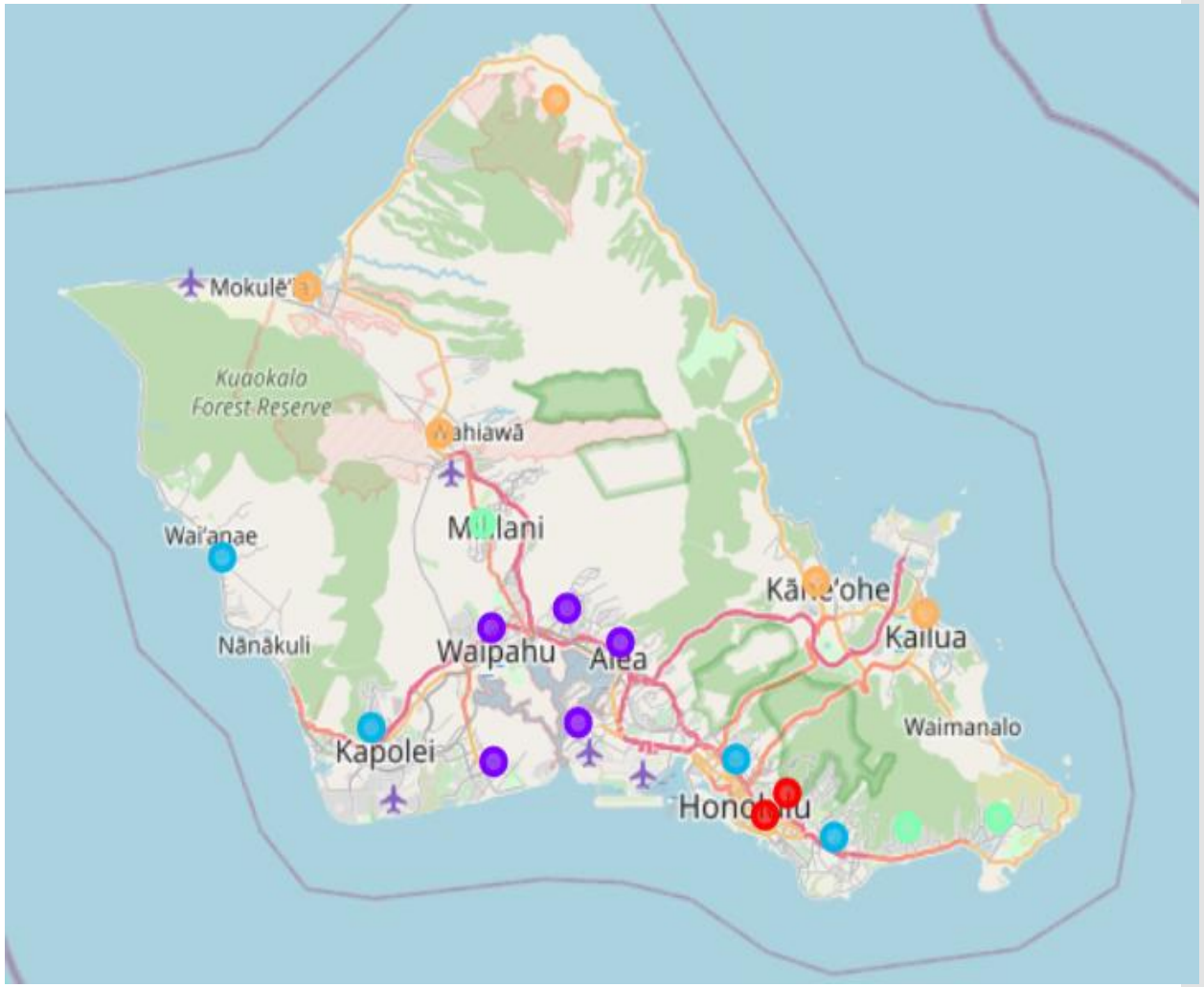
Results
Venue Only

Cluster	Number of Zip Codes	Unique Venues
Red	3	Pizza Place, Golf Course, Bowling Alley
Purple	6	Trail, Scenic Lookout, “Food”
Cyan	1	Post Office, Doctor’s Office
Orange	2	Beach
Green	2	Airport, Yoga Studio, Donut Shop
Blue	42	Asian Cuisine, Coffee shops, Parks



Results
Venue, School, and Crime

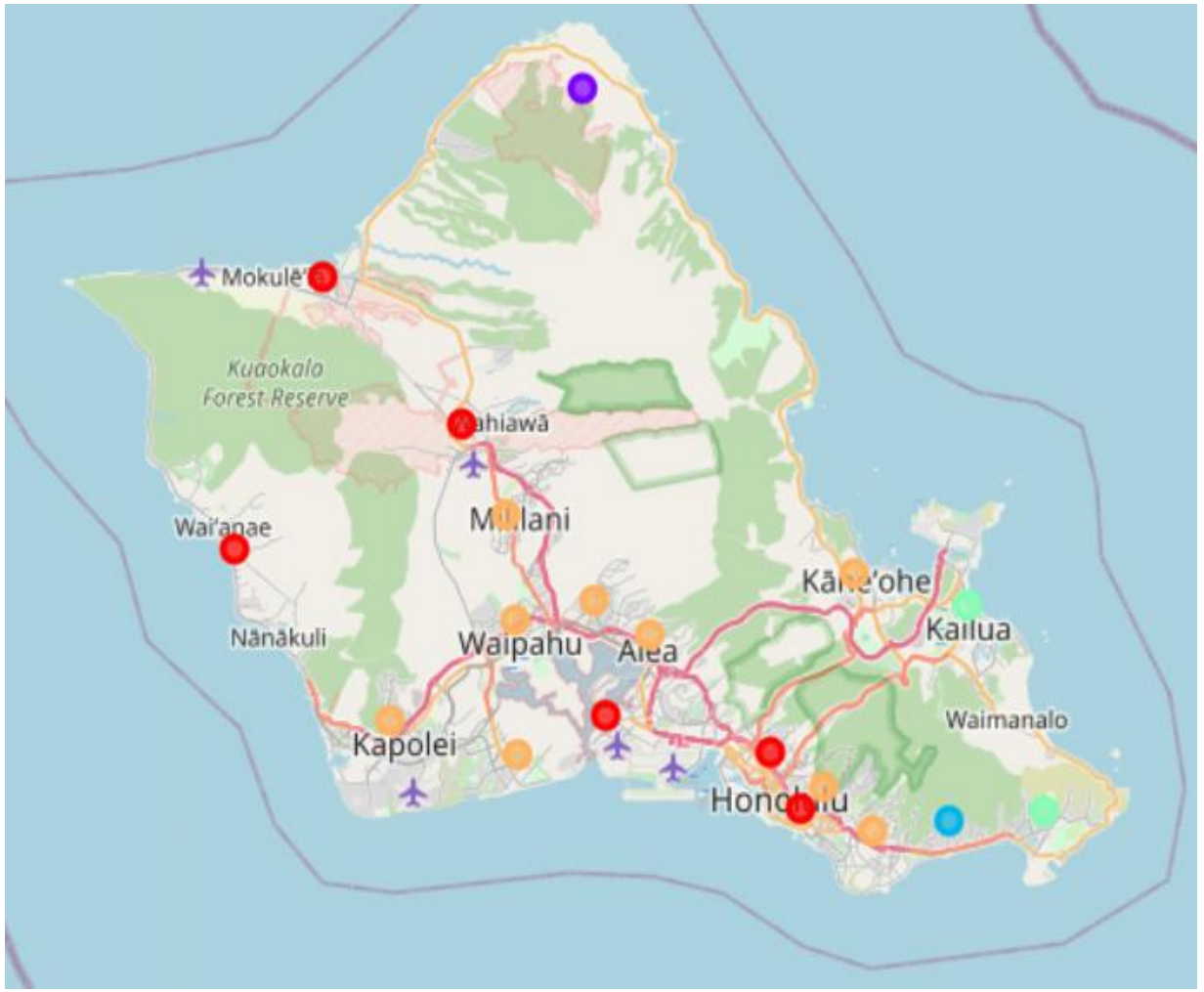
Cluster	Number of Zip Codes	Common Venues	School Scores	Crime
Red	2	Cafe	-	Very Higher
Purple	5	Gold Course, Park	Higher	Higher
Cyan	3	Park, “Food’, Trail	Higher	Lower
Orange	5	Farmer’s Market, Café, Yoga Studio, Electronics	Lower	Lower
Blue	4	Coffee Shop, Pizza Shop, Diner	Very Lower	Higher



Results

Crimes and Wealth Index

Cluster	Number of Zip Codes	Wealth Index	Crimes Per Population
Red	6	Lower	Mixed
Blue	1	Very Higher	Higher
Purple	1	Lower	Much Higher
Green	2	Higher	Lower
Orange	9	Mixed	Lower



Results
Regression

- No significant correlation between total number of crimes in a zip code's district and its high school(s)' average test scores ($r(17) = -0.0187$, $p=0.939$)
- Multiple linear regression model (using ordinary least squares) with the average test scores as a function of a zip code's wealth index and crimes per population yielded a model score of only 0.244
- Significant positive correlation between a zip code's wealth index and its high school(s)' average test scores ($r(17) = 0.494$, $p=0.032$)

Observations

- **Dominant Venue Cluster.** Large, dominant cluster consistent with greater Honolulu downtown area.
- **Unequal Zip Codes.** Varying population density across Oahu caused Foursquare query issues:
  - Rural zip codes return little-to-no venues
  - Urban zip codes overlap with each other
- **Business Examples**. Useful insights for identifying which neighborhoods favor different styles of cuisine
- **Schools, Crime, and Wealth.** Hypothesized that these societal metrics are related to each other. If someone is interested in one, they should consider the others to get the full picture.
- **Generic Venues.** Foursquare returned generic venues such as "Food" and "Restaurants". This likely worsened the performance of the clustering algorithm

Recommendations

- Adjust the Foursquare query radius according to the population density of target area
- Drop or classify Foursquare data that is generic e.g. “restaurant”
- Consolidate Foursquare data into larger categories. There may be value gained by creating categories such as Asian cuisine (Japanese, Chinese, Korean, etc), outdoor activities (trail, beach, park, etc), or white-collar (event planning, doctor’s office, etc
- Develop a consistent technique for identifying venue trends within a cluster
- Two recommendations related to improving the data behind the school-crime-wealth analysis
    - More comprehensive assessment of schools beyond test scores such as attendance, graduation, and college acceptance rates
    - Break down the amount of crime by zip code rather than lumping them into a single crime district

Conclusion

- When a rental unit costs the same in Pearl City as another unit in Hawaii-Kai, what other factors can be considered?
- The K-means cluster algorithms used in this study found that distinct trends could be found between neighborhoods on Oahu
  - Pearl City's Cluster: Diverse restaurant options and parks. Lower than average crime.
  - Hawaii-Kai's Cluster: Trails and scenic lookouts. Lower than average crime and higher than average wealth index.
- Differences between clusters were not always clear
- Much work remains to improve the clarity and utility of the approach laid out in this report, but initial results are promising