

Oakland Wellness Center Analysis

Data Analytics to identify possible geographical areas for marketing and placement of
a new wellness center.

Customer Profile

The client has specified the following characteristics

- Working Age
- Income > \$60,000
- 65% Female
- Household size 2+
- 20% homeowners

Finding and Extracting the Data

- For this analysis, location and demographic data was needed
- Location data was pulled from the Foursquare API that allowed for the analysis of the business environments of the various zip codes in Oakland
- The demographic data was pulled from zip-codes.com which was scraped with the BeautifulSoup library in Python

Cleaning and Organizing the Data

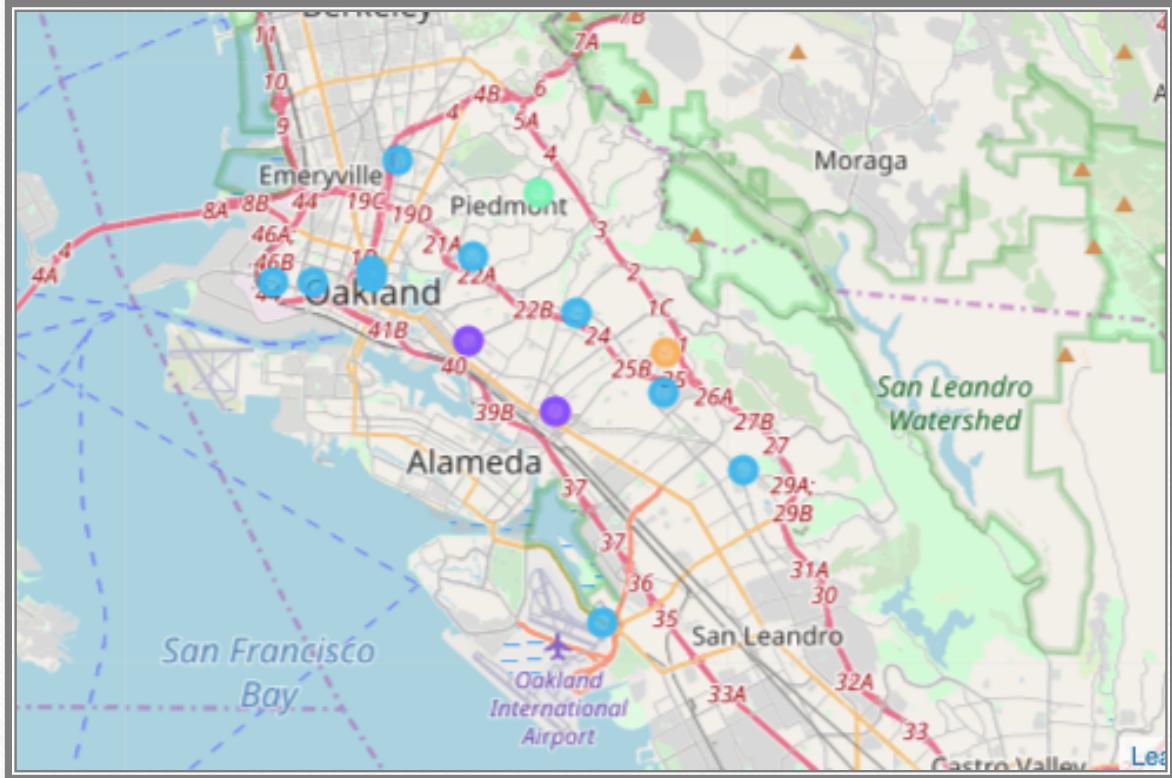
- The data was extracted and loaded into a Pandas data frame for analysis
- 449 venues were located in Oakland , with 28 columns of data points on each venue
- 21 zip codes were located with 21 columns of data points on each zip code

Working With the Data

- To begin clustering the data, the venues were organized by venue category
- A new data frame was created that ranked the 10 most common venue categories in each zip code for clustering analysis
- Five clusters yielded the richest results

Clusters by Venue Type

- The blue, green, and orange clusters showed the most desirable business environments for our proposed business
- These environments contained venues with food, recreation and transit



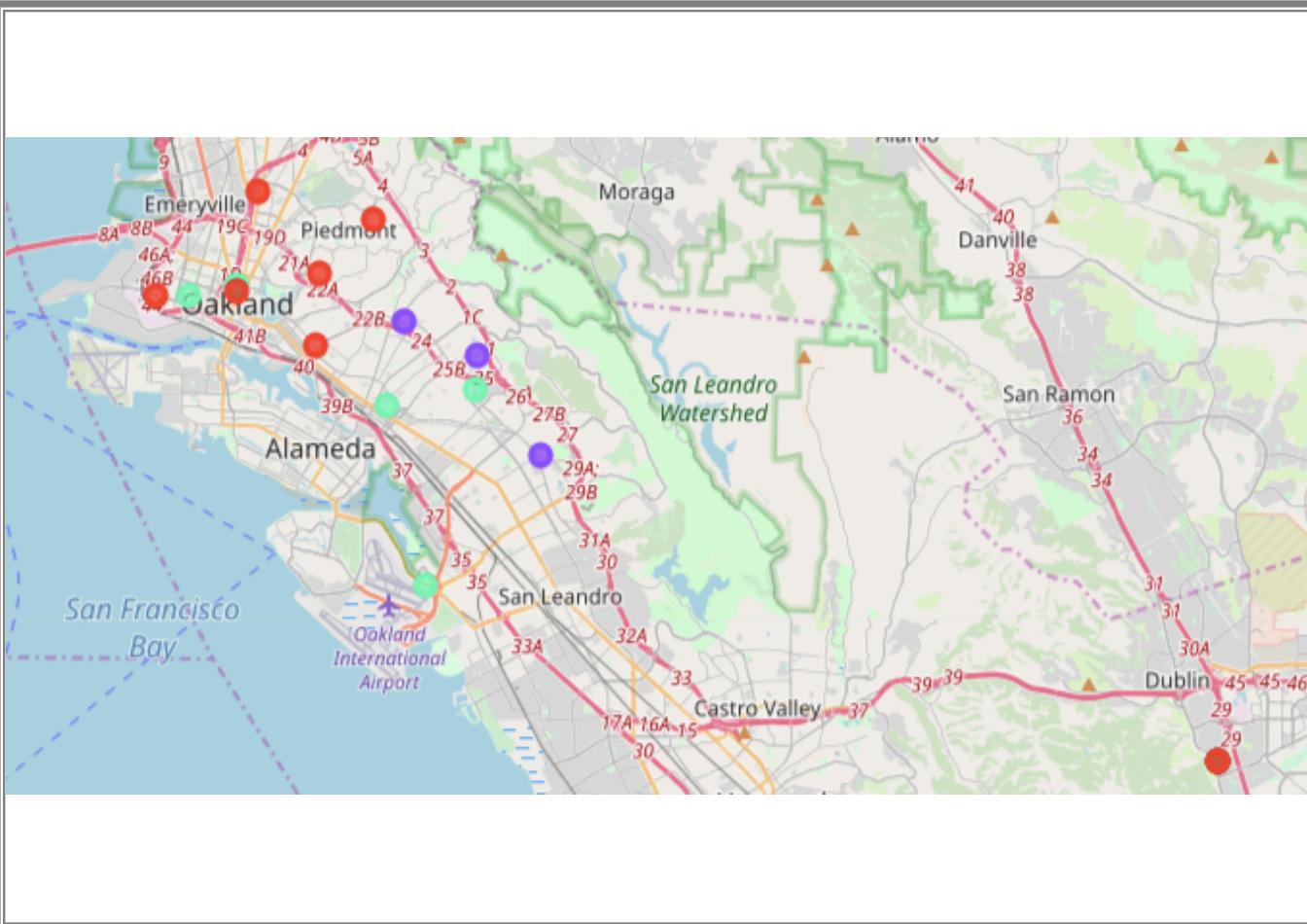
Missing Data

- In the initial clustering, 1 cluster of 7 zip codes was missing
- When further researched, these missing zip codes turned out to be P.O. Boxes
- These were removed from further analysis as they are not suitable for our location project

| Zip | Latitude | Longitude | Cluster Labels | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|-----|-----------|-------------|----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 6 | 37.680181 | -121.921498 | 0 | Trail | Dog Run | Yoga Studio | Dive Bar | Filipino Restaurant | Fast Food Restaurant | Falafel Restaurant | Ethiopian Restaurant | Donut Shop | Deli / Bodega |
| 15 | 37.680181 | -121.921498 | 0 | Trail | Dog Run | Yoga Studio | Dive Bar | Filipino Restaurant | Fast Food Restaurant | Falafel Restaurant | Ethiopian Restaurant | Donut Shop | Deli / Bodega |
| 16 | 37.680181 | -121.921498 | 0 | Trail | Dog Run | Yoga Studio | Dive Bar | Filipino Restaurant | Fast Food Restaurant | Falafel Restaurant | Ethiopian Restaurant | Donut Shop | Deli / Bodega |
| 18 | 37.680181 | -121.921498 | 0 | Trail | Dog Run | Yoga Studio | Dive Bar | Filipino Restaurant | Fast Food Restaurant | Falafel Restaurant | Ethiopian Restaurant | Donut Shop | Deli / Bodega |
| 19 | 37.680181 | -121.921498 | 0 | Trail | Dog Run | Yoga Studio | Dive Bar | Filipino Restaurant | Fast Food Restaurant | Falafel Restaurant | Ethiopian Restaurant | Donut Shop | Deli / Bodega |
| 21 | 37.680181 | -121.921498 | 0 | Trail | Dog Run | Yoga Studio | Dive Bar | Filipino Restaurant | Fast Food Restaurant | Falafel Restaurant | Ethiopian Restaurant | Donut Shop | Deli / Bodega |
| 23 | 37.680181 | -121.921498 | 0 | Trail | Dog Run | Yoga Studio | Dive Bar | Filipino Restaurant | Fast Food Restaurant | Falafel Restaurant | Ethiopian Restaurant | Donut Shop | Deli / Bodega |

Clustering the Demographic Data

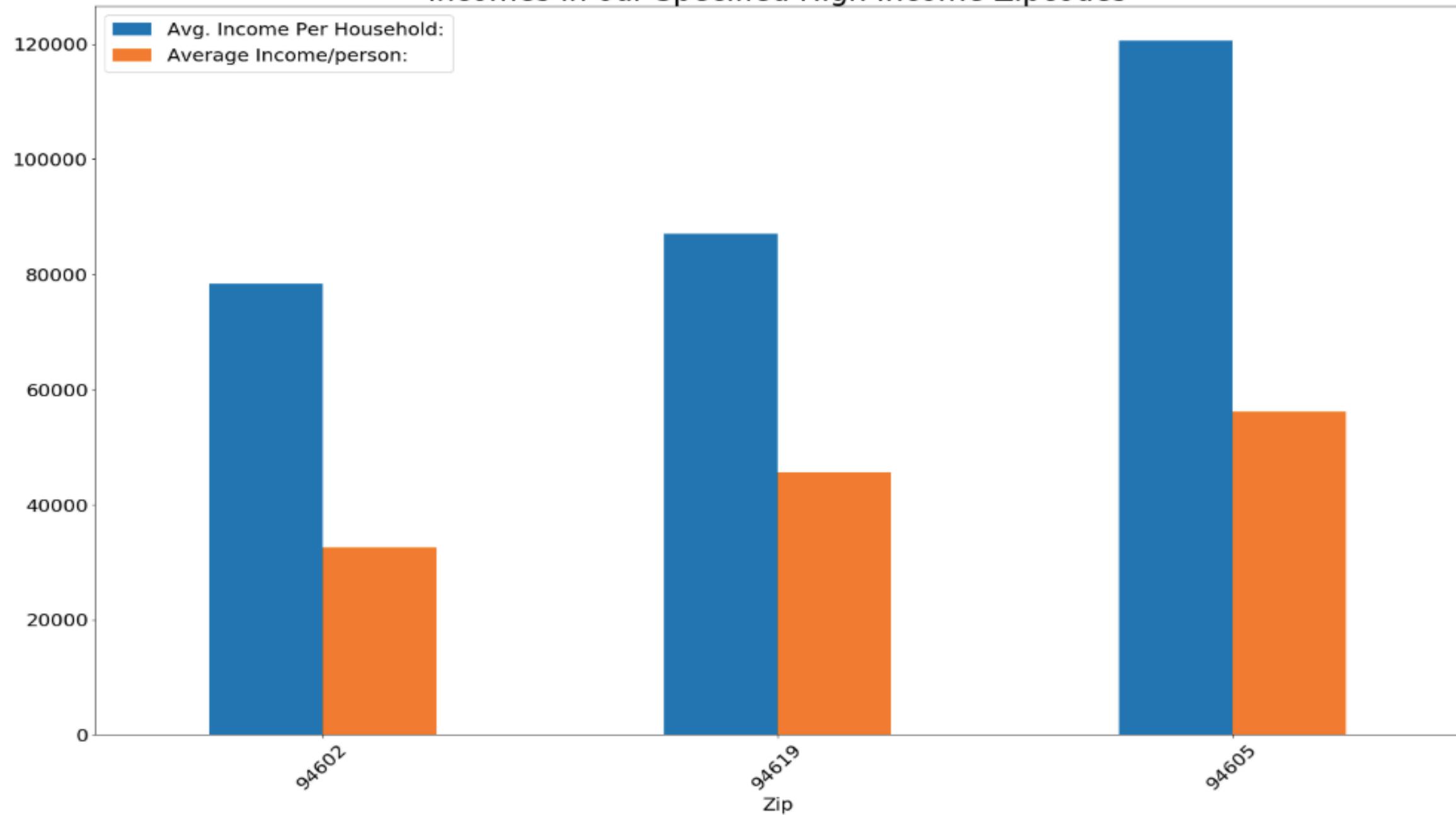
The only areas that fit into our income profile are the purple nodes



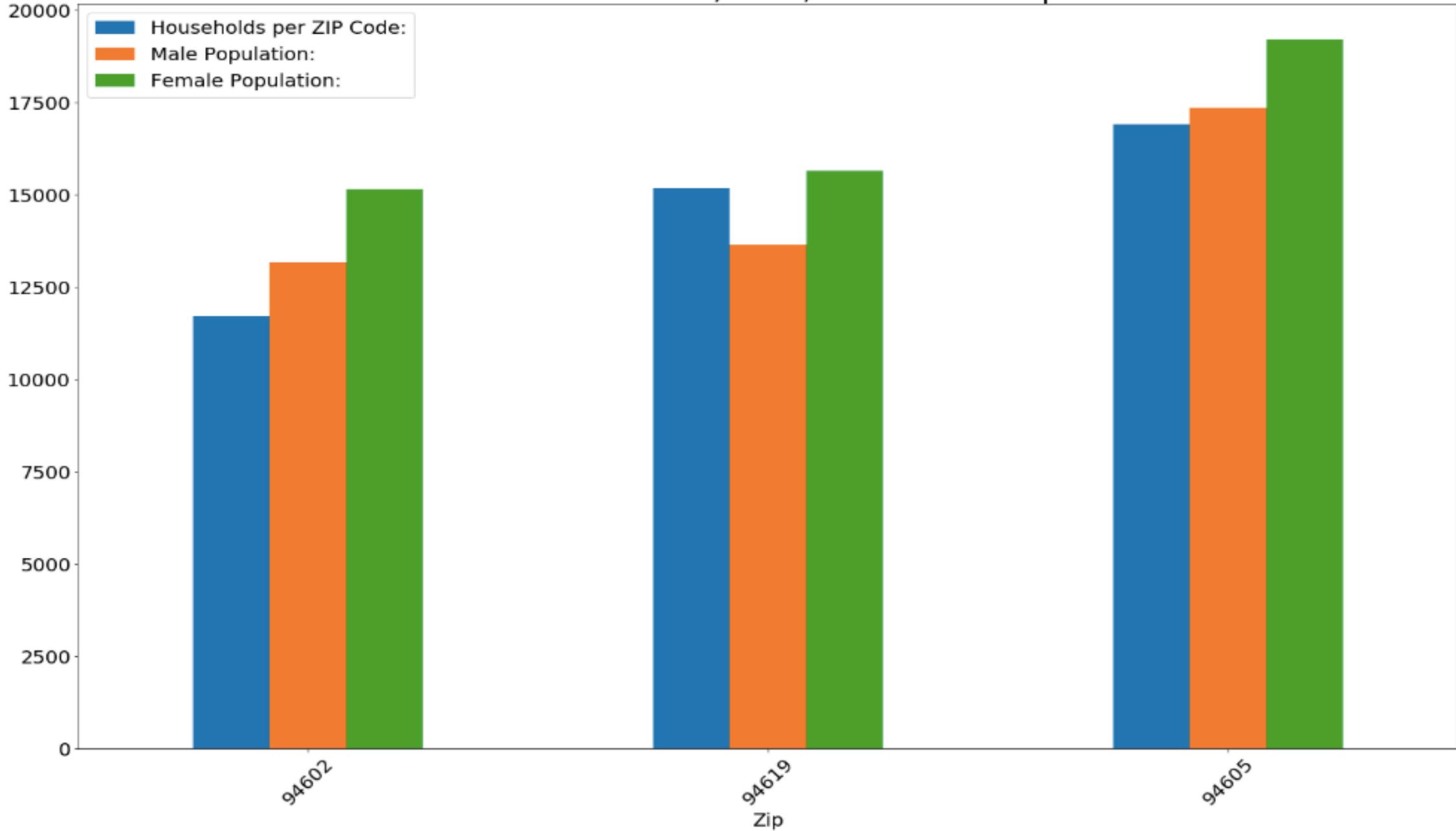
Demographics of Most Suitable Zip Codes (The Purple Nodes)

| Zip | Current Population: | 2010 Population: | Households per ZIP Code: | Average House Value: | Avg. Income Per Household: | Persons Per Household: | Male Population: | Female Population: | Median Age: | Male Median Age: | Female Median Age: |
|----------|---------------------|------------------|--------------------------|----------------------|----------------------------|------------------------|------------------|--------------------|-------------|------------------|--------------------|
| 1 94602 | 29274.0 | 28,329 | 11702.0 | 673100.0 | 78347.0 | 2.41 | 13180.0 | 15149.0 | 41.30 | 40.20 | 42.40 |
| 4 94619 | 30980.0 | 29,287 | 15190.0 | 839100.0 | 87059.0 | 1.91 | 13633.0 | 15654.0 | 39.50 | 39.50 | 39.60 |
| 11 94605 | 39220.0 | 36,565 | 16923.0 | 941900.0 | 120601.0 | 2.15 | 17345.0 | 19220.0 | 43.90 | 43.50 | 44.20 |

Incomes in our Specified High Income Zipcodes



Number of Households, Male, and Female Population



Conclusion

- Out of 21 zip codes analyzed, three were found to fit well with the customer profile provided by the client
- These zip codes are 94602, 94619, and 94605.
- These areas make up a continuous stretch in the south eastern side of the city, making for a large area and population to cover
- These areas could be seen as good for location placement, as well as for marketing