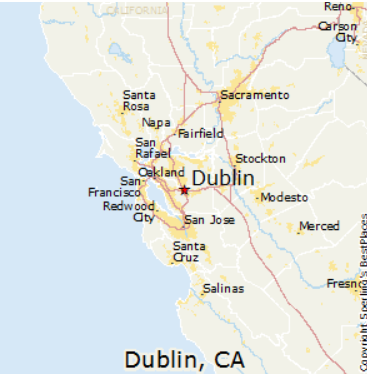
Topic🡪

**Relocation to Dublin, Tri-valley to ramp up your business**

**help find the location for the companies who is moving office to Dublin.**



**Introduction**

**Project background**

Dublin, CA is the suburban city of San Francisco East bay. It is part of Tri-valley region of Alameda County, California. As of 2019, city of Dublin consists of 10 constituent neighborhoods, and is the home of a population of 60,939 people, including me and my family.

We have been living in Dublin, CA for almost 5 years and are consistently seeing significant upside trending of domestic population. More and more highly educated people are relocating to Dublin area from various cities across the bay. Real estate developers and investors have been and are still diving into all neighborhoods, especially east part of Dublin, to construct business and residential properties.

Dublin, together with its sister city Pleasanton, which is adjacent to Dublin right across the highway 580, are currently accommodating or welcoming to host the business of some great corporations, including WeWork, SAP, Oracle, Ross stores, Safeway, etc.

According to the 2018 released report by the Bay Area Council Economic Institute, Tri-valley area, where Dublin & Pleasanton are located, are now home to over 450 tech companies, economically boosted by more than $4 billion investment over last decade.

Regionally, 12 percent job growth for last 4 years is only the start, "We want the next Google and the next Facebook and the next LinkedIn to be created and housed here in the Tri-valley," said Dublin Mayor David Haubert.

There are 3 main reasons that will endorse Mayor’s statement:

* Low rent (Dublin $32/sqft, compared with silicon valley $52/sqft, & SF $72/sqft)
* Increased density of population ( population has been boosted for last several decades as below
* No shortage of talents(61.48% of its adults are holding 4-year or even advanced degree)

At the same time, Dublin residents are spending longer than average time for commuting, specifically, people here spend average 40 minutes each day getting to work, which is significantly higher than the national average.

**Business Problem**

Keeping the background and stakeholders’ interests in mind, I am making reasonable assumption that business relocation will drive out a win-win scenario, both parties, local residents and employers, will have high motivation to facilitate the business/working sites shifting to Dublin.

Specifically, employers will benefit from low rent, and talent pool expansion by re-organizing businesses and hiring locally in Dublin. Dublin residents will gain sufficient exposure to more local job opportunities and take advantage of the short commute in the future.

**Target audience**

For my project, target audience would be employers’ executives who are considering to tag Dublin area as their future working sites. I would like to leverage data science and machine learning methodology to help those potential employers to locate their business address, saying that I would like to run an inter-neighborhood analysis within city of Dublin and help shed some light about the most suitable neighborhood from the business location point of view.

**Data**

Criterion by neighborhood

* Crime rate
* Residential (population)
* Talent (education level)
* Rent/housing price
* Lng & log

Went through the venue categories, and filter & combine by leveraging several most import criteria of working sites

Convenience, productivity & entertainment

Convenience

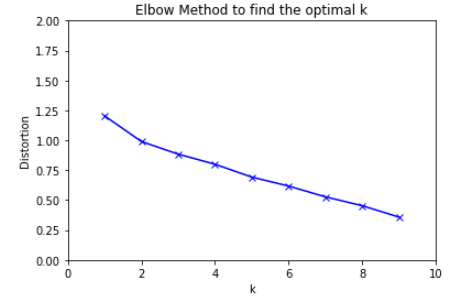
* Food 🡪 café & coffee shop, bakery, fast food, diverse style restaurants, i.e. Asian, American, Italian, Mexican, Afghan food
* Shopping 🡪 shopping mall, grocery stores
* Lifestyle 🡪 health & beauty service, pharmacy, bookstore

Productivity

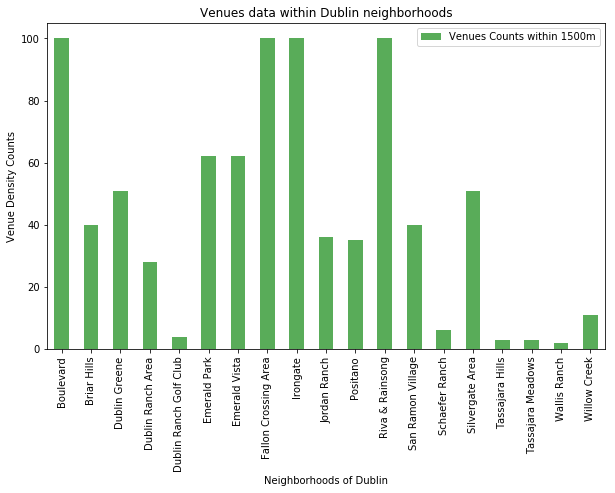
* Commute 🡪 bus station, gas station, ATM.
* Workout 🡪 gym/fitness center, parks, trails
* Entertainment 🡪 bar, movie & music, recreation center

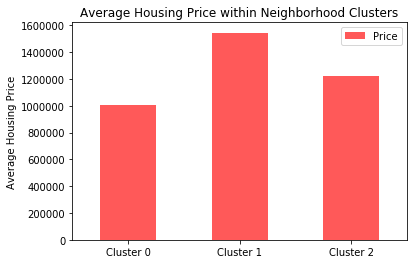
Base on the classification analysis, we picked 17 representative features out of 163 total venue categories(12%) to effectively eliminate the homogeneous factors and comply with our pre-designed criteria: convenience and productivity. 17 features are listed below.

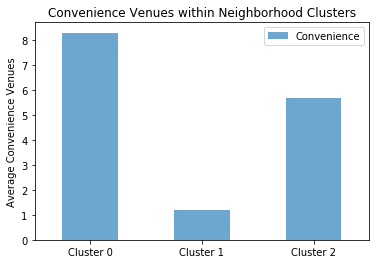
* ATM
* Bakery
* Bar
* Bookstore
* Café
* Coffee Shop
* Fast Food Restaurant
* Gas Station
* Grocery Store
* Gym / Fitness Center
* Health & Beauty Service
* Movie Theater
* Park
* Pharmacy
* Shopping Mall
* Trail

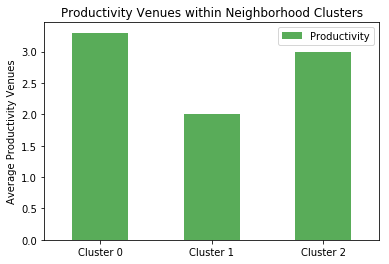
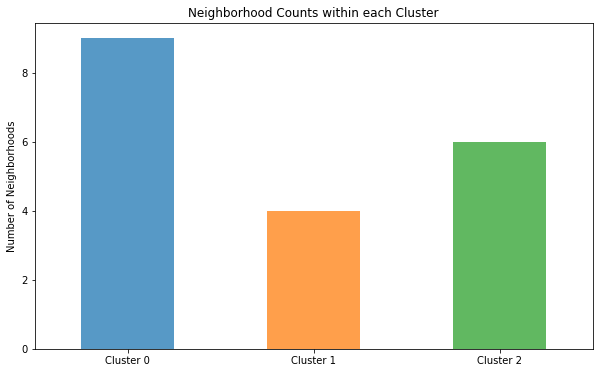


Distoration is starting from around 1.25 when K=1, we know that based on the selected features, the differences between neighborhoods are not as significant as we thought. Meanwhile, when K increases, the centroids are closer to the clusters centroid, the improvement will decline adn the slight elbow shape is created around 3 or 5. We will choose K=3 since when K=3, distortion rate firstly dropped under 1.00









Cluster 0: Neighborhoods

