A Decision Framework for Locating a Professional Services Business, Selecting Between San Francisco, CA and Seattle, WA

IBM Data Science Professional Certification Capstone Project: "Battle of the Neighborhoods"

William Windsor

June 22, 2019

Reference:
IBM Data Science Professional Certification curriculum, hosted on Coursera.com and CognitiveClass.ai

Report Content

- Introduction:
  - Business Location Decision: Project Description
  - Project Focus and Target Audience
- Data Requirements and APIs Referenced
- Decision Methodology
  - Corporate Financial Impact Factors - Analysis
  - Employee Quality of Life Factors - Analysis
- Results
- Discussion
- Conclusion

Introduction: Project Description

- **Part 1(a): Project Description and Business Decision**

  This data science project proposes a decision structure to address the decision:

  Where should I locate a professional services business, selecting between two neighborhoods that each have a concentration of high-tech startup companies?

- **Example possibilities for the professional services business are:**
  - Providing data science and data research services to tech companies (such as motivated by this IBM Data Science Professional Certification program)
  - Facilities planning and facilities location research services
  - Professional human resources contract services
  - On-site health care and monitoring services
  - Security services such as IT software hosting and security, or personnel for on-site security monitoring

Project Focus and Target Audience

- **This project focuses on two neighborhoods noted for their burgeoning growth of technology startups:**
  - San Francisco, CA: South of Market Street region (SoMa)
  - Seattle, WA: Pioneer Square region
- **Target Audience and Their Interest in This Problem:**
  - CEOs, entrepreneurs, and venture capital companies looking to evaluate where to establish a new professional services business.
  - These audiences would care about this data analysis because the study provides a decision framework for where to locate a business, with data sources, with the thesis that professional services businesses generate greater success when they are located near their clients.
- Note: this presentation is best viewed with Page Down / Page Up keys, in order to compare and contrast data visualizations for the two neighborhoods.

Data Requirements: APIs and Datasets Utilized in This Analysis

- Foursquare Venues Database API
- For Commercial Real Estate Availability and Pricing, I accessed three commercial R.E. listing websites, and from these constructed a database:
  - CityFeet.com
  - 42Floors.com
  - Loopnet.com
- State and City Corporate Income Tax References, for California/San Francisco and Washington/Seattle
- For public rail transportation access in SF and SEA: I referenced APIs from their public transportation websites, then constructed a database for the distance from each commercial real estate available property to the nearest public rail transportation station in each neighborhood:
  - San Francisco BART Rail Transit
  - Seattle Link Rail Transit

Business Decision Methodology

- **Part 1(b): Decision Methodology, Description of the Data, and How It Will Be Used to Solve the Business Problem. This business location decision framework focuses on two major sets of factors: Financial Impact to the Company and Employee Quality of Life.**

- **Assess Financial Impact to the Company, focusing on these three factors:**
    1) Commercial Real Estate Rental/Lease Prices for office real estate in each region, to rent/lease a facility for company operations
    2) Business Density Calculation for each region, to calculate and prioritize the neighborhood with the higher business density
    3) Corporate Tax Rates for each state and city

- **Assess Employee Quality of Life, focusing on these three factors:**
    4) Proximity to Public Rail Transportation
    5) Number of restaurants in each neighborhood
    6) Number of exercise facilities in each neighborhood

Corporate Financial Impact Factors for the Business Location Decision

1) **Commercial Office Real Estate: property characteristics, locations, and rental/lease costs, based on current commercial office properties available at the time of this study.**

- **I constructed two commercial real estate listings datasets, one for each city, with these fields for analysis and comparisons:**

    a) Property Address
    b) Neighborhood and Zip Code
    c) Rental Price per Month
    d) Property Area in Square Feet
    e) Rental Price per Square Foot per Year

- **Sources for Commercial Real Estate properties and availability in San Francisco and Seattle:**

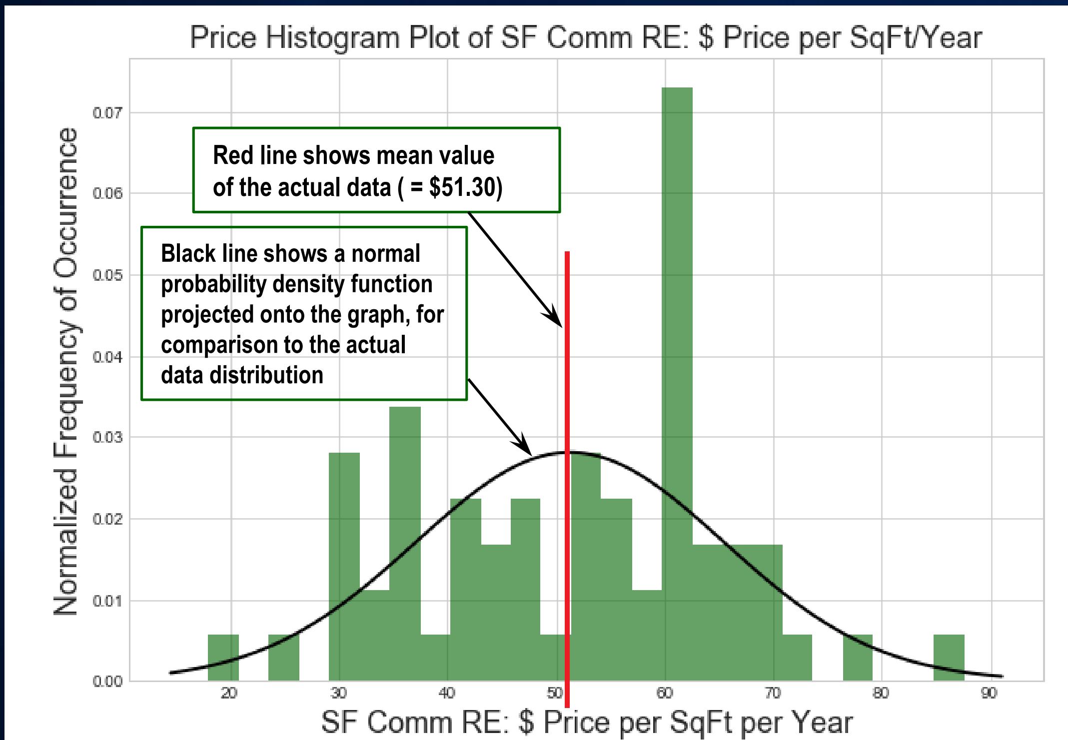
https://www.cityfeet.com, https://42floors.com, https://www.loopnet.com

San Francisco / SoMa: Commercial Real Estate Database (excerpt), and Code to Import and Structure It

	Address	City	State	Zip Code	Neighborhood	Property Sq Ft	List Rental Price/Month	Price per SqFt/Year	Description / Characterization	Rent Range (if provided)	Listing Website
0	1045 Bryant Street	San Francisco	CA	94103	SoMa	4,361	$21,085.00	$60.00	Office	NaN	cityfeet.com
1	1108 Bryant Street	San Francisco	CA	94103	SoMa	1,351	$3,945.00	$35.04	Office	NaN	cityfeet.com
2	1121 Howard Street	San Francisco	CA	94103	SoMa	2,800	$10,500.00	$45.00	Office	NaN	cityfeet.com
3	1169 Howard Street	San Francisco	CA	94103	SoMa	1,400	$6,244.00	$53.52	Office	NaN	cityfeet.com
4	123 10th Street	San Francisco	CA	94103	SoMa	6,000	$18,900.00	$37.80	Office	NaN	cityfeet.com
5	1252 Howard Street	San Francisco	CA	94103	SoMa	4,500	$20,250.00	$54.00	Office	54.00-56.00	cityfeet.com

```
In [15]: # GeoCoders-Nominatum: retrieve Lat-Long coordinates for the San Francisco / SoMa centroid address location.
         #  Note: I define the SoMa center (centroid) from an examination of the SF map as:
         #   508 4th St, San Francisco, CA 94107 (coincides with the location of the "Coin-Op Game Room").
         from geopy.geocoders import Nominatim
         geolocator = Nominatim(user_agent="SF_SEA_Comparison")
         location = geolocator.geocode("508 4th St, San Francisco, CA 94107")
         sfCommPropertyLat = location.latitude
         sfCommPropertyLong = location.longitude

In [16]: # Access the SF SoMa Commercial Real Estate .CSV file, and structure in a Pandas dataframe for computation
         SF_SoMa_CommRE = pd.read_csv(filepath_or_buffer='SF-SOMA-Comm-Real-Estate-Dataset-CSV.csv', sep=',')

         SF_SoMa_CommRE.columns = ['Address', 'City', 'State', 'Zip Code', 'Neighborhood', 'Property Sq Ft', 'List Rental Price/Month', 'Pı
         SF_SoMa_CommRE['Zip Code'] = SF_SoMa_CommRE['Zip Code'].astype(int)
         SF_SoMa_CommRE['Zip Code'] = SF_SoMa_CommRE['Zip Code'].astype(str)
         SF_SoMa_CommRE.head(8)

In [21]: # Plot the SF SoMa Commercial Real Estate rental prices in a histogram
         # Prices = values of 'Price per SqFt/Year' column, from the SF_SoMa Comm R.E. properties referenced in the .CSV file
         y_SFCommRE = SF_SoMa_CommRE['Price per SqFt/Year'].sort_values(ascending=True)
         y_SFCommRE = y_SFCommRE.replace('[\$,]', '', regex=True).astype(float)
         y_SFCommRE = y_SFCommRE.values
         x_UnitSeries = np.arange(1, (len(y_SFCommRE)+1), 1)
```

Commercial Real Estate Rental Prices: San Francisco / SoMa District (June 2019)

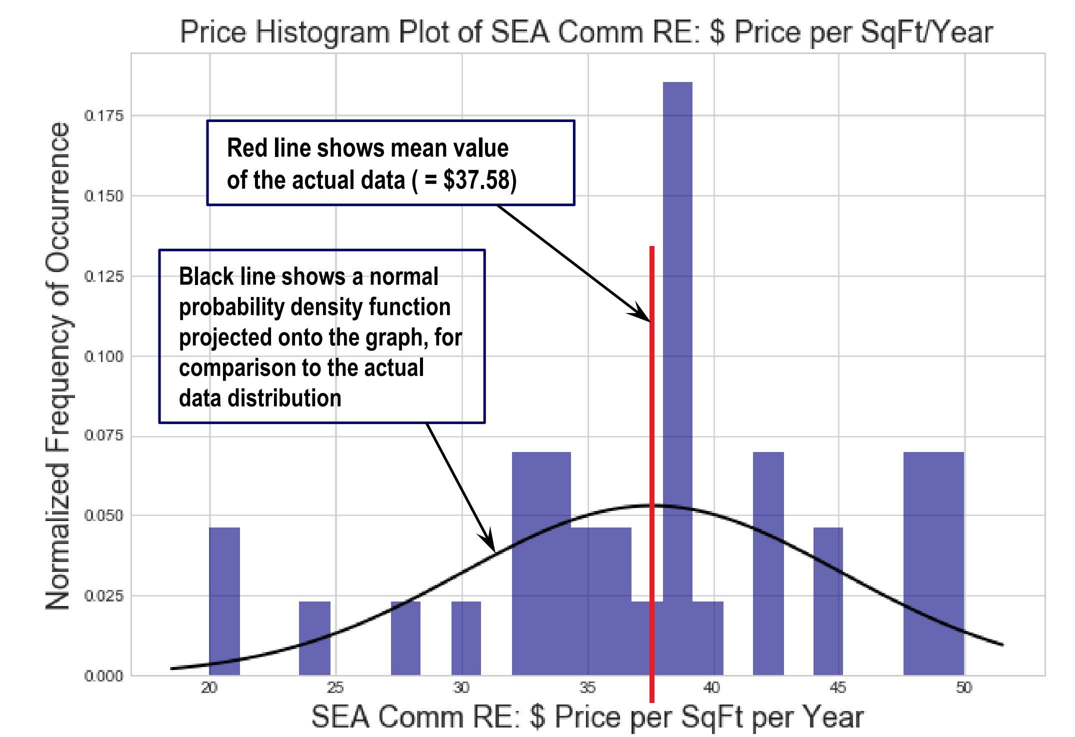



```
Mean    SF Commercial Real Estate Price: $ per SqFt per Year = $ 51.30
Median  SF Commercial Real Estate Price: $ per SqFt per Year = $ 53.82
Std Dev SF Commercial Real Estate Price: $ per SqFt per Year = $ 14.20
```

- San Francisco / SoMa Comm RE Price Statistics:
  - Mean = $51.30
  - Median = $53.82
  - StdDev = $14.20
- Both the SF price mean and standard deviation are higher than for Seattle (see next slide)

Note to dataset: these prices may be skewed on the low side, because these are *published prices* for office commercial real estate; these do not include the more *premium unpublished* real estate location prices.

Commercial Real Estate Rental Prices: Seattle / Pioneer Square Area (June 2019)

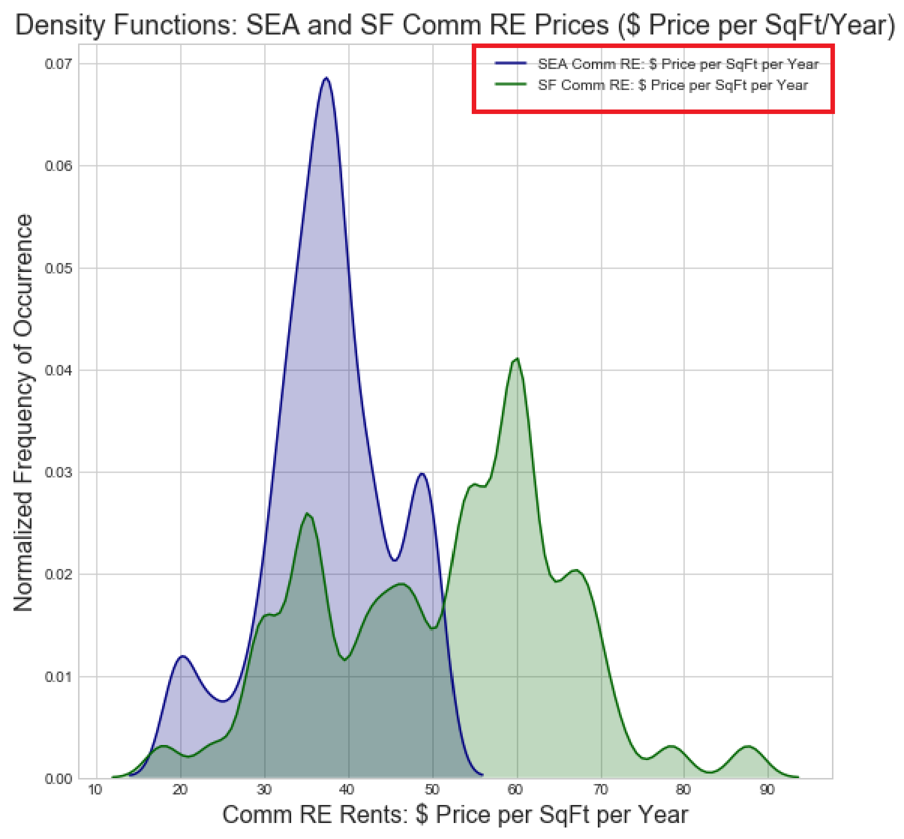


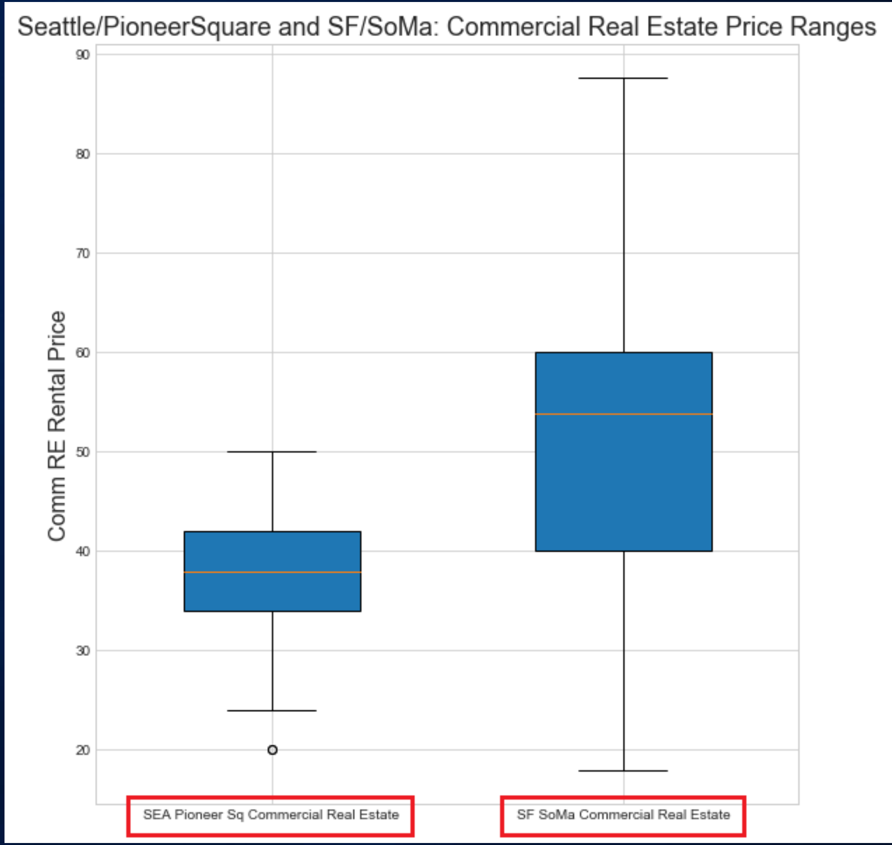

```
Mean    SEA Commercial Real Estate Price: $ per SqFt per Year = $ 37.58
Median  SEA Commercial Real Estate Price: $ per SqFt per Year = $ 38.00
Std Dev SEA Commercial Real Estate Price: $ per SqFt per Year = $  7.53
```

- Seattle / Pioneer Square Comm RE Price Statistics:
  - Mean = $37.58
  - Median = $38.00
  - StdDev = $ 7.53
- Both the SEA price mean and standard deviation are lower than for SF (previous slide)
- Conclusion: Seattle / Pioneer Square commercial office real estate prices are lower than in San Francisco / SoMa, with higher confidence

Note to dataset: these prices may be skewed on the low side, because these are *published prices* for office commercial real estate; these do not include the more *premium unpublished* real estate location prices.

Commercial Real Estate Price Distributions: SEA / Pioneer Square and SF / SoMa District (June 2019)





- Conclusions:
  - Seattle / Pioneer Square shows lower price statistics for mean, median, and standard deviation, relative to prices in the San Francisco / SoMa district.
  - Seattle / Pioneer Square offers the prospect of lower office space location cost, relative to San Francisco / SoMa.

Corporate Financial Impact Factors (cont'd.)

2) Business Density Concentration for Each Neighborhood, Utilizing the Foursquare Dataset

- From Foursquare, I generated the number of business venues in each neighborhood. I calculated and prioritized the highest density of businesses in each neighborhood to drive a higher judgment of success for our proposed company, to engage businesses in our locality.
- Using the machine learning algorithm Density Based Spatial Clustering for Applications with Noise (DBSCAN), I solved for the region with highest density of businesses between San Francisco / SoMa and Seattle / Pioneer Square neighborhoods.

Machine Learning: Why I Selected DBSCAN Clustering for This Study

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is effective for tasks including class identification on a spatial context. Density-based clustering locates regions of high density separated from one another by regions of low density.

→ I selected DBSCAN clustering for this study to generate best access to the business venues for each neighborhood, by grouping the venues into clusters with nearest locations (best geographical proximity). Clusters can be arbitrary in shape.

- - - - - - -

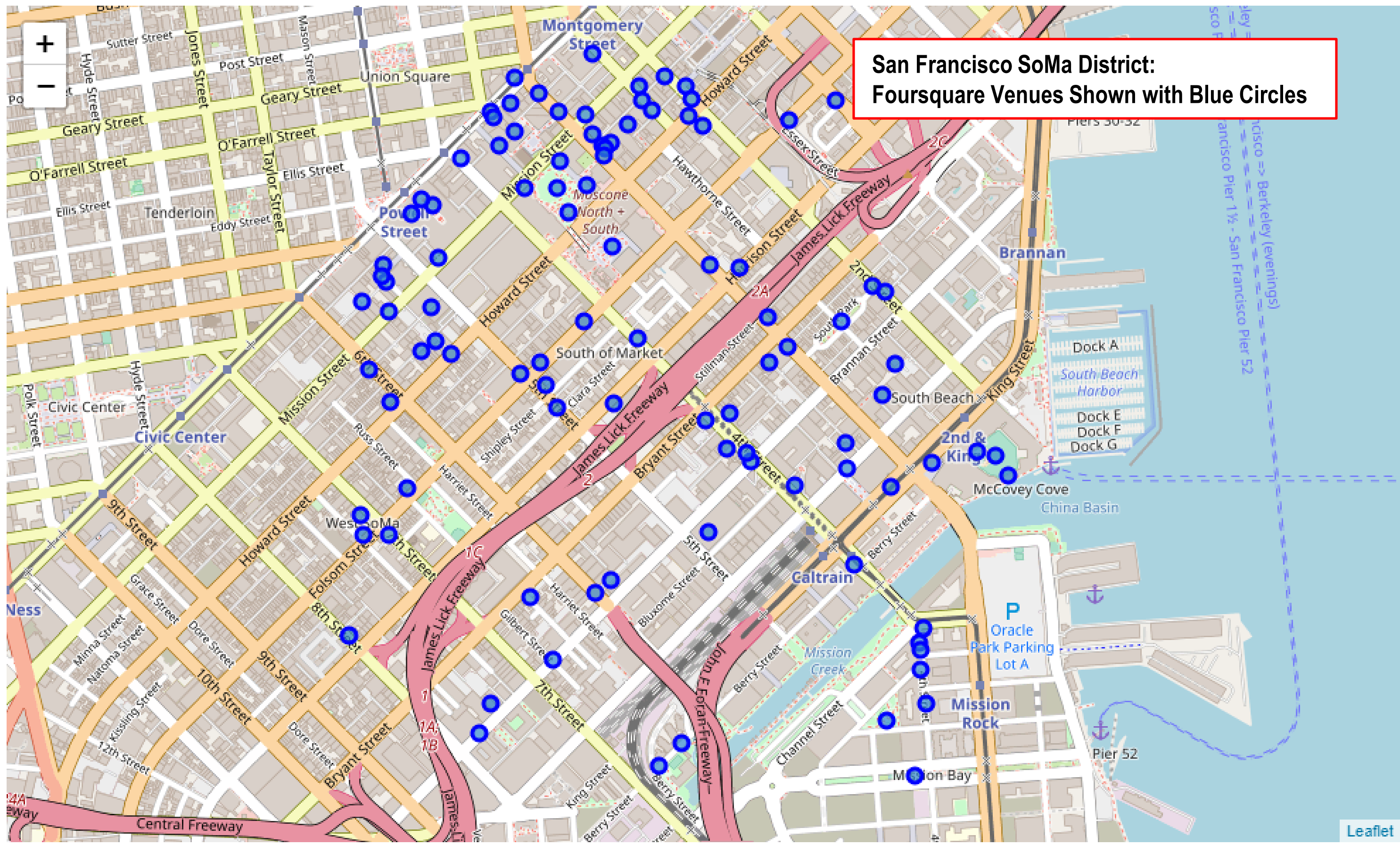
DBSCAN Computation Method:

1) DBSCAN accepts two input parameters: minimum points (M) and radius (R).
   - M is the minimum number of data points we want in a neighborhood to define a cluster.
   - R determines a specified radius that, if it includes the minimum number of points M within it, we call it a dense area.
2) DBSCAN constructs clusters by examining the selected number of minimum points in a cluster (M), then applying the radius (R) to apply to construct continuity of each cluster with the minimum number of points and radius.

- - - - - - -

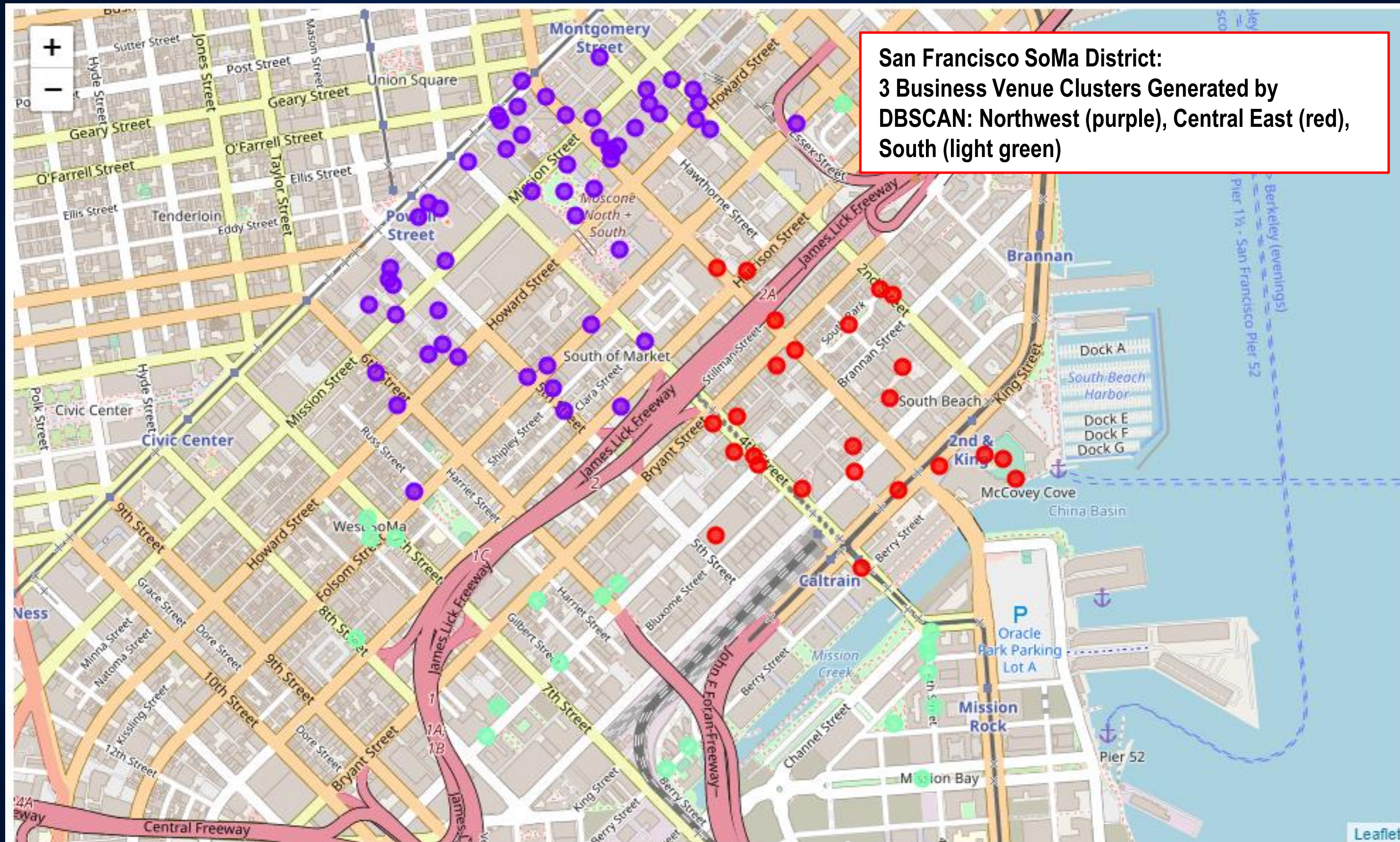
Reference: IBM Data Science content on CognitiveClass.ai

San Francisco / SoMa: Venues Mapped With Folium, and Code to Generate It




```
In [30]: # Access the Foursquare database for San Francisco / SoMa local businesses
         # Analyze venues / local businesses within 1600 meters of the target neighborhood (approx. 1-mile radius)
         latitude_SFSoMa = 37.779190
         longitude_SFSoMa = -122.398102
         radius = 1600
         limit = 100
         url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(CLIEN

In [36]: # Generate map of San Francisco venues in the proximity of SoMa Center using latitude and longitude values
         map_SoMa_venues = folium.Map(location=[latitude_SFSoMa, longitude_SFSoMa], zoom_start=14)
         mapMarkers(map_SoMa_venues, df_SF)
         map_SoMa_venues
```

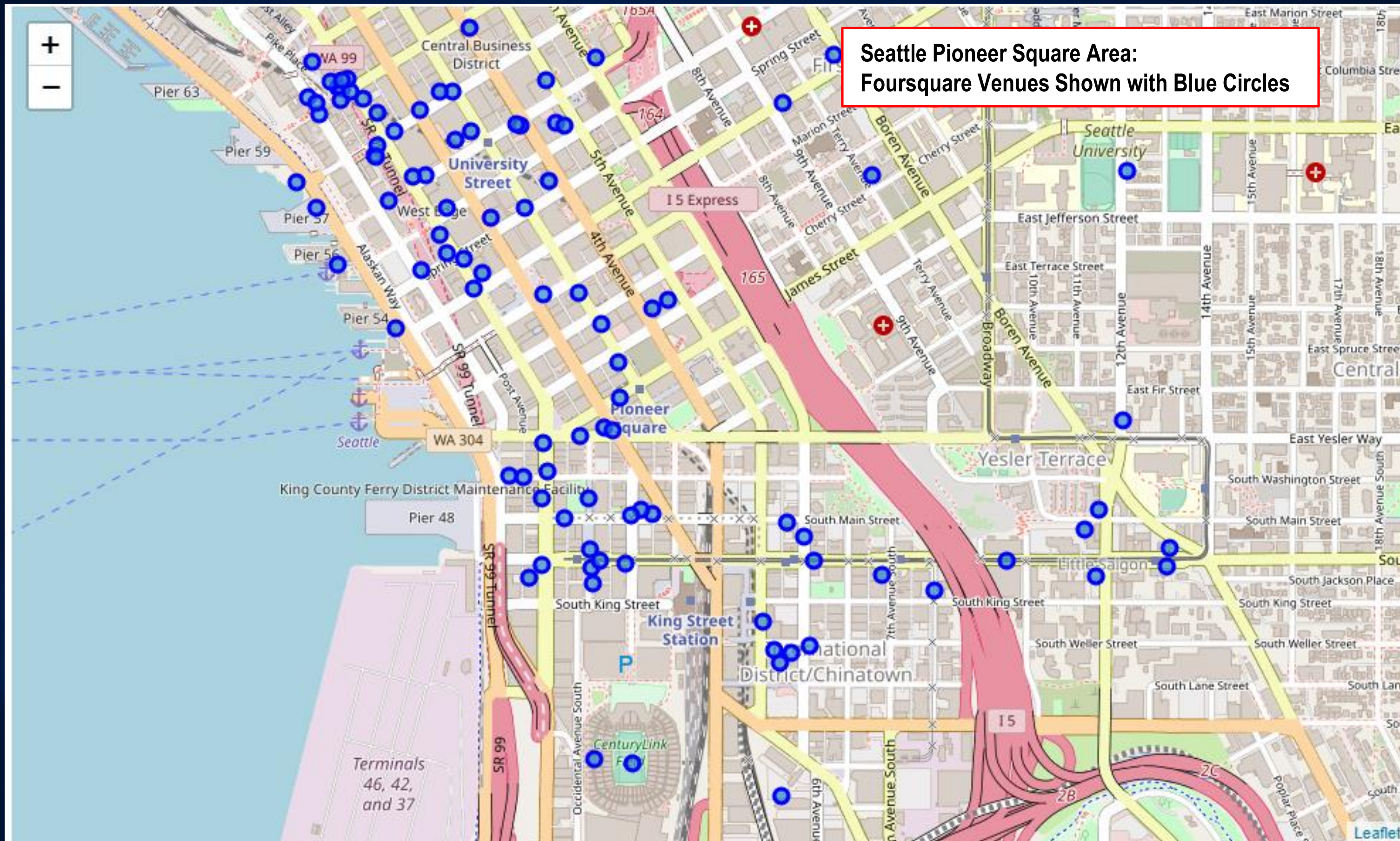
San Francisco / SoMa: Clustered Venues Mapped With Folium, and Code Selections




```
In [38]: # DBSCAN Clustering: Density-Based Spatial Clustering for Applications with Noise

         # First, generate the key two parameters for the DBSCAN Clustering:
         #  minPoints: minimum number of points per cluster
         #  epsilon: proximity ratio
         minPoints = 7
         eps = epsilon(X_df_SF, minPoints)

         # Next, call the DBSCAN clustering algorithm and extract the DBSCAN Cluster Labels
         dbscan_SoMaVenues = DBSCAN(eps, minPoints).fit(X_df_SF)
         df_SF['Cluster Labels'] = dbscan_SoMaVenues.labels_

In [39]: # Now count the number of clustered labels, and count the number of outliers
         df_SF.groupby('Cluster Labels').size().sort_values()
```

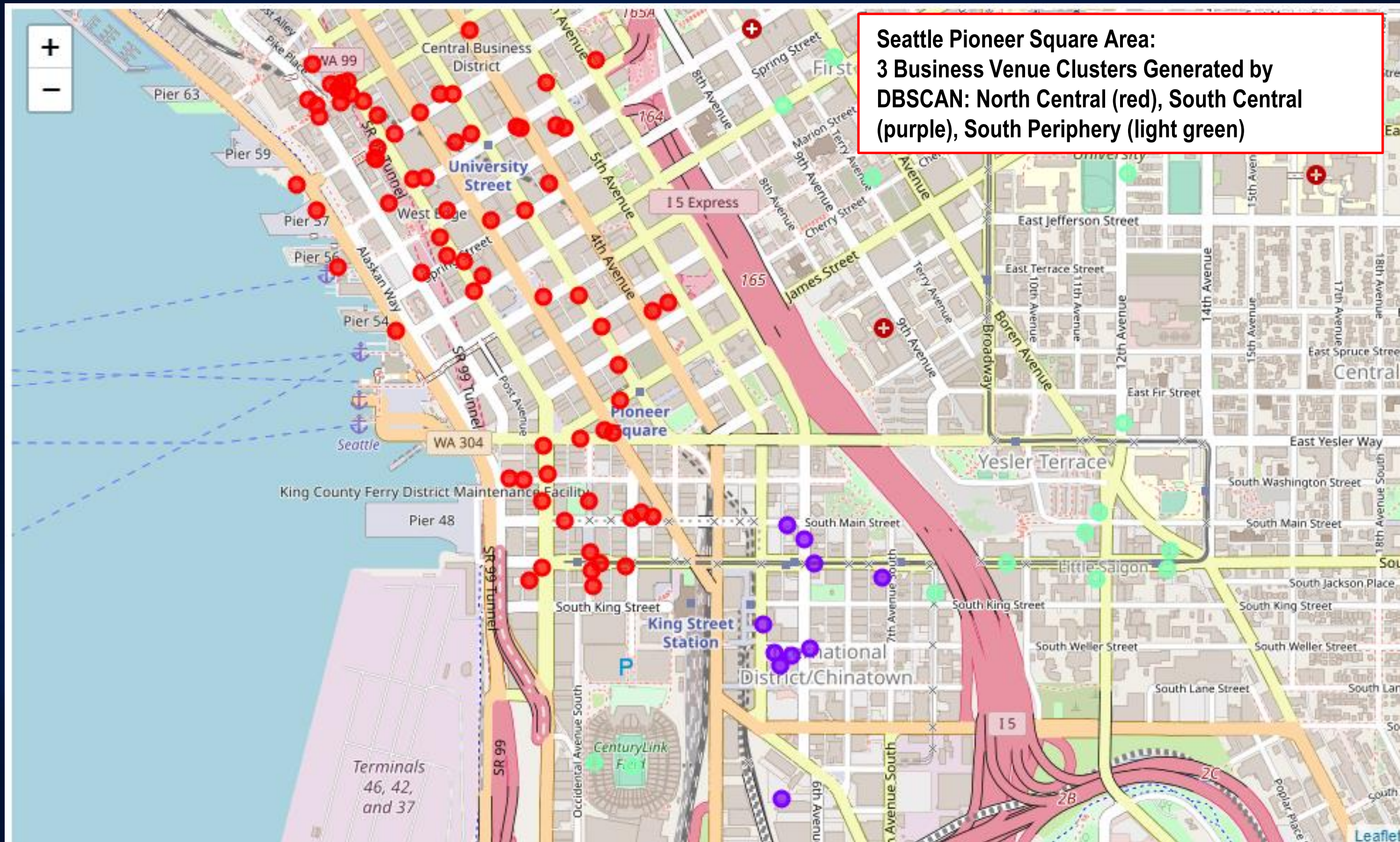
```
Out[39]: Cluster Labels
          2      7
         -1     13
          0     25
          1     55
         dtype: int64
```

Seattle / Pioneer Square: Venues Mapped With Folium, and Code to Generate It




```
In [44]: # Access the Foursquare database for Seattle / Pioneer Square local businesses
         # Analyze venues / local businesses within 1600 meters of the target neighborhood (approx. 1-mile radius)
         latitude_SeattlePS = 47.601954
         longitude_SeattlePS = -122.329204
         radius = 1600
         limit = 100
         url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(CLIEN

In [50]: # Generate map of Seattle venues in the proximity of Pioneer Square using latitude and longitude values
         map_SEA_venues = folium.Map(location=[latitude_SeattlePS, longitude_SeattlePS], zoom_start=15)
         mapMarkers(map_SEA_venues, df_SEA)
         map_SEA_venues
```

Seattle / Pioneer Sq.: Clustered Venues Mapped With Folium, and Code Selections




```
In [52]: # DBSCAN Clustering: Density-Based Spatial Clustering for Applications with Noise

         # First, generate the key two parameters for the DBSCAN Clustering:
         #  minPoints: minimum number of points per cluster
         #  epsilon: proximity ratio
         minPoints = 7
         eps = epsilon(X_df_SEA, minPoints)

         # Next, call the DBSCAN clustering algorithm and extract the DBSCAN Cluster Labels
         dbscan_SEA_PS_Venues = DBSCAN(eps, minPoints).fit(X_df_SEA)
         df_SEA['Cluster Labels'] = dbscan_SEA_PS_Venues.labels_

In [53]: df_SEA.groupby('Cluster Labels').size().sort_values()
```

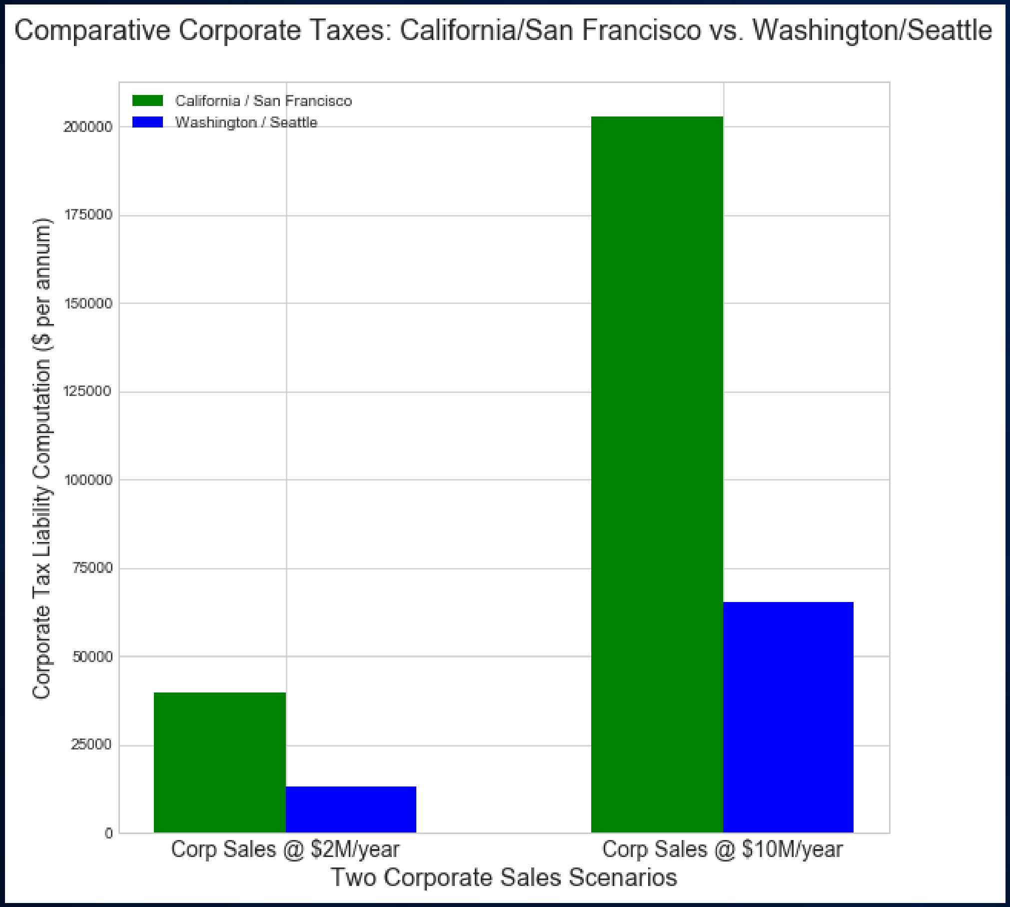
```
Out[53]: Cluster Labels
         -1     7
          2     7
          1    10
          0    76
         dtype: int64
```

Corporate Financial Impact Factors (cont'd.)

3) Corporate Tax Rates by State and City

- To calculate financial impact of corporate business taxes, I calculated two tax planning scenarios for a C corporation located in each neighborhood, based on the percent of gross receipts (gross sales), or percent of net income before taxes, as appropriate for the state and city.
- **State of California and San Francisco City Corporate Tax References:**
  https://www.ftb.ca.gov/businesses/index.shtml, https://sftreasurer.org/business
- **State of Washington and Seattle City Corporate Tax References:**
  https://dor.wa.gov/find-taxes-rates/business-occupation-tax, https://www.seattle.gov/business-licenses-and-taxes

Projected Corporate Tax Computations: SF / SoMa vs. SEA / Pioneer Square



- California and San Francisco tax rates for C corporations are higher than for Washington and Seattle, respectively:
  - California business tax rate = 8.84% on Corporate Net Income
  - San Francisco Corporate Gross Receipts Tax: scales with sales revenues; 0.46% for $2M sales, 0.51% for $10M sales.
  - San Francisco Payroll Expense Tax = 0.38%, for corporate entities with a personnel payroll exceeding $300K/year.
  - Washington business tax rate: 1.5% on Corporate Net Income
  - Seattle City Gross Receipts Tax: 0.427% on Gross Receipts.
- The graphs at right show comparative corporate tax computations for SF and SEA, for two corporate sales scenarios
- Conclusion: locating a business in Seattle should have lower corporate tax liability for comparable sales levels.

Employee Quality of Life Factors for the Business Location Decision

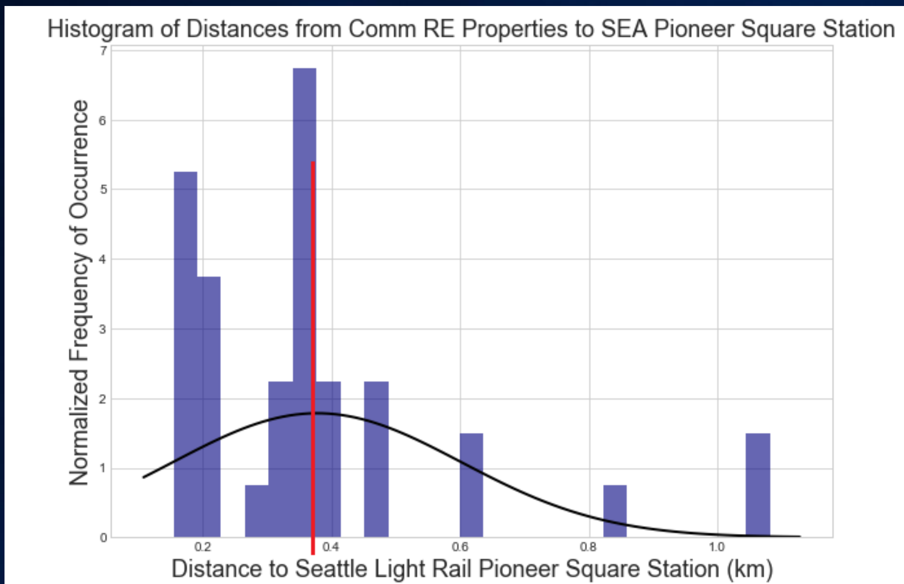
To benchmark Employee Quality of Life, I utilized the Foursquare database and public transportation APIs to generate quantitative measurements to assess:

4) **Ease of Public Rail Transportation Access**
   - Starting from the nearest public rail transportation station for each neighborhood (selected as the station with the minimum mean distance to the centroid of each neighborhood), I calculated distances from the available commercial real estate properties to the nearest public rail transportation station, and compared the results.

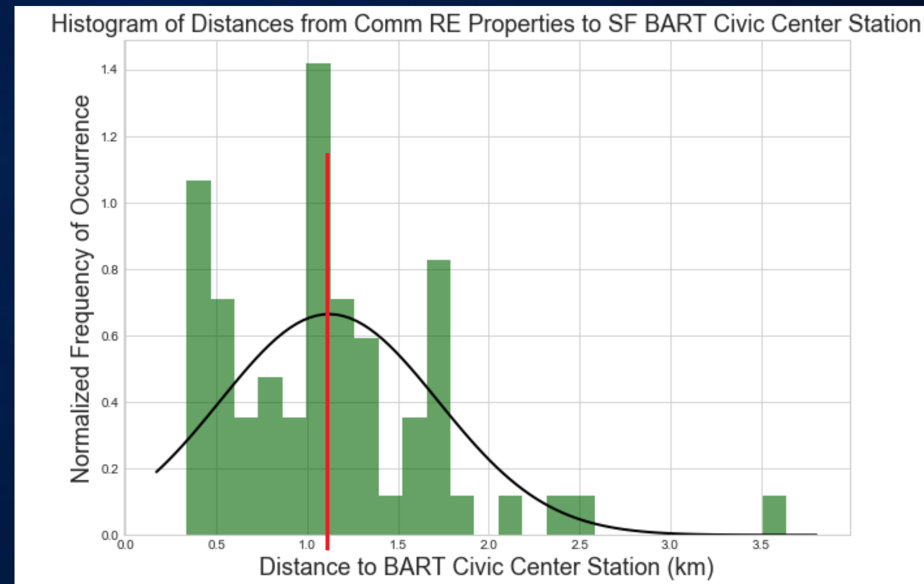
5) **Restaurants Availability**
   - Number of restaurants in each region.

6) **Exercise Facilities Availability**
   - Number of gyms / exercise facilities in each region.

Distances to Public Rail Transportation from Comm RE Properties: SEA / Pioneer Sq and SF / SoMa (p.1 of 2)




```
Mean    Distance to Pioneer Square Station =  0.37739335412958364  km
Median  Distance to Pioneer Square Station =  0.345371181277514  km
Std Dev Distance to Pioneer Square Station =  0.22391925869828758  km
```



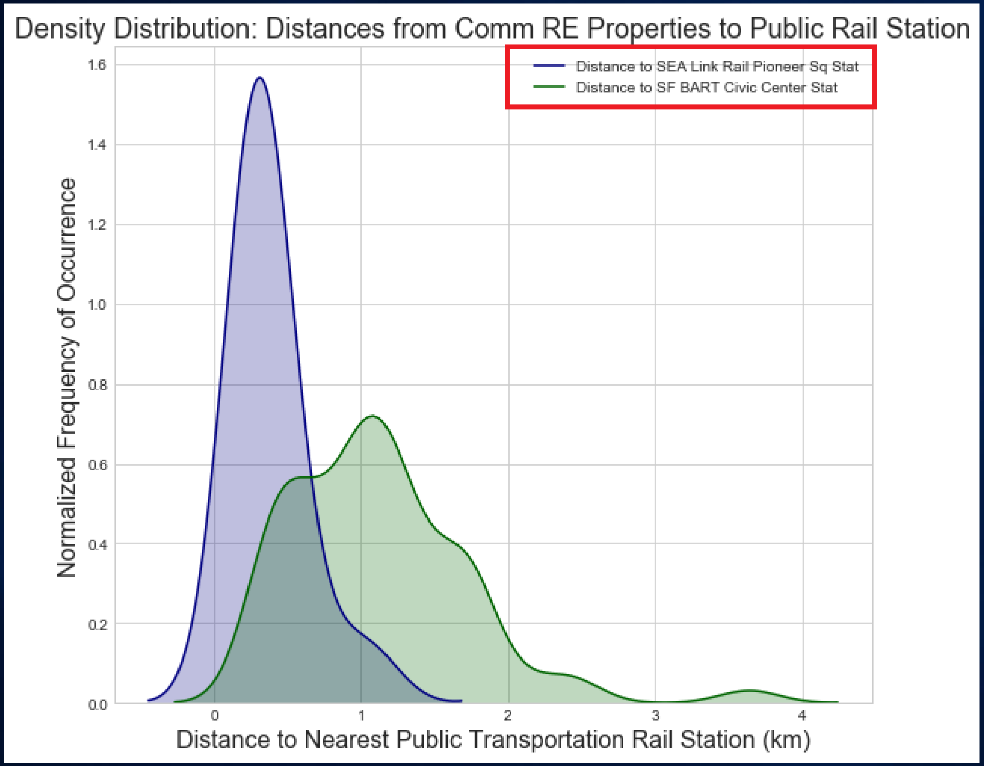

```
Mean    Distance to BART Civic Center Station =  1.1202430127916116  km
Median  Distance to BART Civic Center Station =  1.0487575650479748  km
Std Dev Distance to BART Civic Center Station =  0.6007978504972227  km
```

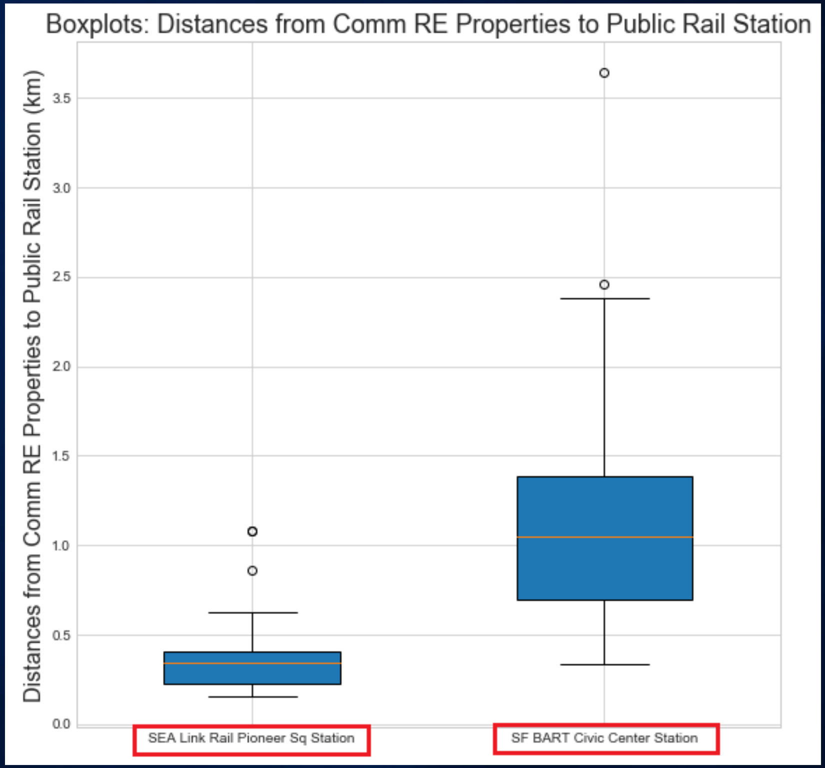
Seattle / Pioneer Square: Mean Distance to Public Rail Transportation = 0.37 km

SF / SoMa: Mean Distance to Public Rail Transportation = 1.12 km

- Conclusion: Seattle / Pioneer Square commercial real estate offices are located nearer to public rail transportation than comparable SF / SoM properties.

Distances from Comm RE Properties to Public Rail Transportation: SEA / Pioneer Sq and SF / SoMa (p.2 of 2)





- Conclusion: these distribution plots show closer proximity of Seattle / Pioneer Square commercial real estate to public rail transportation, relative to comparable to SF / SoMa commercial properties.

Selections from Restaurant Venues Search
(structured into Python Pandas dataframes)

SF / SoMa District

	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	neighborhoo
0	The Chieftain Irish Pub & Restaurant	Irish Pub	198 5th St	US	San Francisco	United States	at Howard St	656	[198 5th St (at Howard St), San Francisco, CA ...	[{'label': 'display', 'lat': 37.78150216218952...	37.781502	-122.404972	Nal
1	Mars Bar & Restaurant	Bar	798 Brannan St	US	San Francisco	United States	at 7th St	795	[798 Brannan St (at 7th St), San Francisco, CA...	[{'label': 'display', 'lat': 37.77324531616279...	37.773245	-122.403121	Nal
2	Canton Dim Sum & Seafood Restaurant	Chinese Restaurant	655 Folsom St	US	San Francisco	United States	at Hawthorne St	604	[655 Folsom St (at Hawthorne St), San Francisc...	[{'label': 'display', 'lat': 37.78461802766915...	37.784618	-122.397991	Nal
3	Restaurant Anzu	Restaurant	222 Mason St	US	San Francisco	United States	at Ellis St	1240	[222 Mason St (at Ellis St), San Francisco, CA...	[{'label': 'display', 'lat': 37.78570272873303...	37.785703	-122.409542	Nal
4	B Restaurant & Bar	Lounge	720 Howard St	US	San Francisco	United States	btw 4th & 5th	684	[720 Howard St (btw 4th & 5th), San Francisco,...	[{'label': 'display', 'lat': 37.78465243491517...	37.784652	-122.401668	Nal

Seattle / Pioneer Square Area

	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	neighborhood
0	Harbor City Restaurant	Chinese Restaurant	707 S King St	US	Seattle	United States	at 7th Ave S	591	[707 S King St (at 7th Ave S), Seattle, WA 981...	[{'label': 'display', 'lat': 47.59837711767561...	47.598377	-122.323382	NaN
1	Lowell's Restaurant	Seafood Restaurant	1519 Pike Pl	US	Seattle	United States	at 1st	1197	[1519 Pike Pl (at 1st), Seattle, WA 98101, Uni...	[{'label': 'display', 'lat': 47.6089372, 'lng'...	47.608937	-122.341346	NaN
2	New Star Restaurant	Chinese Restaurant	516 S Jackson St	US	Seattle	United States	at 5th Ave S	341	[516 S Jackson St (at 5th Ave S), Seattle, WA ...	[{'label': 'display', 'lat': 47.59924168108387...	47.599242	-122.327088	NaN
3	Bush Garden Restaurant	Karaoke Bar	614 Maynard Ave S	US	Seattle	United States	NaN	638	[614 Maynard Ave S, Seattle, WA 98104, United ...	[{'label': 'display', 'lat': 47.59699150849453...	47.596992	-122.324930	NaN
4	The Fisherman's Restaurant & Bar	Seafood Restaurant	1301 Alaskan Way	US	Seattle	United States	On Pier 57	1087	[1301 Alaskan Way (On Pier 57), Seattle, WA 98...	[{'label': 'display', 'lat': 47.60612306667809...	47.606123	-122.342305	NaN

Selections from Fitness Center Venues Search
(structured into Python Pandas dataframes)

SF / SoMa District

	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	neighborl
0	Gym Beacon	Gym	250 King Street, 4th Floor Courtyard	US	San Francisco	United States	NaN	398	[250 King Street, 4th Floor Courtyard, San Fra...	[{'label': 'display', 'lat': 37.77775180480304...	37.777752	-122.393948	
1	Avalon at Mission Bay Gym	Gym	383 King St	US	San Francisco	United States	at 5th St.	541	[383 King St (at 5th St.), San Francisco, CA 9...	[{'label': 'display', 'lat': 37.77593451669104...	37.775935	-122.393527	
2	World Gym	Gym	290 De Haro Street	US	San Francisco	United States	at 16th St	1458	[290 De Haro Street (at 16th St), San Francisc...	[{'label': 'display', 'lat': 37.76646086077708...	37.766461	-122.402019	Show| Sc
3	Google Gym - SPEAR121	Gym / Fitness Center	NaN	US	San Francisco	United States	NaN	1515	[San Francisco, CA 94105, United States]	[{'label': 'display', 'lat': 37.79195025910042...	37.791950	-122.392085	
4	Embarcadero Tower Gym	Gym	88 King St	US	San Francisco	United States	2nd Street	808	[88 King St (2nd Street), San Francisco, CA 94...	[{'label': 'display', 'lat': 37.78046307075421...	37.780463	-122.389059	

Seattle / Pioneer Square Area

	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	postalCode	state
0	The Post Gym	Gym / Fitness Center	888 Western Ave	US	Seattle	United States	Marion & Western	561	[888 Western Ave (Marion & Western), Seattle, ...	[{'label': 'display', 'lat': 47.60318992444650...	47.603190	-122.336452	98104	WA
1	Gym at Hilton Seattle	Gym / Fitness Center	1301 6th Ave	US	Seattle	United States	NaN	910	[1301 6th Ave, Seattle, WA 98101, United States]	[{'label': 'display', 'lat': 47.60887555668738...	47.608876	-122.335657	98101	WA
2	Gym	Gym / Fitness Center	1801 S Jackson St	US	Seattle	United States	NaN	1614	[1801 S Jackson St, Seattle, WA 98144, United ...	[{'label': 'display', 'lat': 47.599087, 'lng':...	47.599087	-122.308114	98144	WA
3	Crowne Plaza Gym	Gym	1113 6th Ave	US	Seattle	United States	NaN	716	[1113 6th Ave, Seattle, WA 98101, United States]	[{'label': 'display', 'lat': 47.60808284363962...	47.608083	-122.332118	98101	WA
4	Harbor Steps Gym	Gym	1221 1st Ave	US	Seattle	United States	University	860	[1221 1st Ave (University), Seattle, WA 98101,...	[{'label': 'display', 'lat': 47.60632437897696...	47.606324	-122.338655	98101	WA

Results: Business Location Decision Analysis for Two Neighborhoods

Decision Factor	San Francisco / SoMa	Seattle / Pioneer Square
Commercial Real Estate Rental Prices	More expensive: mean Comm RE price = $51.30 / SqFt / year	Less expensive: mean Comm RE price = $37.58 / SqFt / year
Business Density Evaluation	Somewhat less dense: business density ratio = 0.87	Somewhat more dense: business density ratio = 0.93
State and City Corporate Taxes	Higher corporate taxes: CA / San Francisco corporate taxes are approx. 3X those in WA / Seattle	Lower corporate taxes: WA / Seattle corporate taxes are approx. 1/3 those in CA / San Francisco
Employee QoL* : Access to Public Rail Transportation	Somewhat more distant, at mean distance = 1.1 km	Better access to public rail, at mean distance = 0.4 km
Employee QoL: Restaurants	Very comparable	Very comparable
Employee QoL: Exercise Facilities	Very comparable	Very comparable

* Employee QoL: Employee Quality of Life

Discussion of Results

With the Results table (above), I constructed a scoring mechanism to judge the business location decision across the six influential factors analyzed, weighting Corporate Financial Impact at 80% and Employee QoL at 20%, and comparing mean scores for each factor:

- Weighting Judgment Measure:

  = (Corporate Financial Impact factors)*80% + (Employee Quality of Life factors)*20%

  = [(Business Density)*50% + (Comm RE Prices)*20% + (Corporate Taxes)*10%
  + (Empl. Access to Public Rail)*8% + (Restaurants)*6% + (Exercise Facilities)*6%]

- SF / SoMa District scoring result = (0.87*50%)+(0.73*20%)+(0.33*10%)+(0.4*8%)+(1.0*6%)+(1.0*6%)
  = 0.77
- Seattle / Pioneer Sq scoring result = (0.93*50%)+(1.0*20%)+(1.0*10%)+(1.0*8%)+(1.0*6%)+(1.0*6%)
  = 0.96

- Based on the above factors considered and the scoring mechanism, Seattle / Pioneer Square scores higher and shows to be a more attractive technology startup business location.
- However, there are clearly additional factors to consider for the location of a technology-based professional services company startup, especially:
  - Access to skilled technology personnel to staff the specific business (this may favor SF).
  - If the firm is venture capital funded, then the VCs' preferences for location are clearly important.

* The 80/20 weighting for Corporate Financial Impact versus Employee QoL factors was also a Bayesian judgment, because I knew that Employee QoL factors were near equal between the two neighborhoods.

Conclusion: Battle of the Neighborhoods - Decision on Where to Locate a Professional Services Business

- Based on the analysis in this data science study on where to locate a professional services company, I would prioritize locating the professional services business in Seattle / Pioneer Square, a location with substantially lower corporate costs than in San Francisco, and with comparable concentration of high tech startup companies and employee quality of life.

- Next steps: a more thorough analysis of locating the business will include:
  - Assess access to skilled technology personnel to staff the specific business.
  - If the firm is venture capital funded, then the VCs' preferences for location are clearly important.
  - Assess one's professional network and contacts in the region to start the business.

- Key Functions and Skills Developed in this IBM Data Science Capstone Project:
  - This project significantly strengthened my skills in several aspects of Data Science project planning and programming. Planning the study content for the target audience and the targeted decisions is critical, as is careful mapping of the science/mathematics methodology to programming language functions and libraries. Furthering my skills in Python Pandas and Scikit-Learn has been valuable for structuring datasets and to design machine learning optimizations. Enhancing my skills with Folium enables delivering powerful geographical visualizations.

- I definitely believe the IBM Data Science Professional Certification Program is high quality. The curriculum and completing the projects builds expertise to design serious data science and data visualization projects.
  Thank you to the team.