



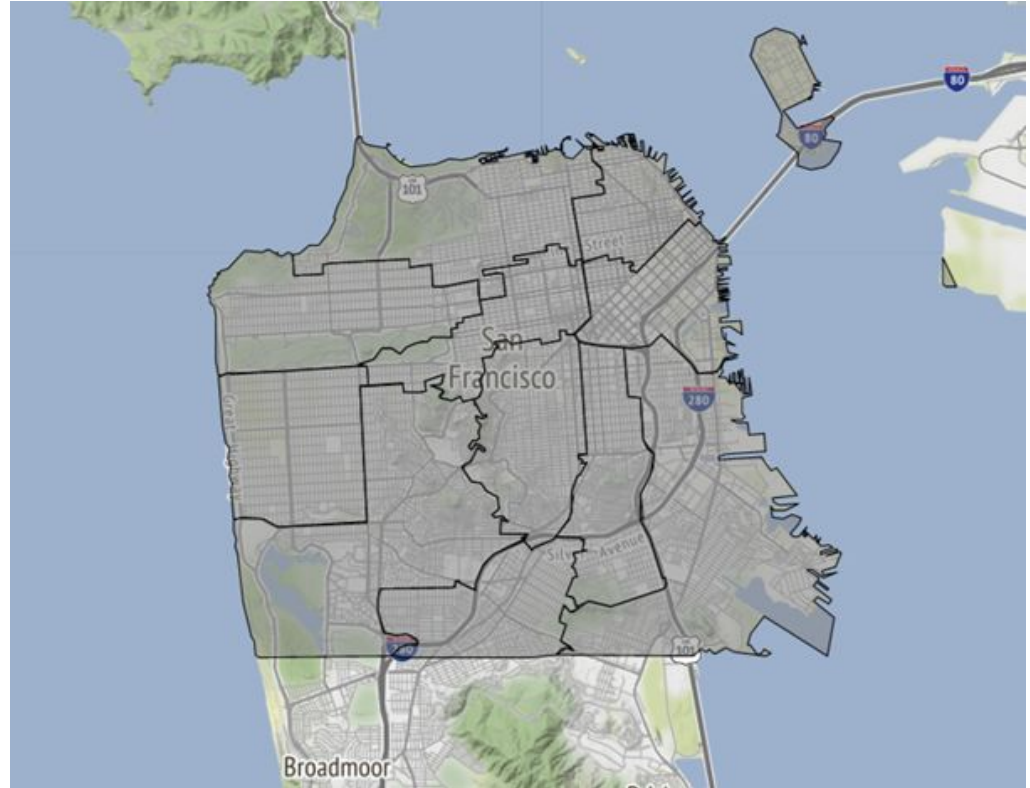
Where to build a pizza place in San Francisco?

A coursera capstone project by Wolfgang Hamer

Data

The analysis is based upon three datasets:

1. Geometries of the districts of San Francisco which are downloaded from the official website data.sfgov.org



Data

The analysis is based upon three datasets:

1. Geometries of the districts of San Francisco which are downloaded from the official website data.sfgov.org
2. Foursquare data for each district which are downloaded via the API

	id	Name	Category	lat	lon
0	4a0e123af964a520c2751fe3	Taquerias El Farolito	Mexican Restaurant	37.721230	-122.437395
0	4ec020bbb8f7963bcdde0f6b	The Dark Horse Inn	Bar	37.716127	-122.440373
0	546960f7498eac74bd5baf47	Tao Sushi	Japanese Restaurant	37.721037	-122.437665
0	4b63b31cf964a520d28c2ae3	Little Joe's Pizza	Pizza Place	37.718478	-122.439856
0	49f796fff964a520c06c1fe3	Roxie Food Center	Sandwich Place	37.726867	-122.441398
...
0	4a64a8f4f964a5206cc61fe3	Spruce	New American Restaurant	37.787551	-122.452777
0	585c8202ca1070180ddb525c	Pearl Spa and Sauna	Bath House	37.785642	-122.429130
0	500088f7d63e64b62bc19e6e	Rich Table	New American Restaurant	37.774891	-122.422736
0	57b3c7c8498e9b9e08349941	Linden Room	Cocktail Bar	37.776503	-122.422794
0	50f21340e4b036c5cc0d7c7d	SFJazz Center	Jazz Club	37.776350	-122.421539

1200 rows × 5 columns

Data

The analysis is based upon three datasets:

1. Geometries of the districts of San Francisco which are downloaded from the official website data.sfgov.org
2. Foursquare data for each district which are downloaded via the API
3. Police Department Incident Reports which are also downloaded from the official website data.sfgov.org

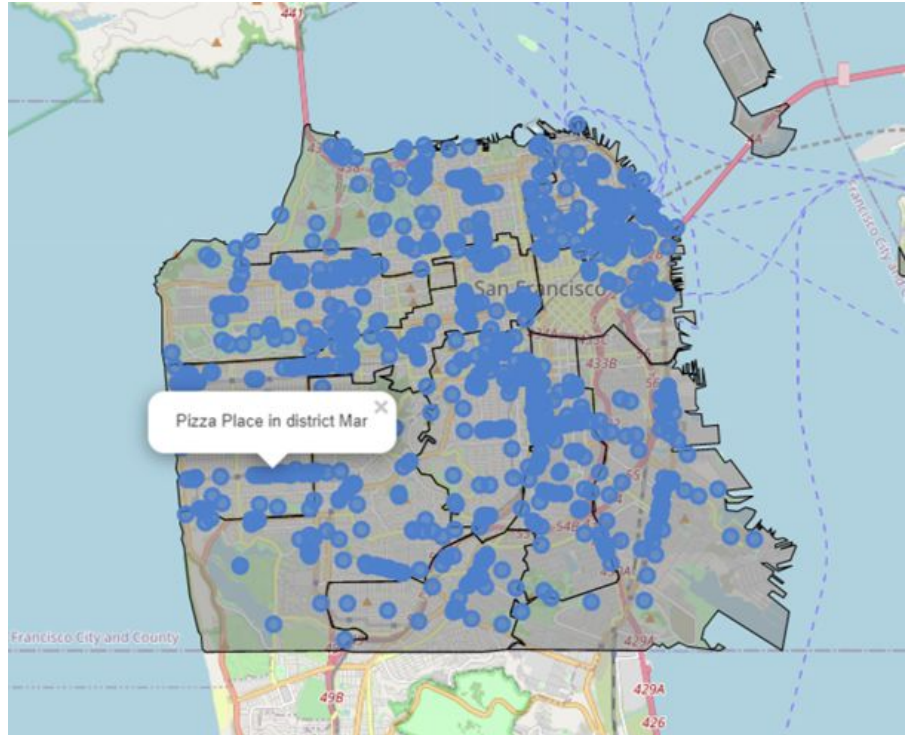
	Incident Category	Incident Subcategory	Incident Description	Latitude	Longitude
0	Offences Against The Family And Children	Other	Domestic Violence (secondary only)	37.762569	-122.499627
1	Non-Criminal	Other	Mental Health Detention	37.780535	-122.408161
2	Missing Person	Missing Person	Found Person	37.721600	-122.390745
3	Offences Against The Family And Children	Family Offenses	Elder Adult or Dependent Abuse (not Embezzleme...	37.794860	-122.404876
4	Assault	Simple Assault	Battery	37.797716	-122.430559
...
356650	Non-Criminal	Non-Criminal	Found Property	37.780927	-122.413676
356651	Larceny Theft	Larceny - From Vehicle	Theft, From Locked Vehicle, >\$950	37.766406	-122.424258
356652	Assault	Simple Assault	Battery	37.759830	-122.425920
356653	Robbery	Robbery - Commercial	Robbery, Chain Store, W/ Force	37.726132	-122.464573
356654	Other Miscellaneous	Other	Resisting, Delaying, or Obstructing Peace Off...	37.784449	-122.416072

356655 rows × 5 columns

Methodology

The methodology is split in three parts:

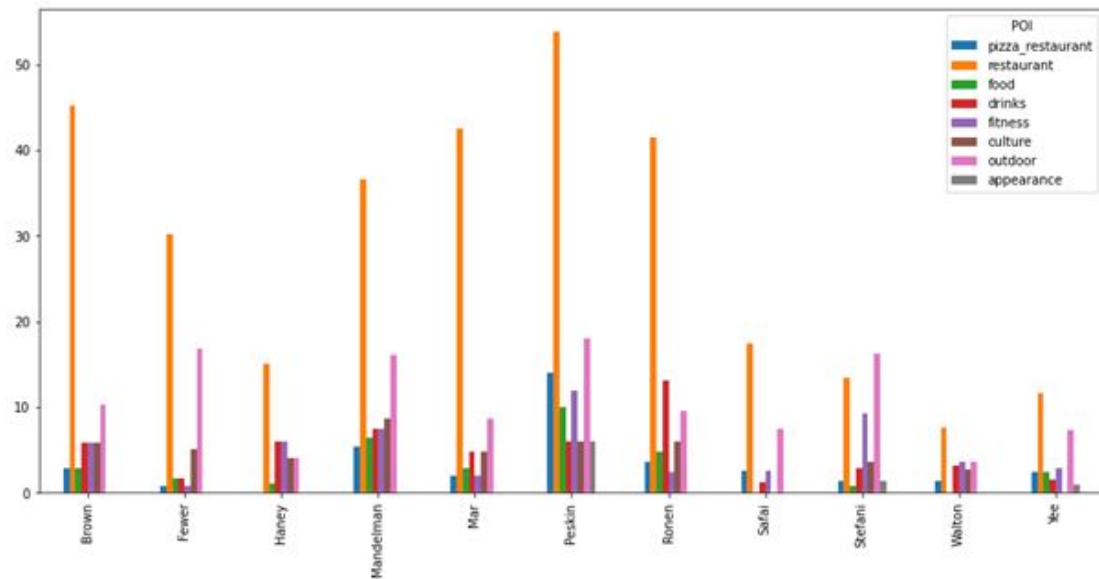
1. Combination of Districts and Foursquare data



Methodology

The methodology is split in three parts:

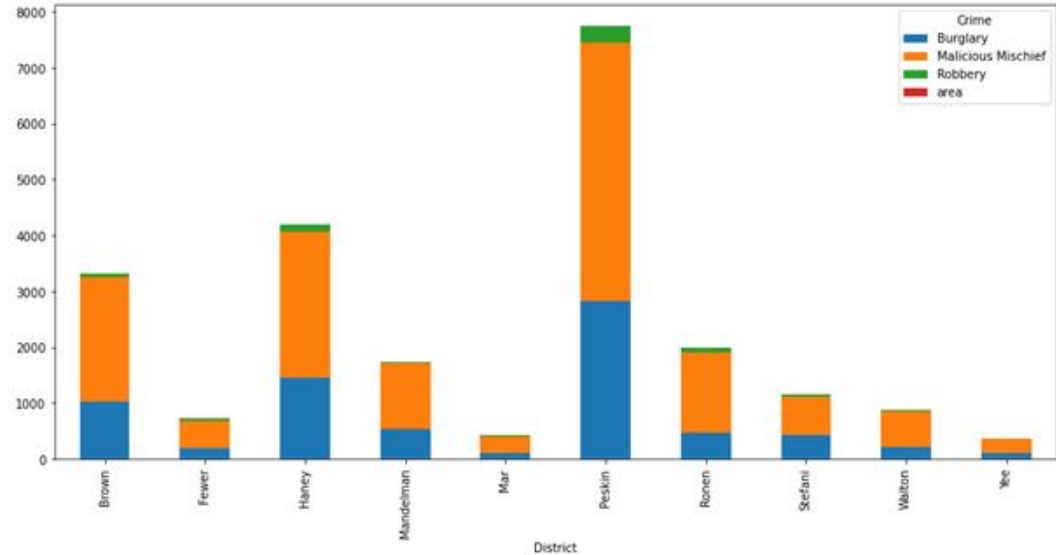
1. Combination of Districts and Foursquare data
2. Combination of the foregone datasets with the crime data



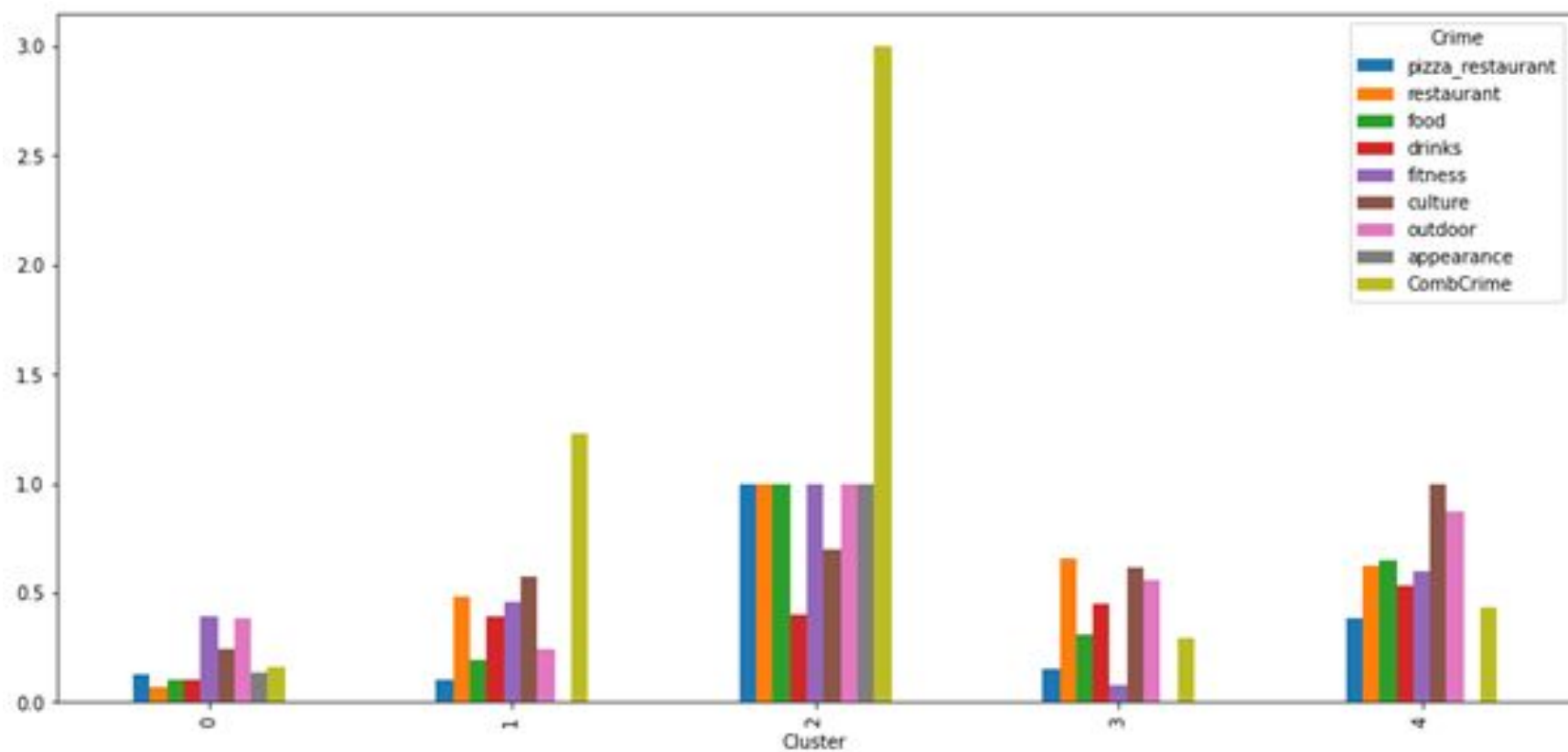
Methodology

The methodology is split in three parts:

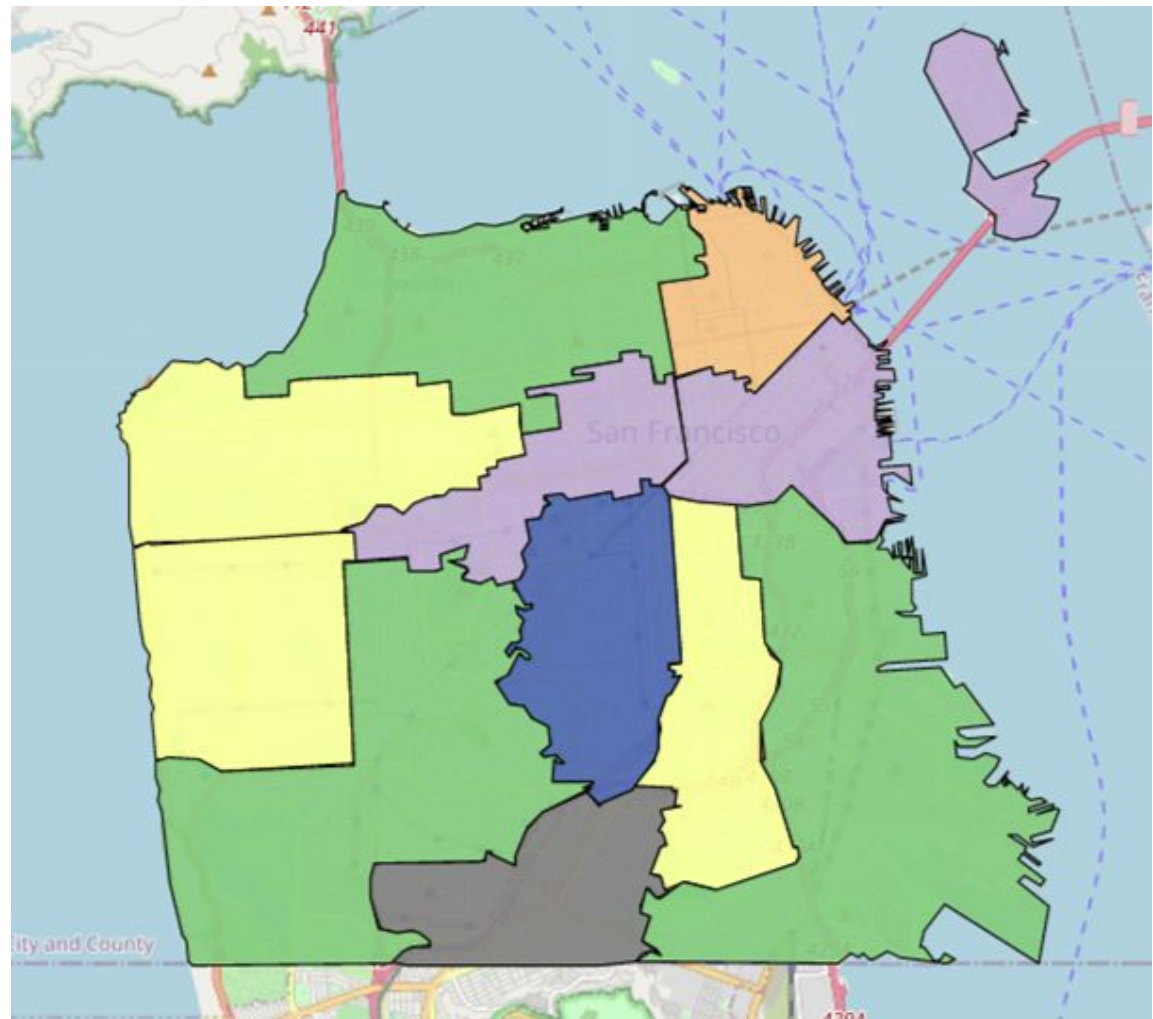
1. Combination of Districts and Foursquare data
2. Combination of the foregone datasets with the crime data
3. Use of the unsupervised kmeans cluster machine learning process to determine the optimal districts for the safe location of a pizza place



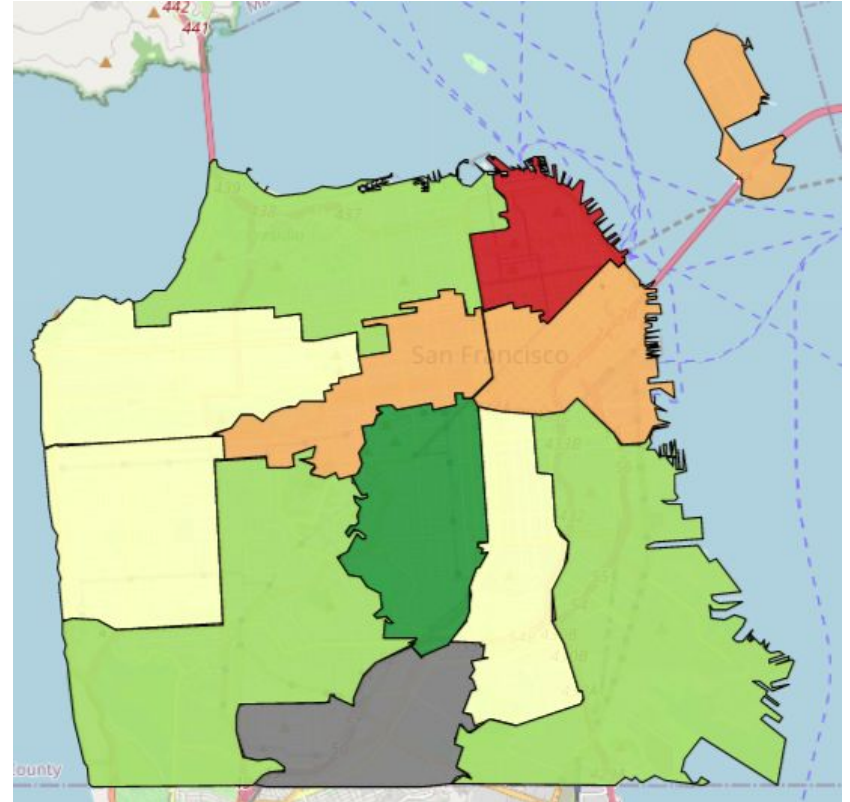
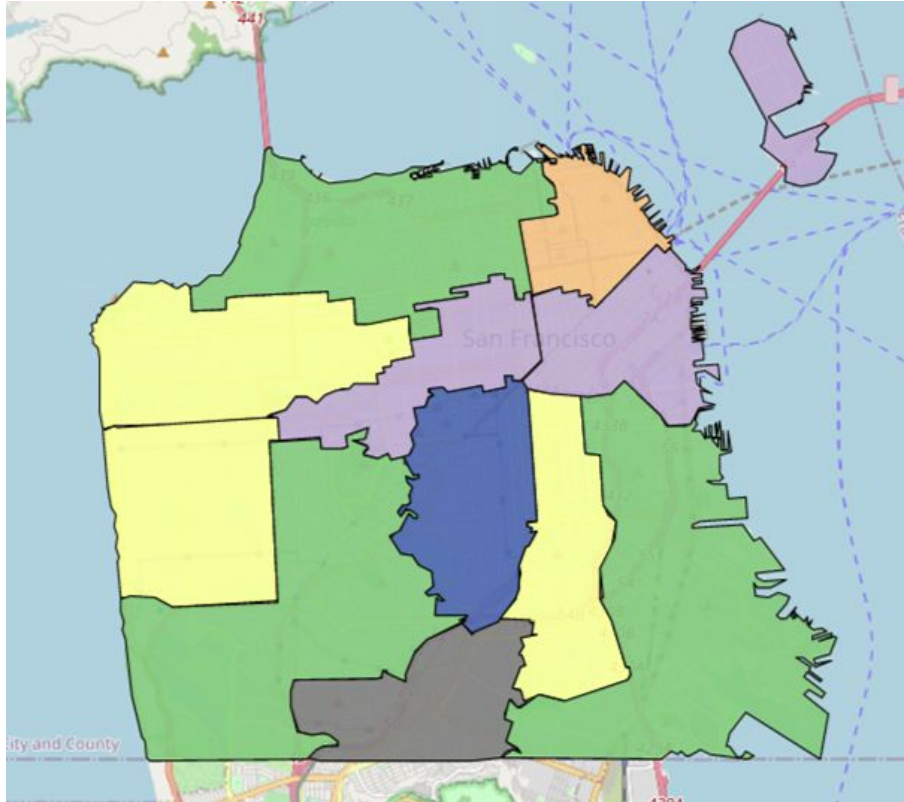
Results



Results



Discussion



Conclusion

- clear **recommendation of the Mandelman** district for the customer
 - **Peskin** seems to be the **least suitable** because of the high crime rate.
 - further information:
 - the targeted shop rent
 - the pizza place has the aim to deliver pizza
 - which income class is expected among the customers
- => further data would be required and **further analysis** could be carried out.