Predicting Best Locations to Move to

Introduction- The problem

People	The problem	Interest
People that live in the district of Nicosia,Cyprus move to Limassol,Cyprus due to a rise in technology jobs.	A lot of people must leave their homes and move to different environments.	Helping workers moving find the best location to move to, where the environment is most similar to where they lived in Nicosia will benefit both the workers and their employers because the toll on their productivity will be less.

Data Acquisition

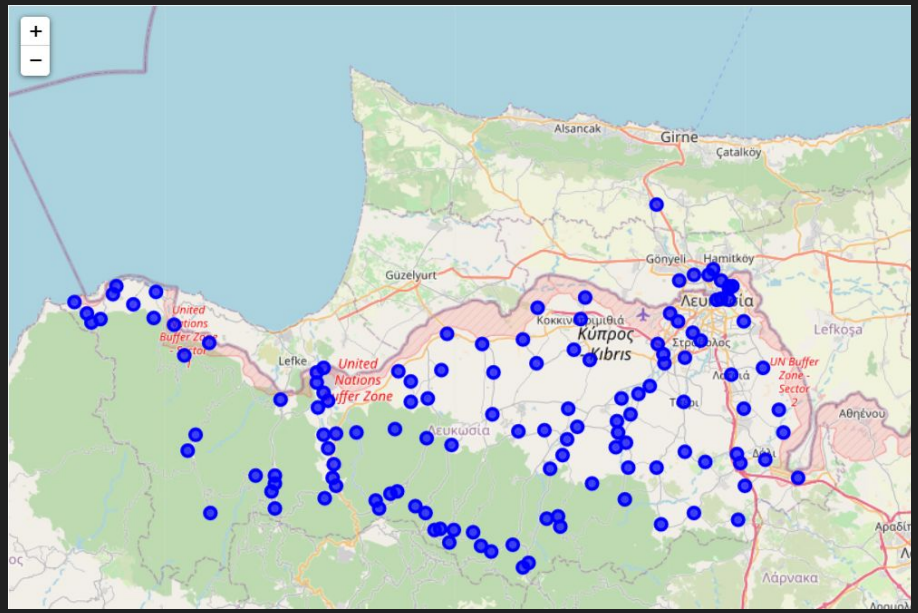
- Municipalities in each district data from the Postal Service of the Republic of Cyprus: https://www.cypruspost.post/en/find-postal-codes
- Coordinates of each municipality obtained through Nominatim API.
- Venues in each municipality acquired through Foursquare API.

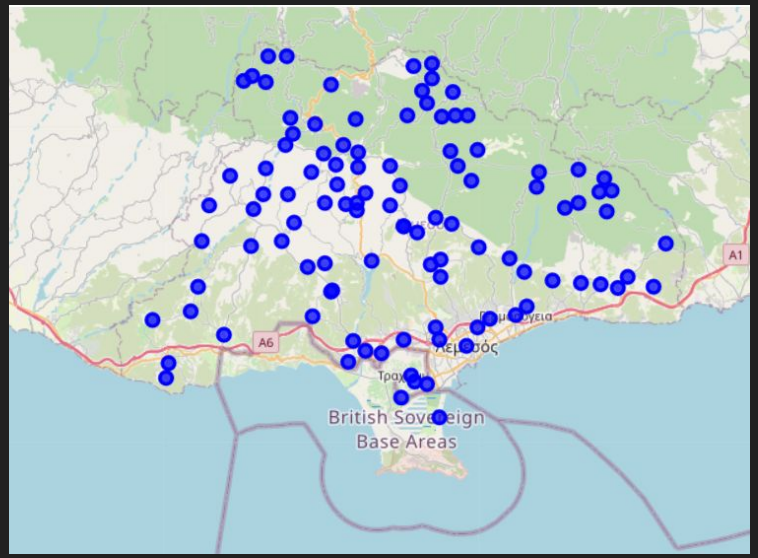
Data Cleaning

- Any unnecessary, missing or duplicate data was removed.
- Municipalities in too rural areas were removed.
- Municipalities with Greek names changed to English versions.

Locations of Municipalities

Locations of Interest in both Districts:





Locations of Municipalities

Locations of Interest in both Districts:

- Spread all across each district, representing a wide range of areas.
- Both rural, coastal and urban areas represented.
- Subcategories of environments are also represented (e.g. areas near bodies of freshwater, areas near main highway roads etc.)

Venues of Municipalities

Using Foursquare API:

- Acquired a sample of 100 venues per municipality.
- Categorized each venue to get the Top 6 Most Common types of venues per municipality.

	1st Most Common	2nd Most Common	3rd Most Common	4th Most Common	5th Most Common	6th Most Common
0	Bakery	Coffee Shop	Café	Greek Restaurant	Gym	Supermarket
1	Café	Bar	Greek Restaurant	Coffee Shop	Historic Site	Bakery
2	Coffee Shop	Bakery	Café	Wine Bar	Greek Restaurant	Gym
3	Bakery	Coffee Shop	Greek Restaurant	Supermarket	Park	Café
4	Bakery	Coffee Shop	Supermarket	Café	Greek Restaurant	Park

Predictive Modeling

Preparing data for use in models:

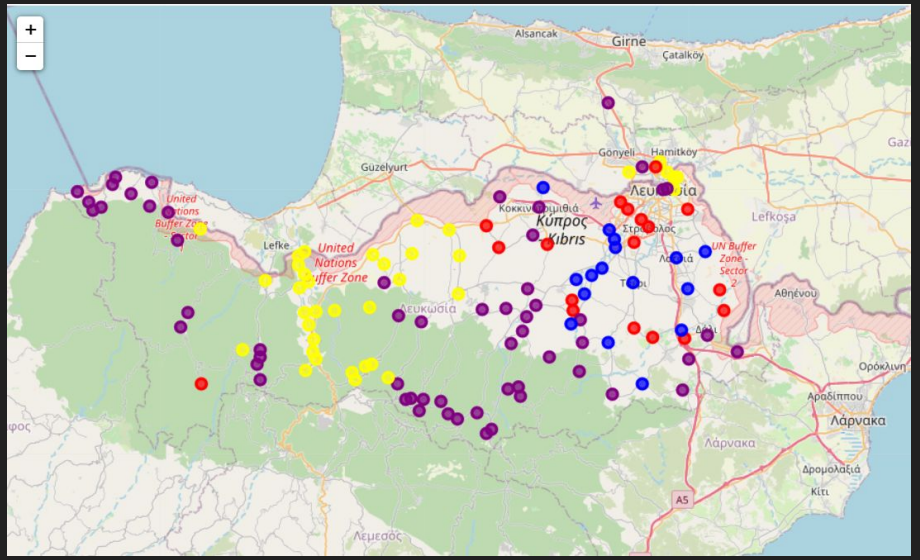
- All of our data is categorical.
- Categorical data is converted into numerical data using the one-hot encoding method.

1st Most Common_Botanical Garden	1st Most Common_Breakfast Spot	1st Most Common_Brewery	...	6th Most Common_Trail	6th Most Common_Turkish Home Cooking Restaurant	6th Most Common_Turkish Restaurant
0	0	0	...	0	0	0
0	0	0	...	0	0	0
0	0	0	...	0	0	0
0	0	0	...	0	0	0
0	0	0	...	0	0	0

Predictive Modeling

Clustering Algorithm:

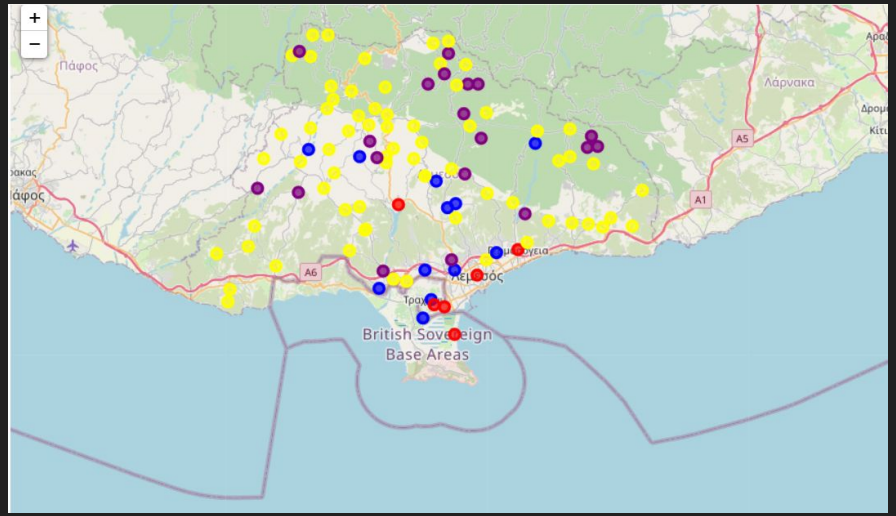
- Locations in Nicosia are clustered based on each municipality's Top 6 Most Common types of venues.



Predictive Modeling

Classification Algorithm:

- Locations in Limassol are classified to the previously generated clusters based on each municipality's Top 6 Most Common types of venues.



Conclusion

- We have clustered our data on Nicosia municipalities using a k-means clustering algorithm, and then the municipalities of Limassol were classified to each cluster by using a k-nearest neighbor algorithm.

- Our final results represent which municipalities in Limassol are better to move to for someone living in the District of Nicosia, based on in what municipality of Nicosia they live.