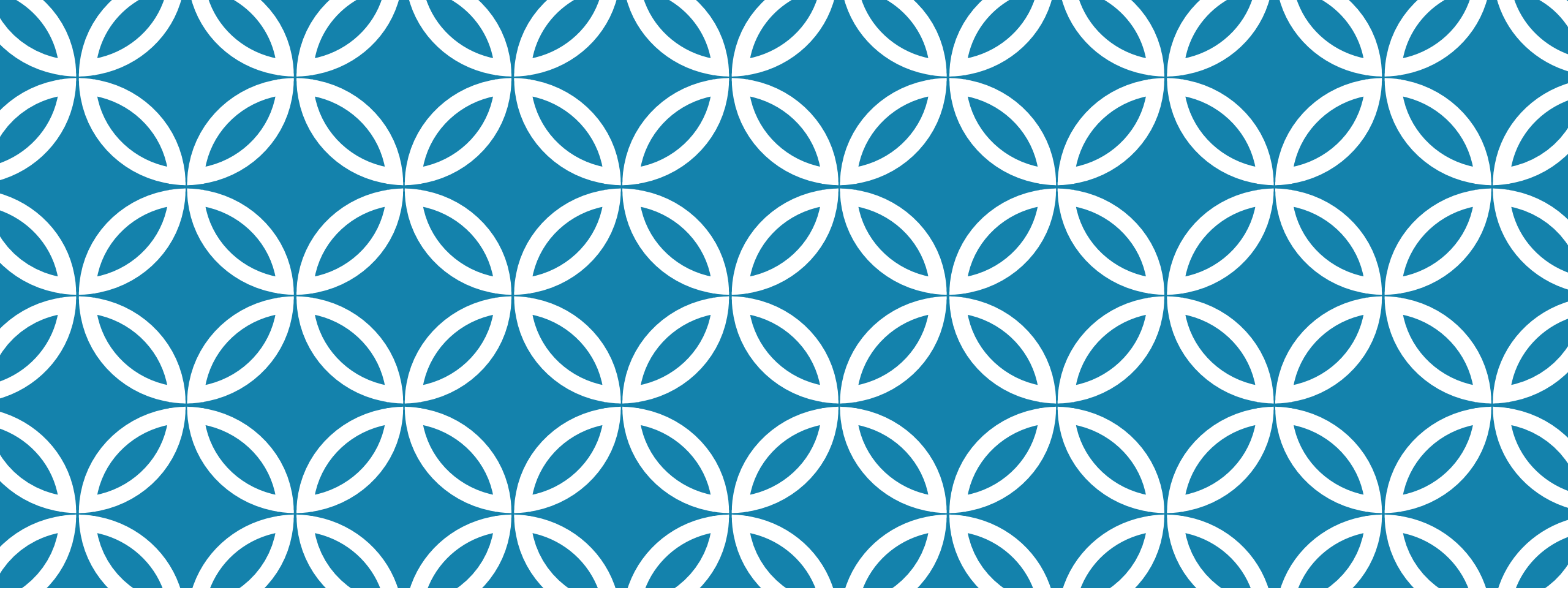


THE BATTLE OF NEIGHBOURHOODS

PRACTICAL QUESTIONS THAT REQUIRE THE COMPARISON ACROSS CITY NEIGHBORHOODS

- A job seeker may wish to focus his/her search on a single neighborhood with jobs that best match his/her qualifications.
- A restaurant looking to expand its locations might perhaps select neighborhoods it wishes to expand into before considering particular sites or neighborhoods
- A person buying or renting a home in a new city might want to be able to compare the neighborhoods of the city.

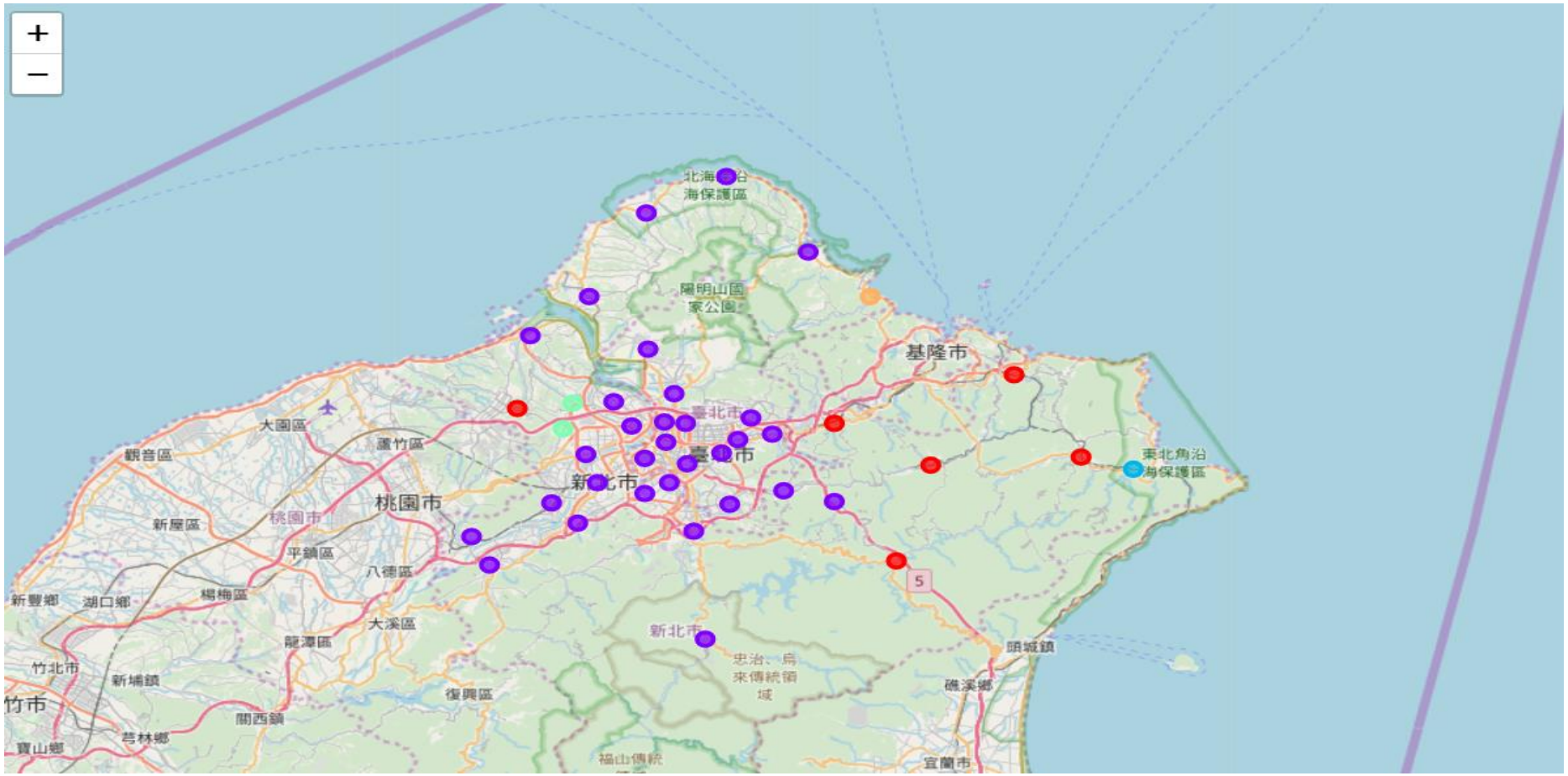
DATA PREPARATION AND ACQUISITION

- **Scraped the Wikipedia page https://en.wikipedia.org/wiki/List_of_townships,_county-administered_cities_and_districts_of_Taiwan**

- **Wrangled the data, cleaned it, and then read it into a pandas dataframe so that it is in a structured format that consist of two columns of 2 Boroughs and 41 Neighborhoods of both cities.**

- **Use Nominatim library to get the latitude and longitude values of all neighborhoods.**

USING K-MEANS CLUSTERING METHOD TO COMPARE NEIGHBORHOODS

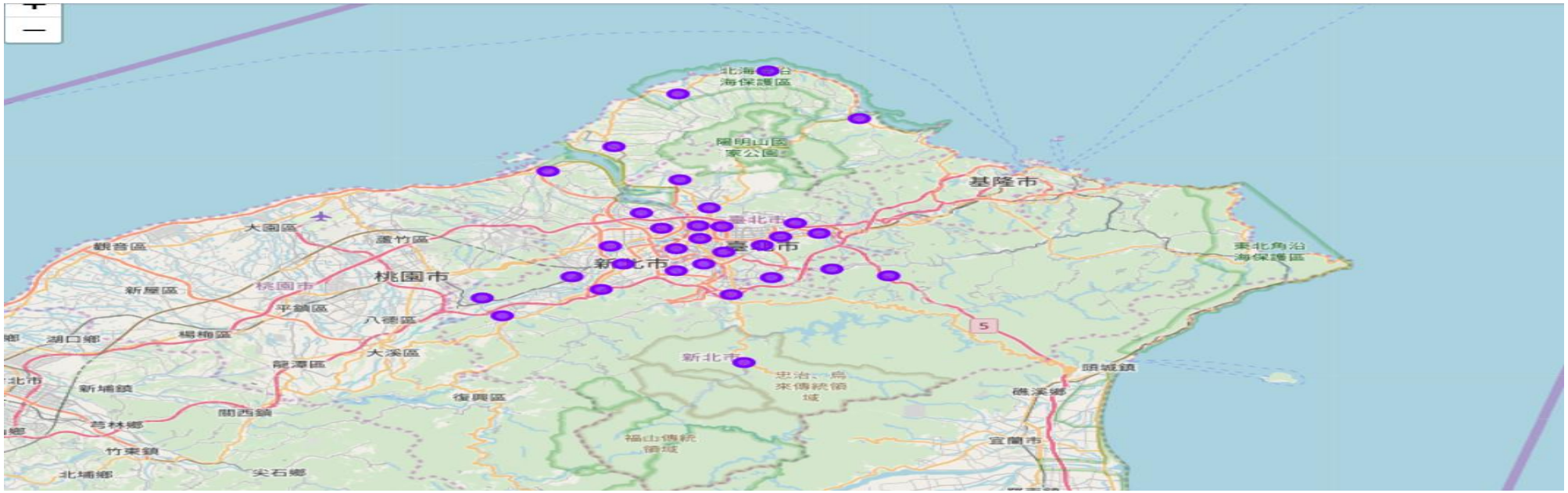


CLUSTER EXAMINATION

Cluster 0 contains the mountain neighborhoods of New Taipei City with historic sites, train stations, convenience stores, and restaurants.

CLUSTER EXAMINATION (CONT.)

Cluster 1 has the heart of both cities. All Taipei neighborhoods and most New Taipei City neighborhoods close to Taipei are included. All these neighborhoods are similar to each other to have access to a variety of venues, including restaurants, hotels, convenience stores and department stores.



CONCLUSION

- For those blue-collar job seekers, neighborhoods in cluster 1 will be a good consideration. These neighborhoods located in city center, has a good choice of restaurants and supermarkets.
- For a Japanese restaurant looking to expand its locations might perhaps select neighborhoods in cluster 0 because there are not many Japanese restaurants there in these districts.
- The limitation on the Foursquare API might has its effect on the result of the neighborhood clustering. If more venues can be returned then the feature of each neighborhood will be highlighted and in return the effectiveness of the result will be improved.