Hanami and Ramen

Antonio Cruz | IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

Outline

- Introduction

- Data

- Methodology

- Analysis

- Results and Discussion

- Conclusion

Introduction

What is Hanami?

- Hanami is the traditional custom of viewing and enjoying the beauty of flowers (mainly cherry blossoms) and springtime in Japan.

- To enjoy Hanami, one can admire from a distance in which the cherry blossoms have been described as appearing as beautiful clouds, have a picnic under the blooming trees, or enjoy a nice stroll through a park under the blooming trees.

- Each year forecasts are made for the *cherry blossom front,* the northward advancement of the blooming of cherry blossoms across Japan. These forecasts are closely followed by those who wish to enjoy Hanami as the blossoms typically last for one to two weeks.

Introduction

What is Ramen?

- While the exact origins of ramen noodles is up to debate (Japan or China origin), what is not up for debate is how Ramen shops across Japan have really made Ramen their own.

- Ramen in simple terms are wheat-flour noodles served in broth.

- With more than 10,000 ramen shops in Japan, each shop provides a different take on the dish of Ramen.

Introduction

Goal of this project

- Identify locations in Japan in which to enjoy both Hanami and Ramen:

  - Date range of: April 15th to May 15th

  - List of Recommended Cherry Blossom viewing spots (RCBs)

  - Number of Ramen Shops around each of the RCBs

  - Distance to Train Station from RCBs

Data
What is needed?

	CITY	AVERAGE_OPENING	AVERAGE_FULL_BLOOM
0	Sapporo	May 3	May 7
1	Hakodate	April 30	May 4
2	Hirosaki	April 23	April 28
3	Sendai	April 11	April 16
7	Matsumoto	April 10	April 15
8	Takayama	April 15	April 20

Cities in Japan Average Cherry Blossom Blooming Dates

	VIEWING_SPOT	AVERAGE_OPENING	AVERAGE_FULL_BLOOM	CITY	PREFECTURE	REGION	LATITUDE	LONGITUDE
0	Maruyama Park and Hokkaido Shrine	May 3	May 7	Sapporo	Hokkaido	Hokkaido	43.055745	141.312607
1	Goryokaku Fort	April 30	May 4	Hakodate	Hokkaido	Hokkaido	41.794670	140.754020
2	Hirosaki Castle	April 23	April 28	Hirosaki	Aomori	Tohoku	40.607452	140.464180
3	Mikamine Park	April 11	April 16	Sendai	Miyagi	Tohoku	38.224822	140.857414
4	Matsumoto Castle	April 10	April 15	Matsumoto	Nagano	Chubu	36.238653	137.968867

Recommended Cherry Blossom viewing spots with reverse geocoding information
(columns not shown: Description, Rating, Rating Description)

VENUE	VENUE_LATITUDE	VENUE_LONGITUDE	DISTANCE_FROM_VIEWING_SPOT_METERS	VENUE_CATEGORY
Matsumoto Station Castle Exit	36.230830	137.965035	936	Train Station
Chinese Soba Maruki Shoten	36.231192	137.968443	831	Ramen Restaurant
Matsumoto Dentetsu Matsumoto Station	36.230737	137.963837	990	Train Station
Platform 6 (Home on track 6)	36.231065	137.963931	953	Platform
Nishimatsumoto Station	36.227216	137.963406	1364	Train Station

5 of the 225 venues around the recommended Cherry Blossom viewing spots

- Average Cherry Blossom Blooming dates for desired date range https://www.japan-guide.com/e/e2011_when.html.

- Recommended Cherry Blossom viewing spots (RCBs) https://www.japan-guide.com/e/e2011_where.html
  - Information for each location obtained: City, Name, Description, Rating

- Reverse geocoding to obtain geographical information for each RCBs
  - Information obtained: Prefecture, Region, City, Coordinates

- Utilizing Foursquare to obtain location data for Ramen Shops and Train Stations near the RCBs.
  - Information obtained: Venue Name, Venue Coordinates, Venue distance from RCB (meters), Venue Category
  - All together, a total of 225 venues were found near each of the RCBs (1.25-mile radius).

Methodology

- In this project, we attempted to find locations in which a person can enjoy the beauty of cherry blossoms and a bowl of delicious Ramen within proximity of a train station.

- The data from Foursquare contained excess train station venues around some of the train station locations. This was remedied through use of Hierarchical Clustering/Agglomerative Clustering and reduced the overall total of train station venues.

- Each of the RCBs were plotted along with the venues in the surrounding area to allow for a visualization of the layout of the venues.

- Compared RCBs with each other using bar charts.

- Finally, choices for which location may be best to enjoy Hanami and Ramen were made based on three different criteria:
    - Number of Ramen Shops
    - Distance to Train Station
    - Rating of each RCB

Analysis
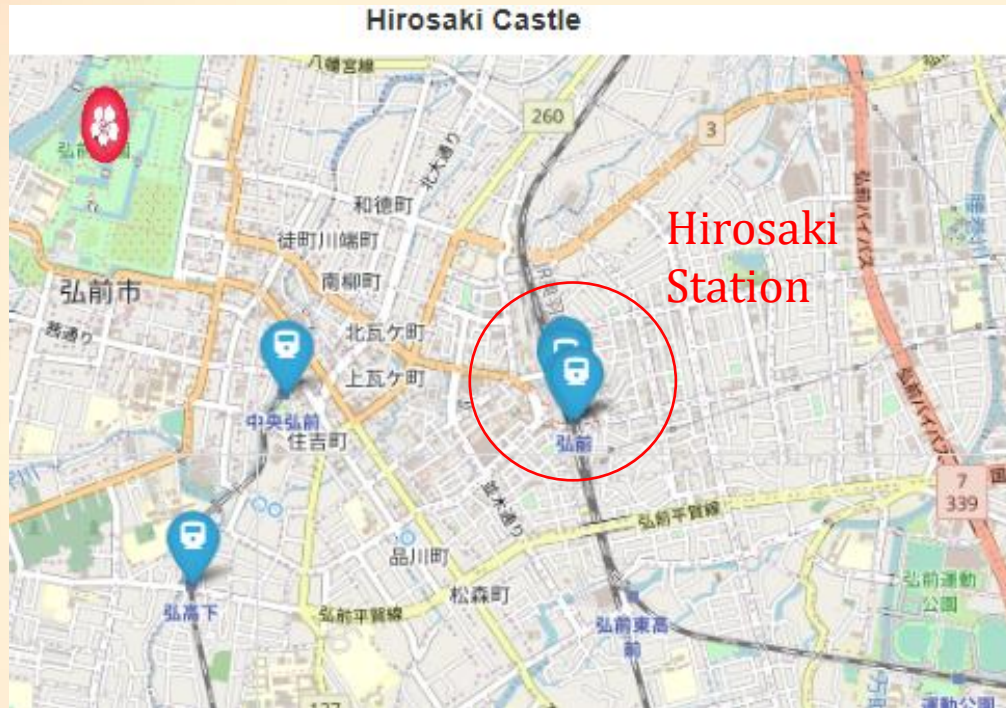
Blue markers with a train icon mark the location of Train Stations.

Sakura marker marks the location of Hirosaki Castle.

Hierarchical Clustering/Agglomerative Clustering

- Hierarchical clustering is a type of unsupervised machine learning algorithm used to cluster unlabeled data points. Hierarchical clustering groups together the data points with similar characteristics. Agglomerative Clustering is a type of hierarchical clustering in which data points are clustered using a bottom-up approach starting with individual data points.

- Excess train station venues around some of the train station locations were returned from the Foursquare API.

- Hirosaki Castle RCB will serve as an example.

- There are 5 train station venues for Hirosaki Station

- The number of train station were reduced by Agglomerative Hierarchical Clustering:

    - Number of clusters determined from dendrograms by use of a threshold line across the 0.002 Euclidean distance, which represent approx. 200 meters.

        - Any venue within 200 meters of another venue are clustered together

    - Cluster labels were used to cluster nearby train station venues into one train station venue.

    - Hirosaki Castle's dendrogram and cluster map are shown in next slide as an example of the process.
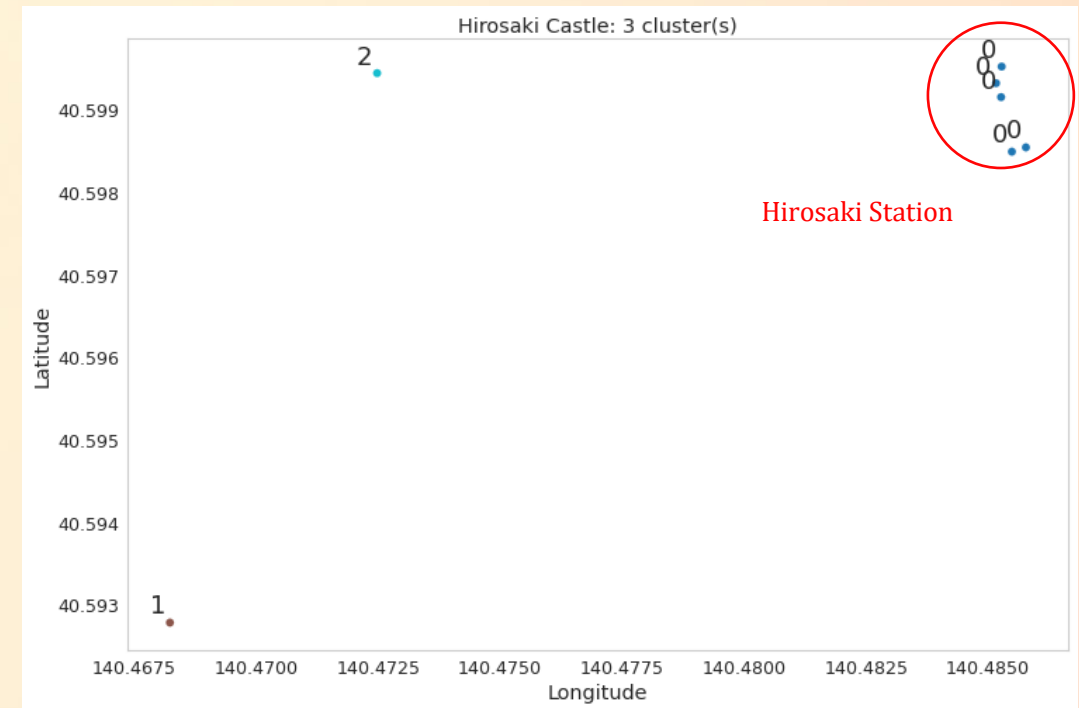
Analysis

Hierarchical Clustering/Agglomerative Clustering

Hirosaki Castle: Train Dendogram

Chūō-Hirosaki Station
Hirokō-shita Station
Hirosaki Station (Hirosaki Station, Konan Railway)
Hirosaki Station Lines 2 and 3 Home Smoking Area
Platforms 2-3 (Home on track 2-3)
Hirosaki Station
Platform 1 (Platform 1 home)

Euclidean distance
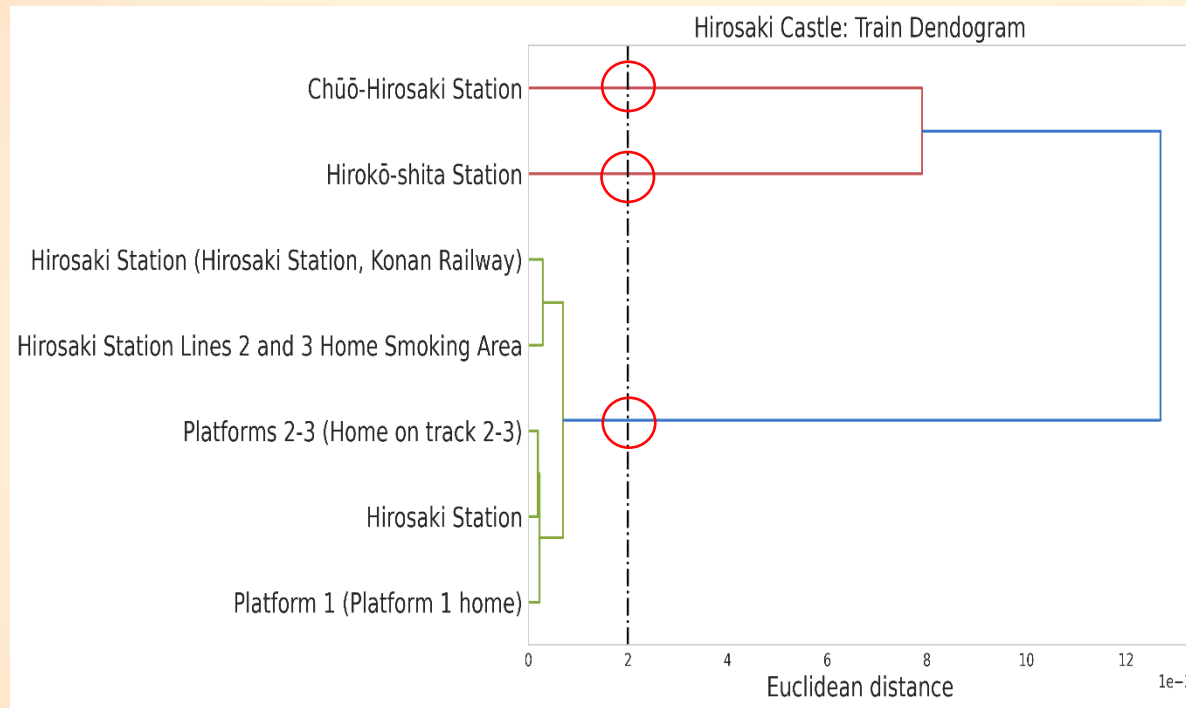
0.002 Euclidean distance threshold line crosses dendrogram at 3 instances, meaning 3 Train Station Clusters

Hirosaki Castle: 3 cluster(s)

Hirosaki Station

Latitude

Longitude

3 clusters: Cluster 0, Cluster 1, Cluster 2

Analysis

Hierarchical Clustering/Agglomerative Clustering
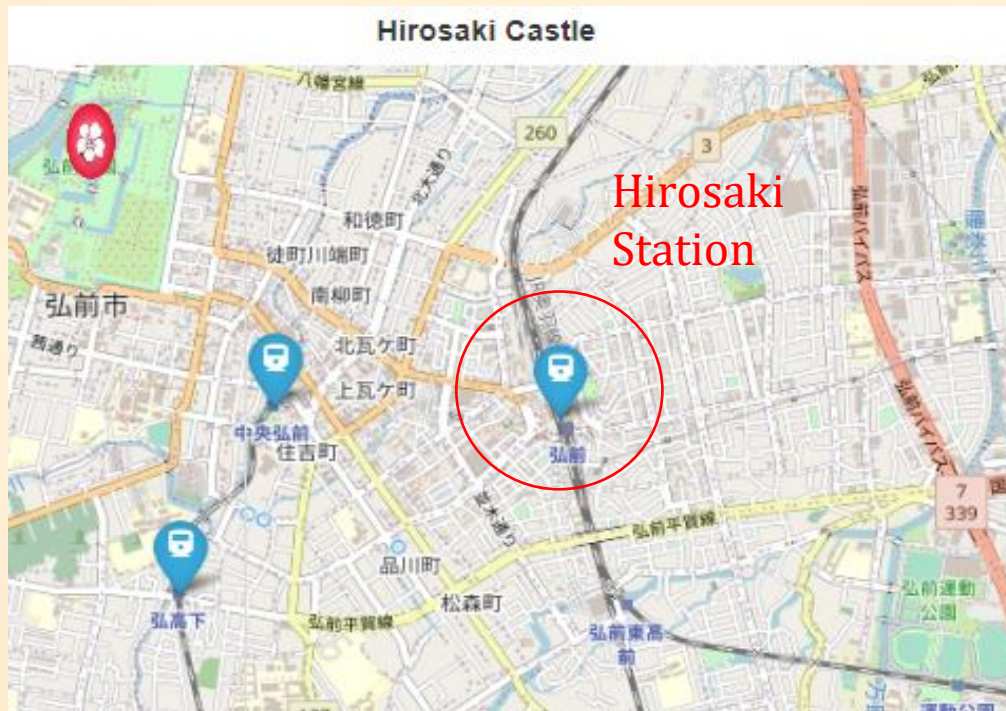
- Hirosaki Station : Originally there were 5 train station venues. With Hierarchical Clustering, these venues were clustered together into 1 train station venue.

- This process was repeated for each of the 5 RCBs.

- This reduced the total amount of train station venues from 42, down to 16 across the areas around the 5 RCBs.

- With the clustering of train station venues completed, the surrounding areas around the RCBs were plotted along with the venues in the surrounding area to allow for a visualization of the layout of the venues.

Analysis

RCBs and surrounding Venues

- Green markers with a cutlery icon mark the location of Ramen Restaurants/Shops

- Blue markers with a train icon mark the location of Train Stations.

- Sakura markers mark the location of each RCB.

- The blue circles each have a radius of 1.25 miles and is centered on the location of their RCB.

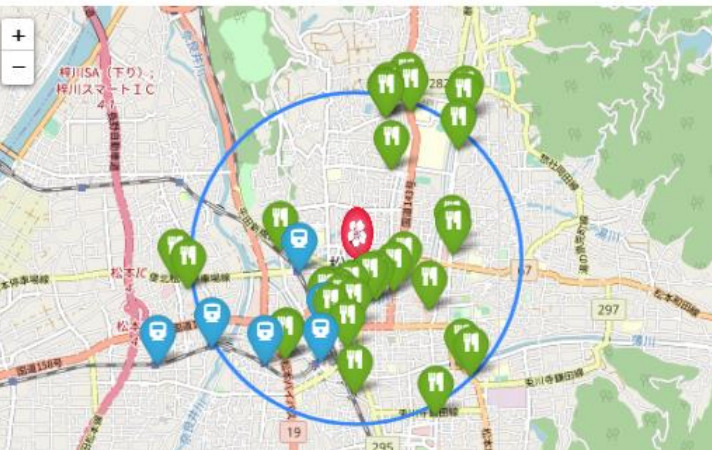- Looking at the maps for each of the RCBs, it seems Foursquare returned venues that where well outside of the desired radius of 1.25 miles.
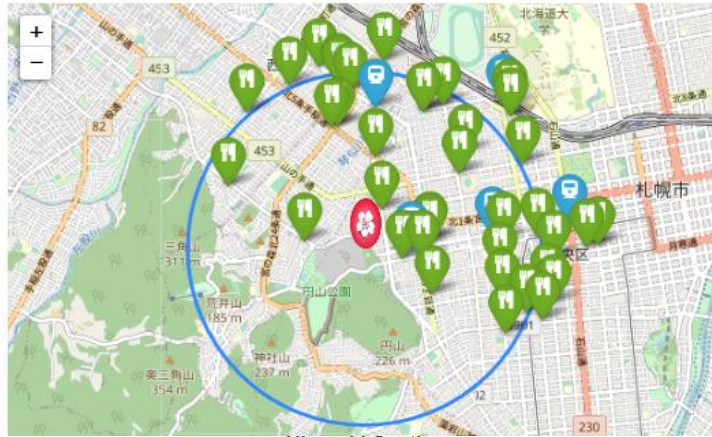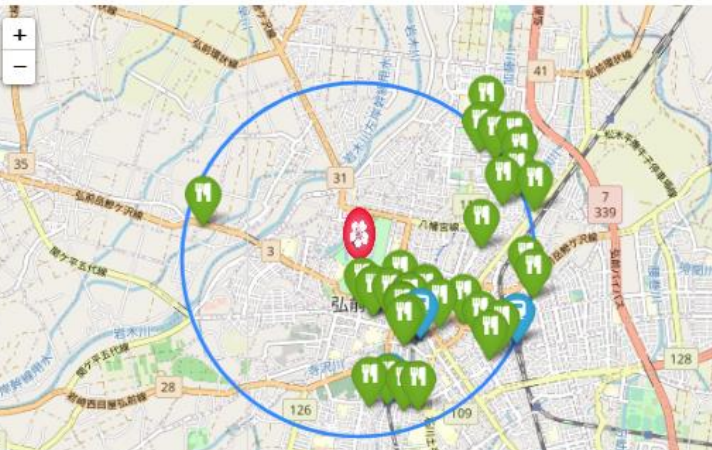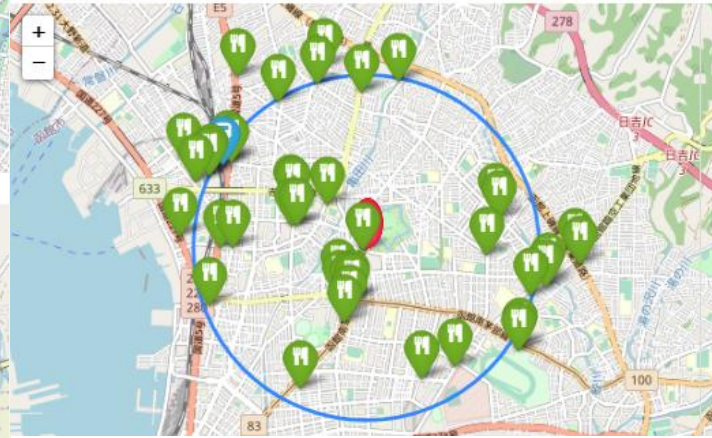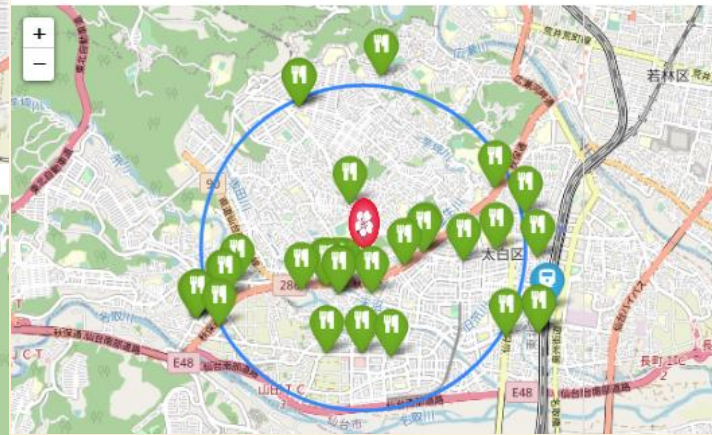
Chart 1

Chart 2

Results and Discussion
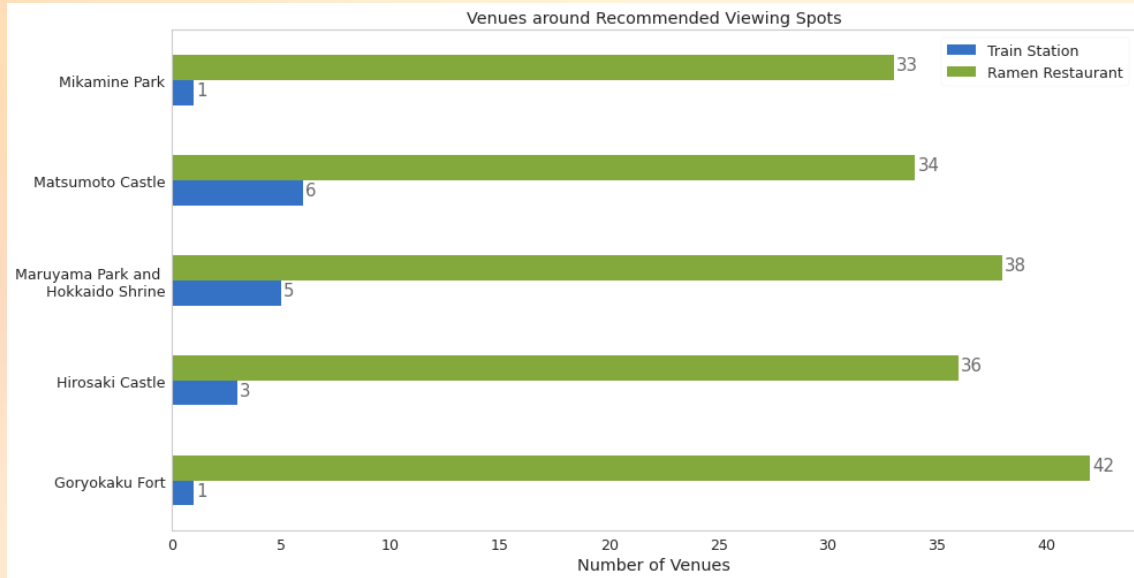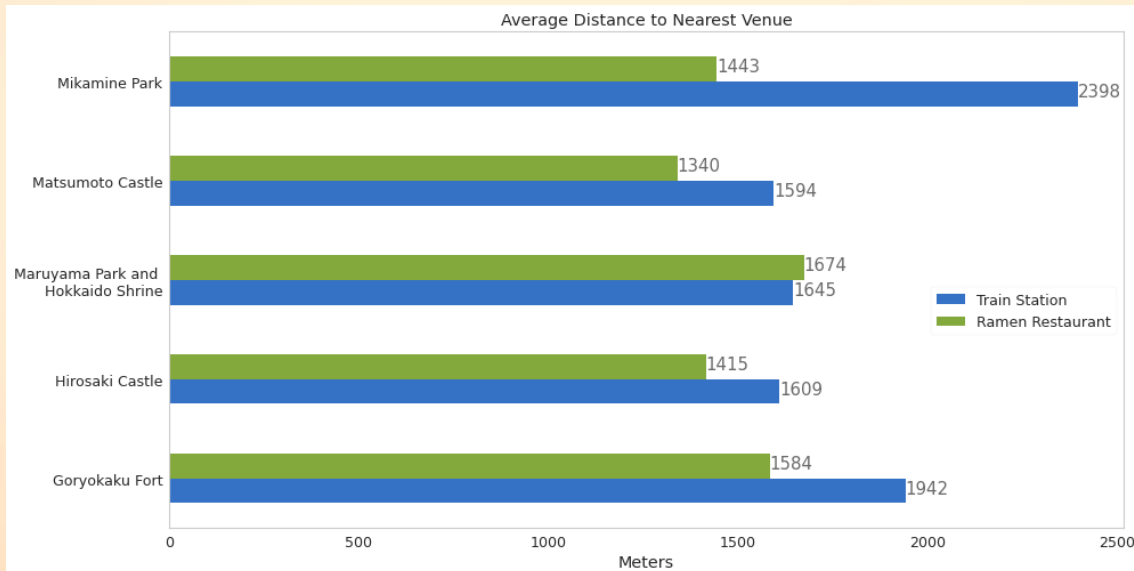
Ramen Shops

- Based off Chart 1 for number of Ramen Shops
  - Goryokaku Fort would be the best with 42 Ramen Shops
  - Mikamine Park would be the worst with 33 Ramen Shops

- Based off Chart 2 for average distance to a Ramen Shop
  - Matsumoto Castle would be the best with an average distance of 1340 meters (~0.83 miles).
  - Maruyama Park and Hokkaido Shrine would be the worst with an average distance of 1674 meters (~1.04 miles)

Train Stations

- Matsumoto Castle
  - 6 Train Stations, all within an average distance of 1594 meters (~0.99 miles).

- Mikamine Park
  - Only 1 Train Station, located 2398 meters (~1.49 miles) away

Results and Discussion

VIEWING_SPOT	AVERAGE_OPENING	AVERAGE_FULL_BLOOM	CITY	LATITUDE	LONGITUDE	RATING	RATING_DESCRIPTION
Hirosaki Castle	April 23	April 28	Hirosaki	40.607452	140.464180	3	Best of Japan
Goryokaku Fort	April 30	May 4	Hakodate	41.794670	140.754020	2	Highly Recommended
Matsumoto Castle	April 10	April 15	Matsumoto	36.238653	137.968867	2	Highly Recommended
Maruyama Park and Hokkaido Shrine	May 3	May 7	Sapporo	43.055745	141.312607	1	Recommended
Mikamine Park	April 11	April 16	Sendai	38.224822	140.857414	1	Recommended

- Another criterion to select the *best* RCB is to use the rating system found at https://www.japan-guide.com/e/e2011_where.html

- The ratings range from 1 to 3:

  1. Recommended

  2. Highly Recommended

  3. Best of Japan

- Based off this system, Hirosaki Castle is the *best* from the RCBs with a **Best of Japan** rating.

Conclusion

- In this project, we identified some of the best locations in Japan in which to enjoy both Hanami and Ramen from April 15th to May 15th.

- These locations would be easy to travel to based off their proximity to train stations.

- Using Hierarchical clustering, multiple train station venues were clustered together to form a single train station.

- We were able to compare the number of venues around each of the RCBs by using bar charts.

- Each of the RCBs have a good number of choices for Ramen Shops located nearby.

- Overall, there are many different locations a person can visit to enjoy springtime and the cherry blossoms in Japan. The biggest factor to keep in mind is the timing of your visit. With cherry blossoms lasting a short time span and the cherry blossom front traveling from south to north, a person needs to keep in mind the time of year when planning a trip to Japan to experience the tradition of Hanami (also to enjoy a bowl of Ramen)