Clustering Philippine Cities

Applied Data Science Capstone
by D. Fenix

Introduction/Business Problem

The year 2020 will be remembered as the year when the COVID-19 pandemic changed the world. Its impact covers all levels and aspects of society. As of December 6, 2020, the World Health Organization reports that over 65 million people have contracted the virus and over 1.5 million have died. Governments continue to impose restrictions on human activity in order to limit the spread of the virus and they are looking for ways to address its economic and societal impact.

On May 6, 2020, Philippine President Rodrigo Duterte issued Executive Order No. 114, instituting the *Balik Probinsya* program. This goal of program is to "develop the quality of life in the rural areas, in effect decongesting the densely populated areas of the country such as Metro Manila by encouraging people to move to the countryside once COVID-19-related quarantine measures COVID-19-related quarantine measures imposed across the country are lifted." (https://en.wikipedia.org/wiki/Balik_Probinsya)

While most migrants to the National Capital Region (NCR) of the Philippines might consider going back to their home cities or provinces as a first choice, a presentation of other options could be invaluable. This is where a clustering exercise of Philippine cities can help. People can identify which cities are potentially good options based on similar areas they are considering. Provincial governments can also use this as a reference for what establishments they can incentivize to make their locations more attractive. Being originally from the Philippines, this project is personally interesting to me. I want to see which areas are similar to Iligan City, where I am from.

Data

The following data will be used in clustering Philippine cities.

- **List of Philippine cities** from https://en.wikipedia.org/wiki/List_of_cities_in_the_Philippines. This data will be read into Python using pandas.
- **Forsquare API** will be used to get the most common venues of each city.
- Google Maps will be used to get the coordinates of each city.

References:

Coronavirus cases and deaths from WHO Coronavirus Disease (COVID-19) Dashboard (https://covid19.who.int/)

List of Philippine cities snapshot from Wikipedia (https://en.wikipedia.org/wiki/List_of_cities_in_the_Philippines)

A · B · C · D · E · F · G · H · I · J · K · L · M · N · O · P · Q · R · S · T · U · V · W · X · Y · Z

City	Population (2015) [4]	Area [5][i]	Density (2015)	Province [5][ii]	Region	Legal class [6]	Charter [iii]	Date of	
								Approval [iv]	Ratification [v]
Alaminos	89,708	164.26 km² (63.42 sq mi)	550/km² (1,400/sq mi)	Pangasinan	I	CC	RA 09025 [7]	March 5, 2001 [7]	March 28, 2001 [8]
Angeles	411,634 [vi]	60.27 km² (23.27 sq mi)	6,200/km² (16,000/sq mi)	*Pampanga*	III	HUC	RA 03700 [9]	June 22, 1963 [9]	January 1, 1964
Antipolo	776,386	306.10 km² (118.19 sq mi)	2,500/km² (6,500/sq mi)	Rizal	IV-A	CC	RA 08508 [10]	February 13, 1998 [10]	April 4, 1998
Bacolod	561,875	162.67 km² (62.81 sq mi)	3,500/km² (9,100/sq mi)	*Negros Occidental*	VI	HUC	CA 326 [11]	June 18, 1938 [12]	October 19, 1938 [13]
Bacoor	600,609	46.17 km² (17.83 sq mi)	13,000/km² (34,000/sq mi)	Cavite	IV-A	CC	RA 10160 [14]	April 10, 2012 [14]	June 23, 2012 [15]
Bago	170,981	401.20 km² (154.90 sq mi)	430/km² (1,100/sq mi)	Negros Occidental	VI	CC	RA 04382 [16]	June 19, 1965 [16]	February 19, 1966 [17]
Baguio	345,366	57.51 km² (22.20 sq mi)	6,000/km² (16,000/sq mi)	*Benguet*	CAR	HUC	Act 1963	September 1, 1909	*September 1, 1909*
Bais	76,291	319.64 km² (123.41 sq mi)	240/km² (620/sq mi)	Negros Oriental	VII	CC	RA 05444 [18]	September 9, 1968 [18]	September 9, 1968