

ANALYSIS OF DISTRICTS OF BANGKOK

ROB EGROT

1. PROBLEM DEFINITION AND BACKGROUND

1.1. Problem Statement. The goal is to support businesses and other organizations in Bangkok by providing district profiles, organizing districts into groups, and analyzing the makeup of venues within each group. We aim to use this analysis to answer two related but distinct questions:

- (a) Given a proposed business venture or community project, for example opening a certain kind of restaurant, or creating a new public space, in which kinds of districts are these venues common/uncommon? Answering this question will allow business owners and community organizers to both target their projects to areas where they are likely to be successful, and also to identify potentially underserved communities.
- (b) Given a district, what kind of venues are common/uncommon? Answering this question will allow potential developers to better understand what opportunities are available in a specific area.

1.2. Background on Bangkok. Thailand is classified as an upper middle income country by the World Bank [1]. It is one of the world's leading exporters of rice [2], and agriculture accounts for 8.6% of GDP [3], and employs 31% of the Thai workforce [4]. As Thailand continues industrializing these percentages decrease, with an increasing share of GDP being taken by service and industry [5]. In particular, Thailand is major regional a center of production in the automotive and electronics industries, particularly for Japanese companies, and also had, before Covid-19, a thriving tourism industry [6].

An estimated 10 million people live in Bangkok, the capital and by far the most populous city in the country [7]. This is roughly 14% of the total population of 69.63 million. Despite rapid growth over the past 30 years, Thailand is one of the most unequal countries in the world in terms of wealth distribution [8]. This inequality is evident at the national level in the divide between urban and rural communities, with some malls in Bangkok consuming more electricity than whole provinces in poorer regions [9], and also within Bangkok itself, as aside from the luxurious malls just mentioned, an estimated 23.7% of the urban population in Thailand lived in slums in 2018 [10]. The majority of these urban slum dwellers live in Bangkok, often in the shadows of the high end malls and apartments frequented by the wealthier citizens.

Bangkok itself is divided into 50 districts, each of which is divided into between 2 and 12 subdistricts. Our analysis will be at the district level, as data for individual subdistricts is difficult to obtain. Bangkok is governed by the Bangkok Metropolitan Administration (BMA), and of course the national government.

2. DATA

2.1. Data summary. To address the problem described in the previous section we use data on the 50 districts of Bangkok. We collected data from multiple sources in 7 broad categories.:

- 1) Basic district data.
- 2) Population data.
- 3) Community data.
- 4) Data on schools.
- 5) Data on new businesses.
- 6) Location of rapid transit stations.
- 7) Venue data.

The data was used in two main stages as follows:

- (i) Data from categories 1) - 6) was used to build profiles of each of the 50 districts of Bangkok, and to arrange these districts into groups using unsupervised clustering techniques.
- (ii) Data from category 7) (venue data) was used to build a venue profile for each of the clusters. These profiles were used to obtain insights relevant to questions (a) and (b) above.

The idea is that categories 1) - 6) provide a broad socioeconomic portrait of the districts, from which it should be possible to distinguish meaningful groupings.

2.2. The details of data collection. The details of data collection are as follows:

- 1) **Basic district data:** This was scraped from Wikipedia [11]. This contains district names in both Thai and Romanized forms, along with geographical coordinates, post codes, population, no. of subdistricts, and map numbers marking locations on an associated map (these map numbers correspond to district codes in official documents, as we shall see later). We do not use the Thai district names, and we replace the population figures with more up to date ones obtained from official data.
- 2) **Population data:** The Bangkok Metropolitan Administration (BMA) makes various datasets available at [12]. We use the data at [13] to obtain more recent district population counts, along with values for the areas and population densities of the districts. Some district areas are missing, so we added values obtained from Wikipedia where necessary. Unfortunately the BMA do not provide Romanized names, so these were added separately. There are several different ways to Romanize Thai words, but it is convenient to use the same method as used by Wikipedia everywhere so we can easily merge dataframes using district names. The BMA data also includes subdistricts, but we dropped these rows from the data table as we do not need this information.

	District	pop_total	area_km2	pop_density	MapNr	Postcode	Subdistricts	Latitude	Longitude
0	Bang Bon	106919	34.745	3077.0	50	10150	4	13.659200	100.399100
1	Bang Kapi	146841	28.523	5148.0	6	10240	2	13.765833	100.647778
2	Bang Khae	193315	44.456	4348.0	40	10160	4	13.696111	100.409444
3	Bang Khen	191323	42.123	4542.0	5	10220	2	13.873889	100.596389
4	Bang Kho Laem	88288	10.921	8084.0	31	10120	3	13.693333	100.502500

Figure 1. Basic population and district data

3) **Community data:** The BMA publishes data on types of community in each district here: [14]. This counts for each district the number of communities in 6 categories. Unfortunately, it is not clear from the column headings provided what the categories are. Moreover, the provided headings seem to be wrong (they look like attempts to Romanize Thai, but e.g. ‘salam’ should probably be ‘slum’, and should be associated with community type 3, not 6, and so on). Correct headings can be worked out by comparing with the data in [15, 2.6] and doing a little translating. A full report on communities in Bangkok is available as [16], but unfortunately this is from a different year so the figures don’t match exactly, and there is no English version available. The correct community types are as follows:

- $C1$ = apartments for government officials and police/military.
- $C2$ = gated community.
- $C3$ = slum.
- $C4$ = dense urban housing (not slum).
- $C5$ = suburbs.
- $C6$ = high rise apartments.

Another complication is that district names are in Thai, but they appear in the same order as those for population data above, so providing Romanizations is straightforward as we can reuse the old method.

	District	C1_gov_aprt	C2_gated	C3_slum	C4_urban	C5_suburb	C6_high_rise
0	Huai Khwang	0	1	18	0	3	1
1	Lak Si	31	6	17	1	14	12
2	Nong Chok	1	1	0	82	13	0
3	Nong Khaem	0	5	2	44	21	2
4	Sai Mai	3	18	4	6	48	0

Figure 2. Community data

4) **Data on schools:** We make a simple count of the number of schools in each district. This is complicated by the fact that in Bangkok there are three different governing bodies for schools. Most schools are registered with the BMA itself, but there are also schools registered with the national Office of the Basic Education Commission (OBEC), and also ‘international’ schools registered with the International Schools Association of Thailand (ISAT). Data on BMA schools was obtained directly from [17], while data on OBEC schools was obtained directly from [18]. For the purpose of this analysis we do not distinguish between e.g. high school and elementary school.

For the BMA schools, the ‘dcodes’ field corresponds to map numbers obtained during step 1), so it was straightforward synthesizing this data with that already obtained. The list of OBEC schools did not provide this, so districts had to be obtained manually. BMA and OBEC schools were combined into a single category ‘government schools’, as we are not aware of a relevant difference between the two categories.

Data on ISAT schools was scraped from [19], and synthesized with previous data using postcodes. When these could not be scraped properly or were incorrect they were added manually. There is a small problem here as sometimes the same postcode is used for more than one district. We chose to count ISAT schools with postcodes used for multiple districts as occurring in *all* districts with that postcode, so several ISAT schools get double counted. This is not ideal, but it seems unlikely to cause big problems.

- 5) **Data on new businesses:** The BMA makes data on new business registrations available. These come in two categories, retail [15, 4.11], and wholesale [15, 4.12]. In each category there is data for the number of new business registrations and total new capital for the years 2017-2019. We record the averages for each district.

No. BMA Schools	No. OBEC Schools		No. Government Schools		No. Government Schools	No. ISAT Schools
9	2.0		11		11	2.0
11	3.0		14		14	6.0
12	3.0	⇒	15	⇒	15	1.0
5	1.0		6		6	0.0
7	0.0		7		7	6.0

Figure 3. Adding school counts

- 6) **Location of rapid transit stations:** There are 3 rapid transit rail systems operating in Bangkok. These are the BTS, the MRT and the Airport Rail Link. We scraped data from the links on the Wikipedia page [20] to count the number of rapid transit stations in each district. Some stations are along the border of two districts. In these cases we counted them for both districts.

District	No. Rapid Transit Stations
0 Bang Khae	2.0
1 Bang Khen	5.0
2 Bang Na	3.0
3 Bang Phlat	4.0
4 Bang Rak	6.0
5 Bang Sue	5.0

Figure 4. Counting rapid transit stations

- 7) **Venue data:** We obtained data on venues in each district using the Foursquare API. We calculated a search radius for each district such that the corresponding circle has the same area as that district. Then for each district center we found the top 50 (this is the maximum allowed by the API) venues within the corresponding search radius in each of the categories ‘food’, ‘drinks’, ‘coffee’, ‘shops’, ‘arts’, ‘outdoors’, ‘sights’. We drop duplicates created by some venues appearing in more than one category. Some venues are still duplicated as they get counted in two different districts due to overlapping circles (since the districts are not actually circular in reality). Some venues may be missed due to being outside the scope of all the approximating circles, and also possibly due to the limit on search results. This is not ideal, but again it should not cause big problems.

REFERENCES

- [1] [Online]. Available: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- [2] [Online]. Available: <https://www.statista.com/statistics/255947/top-rice-exporting-countries-worldwide-2011/>
- [3] [Online]. Available: https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS?locations=TH&year_high_desc=true
- [4] [Online]. Available: <https://www.statista.com/statistics/332341/employment-by-economic-sector-in-thailand/>
- [5] [Online]. Available: <https://www.statista.com/statistics/331893/share-of-economic-sectors-in-the-gdp-in-thailand/#:~:text=>

- ShareofeconomicsectorsintheGDPinThailand2019&text=In2019,theshareof,
sectorcontributedabout58.59percent
- [6] [Online]. Available: <https://aseanup.com/business-thailand/>
 - [7] [Online]. Available: <https://www.cia.gov/the-world-factbook/countries/thailand/#people-and-society>
 - [8] [Online]. Available: <https://worldpopulationreview.com/country-rankings/wealth-inequality-by-country>
 - [9] [Online]. Available: <https://coconuts.co/bangkok/news/bangkoks-luxury-malls-use-more-energy-some-provinces-infographic/>
 - [10] [Online]. Available: <https://data.worldbank.org/indicator/EN.POP.SLUM.UR.ZS?locations=TH>
 - [11] [Online]. Available: https://en.wikipedia.org/wiki/List_of_districts_of_Bangkok
 - [12] [Online]. Available: <http://data.bangkok.go.th/en/dataset/?page=1>
 - [13] [Online]. Available: <http://data.bangkok.go.th/dataset/birthdead/resource/4ea00caf-2e96-40f7-8c2f-b00c36d376b3>
 - [14] [Online]. Available: <http://data.bangkok.go.th/dataset/e57c4e2e-77c8-4994-b9d8-b2d3a67aad82/resource/a41b3558-a441-4bb2-bcde-d51cca91f4c9>
 - [15] [Online]. Available: <http://www.bangkok.go.th/pipd/page/sub/16647/-2562>
 - [16] [Online]. Available: <http://www.bangkok.go.th/upload/user/00000103/KorPorChor/FinalReport.pdf>
 - [17] [Online]. Available: <http://data.bangkok.go.th/en/dataset/bmaschool/resource/3bd7d9f2-61cf-4484-b78c-9573df17fceb>
 - [18] [Online]. Available: <http://data.bangkok.go.th/en/dataset/beckschools/resource/f0dc0de6-5934-48b5-a11d-8ae6b0b241a1>
 - [19] [Online]. Available: <https://www.isat.or.th/member-school>
 - [20] [Online]. Available: https://en.wikipedia.org/wiki/List_of_rapid_transit_stations_in_Bangkok