

Guide for open urban data in Singapore

Our curated inventory of data relevant for geospatial and urban analyses

[Urban Analytics Lab](#)

Last updated on 2020-06-26 · 10 min read



i TL;DR: In the spirit of academia and open science, we're making our notes on open data in Singapore public, and intend to keep them updated. Feel free to visit in future to check for updates as the list grows.

Introduction

In our [research](#) and [teaching](#) activities that are focused on Singapore, we rely almost entirely on open data, enabling reproducibility and fostering open science.

While [Data.gov.sg](#) is thorough and it is the starting and ending point to obtain many useful datasets, it might take time to get an overview and the availability of open data goes beyond that. Furthermore, there are some particularities that may not be evident at first and which we elaborate on in the text (e.g. some datasets are available at multiple locations with slight differences).

This index may be useful to novices to get an overview of what's available in Singapore, but also to seasoned urban scientists who may learn about datasets they might not have been aware of.

The data sources can be grouped into the following categories.

- [Data.gov.sg](#) – the Government's Open Data portal, containing almost 2000 datasets on myriads of topics from dozens of public organisations. Many datasets are regularly updated. There are some GIS datasets too, and also APIs providing real-time data.
- Government resources that are outside the realm of Data.gov.sg, e.g. there may be additional datasets not deposited in the central government repository, some that are slightly different, or those with newer updates. This includes several APIs. For example, [LTA's DataMall](#) and [SingStat](#) have some additional resources or datasets that are arranged in various, potentially more appropriate forms (e.g. detailed time series instead of separate datasets).
- [OpenStreetMap](#) – needless to mention for geospatial data, but surprisingly often overlooked. OSM appears to have a very high level of quality in Singapore and rapid updates.
- Data by research groups, companies, community, ...

This list is by no means a complete inventory of open datasets useful for urban analytics covering the city-state. While there are other datasets, these are the datasets we consider useful for our work, have used in our work, or we bookmarked them to consider using them in future.

The List

Building and housing data

- [HDB Property Information](#) contains data on each public housing block in Singapore (address, number of flats, year of completion, number of storeys, breakdown by flat type, ...). It also includes non-residential blocks such as multi-storey carparks. It does not contain building footprints though. We used this dataset as one of the input datasets to generate [3D building models](#).
- Data on non-HDB buildings (landed houses, condos, commercial buildings...) is not as complete and it is scattered around, but [URA's data portal](#) is a good starting point for exploration.
- For open data on building footprints the best bet is OpenStreetMap, it has nearly 100% completeness with rapid updates, but attribute data may lack. Data.gov.sg contains [a dataset representing building footprints](#), but for some reason it is not complete, covering only a subset of buildings several years ago. It still might be useful though.

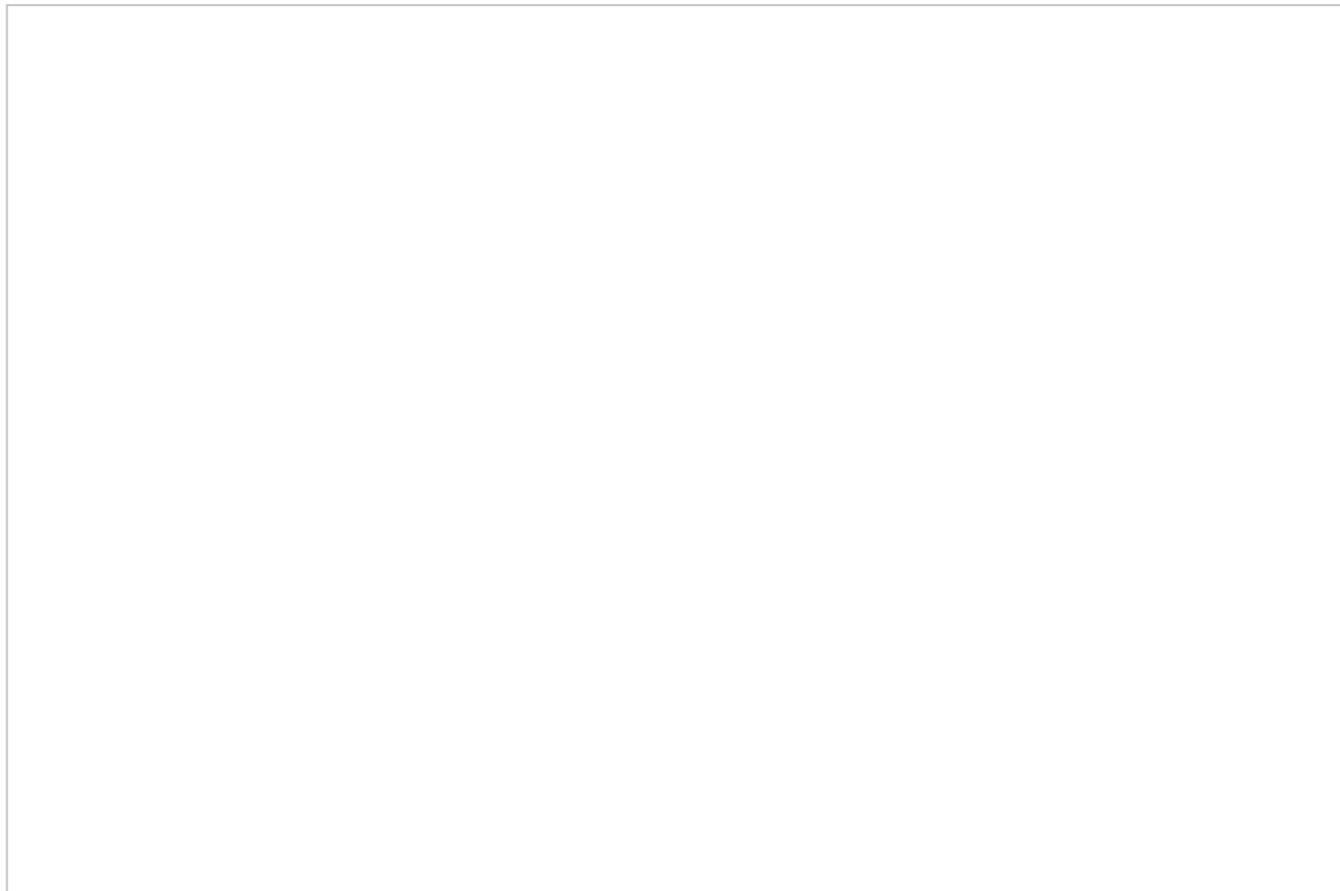


Photo by [贝莉儿 DANIST](#) on [Unsplash](#).

3D city models

Unfortunately, 3D city models are not released as open data, except the one [we generated covering only HDBs](#). We are [working](#) on including other buildings. Worth mentioning is that OpenStreetMap has a relatively high level of completeness of building heights and floors, in comparison to other countries.

Real estate transactions

- There is [a dataset](#) on resale HDB flats transactions, including the address, storey level, price, remaining lease, etc.
- [Median rent by town and flat type \(HDB\)](#), available by quarter since 2005.
- [Data on vacancies](#) at the SingStat's portal.
- Private residential properties transactions (incl. rentals) are [available through URA](#). The same agency also releases [data on commercial properties](#).

Although not open data, it is worth mentioning that NUS staff and students have access to [more detailed data](#) through a subscription.

Demographics

If you need demographic data, you will probably head to Data.gov.sg, where you will find [scores of datasets](#) at different levels (planning area, subzones) and from different years, so it might take time to navigate their landscape. For example, you may find:

- [Households by Monthly Household Income and Household Size](#)
- [Resident Working Persons Aged 15 Years and Over by Planning Area and Gross Monthly Income from Work, 2015](#)
- [Resident Population by Planning Area/Subzone and Type of Dwelling, 2015](#)
- [Resident Population by Single Year of Age, Ethnic Group and Sex, 2015](#)
- [Singapore Residents by Subzone and Type of Dwelling, Jun 2017](#)

Some of them, like the last example, are available in a geospatial format.

However, the best place to get demographic data may be through [SingStat](#), which lists them for a clear overview and has detailed time series datasets, so you don't have to join multiple datasets.

Worth mentioning here is also the SLA's [OneMap API](#) that enables retrieving various demographic data on the planning area level.

Note that most demographic datasets do not include foreigners who are not permanent residents, which represent a sizeable portion of the population.

Energy consumption

To the extent of our knowledge, the most granular dataset available is the Data.gov.sg dataset [Average Monthly Household Electricity Consumption by URA Planning Area & Dwelling Type](#).

Transportation and mobility

There are dozens of datasets in this category, mostly acquired and curated by [LTA](#).

Bus stops, train stations, and routes

The location of bus stops and train stations is available at multiple locations: [OpenStreetMap](#), [LTA DataMall](#), and [Data.gov.sg](#) (note that there are multiple datasets related to this topic, e.g. train stations as [points](#) and [polygons](#), there is even one on [MRT/LRT exits](#)). Furthermore, rail lines are available at [Data.gov.sg](#), but they can also be extracted from [OpenStreetMap](#).

Besides data on bus stops, the [LTA DataMall](#) contains data on bus routes, bus services, and real-time bus arrivals. You may want to check [BusRouter SG](#) (together with its sister project [RailRouter SG](#)) for an awesome web visualisation of this data. Furthermore, there is a [Github repo](#) with the data stored according to the [General Transit Feed Specification](#).

Parking data

Parking data is available in real-time for more than 2000 carparks in Singapore, managed by multiple agencies. One particularity that may go unnoticed is that there are actually two APIs. One is offered at the [LTA DataMall](#) – it returns detailed availability by carpark, and some information about each such as coordinates. The second one, linked on the Developer section at [Data.gov.sg](#) is similar, but it enables querying historical data forgoing some information about the carparks such as location. We used this dataset in our [analysis on mobility during the circuit breaker](#).

You can join the carpark availability data with the dataset [HDB Carpark Information](#) to get a few more columns not returned by the APIs. Note that the location of carparks is simply represented as a point, while the [HDB Map Services](#) shows them as shapes. However, the latter is not available for download.

Origin and destination data, and passenger volume by station/stop

The [LTA DataMall](#) has a few APIs that enable downloading public transport (bus, train) traffic every month. For example, it contains the number of passengers that have travelled between two stations, with a breakdown by type of day (weekday/weekend) and hour. Data is available for the past three months. Do note that the entire trip is not available; it's limited to the transportation mode. For example, if a traveller takes a bus to an MRT station and continues the journey with a train, these are considered as separate trips and cannot be connected in the data.



Photo by [Euan Cameron](#) on [Unsplash](#).

Travel times on roads

Another API available thanks to the [LTA DataMall](#) returns the estimated travel times of expressways. It might be useful for studying the volume of traffic. It doesn't look that it enables querying historical data, though.

Routing

Routing (fetching the distance, estimated travel time, and the geometry of the route) between two points is available through the [OneMap API](#). OpenStreetMap is also useful here, e.g. check out the [Open Source Routing Machine](#) and [Openrouteservice](#). There are interfaces for Python and R, e.g. we used [osrm](#) in [teaching](#).

Although not strictly open, rather commercial (but they offer a free tier), here it is inescapable to mention the trio of APIs under the [Google Maps Platform](#): Directions API, Distance Matrix API, and Roads API, which are of high quality and a lot can be done within the free monthly quota they offer.

Taxi availability

The availability of taxis is also available on the [LTA DataMall](#). The API returns the location of each taxi that is currently available. The data does not include hired/busy taxis. Check out the [TaxiRouter SG](#), which visualises this data in real-time, together with the [taxi stands](#).

Traffic images

Traffic images are available through the [LTA DataMall](#).

Transportation mode

[Data.gov.sg](#) contains several datasets on the usual mode of transport used by residents according to surveys.

Mobility trends

[Apple Mobility Trends Reports](#), [Google Community Mobility Reports](#), and [CityMapper Mobility Index](#) all include Singapore.

Assorted

Both the [LTA DataMall](#) and [SingStat](#) have more datasets worth having a look at, e.g. number of cars in SG at a fine temporal scale (updated monthly).

Map / Geospatial data (general)

Besides OpenStreetMap which is quite complete and of high quality for a wide range of features, well worth mentioning is the Geospatial Whole Island dataset available through the [LTA DataMall](#). It contains a bunch of different features related to transportation, e.g. road crossings, traffic lights, taxi stands, and cycling paths.

Further, [Data.gov.sg](#) contains some datasets such as the [boundaries of administrative areas](#), [master plan land use](#) (containing the [Gross Plot Ratio](#)), and [cadastral land parcels](#). The [series of datasets by NParks](#) hosted on Data.gov.sg deserves special attention: it covers a wide range of [park](#)-related features under their purview, e.g. [boundaries of activity areas](#), [locations of play/fitness equipment](#), [bbq pits](#), [the shape of the park connector loop](#), and [carpark lots](#) (however, do note that the NParks' carparks do not appear to be covered by the LTA's API mentioned above).

For trees, check out [ExploreTrees.SG](#), derived from [Trees.SG](#).

Finally, you may be interested in the [high-resolution map of Singapore's terrestrial ecosystems](#) that was developed by the research team of the [Natural Capital Singapore](#) and released as open data. There is also a [paper](#) published.

Aerial imagery

There are no open data high-resolution resources we are aware of. Satellite imagery is available for academia through the [Planet's Education and Research Programme](#), which we are a member of and which is accessible to other academics as well.

Point clouds (LiDAR), terrain data

None, except terrain data of coarse resolution such as [SRTM](#).

Street-level imagery

Google Street View has pretty good coverage of Singapore (it even includes [hawker centres](#)), and the data is downloadable through their [API](#) (check the T&C though). [Mapillary](#) is also worth considering.

Airbnb

[Inside Airbnb](#) has Airbnb data on Singapore, updated monthly. It includes listings and their reviews.

Other

There are some datasets, which albeit we have not used much so far, are worth mentioning and keeping in mind. The honourable mentions are:

- Weather data, which is available through [Data.gov.sg](#): real-time weather readings, Pollutant Standards Index (PSI), Ultra-violet Index (UVI), etc.
- [Eating establishments by NEA](#) – quite comprehensive dataset on all places allowed to sell food in Singapore. You may also be interested in [data on hawker centres](#) and [supermarkets](#).
- [Skyrise Greenery dataset](#) shows the indicative location of rooftop and vertical greenery.

Notes and considerations

Tabular data / geocoding

Many of the datasets above are CSVs, and even though they are tied to a location, they are not available in a geo-format (e.g. real estate transaction datasets contain an address for each transaction, but not the coordinates). To convert the address into coordinates, use the [OneMap API](#), [Nominatim](#), or [Google Maps API](#).

Web services

There are a few web services containing various interesting datasets (e.g. [OneMap](#), [HDB Map Services](#), [URA SPACE](#), [Trees.sg](#)), but not all of them can be downloaded, so they are not considered as [open data](#). Nevertheless, they may still be useful for viewing.

Social media

[Twitter API](#) enables downloading Twitter data for Singapore, but the service is not very popular here, and the data comes with restrictions (so it is technically not open data).

Licence, validity and quality of data

The usual caveats:

- Check when the dataset has been updated last time.
- Check the licence, e.g. for Data.gov.sg have a look at the [Singapore Open Data Licence](#).
- Do not forget to attribute the data source in your use.
- Some geospatial datasets may not pass all validity checks (e.g. they might have self-intersecting polygons), presenting a problem when they are used in analyses. You can try fixing them using [prepair](#).

Have a suggestion for an entry? Spotted an error?

[Get in touch](#).

[open data](#) [singapore](#)



Urban Analytics Lab

