City Planning Facilities Database (FacDB)

Announcements

The code used to create the City Planning Facilities Database (FacDB) has been completely rewritten to make it more user-friendly and more maintainable. The schema has changed, as have some of the data sources. These changes are summarized below.

FacDB is compiled from approximately 50 source files. Many facilities are available in more than one source file, sometimes with conflicting information. The previous version of FacDB had several fields, including CAPACITY that were formatted as strings, showing each source file that included that facility and the value contained in the source file.

In version 19v1, one source is designated as the primary source for each facility type. All values in the record come from that single source. For some facility types, like public schools or fire stations, there is an authoritative source and no secondary sources are used. For example, the Department of Education is the only source used for public schools. While other source files, such as City Owned and Leased Properties (COLP), contain records for public schools, these are not used for FacDB.

For other facility types, like day care centers, there is not a single authoritative source. One source is designated primary. All the facilities from that source are included in FacDB. If other sources include facilities at locations not in the primary source file, additional records are added to FacDB with data from the secondary source(s).

Schema changes: The schema has been significantly changed for version 19v1. While field names remain the same, some fields have been removed. Some fields, like CAPACITY, which were previously text fields, have been changed to numeric fields. Please review the FacDB Data Dictionary for the new schema.

Data source changes: Some city agency facilities, previously sourced from COLP, have been changed to use a source file from the agency. These agency specific files are updated more frequently than COLP and provide more accurate information. Please review the data sources tab in the FacDB Data Dictionary for a full list of these sources and the date each was updated.

Health and human services facilities formerly taken from HHS Connect are now coming from the Mayor's Office of Economic Opportunity. This has resulted in a number of changes to facility categorization.

Data quality improvements: The geocoding of facilities has been improved and there are fewer duplicate facilities.

About the data

The Department of City Planning aggregates information about 33,000+ facilities and program sites that are owned, operated, funded, licensed, or certified by a City, State, or Federal agency in the City of New York into a central database called the City Planning Facilities Database FacDB). These facilities generally help to shape quality of life in the city's neighborhoods, and this dataset is the basis for a series of planning activities. This public data resource allows all New Yorkers to understand the breadth of government resources in their neighborhoods.

While each source agency classifies its facilities according to their own naming systems, we have grouped all facilities and program sites into the following seven categories to help planners navigate the data more easily:

- Health and Human Services
- Education, Child Welfare, and Youth
- Parks, Gardens, and Historical Sites
- Libraries and Cultural Programs
- Public Safety, Emergency Services, and Administration of Justice
- Core Infrastructure and Transportation
- Administration of Government

Within each of these domains, each record is further categorized into a set of facility groups, subgroups, and types that are intended to make the data easy to navigate and more useful for specific planning purposes. Facility types and names appear as they do in source datasets, wherever possible. A full listing of the facility categories is provided here.

If you have any questions or comments about these data please contact DCPOpenData@planning.nyc.gov.

Common uses

Fair Share Analysis, Neighborhood studies, Facilities planning

Watch-outs

Analysis Limitations. As a result of the data limitations and inconsistencies listed below users should be careful in their use of this database so as to avoid developing suspect analyses. For example, a comparison of the density or accessibility of facilities across neighborhoods should recognize that some of the facilities included are organizational headquarters rather than service sites and that this database is not authoritatively comprehensive. In addition, we rely on source data from other agencies to populate the database, and some of these sources may fall out-of-date. Users can find the date of each source dataset's latest update in the source data dictionary.

Missing Records. Currently, FacDB is the most comprehensive spatial data resource available for facilities run by public and non-public entities in NYC, but it does not claim to capture every facility within the specified domains. Some facilities are deliberately excluded from the data that source agencies provide in order to protect the safety and privacy of their clients. Also, many records could not be geocoded. To learn more about how the data are processed, please review the Data Sources and Compilation Process.

Duplicates. Please be aware that this dataset may include cases of duplicate records for the same facility because several source datasets have content that overlap.

Administrative Addresses. There are known to be cases when the address provided in the source data is for a headquarters office rather than the facility site location. Unfortunately, these could not be systematically verified. For more detailed information on a specific facility reach out to the respective oversight agency.

Public Accessibility of Sites. DCP is unable to verify the public accessibility of all sites. For example, some playgrounds or playing fields may only be accessible to participants in certain programs.

Data Sources and Compilation Process
Since the facility records are aggregated from many datasets designed for different purposes, the data will be transformed over several stages to reach its final state. The stages are described below and all the scripts used are available on the NYC Planning GitHub page.

Data loading. Since the source datasets have been maintained by various agencies and updated with different frequencies, datasets are loaded into Amazon s3 as a centralized datahub preparing for the downstream data processing.

Geoprocessing. When records have address information, spatial data is assigned by taking the centroid of the BIN returned by Geosupport that matches the DoITT building footprints dataset. If a BIN is not available, the latitude and longitude returned by Geosupport is used to create the geometry for the record. If these fields are not available from Geosupport, but the source data has spatial information (i.e. coordinates) the spatial data is created from the source data. If the source data consisted of polygon geometries, the centroid of the polygon was used to assign the geometry for the records in the database. There are cases where the coordinates from the source data fall in the roadbed and not inside a BBL boundary due to the geocoding technique used by the source. Lastly, if a geometry could not be assigned from the BIN, latitude/longitude from Geosupport, or source data, the centroid of the BBL from the clipped MapPLUTO is used. Other geographic information such as the community district is taken from Geosupport if a value is returned, otherwise administrative districts are assigned via spatial joins where the record has a geometry.

Duplicate Record Removal. Several of the source datasets have content that overlaps. Duplicate records were identified by querying for all the records that fall within the same BIN or BBL and have the same Facility Subgroup or Type, same Facility Name, or same Oversight Agency. Where duplicate records were identified all but the primary record was removed from the database.

General Constraints Use Limitations
The facilities database is being provided by the Department of City Planning (DCP) for informational purposes only. DCP does not warrant the completeness, accuracy, content, or fitness for any particular purpose or use of the dataset, nor are any such warranties to be implied or inferred with respect to the dataset as furnished on the website

Legal Constraints Use Limitations

DCP and the City are not liable for any deficiencies in the completeness, accuracy, content, or fitness for any particular purpose or use of the dataset, or applications utilizing Dataset, provided by any third party. The City Planning Facilities Database (FacDB) is only as good as the source data it aggregates, and the Department of City Planning cannot verify the accuracy of all records. Please read more about specific data and analysis limitations before using this data.