

City Planning Facilities Database

Changes between versions 22v1 (Jun 22) and 23v1 (May 23)

1. Subgroup changed – CHEMICAL DEPENDENCY is now SUBSTANCE ABUSE DISORDER TREATMENT PROGRAMS

Changes between versions 21v2 (Aug 21) and 22v1 (Jun 22)

1. Added two new fields: NTA2020 and CT2020
2. Renamed NTA to NTA2010
3. Renamed Census Tract to CT2010
4. Changed the source data for DSNY textile drop-off locations to the DonateNYC Directory, per DSNY's guidance.

Changes between versions 21v1 and 21v2

1. Removed PROPTYPE field. This field was meant to indicate whether a property was city owned, owned by another public entity, or privately owned. Because this information was only available for a subset of the data, the field has been removed.
2. Improved categorization for records sourced from the Mayor's Office of Economic Opportunity (MOEO) for social service site locations and from the New York City Housing Authority for community facilities.
3. MOEO social service site locations are now sourced from NYC Open Data.
4. Removed duplicate records received from a source when the facility name, location, and facility type are the same. Improved the deduplication of records from different sources.
5. Removed records sourced from City Owned and Leased Properties for lots that are not currently in use.
6. Added data from the Department of Sanitation on recycling drop-off sites.

I. Overview

The New York City Department of City Planning (DCP) aggregates information on over 30,000 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.

FacDB is created with data from approximately 50 input files. These files come from city, state and federal agencies, as well as from non-profit organizations. Each source agency classifies its facilities according to its own conventions. DCP has grouped all facilities and program sites into the following seven categories to help users 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 seven domain categories, 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.

If you have any questions or comments about these data please contact the NYC Planning Open Data team at DCPOpendata@planning.nyc.gov

General information

Dataset Name	Facilities Database (FacDB)
Agency Name	Department of City Planning
Update Frequency	Semiannual
Dataset Description	Facilities and program sites that are owned, operated, funded, licensed or certified by a City, State, or Federal agency
Dataset Keywords	Facilities, Education, Child Welfare, Parks, Gardens, Historical Sites, Libraries, Cultural Programs, Public Safety, Emergency Services, Administration of Justice, Health Services, Human Services, Infrastructure, Transportation, Government Administration
Dataset Category	City Government
Additional Information	The Department of City Planning aggregates information about more than 30,000 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.

Each record in FacDB represents a facility site.

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 in order to protect the safety

and privacy of a facility's clients. Also, not all source records could be geocoded; therefore, they cannot be mapped.

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, please reach out to the respective oversight agency.

II. Common uses

Fair Share Analysis, Neighborhood studies, Facilities planning

III. Watch-outs

Analysis Limitations. As a result of the data limitations and inconsistencies listed below users should be careful in their use of this dataset 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 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 source data in order to protect the safety and privacy of a facility's clients. Also, not all records could be geocoded.

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.

IV. Data Sources and Compilation Process

Since the records in FacDB are aggregated from many datasets designed for different purposes, the data is transformed over several stages to reach its final state. The stages are described below and all the scripts used are available in the NYC Planning GitHub organization site.

Data loading. Since input datasets are sourced from various open data sites and agencies, first datasets are loaded into a centralized datahub preparing for the downstream data processing.

Geoprocessing. DCP works to map the point geometry of the record to the most precise location possible based on the input data given. Ideally the point geometry is the centroid of the building footprint from the DoITT building footprints dataset. Secondary, the record is mapped to the centroid of the tax lot. Lastly, if the record cannot be mapped using an available address, BIN, or BBL the geometry will come from the source dataset if available.

Normalization and categorization. DCP maps data from input data fields to the FacDB data schema, where values are cleaned, normalized, and standardized so that they are consistent wherever possible across the entire dataset. Furthermore, DCP categorizes each record to the best of its ability based on the information in each record into seven domains, which are further broken down into groups and subgroups. When possible, the facility type is sourced from the input dataset.

Duplicate Record Removal. Several of the source datasets have content that overlaps. Duplicate records are primarily identified by querying for records with the same geometry, facility type, facility name, operating agency, and data source. Furthermore, records are deduplicated between data sources where DCP has identified overlapping content, for example between child care centers that appear in both the NYCHA community centers dataset and the DOHMH daycare dataset. Where duplicate records are identified all but one record is removed from the database.

V. Data Dictionary

The following table lists and defines each of the fields presented in the Facilities Database

Column Name	Column Description
FACNAME	Facility name
ADDRESSNUM	Address number for the facility's location
STREETNAME	Street name
ADDRESS	Address number and street name for the facility's location
CITY	City name
ZIPCODE	ZIP code
FACTYPE	Facility type
FACSUBGRP	Facility subgroup
FACGROUP	Facility group
FACDOMAIN	Facility domain
SERVAREA	Service area, which may be local or regional
OPNAME	Name of the operating entity
OPABBREV	Abbreviation for the operating entity
OPTYPE	Indicates whether the operating entity is public or non-public
OVERAGENCY	The name of the agency overseeing the facility
OVERABBREV	Abbreviation for the oversight agency
OVERLEVEL	The level of government of the oversight agency
CAPACITY	The capacity of the facility. CAPTYPE contains the units in which capacity is measured.
CAPTYPE	The unit type for capacity, such as beds, seats, etc.
BORO	Borough name
BIN	Building identification number

BBL	Borough, block, and lot number
LATITUDE	Latitude
LONGITUDE	Longitude
XCOORD	X coordinate in NAD 1983 State Plane New York Long Island FIPS 3104 (US Feet)
YCOORD	Y coordinate in NAD 1983 State Plane New York Long Island FIPS 3104 (US Feet)
CD	Community district
NTA2010	Code for the 2010 neighborhood tabulation area
NTA2020	Code for the 2020 neighborhood tabulation area
COUNCIL	City council district
CT2010	2010 census tract
CT2020	2020 census tract
BOROCODE	Borough code
SCHOOLDIST	School district
POLICEPRCT	Police precinct
DATASOURCE	Date source file name
UID	Universal Unique Identifier. When a row is added to the table the uid is automatically generated, enabling database replication.

VI. 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.

VII. 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.