GeoPostcodes

The most comprehensive
postal database in the world

Street database product sheet

Version 2022.02.28

Table of contents

Introduction

GeoPostcodes is a Belgium-based company specializing in postal data sets.

With over 10 years experience in the collection, treatment and enrichment of postal information at worldwide scale, we pride ourselves in offering the most accurate and stable data on the market.

Our products include datasets about administrative regions, ZIP/postal codes, localities, streets and postal boundaries. Each product is the result of a careful data analysis, curation and data quality improvement process combining a multitude of sources using proprietary GeoPostcodes data processing and GIS tools.

Our files are fully geocoded and provided in a plain and simple, consistent format, easily importable in any DBMS or geographic information systems (GIS).

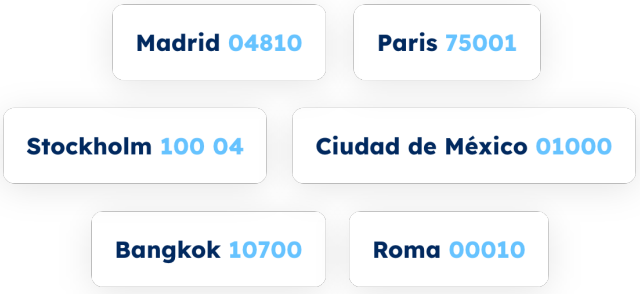
We aim to help our customers achieve their goals by providing them with the most accurate and up-to-date data about the world.

GeoPostcodes currently develops and maintains 4 products:

- a postal database;
- a streets database;
- a postal boundaries database;
- an administrative boundaries database.

Postal database

Our postal database includes **all relevant postal information at the town, village, neighbourhood and suburb level**. For every country in the world, it contains the administrative divisions, the localities and their related ZIP/postal codes, as well as geocodes to link the data with other data sources.



Typical use-cases leveraging the postal database are **ERP integration for services and logistics (including shipping distance estimation) and master data management**. It can also be used to validate forms, create regional reports, etc.

Streets database

Our streets database includes the postal database and extends it with street information. It contains all administrative divisions, localities, streets and related ZIP/postal codes of a country, as well as geocodes to link the data with other data sources.

The streets database is relevant for the same use-cases as the postal database (ERP integration, master data management, address validation), with an extra layer of streets and, consequently, more precise information (streets, house numbers) and coordinates.

Postal Boundaries database

The Postal Boundaries database provides the areas covered by postal codes. It is delivered as vector postal boundaries, coming from a topological model which ensures adjacent polygons perfectly match together (boundaries are shared).

The boundaries database is the product you need if you want to display your own data on a map (e.g., sales figures per postal code) or assign events to their postal code from their coordinates (e.g., geofencing).

Streets database

Our streets database includes the postal database and extends it with street information. It contains all administrative divisions, localities, streets and related ZIP/postal codes (when existing) of a country.

The streets are delivered with their postal code(s). When a street is covered by multiple postal codes, our data contains all of them along with the range of house numbers related to each postal code.

For relevant countries, business and administrative codes are provided as well. These are postal codes directly associated with an administration (e.g., a parliament), a private institution or military (e.g., an airbase or a regiment).

We map the administrative divisions of each country into a uniform structure, using up to four administrative levels. Our data format is hence identical for all countries of the world, making it easy for you to process postal data from different countries.

Additional information such as neighborhoods, suburbs, time zones, elevations, geocodes (ISO, FIPS, HASC, NUTS, and codes from National Statistics offices) is also included. Please check our data design to get a list of all available fields. All data are georeferenced using the WGS84 datum and can be easily imported into any software, database or GIS system.

The data is permanently refreshed: on average 3 updates are delivered every week to always stay synchronized with the local changes. We leverage over 1500 data sources to maintain our database and have partnerships with several postal services.

Languages

The data is provided in the local language (official language(s) of the country). When such a language is not based on a latin-alphabet, an English version is also included, using translations where available and transliterations otherwise. Data is available in up to 3 languages for each country. For instance, Canada is delivered both in English and French, while Japan is delivered in Japanese (non-latin), Katakana (non-latin), and English.

A full version of the database in English is available, even for the countries using a non-English latin-alphabet (e.g., Brazil, France, Germany, Italy, Spain).

Data Design

Our street database is available to all clients in the denormalized format presented in this section. The data will also be accessible in a normalized format soon (i.e., related tables designed to reduce redundancy and favor data integrity). Both formats contain the same data, structured differently. The denormalized format is convenient for reading the data and validating external information against it; the normalized format is preferred for Master Data Management or data enrichment.

Detailed information on the normalized format is available on demand (see contact section).

Tables and Fields description - Denormalized format

In its denormalized version, the street database contains 4 tables: Places, Streets, Regions and Businesses:

- **The Places table** is designed to contain the ZIP/postal codes used for both regular mail delivery (standard house address) and businesses (see Business table). This table is geocoded.

- **The Streets table** extends the Places table with street information. It is geocoded as well. Information is detailed at the street and postal code level: when a street is covered by several postal codes, the table includes one row for each postal code, with the corresponding range of house numbers.
- **The Regions table** defines the administrative divisions hierarchy of a country. An administrative division is a country subdivision with an official status. On top of the country level, we define up to 4 administrative divisions levels per country.
- **The Business table** stores specific postcodes which are not used for regular mail delivery (standard house address). It can be postcodes for governmental organizations, military operations centers, PO boxes, special codes for large companies, etc. This table is not geocoded as these postcodes are not related to a geographical location.

Each table is detailed below, and the indexes are underlined.

Places table

Field name	Field type	Description	Comments
ISO	Char(2)	ISO 3166-1 Country code	The ISO 3166-1 standard is published by the International Organization for Standardization (ISO) and defines a unique code for the name of each country. The country codes are represented as a two-letter code (alpha-2).
Country	Char(50)	Country name	
Language	Char(2)	Language code	When there are several official languages in the country, or for countries using a non-Latin alphabet, we provide a specific version for each record. The language field contains the language code and allows to select a specific version.
ID	Integer	Record identifier	
Region1	Char(80)	Administrative division level 1	These fields contain the name of administrative divisions for each available level. If not applicable, the value '-' will be indicated.
Region2	Char(80)	Administrative division level 2	

Region3	Char(80)	Administrative division level 3	
Region4	Char(80)	Administrative division level 4	
Locality	Char(80)	Locality name	Contains the settlement name. The type of settlement may vary depending on the country (city, town, village, etc..).
Postcode	Char(15)	ZIP / Postal code	In countries/places where there is no postal code in use, this field is left empty. The postcode structure varies for each country.
Suburb	Char(80)	Locality subdivision	A suburb is a residential area or a mixed use area, either existing as part of a city or urban area or as a separate residential community within commuting distance of a city. The suburb is indicated if it is determinant for the postcode.
Latitude	Double	Place latitude (WGS84 coordinates)	These fields contain the geographic coordinates related to the specific locality and postal code of the record. The geographical position is usually precise to about one hundred meters and corresponds to the center of the locality (generally the historical center, and if not, the geographical center). In some countries, the postal code may refer to a block or a street, in which case the coordinates point to that specific location. In case the georeference is not
Longitude	Double	Place longitude (WGS84 coordinates)	

			known precisely, a central point will be defined. Depending on the information available, the point can be defined as a central point of the locality, or as a central point of a higher administrative division. All our files are created using the World Geodetic System standard WGS 84. The corresponding SRID reference is EPSG:4326.
Elevation	Integer	Elevation in meters	
ISO2	Char(10)	ISO 3166-2 Region code	The ISO 3166-2 standard is published by the International Organization for Standardization (ISO) and defines codes for the names of the principal subdivisions of all countries. The codes for subdivisions are represented as the alpha-2 code for the country, followed by up to three characters.
FIPS	Char(10)	NGA Geopolitical code (formerly FIPS PUB 10-4)	The NGA Geopolitical codes (Formerly known as FIPS PUB 10-4) is a standard published by the U.S. National Geospatial Intelligence Agency and contains countries, dependencies, areas of special sovereignty and their principal administrative divisions.
NUTS	Char(12)	European statistical division code	The Nomenclature of Territorial Units for Statistics (NUTS) database is a standard provided by the European Commission for statistical purposes.
HASC	Char(12)	Hierarchical Administrativ	The Hierarchical administrative subdivision codes (HASC) are codes to represent names

		e Subdivision Code	of country primary subdivisions, such as states, provinces, regions. The codes are alphabetic and have constant length for the first level subdivisions.
STAT	Char(20)	National statistics/census code	The statistical codes are codes defined by each National Statistical Offices.
Timezone	Char(30)	Time zone name (Olson)	Following the timezones defined in the IANA tz database (sometimes called Olson database).
UTC	Char(10)	Standard time expressed as an offset with respect to UTC	The Coordinated Universal Time (UTC) is the time standard commonly used across the world to keep time synchronized.
DST	Char(10)	Daylight saving time expressed as an offset with respect to UTC	The Daylight Saving Time (DST) is the practice of setting the clocks forward one hour from standard time during the summer months, and back again in the fall, in order to make better use of natural daylight.
locality_type	Char(20)	Type of Place	'town' for regular places, 'business' where the place is a postal construction to accommodate industrial areas or businesses

			which are not attached to their own region but to a major city in a nearby region
is_postal	Integer	Whether the postal code is used for regular mail	1 when the postal code is used for regular mail (street delivery) in that locality, 0 otherwise
is_business	Integer	Whether the postal code is used for businesses	1 when the postal code is used for businesses in that locality (meaning entities are listed in the Business table for that locality-postcode association), 0 otherwise
is_po_box	Integer	Whether the postal code is used for PO Boxes	1 when the postal code is used for PO Boxes in that locality (meaning PO Boxes are listed in the Business table for that locality-postcode association), 0 otherwise
post_town	Char(80)	Postal Town	Required by Royal Mail when writing addresses for deliveries in Great Britain (GB) and certain dependencies. Only be populated for GB and some dependencies (GG, IM, JE).

Streets table

Field name	Field type	Description	Comments
ISO	Char(2)	ISO 3166-1 Country code	The ISO 3166-1 standard is published by the International Organization for Standardization (ISO) and defines a unique code for the name of each country. The country codes are represented as a two-letter code (alpha-2).
Country	Char(50)	Country name	
Language	Char(2)	Language code	When there are several official languages in the country, or for countries using a non-Latin alphabet, we provide a specific version for each record. The language field contains the language code and allows to select a specific version.
ID	Integer	Record identifier	
Region1	Char(80)	Administrative division level 1	These fields contain the name of administrative divisions for each available level. If not applicable, the value '-' will be indicated.
Region2	Char(80)	Administrative division level 2	

Region3	Char(80)	Administrative division level 3	
Region4	Char(80)	Administrative division level 4	
Locality	Char(80)	Locality name	Contains the settlement name. The type of settlement may vary depending on the country (city, town, village, etc..).
Postcode	Char(15)	ZIP / Postal code	When applicable, this field contains the valid postal code for the specific record. In case there is no postal codes in use, this field remain blank. The postcode structure varies for each country.
Suburb	Char(80)	Locality subdivision	A suburb is a residential area or a mixed use area, either existing as part of a city or urban area or as a separate residential community within commuting distance of a city. The suburb is indicated if it is determinant for the postcode.
Street	Char(100)	Full street name	
Range	Char(50)	Street numbers range	A street numbers range is provided in case it is determinant for the postal code (when a same street is spread across several postal codes). If the entire street is included within the same postcode area, usually this field is left empty.

			When used, the ranges are described using the following notation: - Continuous numbers : 1-10 - Odd numbers range : (1-9) - Even numbers range : [2-8] Block ranges are separated by a comma, ex: (1-5),[2-8],(9-17)
Building	Char(80)	Building name	A building name is provided in case it is determinant for the postal code. If not determinant, the field remains blank.
Latitude	Double	Street latitude (WGS84 coordinates)	These fields contain the geographic coordinates related to the specific locality and postal code of the record. The geographical position is usually precise to about one hundred meters and corresponds to the center of the locality (generally the historical center, and if not, the geographical center). In some countries, the postal code may refer to a block or a street, in which case the coordinates point to that specific location. In case the georeference is not known precisely, a central point will be defined. Depending on the information available, the point can be defined as a central point of the locality, or as a central point of a higher administrative division. All our files are created using the World Geodetic System standard WGS 84. The corresponding SRID reference is EPSG:4326.
Longitude	Double	Street longitude (WGS84 coordinates)	

Elevation	Integer	Elevation in meters	
ISO2	Char(10)	ISO 3166-2 Region code	The ISO 3166-2 standard is published by the International Organization for Standardization (ISO) and defines codes for the names of the principal subdivisions of all countries. The codes for subdivisions are represented as the alpha-2 code for the country, followed by up to three characters.
FIPS	Char(10)	NGA Geopolitical code (formerly FIPS PUB 10-4)	The NGA Geopolitical codes (Formerly known as FIPS PUB 10-4) is a standard published by the U.S. National Geospatial Intelligence Agency and contains countries, dependencies, areas of special sovereignty and their principal administrative divisions.
NUTS	Char(12)	European statistical division code	The Nomenclature of Territorial Units for Statistics (NUTS) database is a standard provided by the European Commission for statistical purposes.
HASC	Char(12)	Hierarchical administrative subdivision code	The Hierarchical administrative subdivision codes (HASC) are codes to represent names of country primary subdivisions, such as states, provinces, regions. The codes are alphabetic and have constant length for the first level subdivisions.

STAT	Char(20)	National statistics/census code	The statistical codes are codes defined by each National Statistical Offices.
Timezone	Char(30)	Time zone name (Olson)	Following the timezones defined in the IANA tz database (sometimes called Olson database).
UTC	Char(10)	Standard time expressed as an offset with respect to UTC	The Coordinated Universal Time (UTC) is the time standard commonly used across the world to keep time synchronized.
DST	Char(10)	Daylight saving time expressed as an offset with respect to UTC	The Daylight Saving Time (DST) is the practice of setting the clocks forward one hour from standard time during the summer months, and back again in the fall, in order to make better use of natural daylight.
locality_type	Char(20)	Type of Place	'town' for regular places, 'business' where the place is a postal construction to accommodate industrial areas or businesses which are not attached to their own region but to a major city in a nearby region

Regions table

Field name	Field type	Description	Comments
ISO	Char(2)	ISO 3166-1 Country code	The ISO 3166-1 standard is published by the International Organization for Standardization (ISO) and defines a unique code for the name of each country. The country codes are represented as a two-letter code (alpha-2).
Country	Char(50)	Country name	
Language	Char(2)	Language code	When there are several official languages in the country, or for countries using a non-Latin alphabet, we provide a specific version for each record. The language field contains the language code and allows to select a specific version.
Level	Integer	Administrative level	This field contains a value ranked from 0 to 4 to define the administrative division level, from the largest to the smallest (0 for the country, 4 for the smallest division). Depending on the country, only some levels may be available.
Type	Char(50)	Type of administrative division	The type of the administrative division (e.g. State, Province, District, Commune)

Name	Char(80)	Administrative division name	The name of the administrative division (e.g. “Illinois”)
Region1	Char(80)	Administrative division level 1	These fields contain the name of administrative divisions for each available level. If not applicable, the value '-' will be indicated.
Region2	Char(80)	Administrative division level 2	
Region3	Char(80)	Administrative division level 3	
Region4	Char(80)	Administrative division level 4	
ISO2	Char(10)	ISO 3166-2 Region code	The ISO 3166-2 standard is published by the International Organization for Standardization (ISO) and defines codes for the names of the principal subdivisions of all countries. The codes for subdivisions are represented as the alpha-2 code for the country, followed by up to three characters.

FIPS	Char(10)	NGA Geopolitical code (formerly FIPS PUB 10-4)	The NGA Geopolitical codes (Formerly known as FIPS PUB 10-4) is a standard published by the U.S. National Geospatial Intelligence Agency and contains countries, dependencies, areas of special sovereignty and their principal administrative divisions.
NUTS	Char(10)	European statistical division code	The Nomenclature of Territorial Units for Statistics (NUTS) database is a standard provided by the European Commission for statistical purposes.
HASC	Char(12)	Hierarchical administrative subdivision code	The Hierarchical administrative subdivision codes (HASC) are codes to represent names of country primary subdivisions, such as states, provinces, regions. The codes are alphabetic and have constant length for the first level subdivisions.
STAT	Char(20)	National statistics/census code	The statistical codes are codes defined by each National Statistical Offices.

Business table

Field name	Field type	Description	Comments
ISO	Char(2)	ISO 3166-1 Country code	The ISO 3166-1 standard is published by the International Organization for Standardization (ISO) and defines a unique code for the name of each country. The country codes are represented as a two-letter code (alpha-2).
Country	Char(50)	Country name	
Language	Char(2)	Language code	When there are several official languages in the country, or for countries using a non-Latin alphabet, we provide a specific version for each record. The language field contains the language code and allows to select a specific version.
ID	Integer	Record identifier	
Region1	Char(80)	Administrative division level 1	These fields contain the name of administrative divisions for each available level. If not applicable, the value '-' will be indicated.
Region2	Char(80)	Administrative division level 2	

Region3	Char(80)	Administrative division level 3	
Region4	Char(80)	Administrative division level 4	
Entity	Char(100)	Entity name	Name of the entity to which the Business & Administrative postal code has been attributed.
Postcode	Char(15)	ZIP / Postal code	
Info	Char(100)	Additional address information	Additional information such as PO Boxes or address details.

How IDs are defined

In the Places table, the ID is unique at the combination of a ZIP/postal code and a suburb (if available) or locality.

In the Business table, the ID is unique for each combination of a ZIP/postal code, a locality and an organization (entity + info).

When data is available in several languages, it keeps the same ID, which enables to link the different denominations of a place or business. Consequently, one table row is uniquely identified through the combination of language and ID.

IDs are stable and never recycled in our data. When a record is deleted, its ID is retired and will not be reused for another record. In cases where slight modifications are done to a record (e.g., a name modification) but it refers to the same place/business, the ID will be preserved. IDs can hence be used to identify changes when data is updated.

File formats

Our postal database is available in several file formats:

Format	Description
CSV	CSV (Comma-separated values) is a delimited text format broadly used for exchanging data as it can be imported/exported by most data management software. Our files use semicolon as separator.
ASC	Normalized ASCII version in CSV format where all accents and diacritics have been removed.
DAT	The DAT format is identical to the CSV format except that it is encoded in UTF-16. This format can be used with Microsoft SQL Server.
GML	The Geography Markup Language is a XML grammar defined by the Open Geospatial Consortium to express geographical features.
GeoJSON	GeoJSON is an open standard format for encoding collections of geographical features along with their non-spatial attributes using JavaScript Object Notation.
KML	The Keyhole Markup Language is a popular geographic format using XML notation. It is maintained by Google.
Shapefile	ESRI Shapefile format is a popular geospatial vector data format which can be imported into most geographic information system software.

Additional datasets

Additional datasets, complementing the Street database described above, are available to customers. Below is a short summary of these datasets. A detailed data design sheet can be obtained on request. Some datasets, marked with *, require the normalized version of the product.

Address formats

Each postal service has its own address formatting requirements (e.g., postcode before or after the locality). GeoPostcodes gathers the standard requirements for each country, linking the fields to the Street database columns.

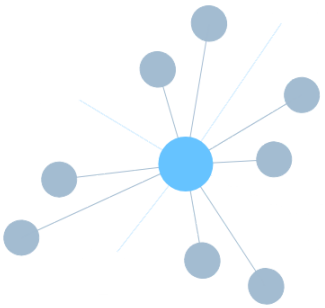
City tags

Many large cities spread over a single, low-level, administrative division. Some of those city boundaries do correspond to higher-level administrative divisions[1] but it is not always the case and the limits of a city may not match with any administrative division[2].

To help customers relate those commonly known cities to our data (administrative divisions and localities composing the city, postal codes associated with it), we have created a dataset of city tags. It links, through ids, places to the bigger city they belong to.

ERP connector

Our data model uses up to 6 levels per country (4 administrative divisions, localities and suburbs). Our customers' ERPs frequently accommodate a smaller number of levels.



To ensure a smooth data integration into your systems, we provide a table indicating which levels should be prioritized, for each country.

For instance, in the USA level 1 (States) will be preferred to level 2 (Counties), while in France level 2 (Departments) is more used than level 1 (Regions).

[1] For instance London corresponds to a region (Greater London) in Great Britain, which is the second administrative GB level, containing 33 boroughs (the 4th administrative level in GB): https://www.london.gov.uk/in-my-area.

[2] For example the city of New York groups 5 boroughs, corresponding to 5 counties (second administrative level in US) among the 62 counties of the State of New York (first administrative level in the US): https://en.wikipedia.org/wiki/Boroughs_of_New_York_City

*English version

In addition to the data in local language(s), we deliver an English version for the whole world, using official or popular translations where available, local names (transliterated in the case of non-Roman alphabet) otherwise, as many places/regions in the world don't have an English translation.

Roma	→	Rome
Sevilla	→	Seville
Bruxelles	→	Brussels
Lisboa	→	Lisbon
München	→	Munich

Exonyms

Still on the language topic, we deliver a list of alternative names (also called "exonyms") for all places/regions in the world, and all languages. This enables you to know what London is called in Arabic, for instance, so you can allow your users to query the data in their own language.

*Primary cities

A postal code is often shared between different localities. In the USA, the concept of "Primary Locality", i.e. the main locality associated with a ZIP code, is clearly established and US Postal delivers such information.

We have developed an algorithm to identify the primary locality of every postal code worldwide, and make that information available to you.

Timezones Daylight Saving Times

We provide the exact dates and times of switching from standard time to Daylight Saving Time, and back, for all the IANA time zones. These times are provided for the current and future years, according to the latest information available from IANA.

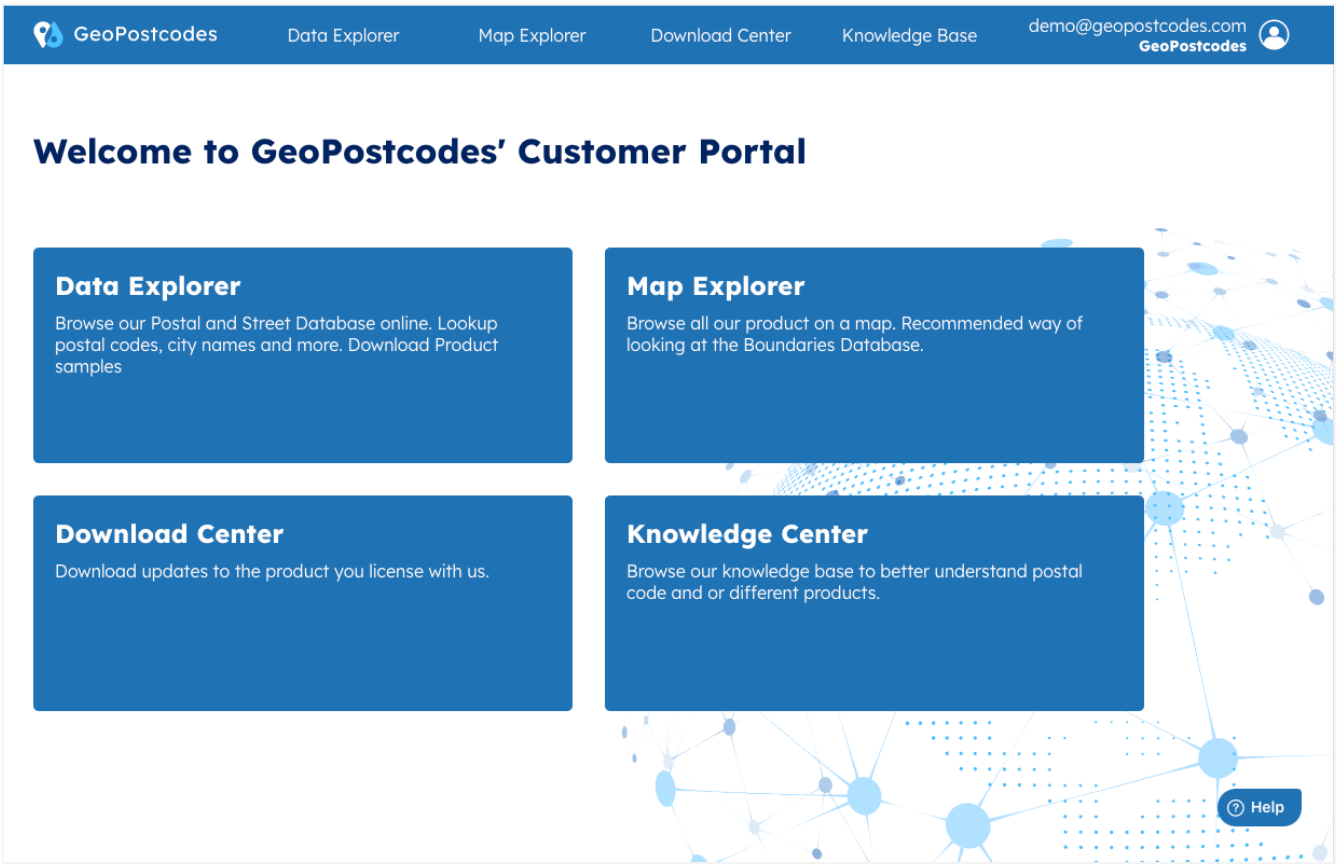
Our convention is to consider that Daylight Saving Time is always ahead of standard time.

*Town weights

To help auto-completion and partial search, we output weights together with places and regions. If one starts typing "Boston" for instance, or completely enters it, the weights enable to rank the search results and favor Boston, Massachusetts over other localities called Boston (in the state of New York, etc.).

GeoPostcodes Customer Portal

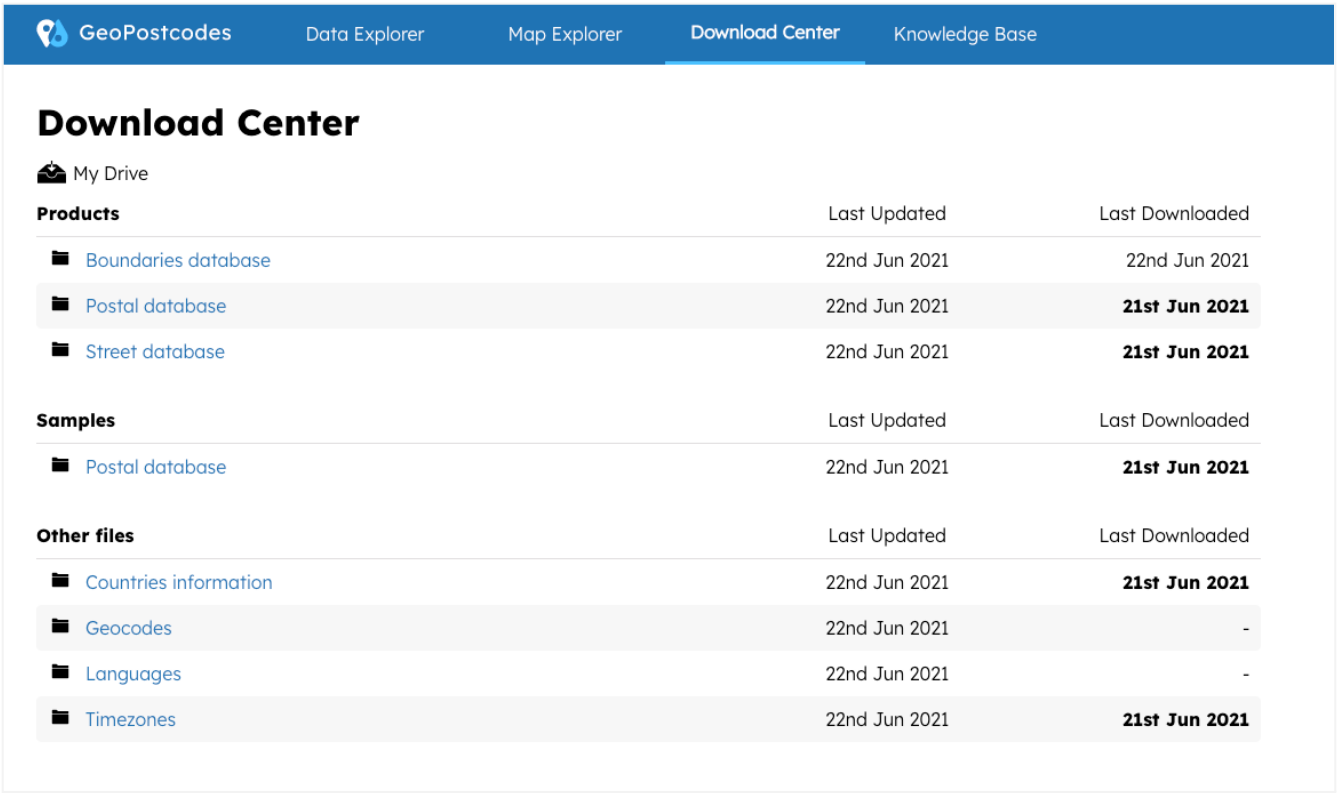
As a GeoPostcodes customer you'll have access to our Customer Portal at https://portal.geopostcodes.com.



Accessing your licensed products

Through the Download Center

The Download Center gathers all the files that you have licensed with us. It shows which files were updated when and highlights the files for which you don't have the latest version yet. You can select the format you want to download.



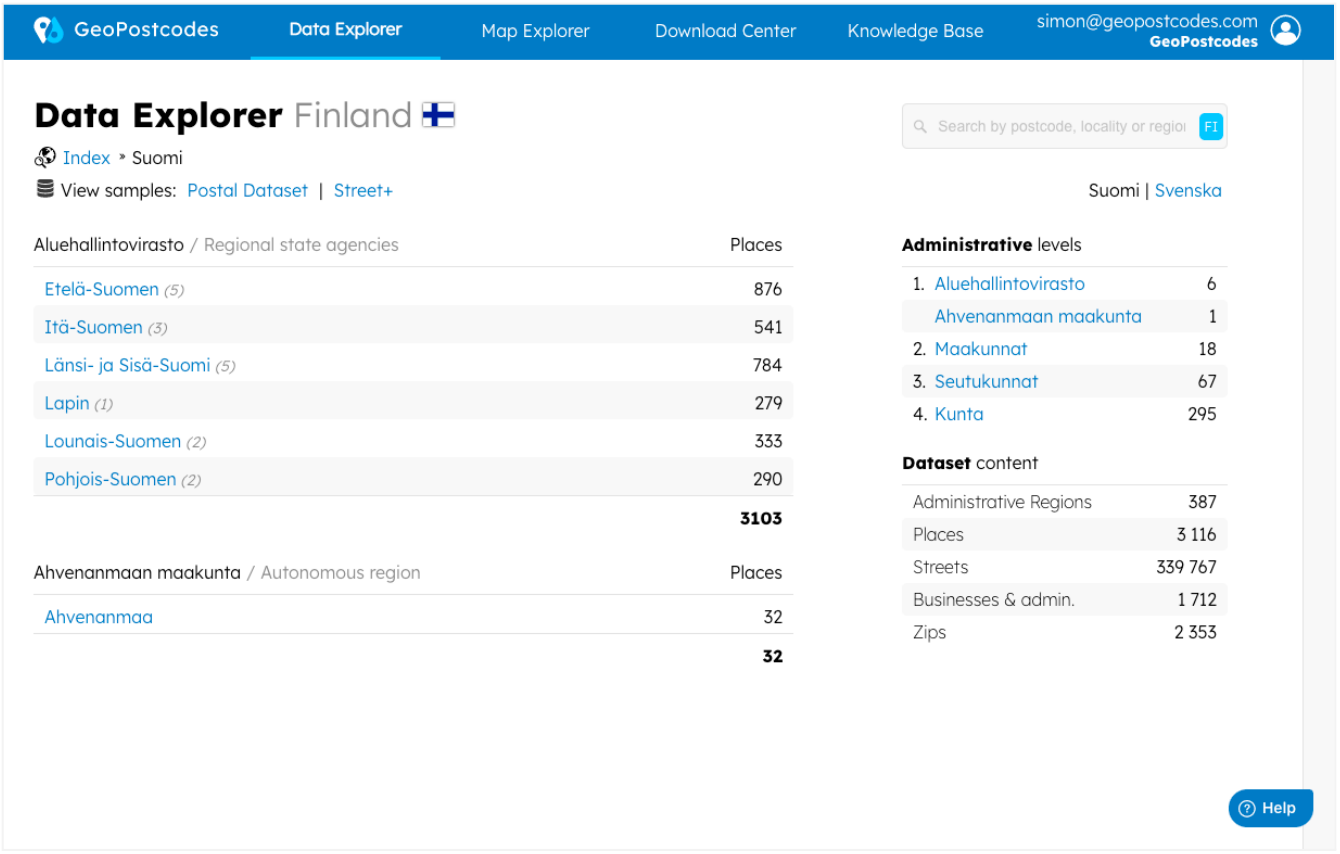
Through our download API

Files can also be pulled from our download api using the same credentials as the ones used for the Customer Portal. The API and documentation are accessible on

https://download.geopostcodes.com/

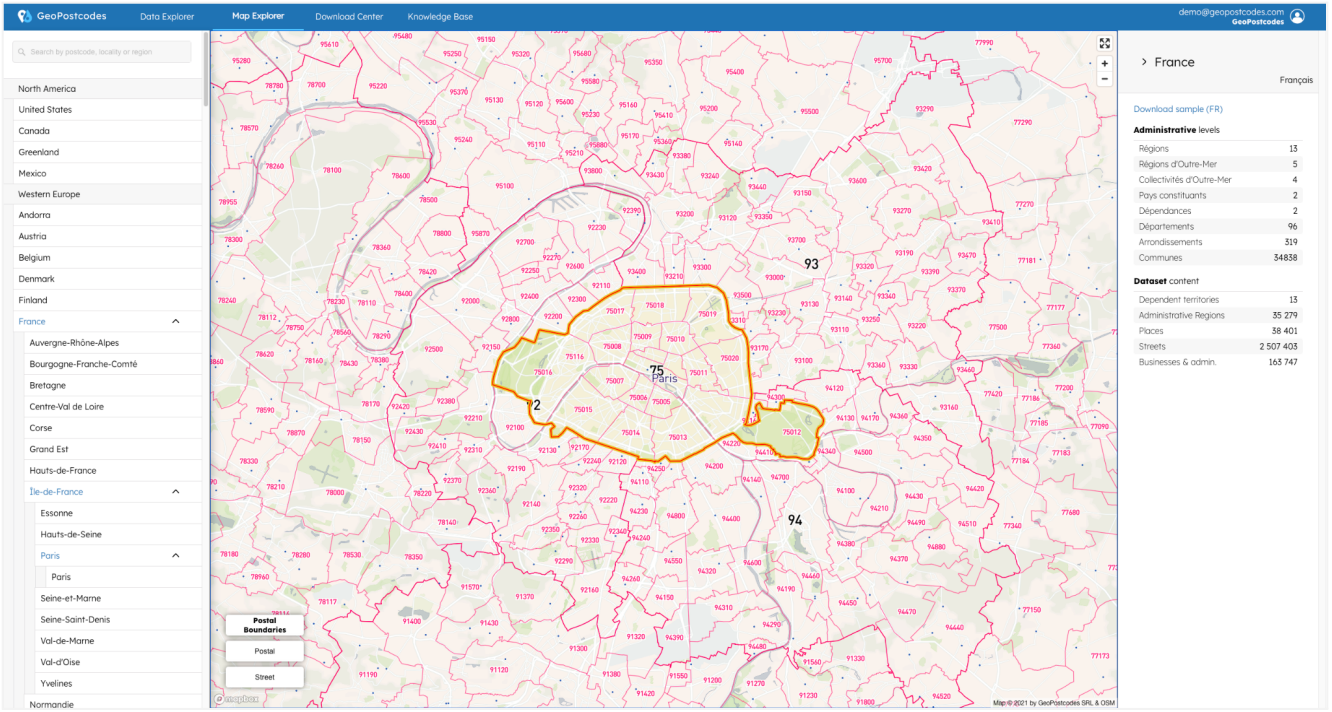
Using the Data Explorer

The Data Explorer allows you to browse the latest version of our Postal and Street Database in an intuitive table layout. You can drill down from the top (country level) all the way to a city, postal code or street. Alternatively, you can search directly for a specific piece of data.



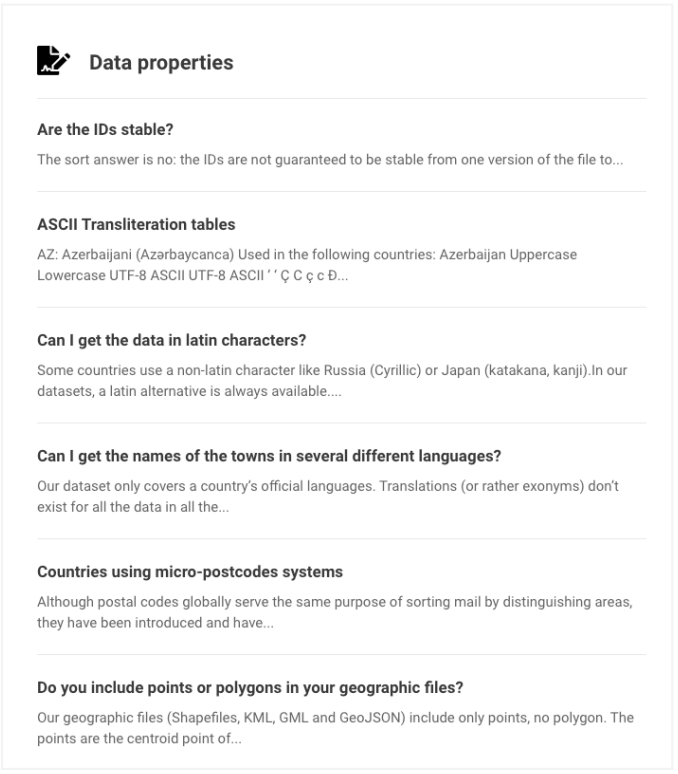
Using the Map Explorer

The Map Explorer is a lightweight GIS viewer allowing you to quickly investigate geographic relationships between regions, postal areas and cities. It is the perfect tool to visualize our upcoming Postal Boundaries Database product (release planned in September 2021).



Getting answers to your questions

Integrating postal codes on a worldwide scale often turns out to be more complex than many customers initially thought. And while we are always available to help solve tricky questions, we also make an extensive knowledge base available to support your teams in their integration efforts. It covers everything you need to know about our products as well as general information about countries and postal code systems.



Integration assistance and Support

Integration assistance

At GeoPostcodes we do not simply provide a dataset, we provide dedicated assistance until your project is up and running. Your onboarding starts with a kick-off call between your integration team and one of our data specialists. During that meeting we'll identify the pain points of your use-case and guide you to the best dataset formats for your situation. Our data specialist will then assist your team until your project is complete.

Our intention is to take the entire complexity of the location master data aspect of your project on our side so your team can focus on the project at hand.

Support

If, at any point you need assistance with a data issue you can contact our support team through the customer portal.

Appendices

This product sheet covers the Street Database core data set. For more information on the additional datasets and data formats just ask our team.

- Normalized format documentation
- Steet Database Datasheet
- Admin division names
- Address Formats
- City tags
- ERP connector
- English version
- Exonyms
- Primary cities
- Timezones DST
- Town weights