ONS API Glossary

APIKEY

Each request to the API must have an API key which is statelessy checked each time. This key is used to record usage statistics for ONS. These statistics will not be shared with others and are stored anonymously. The API administrator can run a secure application which will reveal the identity of users if necessary, such as in the case of excessive usage. In some cases key sharing will be allowed such as third party visualisation tools.

AREA TYPE

A type of area such as Local Authority, Output Area or Ward. Each of these has a standard abbreviation and when you view the details of a geographic hierarchy the area types used are shown.

CLASSIFICATION

A classification is a categorisation scheme for a particular characteristic of a person or household, such as age or social class. There can be more than one classification for the same characteristic e.g. AGE_10YRS, AGE_5YRS. Each dataset has a number of dimensions and each of these dimensions implements a classification.

COLLECTION

A collection is a series of datasets identical apart from the geographic hierarchy used. The collection id is by convention the same as the dataset(s).

CONCEPT

A concept is a high level category of classification. For example a concept of AGE might have classifications AGE_10YRS and AGE_5YRS with different bands. SDMX also requires concepts to be defined for some attributes such as observation status.

CONTEXT

To allow for future expansion, the underlying dataset in WDA is partitioned. Each partition corresponds to a different data group, currently Census, Economic, Social and Socio-economic.

CSV

Comma Separated Values http://en.wikipedia.org/wiki/Comma-separated_values

DATASET

A dataset is a set of related data items. Each cell in a dataset is uniquely identified by a different combination of dimension values. Datasets in WDA are all aggregate data, so for each valid combination the number of persons, percentage of persons, number of households is held. Unit data is not available via this API.

Datasets can vary in complexity. Most of the Census datasets are quite simple (less than 5 dimension) and represent a single “table”. There will later be some “hypercubes” which have many dimensions and are intended for more flexible usage. Datasets can be downloaded in full or can be dynamically “sliced”

DIFFERENTIATOR

Datasets in a collection are usually identical apart from geography, but in some cases there is more than one version of the dataset for the same geography. In this case the datasets are marked with a differentiator. This can be anything but by far the most common is a year .

DIMENSION

Each dataset has a number of dimensions. Most of these will be linked to a particular classification. For example a simple dataset might have the dimensions LOCATION, SEX and AGE. LOCATION is the geographic dimension which will have a hierarchical geographic classification, SEX and AGE will each implement a simple classification such as SEX01 and AGE_5YRS.

DIMENSION ITEM

Each dimension has a number of items according to the classification it uses. For example SEX might be "All Persons", "Male" and "Female".

DISPLAY ORDER

For Census data, classifications are supplied with a recommended order to show dimension items. For example an age classification would normally be shown from the youngest to oldest age band. API users do not have to obey this order, but it is recommended to use it by default.

DSD

This stands for Data Structure Definition and is the SMDX term for the structure of a dataset. This comprises codelists for all the dimensions, and the key family. Datasets in a collection will have identical DSDs with the exception of the geography dimension.

GROUP

Data cells returned by the API are structured using three nested SDMX entities: Group, Series and Observation (see example SDMX output). Dimensions and attributes can be applied at any of the three levels. For multi-segment datasets, each segment has its own group.

HIDDEN

The ONS Web Data Access repository has a function to hide and unhide datasets. This allows data to be temporarily made unavailable if a problem is suspected, the dataset is later restored after correction (or an investigation has shown no correction is required). New datasets that are not market sensitive can also be uploaded in advance of their release date as hidden, then unhidden at publication time (e.g. 9.30am). Market sensitive datasets cannot be physically loaded in advance.

HIERARCHY

ONS normally holds its geographically coded data based on hierarchies. Each hierarchy consists of a number of levels. Members of each level are contained within a parent area of the level above. For example each Local Authority contains a number of wards. Examples of hierarchies are Administrative, Statistical, Parishes, Parliamentary Constituencies and Health Areas. Data explorer type applications will often allow the user to navigate hierarchies via a tree.

HTML

Hypertext Markup Language http://en.wikipedia.org/wiki/HTML

HTTP (GET, PUT, POST DELETE)

Hypertext Transfer Protocol http://en.wikipedia.org/wiki/HTTP

JSON

Java Script Object Notation http://en.wikipedia.org/wiki/JSON

JSON-Stat

A newish standard for representing statistical data in JSON see http://www.json-stat.org

KEY FAMILY

As part of a DSD, the Key Family is a kind of skeletal representation of the dataset, listing the dimension and attributes. It links dimensions to codelists and concepts.

LEVEL

Each geographic hierarchy comprises a number of levels. It is possible to have more than one type of area at the same level. For example in the Admin hierarchy counties and unitary authorities both exist at level 4 (beneath regions).

NODE

An object in the API's entity model.

NOOBS

Number of observations. As a query string parameter it is used to restrict the size of the response, as part of the URL it is a request to return the number of cells in a dataset.

PRESENTATION

All datasets are furnished with presentation information regarding the layout of the table which the dataset represents. This includes which axis each dimension resides upon, and if the table has more than one segment, the order that the segments are displayed.

RDF

Resource Description Framework http://en.wikipedia.org/wiki/Resource_Description_Framework

REST

Representational State Transfer - http://en.wikipedia.org/wiki/REST

REFERENCE METADATA

There are two types of metadata held for entities within the WDA database. Structural metadata items are entities in their own right, for example dimensions. Reference metadata items are obtained by adding the word 'metadata' to the URI and are notes and commentaries about a particular entity.

ROOT

The top level node of the API. This is the start point for discovery of the API's nodes, sometimes called HATEOAS (Hypertext As The Engine Of Application State).

SDMX

Statistical Data Metadata Exchange - http://sdmx.org/ - International XML standard for the exchange of statistical data.

SEGMENT

Some datasets are completely regular in terms of which dimensions are cross-tabulated. In other cases a dataset may consist of two or more different cross-tabulations, in which case each of these is labelled as a segment. For example LOCATION by AGE by SEX by HEALTH as a single segment, or LOCATION by AGE by SEX and LOCATION by SEX by HEALTH has two segments. In the latter case it would not be possible to ascertain the breakdown of AGE and LOCATION.

SLICE

A selection of data from a dataset. A slice is defined by limiting the items included for one or more dimensions. For example limiting the cells to those for a single administrative area.

STARTOBS

Start observation, used in conjunction with noobs.

TOPIC

Topic is used synonymously with a concept of the first kind e.g. “Ethnic Group”. It is used by the ONS Data Explorer.

TOTAL

This is a keyword which enables the API user to choose how unspecified dimensions are dealt with in a slice. If set to true, the default, it is assumed that only the total item is required. If set to false, all items will be returned.

URI

Uniform Resource Identifier http://en.wikipedia.org/wiki/URI

URI Type

There are six types of URI: Identifier, Document, Representation, List, Set and Ontology. See https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60975/designing-URI-sets-uk-public-sector.pdf

URL

Uniform Resource Locator http://en.wikipedia.org/wiki/URL

XML

Extensible Markup Language http://en.wikipedia.org/wiki/XML