ORDNANCE SURVEY GB

CODE-POINT WITH POLYGONS OVERVIEW

Version history

Version	Date	Description
1.0	10/2021	Initial release of this Overview document. New data format.

Purpose of this document

This document provides information about and insight into the Code-Point with Polygons product and its potential applications. For information on the contents and structure of Code-Point with Polygons, please refer to the Getting Started Guide and Technical Specification.

The terms and conditions on which Code-Point with Polygons is made available to you and your organisation are contained in that Ordnance Survey customer contract. Please ensure your organisation has signed a valid current customer contract to be able to use Code-Point with Polygons.

We may change the information in this document at any time, giving you the notice period set out in your contract. We do not accept responsibility for the content of any third-party websites referenced or accessed in or through this document.

This document has been screened according to Ordnance Survey's Equality Scheme. If you have difficulty reading this information in its current format and would like to find out how to access it in a different format (braille, large print, computer disk or in another language), please contact us on:
+44 (0)3456 05 05 05.

Copyright in this document

© Ordnance Survey Limited 2021. This document (including for the avoidance of doubt, any mapping images reproduced within it) is protected by copyright and apart from the rights expressly granted within this document to use the content, all rights are reserved. Any part of this document may be copied for use internally in your organisation or business so that you can use Code-Point with Polygons under the terms of your licence (but not otherwise).

No part of this document may be reproduced or transmitted in any form or by any means (including electronically) for commercial exploitation, onward sale or as free promotional material without getting the written consent of Ordnance Survey beforehand.

Trademarks

Ordnance Survey, OS, the OS Logos and Code-Point are registered trademarks, and Code-Point with Polygons is a trademark of Ordnance Survey, Britain's mapping agency.

PAF, PO Box and Royal Mail are registered trademarks of Royal Mail Group Ltd.

Contact details

OS website 'Contact us' page (https://www.ordnancesurvey.co.uk/contact-us).

Contents

1.   Introduction

Code-Point with Polygons is a dataset that contains the notional area of postcode units, allowing customers to display and analyse any data collected at the postcode level.

The polygons within the product are derived from georeferenced Royal Mail Postal Address File (PAF) delivery addresses. A process is undertaken to create a set of polygons around individual address records within a postcode. This is called a Thiessen process and the polygons are the result of a mathematical computation that creates polygons from point data. In this way, mathematically-consistent boundaries are created between distinct postcode groups, creating this notional boundary set.

Postcode unit boundaries are, by definition, only the delivery point or collection of delivery points that constitutes the postcode units. The boundary is therefore a notional one, the position of which is arbitrary. What has been created, however, is a set of boundaries that follows a consistent logic and portrays the notional footprint of each postcode unit. The boundary encloses every delivery address for which positional data of sufficient quality is available, and which follow major physical features that could reasonably be regarded as part of the postcode boundary.

1.1   Key features

The key features of the Code-Point with Polygons product are:

- A set of 120 postal-area-based files that provide a set of boundaries for the postcode units in Great Britain for SHP, TAB and MID/MIF supply formats. For Vector Tile supply, a single national file is provided.

- Vertical_streets: a list of polygons for locations that contain more than one postcode, e.g. office blocks and flats.

- Corresponding Code-Point information providing the number of delivery addresses and the health and administrative area codes related to each postcode.

1.2   Uses of Code-Point with Polygons

The quality of this polygon creation allows the polygons to be used for a wide range of applications. This will include analysis of geographically-based information or statistics by postcode, and the pictorial display of information that has been analysed or sorted by postcode.

2.   Product details

2.1   Postcodes

There are two main components of a postcode:

- **The outward code (also called outcode).** The first two to four characters of the postcode constitutes the postcode area and the postcode district. It is the part of the postcode that enables mail to be sent from the accepting office to the correct area for delivery.

- **The inward code (also called incode).** The last three characters of the postcode constitutes the postcode sector and the postcode unit. It is used to sort mail at the local delivery office.

For example:

Outward		Inward	
NW	6	4	DP
			Unit
		Sector	
	District		
Area			

When used in an address, the incode should be separated from the outcode by a single space. The following is a list of the valid formats of postcodes (an A indicates an alphabetic character; an N indicates a numeric character).

Outcode	Incode	Example postcode
AN	NAA	M2 5BQ
ANN	NAA	M34 3AB
AAN	NAA	DN5 7XY
AANN	NAA	DN16 9AA
ANA	NAA	W1A 4WW
AANA	NAA	EC1A 1HQ

2.2   Postcode polygon geometry

Postcode polygons are produced by the tessellation of georeferenced PAF coordinates for individual Royal Mail delivery addresses. Only addresses having a positional quality value indicating the location is within a building are used to create the polygons file. Postcodes of addresses of lower quality will be included in the discard files.

Due to the nature of postcode geography, the polygons representing some postcode units are unavoidably split. Every effort has been made to ensure the absolute minimum of postcodes is represented by multiple polygons. These split polygons representing a single postcode remain a single object with one set of attributes.

Each polygon is assigned a unique identifier. The identifier will be a 16-digit series. These identifiers are not reused should a polygon be deleted.

The polygon dataset contains non-overlapping polygon coverage of Great Britain, originally constrained by the extent of realm (EOR) coastline from Ordnance Survey's Boundary-Line data and postcode polygons. Should any addresses fall outside the constraining datasets, the postcodes should be included in the discard files.

The data is divided into 120 postcode area files, each file named with a one- or two-letter postcode area code.

2.3   Vertical Streets

Postcodes that are vertically stacked, that is, two or more postcodes within a single building that are represented by a single large-scale building seed. In these situations, a single square polygon represents all the postcodes attributed to the single building seed. These polygons have a special series of identifiers, all commencing with the letter V.

A separate vertical streets lookup table is provided with the polygons and lists the postcodes with the 20-digit unique identifier that are represented by each special polygon. Where these distinctive polygons are crowded closely together, they are reduced in size to prevent overlaps hiding some of the polygons.

2.4   Data capture

The polygon set contains a polygon for every postcode in England, Scotland and Wales that is contained in Royal Mail's PAF product, with the following exceptions:

•   Postcodes for which there is no location data of sufficient quality.

•   Postcodes for which there is no data that lies within the extent of the realm coastline.

•   Postcodes that relate to PO Boxes.

•   GeoPlace geocode the PAF data from Royal Mail, using source coordinates from Local Authorities in England, Wales & Scotland and Ordnance Survey. GeoPlace then provide the georeferenced PAF data to Ordnance Survey.

2.5   Coordinate Reference Systems

The Coordinate Reference Systems (CRS) of the polygon data and the coordinates in the Code-Point tables is provided in British National Grid (BNG).

BNG uses the OSGB36 geodetic datum and a single Transverse Mercator projection for the whole of Great Britain. Positions on this projection are described using Easting and Northing coordinates in units of metres. The BNG is a horizontal spatial reference system only; it does not specify a vertical (height) reference system.

2.6   Currency

Updates will normally be at 3-monthly intervals and are a complete resupply of the national dataset.

Each edition of Code-Point with Polygons will have a version number showing the release month for the year (e.g. April) followed by the release year (e.g. 2020).

Starting from April 2020, the version for each quarterly release will be in this format:

• April 2020
• July 2020
• October 2020
• January 2021

Within the Code-Point with Polygons product set, the Code-Point data and the polygons will have their own version numbering because they are released at different times. For example, Code-Point with Polygons April 2020 release will be supplied alongside a set of Code-Point version 2020.1.0. Both Code-Point and the polygons are initially sourced from the same version of address information.

3.   Product supply

3.1   Supply format and media

Code-Point with Polygons is available in the following formats:

- Shapefile
- MapInfo MID/MIF
- MapInfo TAB file
- MBTiles

Code-Point with Polygons are supplied as a zip file in the Data folder on the DVD. The Code-Point polygon files (only) contain substantial amounts of information, which in both Shapefile and MID/MIF formats necessitate that file compression be used.

3.2   Coverage and file sizes

The polygon data coverage is Great Britain. Each format is split into 120 files where each file represents the notional geometry for a postcode area.

File sizes for Great Britain are approximately:

- Shape: 605 Mb
- Tab: 524 Mb
- MID/MIF: 564 Mb
- MBTiles: 434 Mb

The following two CSV text files accompanying the polygon data are:

- **Vertical_streets** – a list of polygons, identified by a serial number that is prefixed by the letter V, which contain more than one postcode. This situation can occur in, for example, a block of flats where there is more than one postcode within a single building.

- **Discard_files** – a list of the postcodes for which polygons have not been included because there is no data of sufficient quality to use in the polygon creation, or because their constituent addresses lie outside the extent of the realm (coastline). The discard file also contains a list of PO Box postcodes as none of these will have been used in the creation of the polygon set.

The associated Code-Point coverage is for United Kingdom, provided as 121 comma-separated value (CSV) files, as it includes Northern Ireland postcodes. Polygons are not provided for Northern Ireland postcodes. Code-Point contains georeferenced postcode unit data, with associated metadata such as address counts and quality indicators. It also provides the health and administrative area codes related to each postcode.

The following text files that are associated with the Code-Point data provide:

- The full text equivalents of the administrative area codes.
- The number of postcode units in each postcode area.