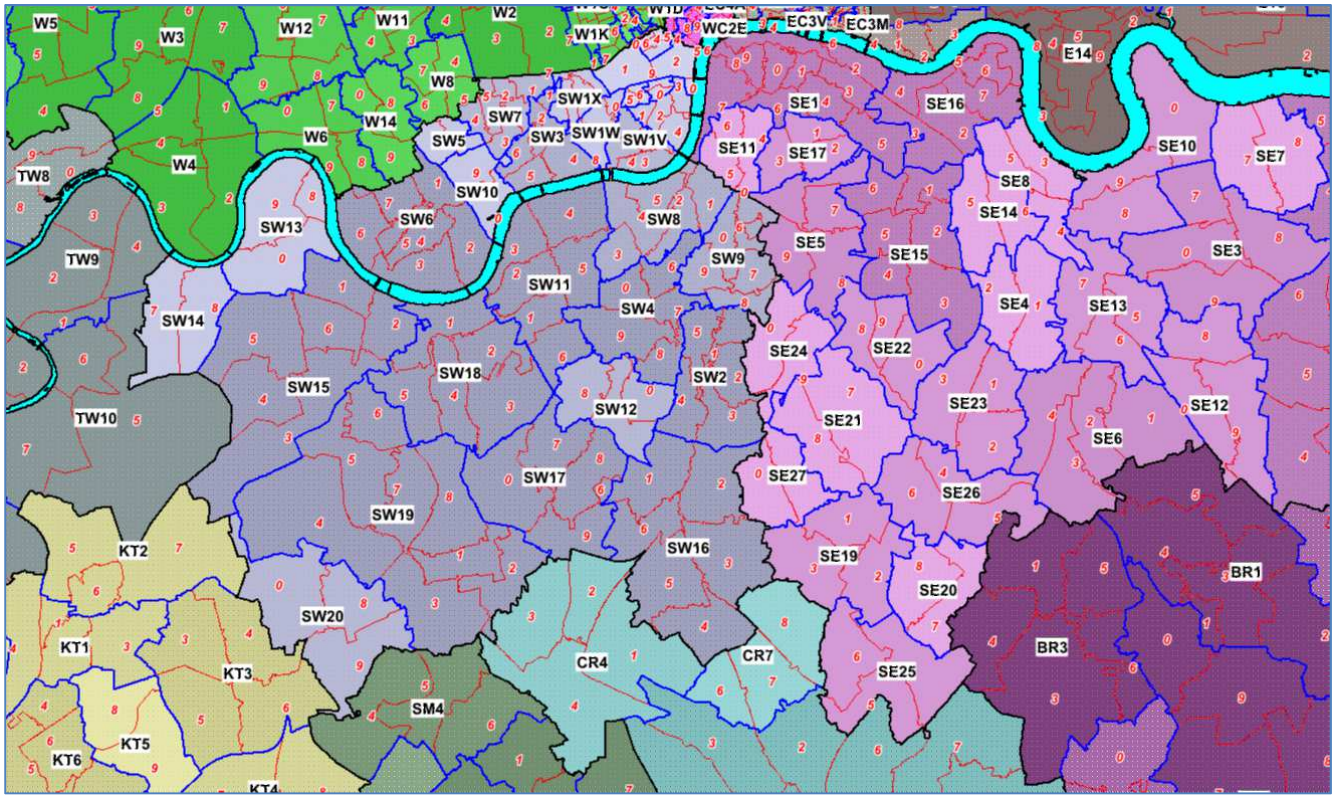
GeoData

Postal Boundaries Open

User Guide October 2012



Preface

Document Title	Postal Boundaries Open User Guide 2012
Version No	Version 2.2
Date	October 2012
Trademarks & Copyright	Postal Boundaries Open User Guide is copyright Spatial Intelligence Solutions Ltd 2012 trading as GeoLytix. All other trademarks are the property of their owners and are used in this document for descriptive purposes only.
Quality Statement	The Postal Boundaries Open data are derived from multiple sources including OS OpenData™. As such the data is dependent on the accuracy of those input datasets. GeoLytix have taken reasonable care to standardise and quality check all the input data, however there may be errors carried forward from the input datasets. Additionally the processing and boundary creation processes may have introduced additional errors. The data is provided as is. All the information providers exclude all representations, warranties, obligations and liabilities in relation to the Information to the maximum extent permitted by law. This Open version of postal boundaries is a frozen product and no updates or corrections will be released.
Product License Terms	Postal Boundaries Open' is licensed under the same terms as the OS OpenData license with the addition of GeoLytix to the attribution statement.
Contact Details	Email. Blair.freebairn@geolytix.co.uk or Sarah.Hitchcock@geolytix.co.uk

Document History

1.0	Initial Release delivered with June 2012 Dataset.
1.1	Amendments following user feedback from first release.
2.0	Update to August data and corrections based on feedback.
2.1	Amendments based on clip to high-water mark.
2.2	Creation of documents for Open Data version.

Contents

Overview

This document details the content of the Postal Boundaries Open 2012 dataset; the methodology used to create it and shows examples of the data. Postal Boundaries Open 2012 is a spatial dataset detailing the extent of the 9,232 geographic postal sectors (eg H21 8) covering Great Britain as of October 2012.

Sectors are grouped together to form Postal Districts (eg HP21) which in turn group together to form Postal Areas (eg HP). Each Postal Area is given its formal name; districts and sectors are given names based on the dominant GeoLytix town region and within the largest cities the dominant suburb. The methodology used places an emphasis on following 'natural' boundaries; avoiding, where possible, enclaves, exclaves and split Sectors.

In addition to the boundaries each sector, area and district is represented by a point, derived from the postcode weighted centroid and snapped to the nearest postcode point within the appropriate region.

The 1,666 non-geographic sectors across the whole of the UK and crown dependencies; Guernsey, Jersey and Isle of Man, are listed in a table showing the sector typology, large user or PO Box, and where relevant the 'large user name' of the sector.

Methodology

The principle objective was for boundaries to follow 'natural' boundaries where possible, candidate boundary lines for Great Britain were sourced from a range of OS OpenData™; these included rail, road, the coastline, river lines, administrative boundaries and open water boundaries.

The outputs also required the creation of boundaries via manual digitisation.

To create boundaries a series of spatial algorithms were developed to generate 'boundary corridors' between sectors and then a subsequent set of protocols were used to construct optimal boundaries from the boundary candidate pool. Where no boundary candidates were appropriate or selected boundaries failed acceptance tests, boundaries were created by manual digitisation. The manual digitisation was conducted with a backcloth of the 1:10,000 OS VectorMap® District raster data for Great Britain. The requirement for manual digitisation was most marked in the two extremes of urbanity; in very remote areas there tended to be too few boundary candidates, and in very dense urban areas where fine scale postal sector discontinuities and dense candidate boundaries created issues. For example the sector boundaries in Central London and in the Scottish Highlands both tend to have a higher incidence of manually created boundaries. The boundaries are generally clipped to the high-water mark of Great Britain as defined by the VectorMap ® District vector data, however minor tidal rivers have been 'in-filled', generally upstream of their first crossing point. In a number of instances manual corrections had to be made to the high-water mark to ensure all postcodes fell within a postal sector, for example for postcodes referencing piers, houseboats or on recently reclaimed land.

The object creation algorithms generated approximately 95% of the boundaries with manual digitisation accounting for the remainder. The objects have undergone visual as well as programmatic quality assurance.

As of this release out of ~1,700,000 postcode points only ~77,000 (4.5%) have a recorded position that does not fall within the postal sector object that their text indicates they 'should'. Of these only 900 (0.05%) are genuinely anomalous. These anomalies occur for two major reasons, and one minor reason:

1 – **Co-incident postcodes with different postal sectors**. Occasionally a postcode point shares co-ordinates with another postcode even though its text indicates it is in a different sector. In these cases the dominant sector for the point is adopted and used in subsequent processing. Approximately 31,000 'out of sector' postcodes exist due to this 'vertical street' reason.

2 – **Non-geographic postcodes with co-ordinates**. Approximately 45,000 postcodes exist with geographic co-ordinates even though the sector of the postcode is classified as non-geographic. These are sometimes the location of a postal sorting office, and sometimes the actual delivery location of large users that have been assigned their own postal sector. These non-geographic postcode points were not used in the sector creation process and so there 'pseudo-location' is often in the 'wrong' sector region.

3 – **Legacy or outlier postcodes**. When a new postal sector is created or an existing sector radically altered due to new development, there may be existing properties with a postcode that pre-date the creation of the new sector or be in a location that becomes 'orphaned' by the development. These properties are generally allowed by the Royal Mail to keep their 'old' sector postcode and so appear as

isolated ‘islands’ within the sea of the new postal sector. Constructing a connecting path from the predecessor sector through the new sector to the island is often impractical, and maintaining a series of legacy postcodes as a chain of islands is unsatisfactory. Where the orphaned enclave consists of multiple postcodes then an island is viewed as acceptable and the island object is retained. Where this island is a single postcode, the postcodes are generally allowed to anomalously fall within the ‘new’ postal sector. These outliers account for ~900 of the ‘out of sector’ postcodes. No single postal sector contains more than four such outlier postcodes.

After production all objects underwent spatial cleaning to remove processing artefacts and objects contain no overlaps or self-intersections. Node-to-node spikes and ‘kick-backs’ with an interior angle below 3 degrees have been removed. All postcodes fall within one and only one postal sector object. Each sector record contains a single postal sector region object, where postal sectors are made from discrete separated region objects (physical islands or exclaves/enclaves) they are represented as a single record with a single multi-part object.

Product Description

The product consists of eight file sets. Three sets of boundaries and points for the main geographic Postal Sectors, Districts and Areas, a single non-spatial dataset detailing the non-geographic Postal Sectors, and a point dataset of all the ‘out-of-sector’ postcodes. The spatial files are provided in BNG projection (EPSG 27700) in MapInfo tab format or Esri shapefile. The input data is of variable scale ranging from 1:1,250 to 1:10,000. Only geographic postcodes with the highest positional quality indicator were used to in the creation process. Manual digitisation was carried out a scale dependant on the urbanity of the region, ranging from 1:2,500 to 1:20,000. The following section contains a brief overview and a field description table for each dataset. All column names are capped at eight characters and are made up of only the letters A-Z (no spaces, punctuation or underscores). All file names use ‘CamelCase’ conventions and only contain the characters A-Z; a-z.

Geographic Postal Sectors Regions

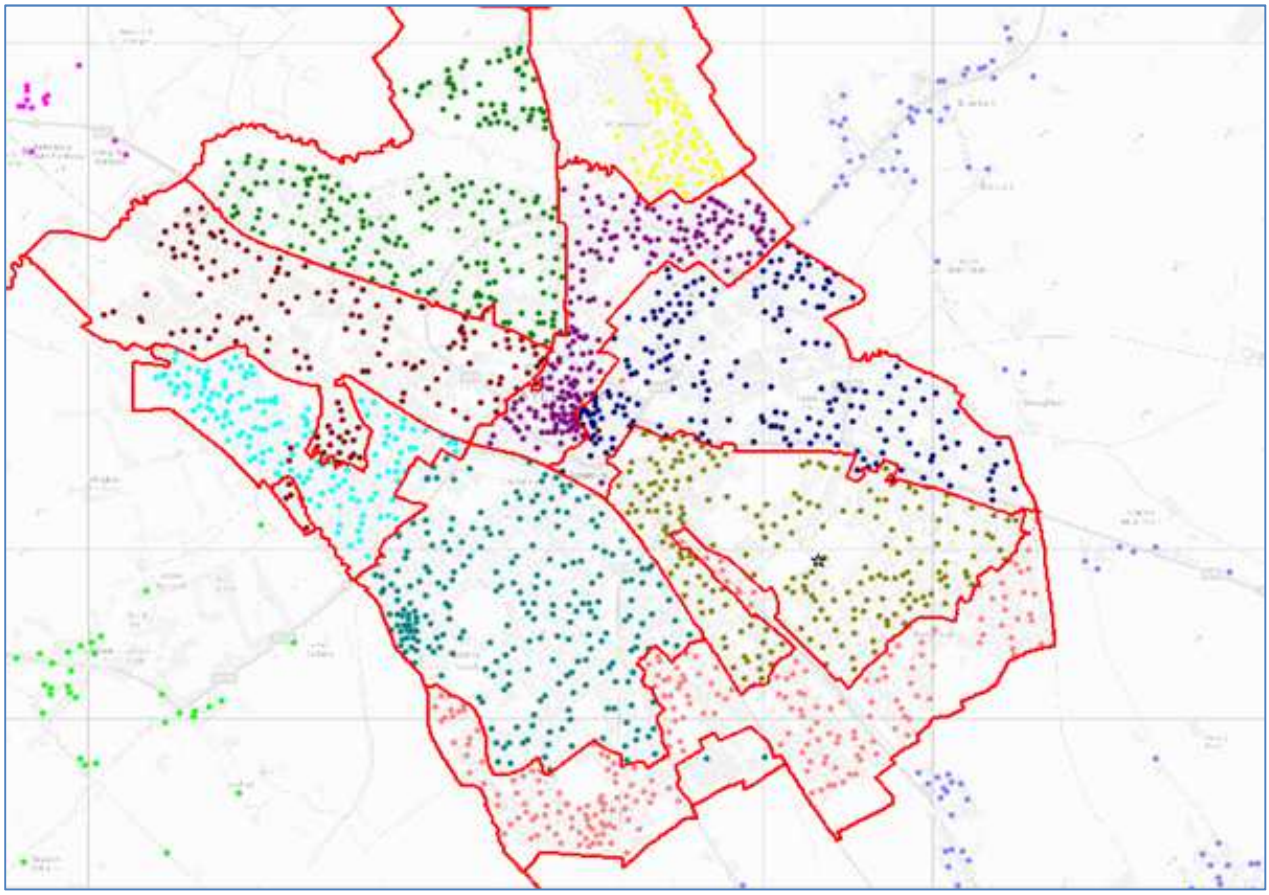
File Name: PostalSector

Record Count: 9,232

This file contains all the geographic sectors covering Great Britain as of October 2012. For reference the Olympics postal sectors (E20 1, 2 & 3) are not included as they remain non-geographic sectors pending their introduction into the live data in 2013.

Column Name	Type	Description
SectID	int	Unique ID assigned to the postcode sector
RMSect	Char(6)	The Postcode Sector text in Standard 'Royal Mail' format, this is how a sector would naturally be written and is made up of the Postcode District followed by a single space and then the sector number. All text is capitalised.
GISSect	Char(6)	The Postcode Sector text in Standard 'GIS' format, this is designed to aid extraction of Postal Areas and Districts from a Sector string and also forces sectors to 'sort' in an intuitive manner. The text is of the form AANN<space>N. Where AA represents the postal Area in character format, NN the postal district number, <space> is a space and N is the sector number. Note where an area is a single letter AA takes the form A<space> and where the district is a single digit it is represented by <space>N. Therefore the postal sector written G1 1 in RM format is written G<space><space>1<space>1 in GIS format. For the Central London Sectors with an Area using a letter suffix NN is replaced with NA, eg WC1X
StrSect	Char(6)	The postcode sector text with all spaces removed. This should be the default style used when matching across multiple datasets as it removes reliance on any text formatting.
PostDist	Char(4)	The Postal District of the sector without any spaces
PostArea	Char(2)	The Postal Area of the sector without any spaces
DistNum	Char(2)	The number of the Postal District, eg for sector RM20 1 this field would contain the number 20. This is included to aid with cartographically pleasing labelling at large to mid-scale. For central London postal Districts the number is replaced with two characters of the form NA.
SectNum	int	The number of the Postal Sector, eg for sector RM20 1 this field would contain the number 1. This is included to aid with cartographically pleasing labelling at small to mid-scale.
PCCnt	int	Count of live postcodes within the sector object as of October 2012.
AnomCnt	int	The count of postcodes within the sector objects that are of a different sector to the containing sector object and are not vertical street or non-geographic postcodes.
RefPC	Char(10)	The postcode from the sector that is closest to the postcode weighted centroid of the sector.
X	int	The x co-ordinate of the sector reference postcode point in the British National Grid co-

		ordinate system
Y	int	The y co-ordinate of the sector reference postcode point in the British National Grid co-ordinate system
Sprawl	Char(100)	The name of the dominant GeoSprawl object that the sector overlaps as measured by shared postcodes. Thirty two Sectors contain only postcodes that do not fall within a sprawl object; these sectors contain empty strings in the GeoSprawl field.
Locale	Char(100)	Within the largest 24 urban sprawls Britain (Bristol, Birmingham, Edinburgh, Glasgow, Leeds, Liverpool, Manchester, Nottingham, Sheffield, London, Cardiff, Hull, Bradford, Newcastle, Leicester, Belfast, Black Country, Brighton, Coventry, Derby, Reading, Southampton, Potteries) the name of the dominant GeoLocale (suburb) that the sector overlaps as measured by shared postcodes. Outside of these Cities the GeoLocale field contains an empty string.



Postal Sectors in Aylesbury area. Map contains Ordnance Survey data © Crown copyright and database right 2012, Royal Mail data © Royal Mail copyright and database right 2012, National Statistics data © Crown copyright and database right 2012, Postal Boundaries © GeoLytix 2012 copyright and database right 2012.

Geographic Postal Sector Points

File Name: PostalSectorP

Record Count: 9,232

For every record in the file PostalSector a complementary record with the postal sector represented as a single point. The point is taken from the single unique postcode of the sector that falls closest to the postcode-weighted centroid of the sector.

Column Name	Type	Description
SectID	int	Unique ID assigned to the postcode sector
RMSect	Char(6)	The Postcode Sector text in Standard 'Royal Mail' format, this is how a sector would naturally be written and is made up of the Postcode District followed by a single space and then the sector number. All text is capitalised.

Geographic Postal District Regions

File Name: PostalDistrict

Record Count: 2,736

This file contains all the geographic Districts covering Great Britain as of October 2012. For reference the Olympics postal sectors (E20) is not included as it remains a non-geographic district pending its introduction into the live data in early 2013.

Column Name	Type	Description
DistID	int	Unique ID assigned to the postcode sector
PostDist	Char(4)	The Postal District of the sector without any spaces
PostArea	Char(2)	The Postal Area of the sector without any spaces
DistNum	Char(2)	The number of the Postal District, eg for sector RM20 1 this field would contain the number 20. This is included to aid with cartographically pleasing labelling at large to mid-scale. For central London postal Districts the number is replaced with two characters of the form NA.
PCCnt	int	Count of live postcodes within the district object as of October 2012.
AnomCnt	int	The count of postcodes within the district object that are of a different district to the containing district object based on their text and are not Vertical Street or non-geographic postcodes.
RefPC	Char(10)	The postcode within the district that is closest to the postcode weighted centroid of the district.
X	int	The x co-ordinate of the district reference postcode point in the British National Grid co-ordinate system
Y	int	The y co-ordinate of the district reference postcode point in the British National Grid co-ordinate system
Sprawl	Char(100)	The name of the dominant GeoSprawl object that the District overlaps as measured by shared postcodes. Nine Districts contain only postcodes that do not fall within a sprawl object; these sectors contain empty strings in the GeoSprawl field.
Locale	Char(100)	Within the largest 24 urban sprawls Britain (Bristol, Birmingham, Edinburgh, Glasgow, Leeds, Liverpool, Manchester, Nottingham, Sheffield, London, Cardiff, Hull, Bradford, Newcastle, Leicester, Belfast, Black Country, Brighton, Coventry, Derby, Reading, Southampton, Potteries) the name of the dominant GeoLocale (suburb) that the District overlaps as measured by shared postcodes. Outside of these largest 24 Cities the field contains an empty string

Geographic Postal District Points

File Name: PostalDistrictP

Record Count: 2,736

For every record in PostalDistrict this dataset contains a complementary record with the postal district represented by a single point. The point is that of the single unique postcode within the district that is closest to the postcode-weighted centroid of the district.

Column Name	Type	Description
DistID	int	Unique ID assigned to the postcode sector
PostDist	Char(4)	The Postal District of the sector without any spaces

Geographic Postal Area Regions

File Name: PostalArea

Record Count: 120

This file contains regions representing all the Postal Areas covering Great Britain as of October 2012.

Column Name	Type	Description
AreaID	int	Unique ID assigned to the postcode Area
PostArea	Char(2)	The Postal Area of the sector without any spaces
AreaName	Char(50)	The Postal Area Name
PCCnt	int	Count of live postcodes within the Area object as of October 2012.
RefPC	Char(10)	The postcode within the district that is closest to the postcode weighted centroid of the Area.
X	int	The x co-ordinate of the Area reference postcode point in the British National Grid co-ordinate system
Y	int	The y co-ordinate of the Area reference postcode point in the British National Grid co-ordinate system
Sprawl	Char(100)	The name of the largest GeoSprawl object that the Area overlaps as measured by shared postcodes. For example the sprawl with the most 'RM' postcodes within it is London and consequently the RM postal area is tagged with the 'London'.

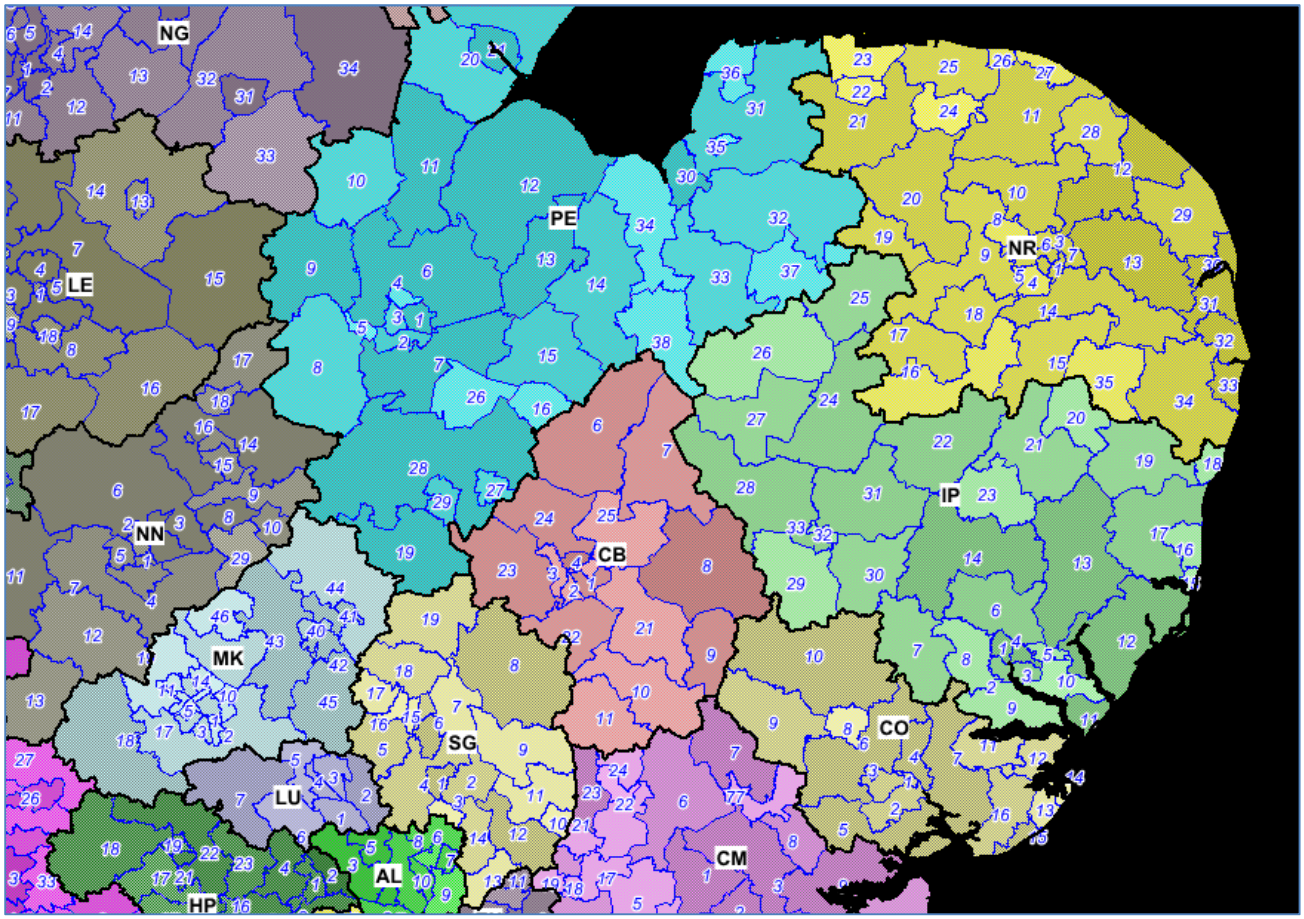
Geographic Postal Area Points

File Name: PostalAreaP

Record Count: 120

For every record in PostalArea a complementary record with the postal Area represented as a single point to represent the Area. The point is that of the single unique postcode within the Area that is closest to the postcode-weighted centroid of the Postal Area.

Column Name	Type	Description
AreaID	int	Unique ID assigned to the postcode Area
PostArea	Char(2)	The Postal Area of the sector without any spaces



Postal Areas and Districts in East England. Map contains Ordnance Survey data © Crown copyright and database right 2012, Royal Mail data © Royal Mail copyright and database right 2012, National Statistics data © Crown copyright and database right 2012, Postal Boundaries © GeoLytix 2012 copyright and database right 2012.

Non-Geographic Sectors

File Name: PostalSectorNG

Record count: 1,666

Within the United Kingdom's postal system certain postal sectors are reserved for special 'large users' and for PO Boxes. These are included as a csv file for the sake of completeness. The non-geographic sector file contains data for the whole of the United Kingdom together with Jersey, Guernsey and the Isle of Man.

Column Name	Type	Description
NGSectID	int	Unique ID assigned to the postcode sector, unique both within this file and also unique if merged with the geographic postal sectors file, UID's begin at 20,000.
RMSect	Char(6)	The Postcode Sector text in Standard 'Royal Mail' format, this is how a sector would naturally be written and is made up of the Postcode District followed by a single space and then the sector number. All text is capitalised. Unique if merged with the geographic postal sector file – no sectors appear in both this file and geographic file.
StrSect	Char(6)	The postcode sector text with all spaces removed. This should be the default style used when matching across multiple datasets as it removes any reliance on text formatting.
DateLive	Char(10)	Date the postal sector went live where recorded.
Type	Char(20)	Type of non-geographic sector, contains either the text 'PO Box' or 'Large User'.
Desc	Char(250)	Description of the non-geographic sector, for large user sectors a description of the user is usually given together with a rough indication of where the user is; for PO box sectors an indication is given of the towns and Postal Areas/Sectors covered by the PO Box service.

'Out of Sector' Postcodes

File Name: OutOfSectorPostcode

Record count: 76,683

This dataset contains all postcodes from the latest release of Code-Point™ Open covering Great Britain where the text of the postcode implies a postal sector different to that recorded by a point-in-polygon query conducted on the latest Postal Sector objects. This file also includes 248 postcodes where no co-ordinates are attributed to the Code-Point™ Open record, for these records the GeoSect is an empty string.

Column Name	Type	Description
Postcode	Char(7)	The postcode from Codepoint with all space characters removed.
TextSect	Char(5)	The sector based on the text of the postcode. Derived by trimming the final two characters from the full postcode.
GeoSect	Char(5)	The sector object, within which the postcode falls, based on a point-in-polygon query using the latest sector region objects.
PQI	int	The PQI indicator from within Codepoint Open. These values have the following meanings as described in the OS Code-Point™ Open User guide. 10 Grid reference is within the building of the matched address closest to the postcode mean - determined automatically by GB OS 20 Grid reference is within the building of the matched address closest to the postcode mean - determined by visual inspection by GROS (Scottish postcodes) 30 Grid reference approximate to within 50m of true position (100m for some postcodes relating to developing sites) 40 Grid reference is the mean of the positions of previously matched addresses, which have subsequently been deleted or recoded in ADDRESS-POINT (RARELY USED) 50 Estimated position based on surrounding postcode coordinates, usually to 100m resolution but 10m in Scotland 60 Grid reference based on Postcode sector mean - mainly PO boxes 90 No Grid reference coordinates available
OOSFlag	int	A flag indicating if the postcode is an 'out of sector' postcode. Set to 1 if true.
x	int	The x co-ordinate of the postcode point in the British National Grid co-ordinate system
y	int	The y co-ordinate of the postcode point in the British National Grid co-ordinate system
NGFlag	int	A flag indicating if the sector indicated by the postcode text is recorded as a 'non-geographic' sector as defined by the Royal Mail. Set to 1 if true.
VSFlag	int	A flag indicating that the postcode point of the record is co-incident with another postcode point. Commonly referred to as 'vertical streets'. Set to 1 if true.
AnomFlag	int	A flag indicating that the postcode point appears to be 'anomalous', that it is a single postcode outlier surrounded by many other postcodes of a different sector and is significantly spatially displaced from other postcodes that it shares a sector with. Set to 1 if true.