UK Data Archive Study Number 9183 - Acorn Postcode-Level Directory for the United Kingdom, 2023

Acorn Technical Guide

Version: Document release:

2023 11/07/2023

© CACI 2023

CACI Ltd., Kensington Village, Avonmore Road, London W14 8TS





Table of Contents

Introduction	4
What is Acorn?	4
Using Acorn	4
The Acorn Solution	6
Objectives	6
Outline of Methodology	7
Segmentation using the latest advances in AI	8
Acorn Data Sources	9
Acorn Data Flows	
Examples of Additional Algorithms	11
Age-Limited Housing	11
Manual Allocations	12
The Acorn Segmentation of Newly Built Housing	12
Acorn Driver Variables and their Sources	13
Acknowledgements	14
Quality Management	15
Quality Control of Input Datasets	15
Assessment of Outputs	16
Quality Control of Process for Annual Updates	17
Additional information	18
The UK postcode system	18
Special populations	18
Data Sources	19
Postcode Spine	19
Address Spine	19
PAF	19
The Land Registry – England & Wales	19
Registers of Scotland	
Housing for Older People	
Private Rental Information	
Social Housing	
Planning Data	
Population Density Indicator	





9	Student Accommodation	
I	EPC	20
(Other Communal Populations	21
(Geospatial data	21
Other P	Postcode Spine data	21
I	Benefits Data	21
I	Paycheck	21
I	Paycheck Disposable Income	21
9	StreetValue	21
ſ	Population projections	21
The Cer	nsus	22





Introduction

Acorn is the leading geodemographic segmentation of residential neighbourhoods in the UK. It classifies each postcode in the country into one of 65 types that give a distinctive picture of the kinds of people who live in an area, their attitudes and how they behave.

The Acorn segmentation has a hierarchical structure. The 65 types aggregate into 22 Acorn groups which lie within 7 descriptive Acorn categories at the top level. Acorn types are also further subdivided into 282 detailed micro-segments. These micro-segments provide a finer level of resolution that is useful for more specialist analysis.

This document complements the Acorn User Guide, the Acorn Knowledge sheet, and the Acorn microsite and gives further technical background on the Acorn segmentation.

What is Acorn?

Acorn is a powerful targeting tool that combines geography with demographics and lifestyle information, and the places where people live with their underlying characteristics and behaviour, to create a tool for understanding the different types of people in different areas throughout the country. It enables users to understand the kind of people living in their area, buying their goods, or using their services.

Acorn was the first commercial geodemographic segmentation and has maintained its position as the leading segmentation tool in the UK for over 40 years. It has been continually developed across its many versions, taking advantage of new data sources as they become available together with regular reviews of methodological approaches, and has maintained its leading position among small area segmentations in the UK.

In 2023 the Acorn classification was completely rebuilt to create a new segmentation, having new categories, groups and types that reflect recent social, demographic and economic changes. This new product uses leading edge AI techniques to achieve a level of discrimination not previously attainable.

The present version of Acorn draws on a wide range of data sources, both commercial and public sector open data and administrative data. See the section entitled Data Sources for more details.

Using Acorn

The analysis of significant social factors and population behaviour enables Acorn to provide an in-depth understanding of different types of people. Whilst some names refer to types of residential property as a label it should be remembered that Acorn is essentially a segmentation of people and their characteristics. Acorn provides a detailed understanding of the people who interact with your organisation. It helps you learn about their relationship with you. This knowledge gives you the opportunity to target, acquire and develop profitable customer relationships and improve service delivery.





Geodemographic targeting helps government and businesses pinpoint the people who are most likely to benefit from their products or services, and avoid those who will not, allowing them to improve their understanding of customers, target markets and determine where to locate operations.

- Financial organisations use Acorn to understand their customers, cross-sell their product range, set branch targets, identify loyal customers, support marketing strategy and marketing activities, and plan their network strategy.
- Retailers use Acorn to locate stores, plan merchandising and product ranges, assess refurbishments, support marketing strategy and marketing activities, forecast turnover and demand, and target local markets for stores.
- Media owners use Acorn to support advertising sales, evaluate sales potential, and develop new markets.
- Digital Marketers use Acorn to target display advertising, direct mail campaigns and other digital campaigns.
- In Fast Moving Consumer Goods (FMCG), Acorn is used to drive customer communication, instore marketing, merchandising and product distribution.
- The Public Sector uses Acorn to target services to areas of need and inform policy decisions.
- Residential developers and property companies use Acorn to understand consumers, neighbourhoods and inform strategic decisions regarding investment using robust evidencebased data.
- Charities use Acorn to identify and support those in need and understand the best way to communicate with different types of supporters, improving the effectiveness of fundraising.





The Acorn Solution

Acorn segments postcodes into 7 categories, 22 groups and 65 types. Types are further subdivided into 282 micro-segments that can be used to add an extra level of precision to the segmentation for specialist analyses.

Most people in the UK live in private households. Six of the Acorn categories, comprising 21 of the groups and 61 of the types, represent the population in private households. The twenty-second group represents other kinds of postcode, primarily communal population who live in various kinds of institutions rather than in private households, and postcodes with no resident population.

The communal population occupies types 62, 63 and 64.

- Type 62 represents students living in communal establishments such as halls of residence.
- Type 63 represents active communal populations, such as military personnel, hotels, and hostels.
- Type 64 includes less active communal population such as those in care homes, medical establishments, and prisons. These various kinds of establishments are distinguished at the micro-segment level.

Type 65 represents postcodes which have no resident population, either in private households or communal.

Objectives

The Acorn product in the UK is a segmentation of full postcodes and is updated on an annual basis (or more frequently when appropriate). The types and the segmentation methodology were completely rebuilt in 2023 to reflect the demographic, social and economic changes in the UK, to take maximum benefit of the changing data landscape in the UK both now and into the future, and to utilise the latest advances in AI techniques.

Many new data sources are now available that allow greater precision and geographical detail than has ever been possible with census data. Geographically detailed data about house sales and values, property details, rental properties, care homes, accommodation for the elderly and other attributes are now available from non-census sources. The present drive towards the release of ever more government information as Open Data is likely to lead to continually increasing availability of new data sources on the characteristics of the population of the UK for small geographical areas.

CACI has put in place a methodology that takes full advantage of the data sources available now and is designed to be adaptable to the rapidly changing data landscape of the future, adding data sources when they give benefits and down-weighting or dropping those that become out of date.

The Acorn methodology is based around a model that uses the individual postcode as its basic building block. It combines data from a range of data sources – house sales, house rentals, accommodation designed for elderly people, high rise social housing, other housing lists, care accommodation, student accommodation, information about residents, benefits claimants, census, etc. – to produce accurate and up to date estimates of the characteristics of the people who live in that postcode. These estimates are used to classify the households and postcodes into the 7 Acorn categories, 22 groups, 65 types and, at a finer level than types, 282 micro-segments.





Outline of Methodology

CACI's approach to creating the Acorn classification uses a methodology that responds to limitations inherent in traditional approaches to segmentation and allows agile future adaptation to the rapidly changing data landscape in the UK.

The traditional approach to geodemographic segmentation is in broad terms similar for all classifications. The common themes are:

- 1. Data is compiled for the local geographic units. Either aggregated from addresses into postcodes or from census output areas.
- 2. The data is fed through statistical software to perform a multivariate segmentation.
- 3. The resulting segmentation is analysed, insight collected for the types, which are then given labels and described at more length.



Figure 1. The traditional approach to geodemographic segmentation

Our approach to geodemographics starts by separating the process of defining the segments used to describe society (segmentation) from the process of assigning postcodes to these segments (classification). The approach then allows the process of assignment (classification) to be done using many different algorithms.







Figure 2. The modern approach to geodemographic segmentation

The first benefit of the new approach is that any future data can be incorporated into the classification process. This brings about improved updating and 'future-proofs' the solution.

The second benefit is that it is not essential to have data for all the United Kingdom. If one can improve the Acorn solution for only part of the country without losing anything elsewhere, then it is clearly advantageous to do so. With devolved government a great deal of Open Data is released covering only England, Scotland, or Wales. The traditional approach could not use this data since it required the same data for every postcode. Our approach allows this data to be used effectively.

Segmentation using the latest advances in AI

The process of constructing geodemographic segments, "segmentation", essentially involves finding patterns of similarity among postcodes. In particular, each postcode is quantified by a collection (vector) of characteristics, which can be represented as a point in a high dimensional mathematical space in such a way that postcodes with similar characteristics will be represented by points nearby one another. High density clusters of points define the segments, i.e., all those points are from postcodes with a similar pattern of characteristics.

Traditionally clustering algorithms used to find these high-density point clusters difficult to operate optimally in these high-dimensional spaces due to the "Curse of Dimensionality", wherein the mathematical space becomes increasingly sparse and mathematical distances between points in the space become increasingly similar. However, in common with most datasets, the points represented by





the postcodes have a natural tendency to lie on a lower-dimensional surface within the higherdimensional space.

Recent advances in Artificial Intelligence have created new Deep Learning algorithms that are able to effectively produce a "map" of these surfaces. The map captures all the meaningful relationships in the distribution of points at a much lower dimension, which means we can apply traditional clustering techniques to the map, instead of the original high dimensional space, to generate more optimal segments.



Figure 3. Creating optimal Acorn segments through Deep Learning and Clustering

Acorn Data Sources

The two main sources of data used to build Acorn are CACI's proprietary Postcode Spine and the Census data.

The Postcode Spine is comprised of CACI's Address Spine data which has been aggregated to postcode level and combined with other postcode level datasets.

The Address Spine has been built by combining a number of property level datasets at an address, as defined by PAF's UDPRN, as well as using Geospatial data to infer information. Data includes the Land Registry, Registers of Scotland, Energy Performance Certificate data, benefits data, population density, planning data, care homes, social housing, other rental property information and geospatial data.

This data has been combined with other postcode level data such as CACI's proprietary products Paycheck and Paycheck Disposable Income (our income models) and StreetValue (a house price model) as well as population projections and benefits data. In addition, CACI has created enhanced databases, including the location of prisons, age-restricted housing, high-rise buildings and student halls.

The other main source of data is the UK Census, conducted every ten years across the UK population, it covers a wide range of topics, and provides a useful and detailed dataset. It provides an excellent means of calibrating alternative information sources that are incomplete.





For more information see the section on Data Sources under Additional Information.

Acorn Data Flows

The below diagram represents the building blocks for the latest version of Acorn. The postcode spine has been developed using a number of sources of data aggregated to postcode level and often from multiple different sources e.g., house type and tenure. This data is combined with the Census to create the data input file which is then used to build the Acorn segmentation.







Examples of Additional Algorithms

CACI makes use of a great many data sets, each of which might only cover part of the country or apply to only certain kinds of location. The following examples illustrate some of the concepts we apply to get the most out of this data.

Age-Limited Housing

CACI has a database of over 2,000 social housing developments built to exclusively house older people and over 6,000 private sector developments restricted to elderly people. This has enabled us to accurately identify and more accurately classify this important sector of the population.

The data is used in two ways. Firstly, the data provides information not only about age, but also about the type of housing, its tenure and its size. Other sources (discussed later) tell us the value of those that are owner-occupied, how long they have been occupied and similar factors. In summary a substantial proportion of the input variables traditionally used in segmentation are known for each individual household.

Secondly the data is used to improve the information about surrounding households and postcodes. In many circumstances the census data can lead to false impressions because, by reporting data only for census output areas, it merges together disparate types of housing or people.

This problem can be corrected. Consider the example below.



The dark blue lines are the boundaries of census output areas. The pale blue streets are known to be owner occupied housing that is restricted to elderly residents.

To the east there is detached housing that is owner occupied and relatively affluent. Various data sources indicate these are wealthy couples, and not elderly people. They form part of the same census OA and the census merges these couples with the elderly people.





To the south there is a census output area with mainly less affluent terraced housing, often occupied by families with children. However, a small part of these streets is occupied by elderly people.

Clearly the census data in both cases gives a distorted picture. However, since details are known for each of the addresses forming the elderly housing it is possible to correct for this and derive true input variables to accurately segment the surrounding postcodes.

Similar corrections can be made to improve data for postcodes in the vicinity of prisons, student halls of residence, and other populations with known characteristics.

Manual Allocations

There are a small number of situations where the manual allocation of Acorn types is appropriate. Manual allocations are made for some student halls of residence, for prisons, and for postcodes we have identified as containing anomalous data such as rare cases of misleading indications of house price on the Land Registry.

The Acorn Segmentation of Newly Built Housing

Census data provides little reliable information about the characteristics of significant housing developments built after census day.

In these cases, the traditional approach to geodemographics must be replaced by an alternative that makes best use of all information sources available. The broad approach is similar; estimates are made of the variables that drive allocation to Acorn types, and this demographic profile of the postcode is used to allocate it to the appropriate type.

Housing newly built since the census ('new housing') is identified as being either infill housing, where typically a small quantity of new housing is intermingled with existing housing, or entirely new housing, where typically a larger quantity of new housing is built in a locality not previously populated.

The method of distinguishing between these two cases works at the level of a full postcode. If a postcode consists mainly of new housing, it is classed as being new, while if only a minority share of the postcode consists of new housing it is classed as infill housing.

Infill housing is provisionally assumed to be of similar type to the existing surrounding housing. Then, wherever possible, modifications are made through the use other data sources. Following such modification of the data, postcodes with infill housing are allocated to the Acorn segmentation using the same update procedure as for housing which existed at the time of the census.

For new postcodes the modelling approach is independent of census data. Rather than multivariate segmentation techniques, entirely different regression modelling approaches are used. This is based upon known information about the actual postcode, typically taken from the Land Registry, population density and other up-to-date small area information sources.





Acorn Driver Variables and their Sources

The below is a list of the key data inputs which drive the Acorn segments and the sources where the data has been derived from where Spine is CACI's Postcode Spine.

Label	Sources
StreetValue: ratio to UK mean	Spine
StreetValue: compared to surrounding 10km	Spine
Qualifications: none / level 1	Census
Qualifications: degree or higher degree	Census
Higher managerial/professional	Census
Routine or semi-routine occupation	Census
Full-time students	Census
Owner occupied	Census
Privately rented	Census, Spine
Social renting	Census, Spine
Persons per household	Spine
Rooms per household	Census, Spine
Paycheck Income	Spine
Paycheck Disposable Income	Spine
Universal Credit claimants as proportion of population	Spine
Proportion of population aged 0-4	Spine
Proportion of population aged 5-11	Spine
Proportion of population aged 12-17	Spine
Proportion of population aged 18-34	Spine
Proportion of population aged 35-69	Spine
Proportion of population aged 70+	Spine
Family households: couple with dependent children	Census
Family households: couple with no children	Census
Households with all (1+) resident children non-dependent	Census
Empty nesters	Census





Single person household: Pensioner/Older person	Census
Single person household: under 65	Census
Single parent with dependent children	Census
Multiple Occupants (not related or in a relationship)	Census
Proportion of addresses that are student private rental	Spine
Detached	Census, Spine
Semi-detached	Census, Spine
Terraced	Census, Spine
Flat converted	Census, Spine
Proportion of bungalow	Spine
Population density: population within 2km	Spine
High rise	Spine
House price: rate of change	Spine
South Asian population	Census
Chinese population	Census
Black African population	Census
Black Caribbean population	Census

Acknowledgements

We acknowledge the support and assistance of local authorities, other public sector bodies and academic institutions in the supply of data in the Acorn development process. Their help has been invaluable in ensuring that Acorn uses the latest relevant data.





Quality Management

This section briefly outlines some of the quality assurance processes used in the build and maintenance of Acorn. The quality process is split into three broad areas:

- 1. Assessment of input data sources
- 2. Quality control of the build process
- 3. Assessment of the final classification, and other associated data

These three areas are described below.

Quality Control of Input Datasets

The main methods used to assess the quality of data sources are:

Abbreviation	Description
OQ	Supplier's own documentation of their quality control systems
IQ	Independently conducted quality assessments
OS	Cross-check against other suppliers of the same information
IA	Cross-check against independent national or regional aggregate data sources
VC	Visual checks (using internet searches, maps, or similar)
IP	Check against independently collected panel information
со	Consistency checks with other data sources
IC	Internal consistency checks

The methods used for each of the main data sources used to build Acorn are:

Data Source	Methods Used	Notes
The Land Registry (Price Paid data)	OQ, IP, IA, CO, VC	Merge and check against independent panel, to measure agreement on tenure and house type. Validate 'newly built' indicator against house build statistics
Energy Performance Certificate	OQ, IP, IA, VC, CO	Merge and check against independent panel, to measure agreement on tenure, number of rooms and house type. Visual checks on floor counts.
Housing for older people	OQ, VC, CO	Check against sources of age information, and online checks.
Private rental information	OQ, IP, VC	Check tenure, number of rooms against that collected on independent panel. Map checks for student rent hot spots
Social housing	OQ, OS, IA, IP	a. Cross-check of counts with census and other aggregate information
		b. Merge and check against independent panel





High-Rise Buildings	VC	Manual checks, either online or by location visits
Population density indicator	OQ, VC	Mapping checks
Geospatial data	0Q, IQ, VC	Mapping checks, for example, to see if addresses fall within building outline
The Census	OQ, IA	Cross-checks against, for example, supply-side statistics for social housing.
Student accommodation	VC	Internet research: check that postcodes identified as student accommodation by automated methods are correctly allocated
Other communal populations	VC, IQ	PAF, mapping checked (for, e.g. presence of prisons) Use of ONS pre-census assessment of publicly available databases of care homes. Online validation of a sample of care home data

Assessment of Outputs

Two main methods were used to evaluate and monitor the segmentation:

- Peer review from a wide panel of internal CACI experts, with a broad and deep experience of client use of Acorn segmentation.
- An analysis of gains scores specifically GINI scores which measure the effectiveness of Acorn in discriminating across a wide range of variables.

The gains score tests take data sourced from Market Research Surveys, which measure multiple aspects of consumer behaviour.

For each type of behaviour, a GINI (Gains) score is calculated. The general approach to calculating gains is based upon determining a propensity ordering of Acorn types, separately for each of the target measures. Once the preference ordering of types is determined for each target measure, a graph can be drawn showing the relationship between the proportion of target (those people who have the target characteristic) reached against the proportion of non-target (those people who do not have the target characteristic) reached, as Acorn types are progressively selected according to the priority order for this characteristic. The GINI score is calculated from the area under this curve and gives a measure of discrimination of the segmentation that is independent of the global penetration of each characteristic.

The GINI calculation incorporates a process of splitting the file randomly into two separate parts, one of which is used to determine the priority ordering and the other of which is used to calculate the gains based upon this priority ordering, in order to avoid upward sampling bias on the score.

The new Acorn segmentation out-performs previous segmentations in the key type of variables tested, such as occupation and lifestyle attitude.





Quality Control of Process for Annual Updates

Each stage in the processing is carefully checked. Process checks include:

- Check that postcode updates have been applied correctly.
- Check stability compared to previous version, with a target that change at the level of Acorn Types should normally be less than 10% of population.
- Check the level of coding, with a target that all postcodes with population have a valid Acorn code.
- Check that the majority of changes to Acorn type are to similar Acorn types, and in particular that change of group should be much lower than the level of changes between types, and changes between categories are lower still.
- Check that all new postcodes are allocated using the CACI newbuild methodology.





Additional information

The UK postcode system

Full postcodes in the UK provide a very detailed and precise geographical breakdown. The median size of a residential postcode is just 13 households and 31 residents. While the size of postcodes does vary significantly, only 1% of residential postcodes have more than 64 households and 154 residents.

Postcodes in the UK are updated by Royal Mail, typically two or three times per year, in response to changes in the housing stock. Significant developments of new housing will normally have new postcodes allocated to them. Postcodes that become too large because of infill housing developments may be split. Postcodes that represent demolished housing may be withdrawn from use.

Reorganisations of local subsections of the postcode system are also sometimes required if, for example, no more postcodes are available to be allocated within a postcode sector. As a result, the postcode system is maintained in such a way as to represent the current distribution of addresses in the UK without, for example, any individual postcode being allowed to grow excessively large because of new housing developments.

Special populations

Acorn Group V comprises postcodes or output areas which contain predominantly communal establishments, and those that do not contain residential population.

After student population has been reclassified as being in private households, any postcode estimated to have more than 50% of its population living in communal establishments is allocated to Group V. Postcodes containing communal population are almost always defined in such a way that only one kind of communal establishment is present. The dominant type of communal population within the postcode is used to classify the postcode as:

Type / Segment		Micro-segment number
Type 62: Students Communal Population		275
Type 63: Active	e Communal Population	
	Segment 63.1 Defence Establishments	276
	Segment 63.2 Hotels and Holiday Accommodation	277
	Segment 63.3 Other Homes and Hostels, Travellers	278
Type 64: Inacti	ve Communal Population	
	Segment 64.1 Care Homes	279
	Segment 64.2 Medical and Nursing Establishments	280
	Segment 64.3 Prisons	281





There is a further type in Group V, which is used to classify business addresses, PO Boxes etc.:

Type / Segment	Micro-segment number
Type 65: Without resident population	282

Data Sources

Below we describe the data sources used in building Acorn. This should not be considered a definition of all inputs to Acorn. The list may change. As new sources, particularly of Open Data, become available they may be added to the process.

It is important to remember that not all these sources are used in every instance. Moreover, different parts of a data source may well be used in different ways in different algorithms.

Many of these data sources provide substantial information for addresses or households. Some provide aggregate information for households and some for larger geographic areas. They are not all used in the same way. Clearly, when building a postcode segmentation, sources of data at these levels of geographic detail will assume greater importance.

Postcode Spine

The following data sources are used to build CACI's postcode spine.

Address Spine

The Address Spine is a CACI product which has been built by combining a number of property level datasets at an address as defined by PAF's UDPRN, as well as using Geospatial data to infer information. The list below covers the datasets used to develop the Address Spine:

PAF

The Postcode Address File (PAF) is a database that contains all known Delivery Points (addresses) and postcodes in the UK. PAF is a collection of over 31 million Royal Mail postal addresses across 1.8m postcodes. Each one has a unique reference called a UDPRN (Unique Delivery Point Reference Number) which can be used to tie pieces of information together about the same property.

The Land Registry – England & Wales

The Land Registry for England and Wales provide information at address level about housing sold in each postcode. It thus covers owner occupied housing or housing that has been purchased and privately rented. It does not include social rented housing.

Information from the database includes the date the purchase completed, the price paid and other attributes such as the type of house. Data is available going back to 1995.





The data is used to determine property types as well as an affluence indicator based on house prices. In the current version we are also using the relative house price increases as a measure of gentrification of an area.

Registers of Scotland

Registers of Scotland provide similar, but not identical, information to the Land Registry. The respective organisations operate within different legal systems and their administrative data are thus not identical. CACI use data going back to 2001.

Housing for Older People

CACI uses a database of age-limited housing, which is housing required to have an elderly occupant. The information gives the postcode, the name of the facility, and information on the number and kind of properties.

Private Rental Information

We use data giving the address of private rental properties advertised through a number of major internet property sites. This is used to improve estimates of the level of private renting in each postcode to improve tenure information.

Social Housing

Information on social housing is obtained from two sources:

The *National Register of Social Housing* (NROSH) is an openly available dataset of social housing in England and Wales. Data items included in the dataset are address, together with property attributes such as type of property, room count and floor count.

The Land Registry Commercial file is a list of companies which are the owners of commercial property. It is used in combination with a list of Social Housing Providers to understand the location of social housing and assist with the identification of tenure.

Planning Data

CACI have licensed planning data which provides regularly updated information on planning applications and consents including house extensions which help us identify areas of gentrification.

Population Density Indicator

We use a CACI proprietary population density measure calculated at individual postcode level using information about the surrounding population. The basic measure used is a count of the residential population within 2km of each unit postcode.

Student Accommodation

Student accommodation is defined using a combination of data from PAF showing Halls of Residence and data licensed from a property portal that shows properties which have been advertised for rent to students. High concentrations of students in private rental properties are split across a number of Groups and Types also containing high proportions of young adults.

EPC

The Energy Performance Certificate (EPC) data is available at a property level and provides information related to energy efficiency and usage when a property is sold or rented. The data is used to help understand property information, such as property type, size of property, number of rooms etc and is used amongst other things, as an indicator of affluence.





Other Communal Populations

CACI uses a database of care homes in order to identify those postcodes that contain this particular kind of communal population. Other types of communal populations such as prisons were identified in the 2011 census data and have since been updated using web research such as data published by the Ministry of Justice. These postcodes are used in the initial process of splitting census data down to individual postcodes and are assigned to Acorn Type 64 (Inactive Communal Population).

Geospatial data

Geospatial data is used for enhancing the property data to infer property type and number of floors where partial information is known about a street from sources such as the Land Registry and EPC. We link address level data from the Address Spine with boundary data and land parcel data to infer for example house type and high-rise buildings.

Other Postcode Spine data

In addition to the data aggregated to postcode level from the Address Spine, CACI also use a number of data sources which are available at postcode or Output Area (OA) level to develop our Postcode Spine. These are as follows:

Benefits Data

Universal Credit data is being used in Acorn this time to replace the Job Seekers Allowance (JSA) data used previously. The data is available at OA level and used as an indicator of deprivation.

Paycheck

Paycheck is a CACI proprietary postcode data model that models gross household income at postcode level across the UK. It is based upon government data sources, including Census and Land Registry information, together with research from ISPOS's Financial Research Survey as well as CACI's Ocean database and StreetValue.

Paycheck Disposable Income

Paycheck Disposable Income allocates 16 non-discretionary outgoings to the Paycheck average household income. This results in a net disposable income per household at a postcode level. Again, this is a proprietary product that uses government sources to determine likely outgoings.

StreetValue

StreetValue is a proprietary data model that estimates property values for all individual residential postcodes in the UK. It is based upon analysis of Land Registry data in England and Wales since 1995 and data from the Registers of Scotland since 2005, analysed together with census data, other government data and CACI's range of geographic data sources.

Population projections

CACI's proprietary population projections are used to create postcode level age profiles, i.e., how many people in a postcode are in each age band. They are based on small area population profiles using census data and updated annually based on published high level control figures.

The Postcode Spine is then combined with the census data to create the Acorn base data which is used to build the segmentation.





The Census

All UK residents and visitors are asked to fill out a questionnaire every ten years covering a wide range of topics, designed primarily to be of use in the development and allocation of public services.

The census provides a useful and detailed data set due to its completeness. It also helps as a means of calibrating alternative information sources that are incomplete.

Census questions include information about family composition, qualifications, type of occupation and ethnicity as well as questions related to housing.



