GLAINTELLIGENCE

GLA Yield Calculator

September 2014, updated June 2019

Introduction

Existing methods of predicting child yield from new developments have a number of limitations. The most commonly used approaches are based on either:

1. Census commissioned tables on housing characteristics of the population. Tables exist by local authority for both the population as a whole and those who have changed address within the previous 12 months.

2. Surveys of new housing.

Both methods have drawbacks, and neither is considered to give a picture of yields that can be reliably used across London. The GLA Intelligence Unit has received a number of requests to assist in improving estimates of yields across London. To that end a population yield calculator has been developed which aims to provide users with geographically-specific estimates of yield from new development. The calculator makes use of 2011 Census data in concert with the GLA's London Development Database (LDD).

The approach taken assumes that past observed trends will continue into the future.

The yield calculator is designed to provide estimates of the number of people who can be expected to live in a new housing development once it is completed. It also provides some detail on the age structure of that population. However, the calculator is not intended as a population projection tool and does not indicate how the development impacts population in the wider context. Where the calculator is used as part of work to assess additional service demand arising from new development users should consider the impact of the development on the population of surrounding areas to ensure the net change in demand for the wider area is properly accounted for.

Estimates derived using this calculator should be referenced as 'GLA Experimental Statistics'.

Note: The calculator is only available in the XLSX format (Excel 2007 and later). The calculator uses formulae and features not available in earlier versions of Excel and therefore cannot be converted.

Calculating Yield

The yield calculator works as follows:

a) User inputs number of 1, 2, 3 and 4 market and intermediate and social units into the calculator.

NOTE: Intermediate housing is included with the market rather than social units. This is because the census dataset includes shared ownership under the owner-occupied grouping which itself is market tenure.

b) User selects method of aggregation by sub-region and PTAL rating.

NOTE: PTAL 3-4 should be used to calculate yield for developments in outer London with PTAL 5-6. This is because PTAL 5-6 in outer London includes only 4 sites, a sample size too small to reliably calculate an average yield.

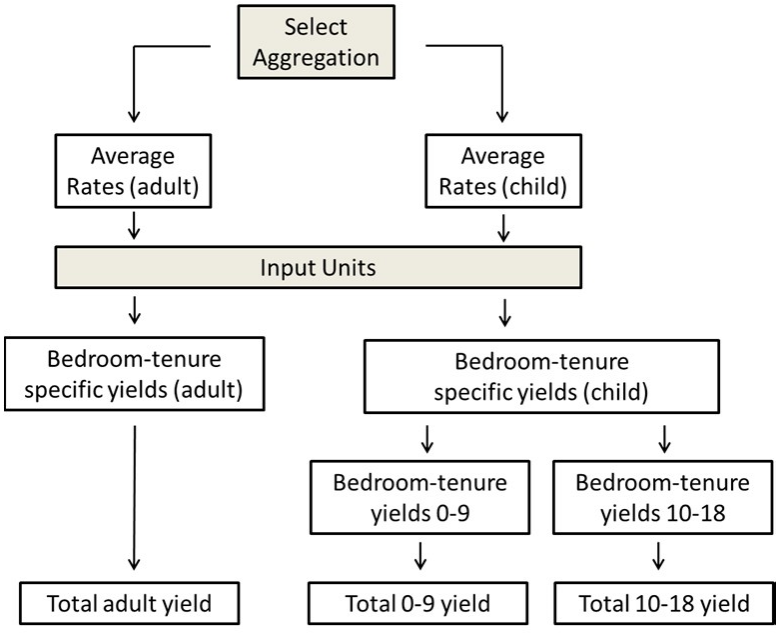
c) Calculator identifies those output areas (out of 112 previously defined) which meet the aggregation criteria (for example -are in outer London and have a PTAL of 2)

d) For each of the eight different bedroom-tenure pairs an average yield rate is calculated from the output areas identified in step (c). The output areas are weighted according to the number of units of each type they contain.

e) The rates calculated in (d) are multiplied by the units input in (a) to give a total child yield for each bedroom-tenure pair.

f) This total yield is then disaggregated into 0-3 year old's, 4-10 year old's, 11-15 year old's, 16-17 year old's, 18-64 year old's and 65+ using rates derived from census data.

g) Steps (d) and (e) are completed for the adult population



Limitations

The calculator is limited by the available census data. The age groupings in table LC4103EW dictate the age groupings for the calculator output. The GLA requested from ONS data for 5-year age bands for the entire population, however this level of detail was deemed disclosive and the request was denied. The ONS would only provide the total adult population (19 and over) and no more detail for the child population.

Sample sizes, particularly when using some the present aggregations, can be quite small. The methodology identified 112 output areas to use as its base, however, these are not equally distributed across London and so some aggregations will have fewer output areas to create the average rates than others. Users should be particularly aware of this when using the borough selection tool and when creating their own aggregations.

Methodology

Overview of methodology:

1. Identifying sites using the London Development Database to identify new developments built between 2001 and 2011.
2. Identifying new 2011 Census output areas that consist wholly or primarily of new development.
3. Using census data to identify housing characteristics and population yield.
4. Creating geographical aggregations of output area data.
5. Building a tool for estimating child yield from new development based on characteristics of the development.

1. Identification of sites

The London Development Database (LDD) is a database of residential development in London since the mid-1980s. The 33 London Boroughs are required to submit, on a monthly basis, details of planning permissions granted for residential units. The system also monitors the progress of developments and records a date of completion when units are ready for occupation. The database encompasses all residential development including house-to-flat conversions through to large-scale housing developments and estate renewal projects. In March 2011 the LDD contained almost 56,000 developments of varying sizes across London.

The population yield project matches the LDD to census data to identify population characteristics in new housing development. It is therefore necessary to filter the records in the LDD to include only those completed within the intercensal period (April 2002 to March 2011). Within this range there were 42,500 developments completed. A further requirement of the dataset is that developments be of a sufficient size that their resident population is identifiable in the census data. A threshold of 50 units was chosen to identify large-scale developments. A total of 1,018 sites met these two criteria.

2. Matching sites to Output Areas

The methodology is predicated on the idea that census population data can be matched to LDD development sites. This is achieved by first identifying areas of significant population growth over the intercensal period. Significant population growth is indicated by the creation of new output areas. An output area is the smallest unit for which census data are produced and is a statistical geography. This means that its boundaries are defined by statistical (i.e. population numbers) rather than administrative (i.e.

electoral) concerns. As a result, when population increases in an output area beyond the allowed threshold the output area must be split to accommodate the larger population.

Therefore, a map of new output areas created following the 2011 Census gives a spatial distribution of concentrated population growth over the intercensal period. These output areas can then be cross-matched with the residential development sites previously identified from the LDD to produce a list of sites where residential development has taken place and population has increased. The assumption is that the population growth is a result of the development.

Of the 1,018 sites initially identified from the LDD, 267 coincide with one or more of the new output areas. In total there were 664 output areas with some part of a new development within its boundary.

Finally, in order to confidently infer that the population of the output area was representative of the population of the residential development a coverage threshold was established. Only output areas where 60 per cent or more were covered by a development site were moved forward. This reduced the number of output areas to 112; this is the final selection used in the yield calculator.

All subsequent processes in the calculator use the output area, rather than the LDD development site, as their unit of operation.

3. Incorporating Census data

For each of the identified output areas the following census data were extracted:

- The number of 1, 2, 3 and 4 bedroom properties in market[1] tenure
- The number of 1, 2, 3 and 4 bedroom properties in social[2] rented tenure
- For each of the above eight bedrooms-tenure pairs the number of children aged 0-9
- For each of the above eight bedrooms-tenure pairs the number of children aged 10-18
- For each of the above eight bedrooms-tenure pairs the number of adults aged 19 and over

These data were obtained from the following 2011 Census tables:

- LC4103EW – Number of bedrooms by tenure by age of dependent children
- LC4405EW – Tenure by household size by number of bedrooms
- CT0279 – Number of bedrooms by tenure for adults (commission by the GLA for this project)

Development specific-yield rates were calculated for each bedroom-tenure pair for all 112 output areas.

For example:

Output area E00000042 has 15 children aged 0-18 in 1-bedroom market households. There were 27 1-bedroom market households in the output area, so: 15 ÷ 27 = 0.481

Once yields for 0-18 year old's have been calculated for each of the bedroom-tenure pairings the total is divided into 0-9 and 10-18. This is done for each bedroom-tenure pair separately using the proportions outlined below. These proportions are derived from the census data for the 112 output areas by simply calculating how many children in total live in each of the bedroom-tenure pairs and then calculating what proportion are aged 0-9 and 10-18.

Proportion aged 0-9:

	1 bed	2 bed	3 bed	4 bed
Market	92%	86%	64%	60%
Social	78%	78%	53%	41%

[1] Market tenure includes owner occupied, shared ownership and private rented
[2] Social rented tenure includes RSL and Local Authority

4. Creating geographical aggregations

Data for individual output areas are aggregated together on the basis of common characteristics to give more statistically robust results. There is a trade-off between the number of groupings that are used and the sample sizes in each group. This in turn means there is a trade-off between specificity and robustness of the results.

This aggregation has been provided in the calculator in three 'off-the-shelf' formats. The raw data are also included, enabling users to create their own aggregations. The aggregation methods provided are:

a) London Plan sub-region
b) Public Transport Accessibility Level (PTAL)
c) Borough selection tool

The output areas included in the aggregation are weighted according to the number of units so that sites with more units are given greater prominence in the average rate than smaller sites.

Sample Sizes

M1 = Market tenure, 1 bedroom

S1 = social rented tenure 1 bedroom

Below are tables outlining the sample sizes for each method of geographic aggregation.

The calculator has been produced on the basis of sub-regions and PTAL level.

Sample sizes: Number of Sites for PTAL method

PTAL level	M1	M2	M3	M4	S1	S2	S3	S4
0 to 2	33	45	37	38	21	42	40	45
3 & 4	34	38	28	32	22	29	36	38
5 & 6	21	25	23	17	19	22	25	24

Sample sizes: Number of Units for PTAL method

PTAL level	M1	M2	M3	M4	S1	S2	S3	S4
0 to 2	1220	2310	425	252	380	591	296	59
3 & 4	2054	2243	247	75	373	554	268	74
5 & 6	1624	1685	288	37	364	395	130	66

Sample sizes: Number of Sites for sub-region method

Sub-region	M1	M2	M3	M4	S1	S2	S3	S4
Central	30	35	29	29	28	30	37	35
North	5	5	3	5	1	5	4	5
East	25	29	23	25	14	25	25	29
South	9	11	9	7	4	9	9	11
West	19	27	23	20	14	23	25	26

Sample sizes: Number of Units for sub-region method

Sub-region	M1	M2	M3	M4	S1	S2	S3	S4
Central	1649	1575	213	42	352	433	166	52
North	193	301	22	1	27	61	17	1
East	1715	2158	356	107	371	495	240	91
South	620	666	70	13	100	55	20	5
West	721	1538	299	201	267	496	251	50

Consultation Response

In April 2014 the GLA distributed a development version of the calculator to contacts in local authorities in London. Users were asked to comment on the calculator and suggest any changes or improvements they believed would make the tool more robust and more effective.

On the whole feedback was positive but a number of concerns were raised. Chiefly that the utility of a product would be greatly increased if it identified school-age groupings rather than 0-9- and 10-18-year old's. In response to this the GLA attempted to commission a table which would provide the data necessary for greater flexibility in the age-specific outputs. However, as described above this request was unsuccessful.

A second source of concern was the use of the PTAL level as a potential proxy for density. Users commented that PTAL levels may vary significantly across a site and that the levels themselves are often the source for debate during development negotiations. While these comments are accepted the final version of the calculator retains the PTAL-based aggregation (albeit in a more condensed form). This is because the intention of the tool is not to provide a single GLA-approved yield figure but rather to provide estimates which may inform discussion. As such it was felt that retaining the PTAL element was beneficial for context.

Finally, users noted that some form of bespoke aggregation tool, rather than pre-defined groupings, would be useful. To that end this version of the calculator contains underlying data that would allow a borough to analyse at borough-level or through a bespoke grouping of boroughs; however, this has not been included within the calculator sheet given the relatively limited amount of data that is available for some boroughs.

Appendix A: GLA Population Yield Calculator Workbook Metadata

This appendix contains detailed metadata about the yield calculator workbook. It lists all of the sheets that can be found in the workbook and explains in detail their contents and calculations. The sheets are as follows:

- Calculator
- LDD Data
- Census – All people
- Census – Children
- Census – Adults
- Age Split
- Calculations
- Borough Rates
- Sample Sizes

Calculator

User interface allows details of the development to be entered and a method of aggregation chosen. This data is used in the calculations described below in the 'Calculations' sheet. The resultant yield estimates are passed back to the 'Calculator' sheet and presented on the right side of the screen. A graph outlines the age distribution of the yield estimate and shows how yields differ for the three off-the-shelf aggregations.

LDD Data

This sheet contains an extract of the London Development Database. The data here constitute the 112 sites which coincide with new output areas and which cover at least 60 per cent of the relevant output area.

Fields:

Output Area	The area in which the development sits
Permission No	The LDD permission number
Comp Date	The completion date of the development
Borough	The local authority of the development

Census – All People

This sheet contains the 2011 Census commissioned table CT0279. This table was commissioned by the GLA specifically for this project. The table population is all usual residents and data are provided for local authorities and Output Areas in London.

Fields:

Tenure (Owned/Shared Ownership, Social Rented, Private Rented or Rent Free)

Number of bedrooms (1, 2, 3, 4 or more)

Census – Children

This sheet contains 2011 Census Local Characteristics table LC4103EW. This table provides the same data as CT0279 but the table population is dependent children only (further disaggregated into 0-9 year old's and 10-18 year old's). These data are found in columns B to I and K to R.

Fields:

Tenure (Owned/Shared Ownership, Social Rented, Private Rented or Rent Free)

Number of bedrooms (1, 2, 3, 4 or more)

Additionally, data on the number of households (units) by tenure and bedrooms are included. These data are taken from 2011 Census table LC4405EW. These data are in columns T to AB.

Fields:

Tenure (Owned/Shared Ownership, Social Rented, Private Rented or Rent Free)

Number of bedrooms (1, 2, 3, 4 or more)

Finally, a rate of dependent children per household is calculated for each bedroom/tenure pair. These are found in columns AD to AK. Where there are no units the calculation returns a #DIV/0! Error.

Fields:

Tenure (Owned/Shared Ownership, Social Rented, Private Rented or Rent Free)

Number of bedrooms (1, 2, 3, 4 or more)

Census - Adults

This sheet subtracts the total children in each bedroom/tenure pair (Census – Children) from the overall population (Census – All People) and calculates a rate using the data from LC4405EW as described above.

Age Split

For each bedroom/tenure pair this sheet calculates the total number of children. The for each bedroom type it adds the social to the market and calculates the proportion in each.

Calculations

This sheet undertakes the calculations necessary to produce the population yield estimates.

Each row represents one of the 112 Output Areas earlier identified as containing a new housing development.

Columns B-E provide the lookup information for the geographical aggregations (sub region, PTAL or borough).

Columns F-M are the child yield rates calculated in the sheet Census- Children.

Columns O-V are the adult yield rates calculated in the sheet Census- Adults.

Columns X to AE weight the child yield rates based on the size (number of units) of the development.

Columns AG to AN weight the adult yield rates based on the size (number of units) of the development.

Columns AP-AW contain the number of units which are used in the above calculations.

Columns AY-BD For each bedroom/tenure pairing the following calculations are undertaken:

1. Rate – An average child yield rate is calculated. This is done by summing all of the rates for areas which meet the geographic criteria specified in the calculator and then dividing by the number of units in those areas.

2. Yield – The rate above is multiplied by the number of units as specified in the calculator.

3. Market – The yield is split between 0-9 year olds and 10-18 year olds. The rates used to do this are taken from the Age Split sheet.

4. Social – As with Market above.

5. Calc – the individual yields are combined to give totals which are returned to the Calculator sheet.

The above is done for the sub-region selection and the PTAL selection. The Borough Selection is slightly different in the rates are calculated in the Borough Rates sheet (see below). The remainder of the Borough Selection section works I the same way.

Columns BF-BC The same calculations are undertaken for the adult population. There is no need for the age distribution step with this population.

Columns BN – BT contain lookup information used elsewhere in the sheet.

Borough Rates

The following is done for each bedroom-tenure pair and for each borough.

Columns E-L weighted rates for each borough (from the calculations sheet) are summed.

Columns O-V the boroughs selected in the Calculator sheet are included and the other are excluded.

Columns Y-AF total units in each borough are summed.

Columns AI-AP the boroughs selected in the Calculator sheet are included and the other are excluded.

Rows 69-78 average rates for the selected boroughs are calculated and returned to the Calculations sheet.

Sample Sizes

This sheet provides information on the sample sizes for each of the geographical aggregations. It shows the number of sites (output areas) in each bedroom-tenure pair.

For more information please contact Demography Team, GLA Intelligence
Greater London Authority, City Hall, The Queen's Walk, More London, London SE1 2AA
e-mail: demography@london.gov.uk