

Data from Pop247NRT Project (ES/P010768/1):

Data library, sample outputs and batch files for England, 2011

Version 2-0 20/04/21

This data deposition comprises three groups of files: a data library, a set of sample outputs, and the batch files employed to create the sample outputs. The data deposition is for England, 2011.

Version 2-0 is a fully open version, with all input datasets and any derived output datasets being available for re-use under the [Open Government Licence](#) or the [CC-BY 4.0 licence](#).

1. Data library

The data library contains a set of pre-prepared input files in a bespoke textfile format, which can be interrogated by the SurfaceBuilder247 software to create population estimates for the time of day specified by the user. This data library is for England, 2011. It was prepared using the input data sources detailed in Table 1, and was employed to create the sample outputs provided in this data deposition. The folder structure employed in the data library is that required by SurfaceBuilder247; thus, it can be employed by other users to create population estimates for alternative target times as required.

There are four key types of files in the data library:

- (1) **Origins:** these contain the counts of total population available at residential locations, which are available for redistribution at the specified target time. There are two versions: one representing university term-time (with students at their term-time address) and one for out of term-time (where term and out of term are as per the Office for National Statistics (ONS) definitions). The England 2011 data library contains residential households and communal establishments, such as prisons, University halls of residence and care homes.
- (2) **Destinations:** locations/containers which can contain population but which are not origins. Population is redistributed to these locations for the modelled target time according to the specified time profile. The England 2011 data library contains workplaces (subdivided into six Standard Industrial Classification (SIC) groups – see Table 2), educational establishments (primary and secondary schools, Further Education colleges and Universities) and healthcare establishments (inpatients, outpatients and accident and emergency).
- (3) **Time Series:** time profiles specifying what percentage of the population is in transit or on site at the specified time, for a given type of destination. The temporal granularity of the profiles varies depending on the input data sources, and is shown in Table 3.
- (4) **Background layers:** a gridded set of weights representing the relative number of people in transit at the given time in each cell. This is used by SurfaceBuilder247 to redistribute population in transit at the specified time into relevant cells. The England 2011 background layers are derived from annual average daily flows (AADF) and other data from the Department for Transport National Transport Model (NTM) interpolated onto the major road network, with sea and inland water cells masked out to ensure that they do not receive any population. There are 19 background layers, one for each of the Department for Transport NTM time and day of week periods, as shown in Table 4.

Table 1 Pop247NRT England 2011: input data sources

Population sub-groups/ containers	Data source details	URL and license details (accessed 15/04/2021)	Time profile data
Origins			
Residential locations, including communal	2011 census tables: QS103EW Age by single year (usual residents); LC4411EW - Student accommodation by age; QS421EW - Communal establishment management and type – People	https://www.nomisweb.co.uk/census/2011 [Contains data from the Office for National Statistics licensed under the Open Government Licence . Under the terms of the Open Government Licence (OGL) and UK Government Licensing Framework (launched 30 September 2010), anyone wishing to use or re-use ONS material, whether commercially or privately, may do so freely without a specific application for a licence, subject to the conditions of the OGL and the Framework].	Term time/out of term time based on ONS census definitions.
Destinations			
Workplaces	2011 census Workplace Population tables: WP1101EW Sex by single year of age, WP605EW Industry, WP601EW Employment Status, WP702EW Distance travelled to work	https://www.nomisweb.co.uk/census/2011 [Contains data from the Office for National Statistics licensed under the Open Government Licence . Under the terms of the Open Government Licence (OGL) and UK Government Licensing Framework (launched 30 September 2010), anyone wishing to use or re-use ONS material, whether commercially or privately, may do so freely without a specific application for a licence, subject to the conditions of the OGL and the Framework].	Working patterns by bespoke standard industry classification (SIC) groups, derived from Gershuny, J., Sullivan, O. (2017). United Kingdom Time Use Survey (TUS), 2014-2015. Centre for Time Use Research, University of Oxford. [data collection]. UK Data Service. SN: 8128, http://doi.org/10.5255/UKDA-SN-8128-1 .
Primary and secondary schools	School Census and Independent School Census datasets, pupil numbers, postcode	https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2011 [All content is available under the Open	Terms and hours estimated clerically from education authority websites. Average attendance rates from https://www.gov.uk/government/up

		<p>Government Licence v3.0, except where otherwise stated];</p> <p>https://webarchive.nationalarchives.gov.uk/20130107054734/http://www.education.gov.uk/schools/performance/2011/download_data.html</p> <p>[Crown Copyright 2012: The Crown copyright-protected material featured on this website (other than Departmental or agency logos and visual media) may be reproduced free of charge in any format or medium, under the terms of the Open Government Licence.];</p> <p>https://get-information-schools.service.gov.uk/ (previously Edubase) [All content is available under the Open Government Licence v3.0, except where otherwise stated v 1.1.19289.1].</p>	<p>loads/system/uploads/attachment_data/file/219292/sfr04-2012nt.xls</p> <p>[All content is available under the Open Government Licence v3.0, except where otherwise stated].</p>
Further Education colleges	Identified from combination of schools census and Higher Education Statistics Agency (HESA) data, based on age profile of students (100% in 16-17 sub-group), enrolled student numbers, postcode	Schools census and performance data as above; HESA data as below.	<p>Assigned clerically based on planned and funded hours</p> <p>(https://www.gov.uk/guidance/16-to-19-funding-planned-hours-in-study-programmes) [All content is available under the Open Government Licence v3.0, except where otherwise stated])</p> <p>and on proportion of full-time/part-time students</p> <p>(https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/208903/SFR_Tables_Numbers_-_FINAL.xls) [All content is available under the Open Government Licence v3.0, except where otherwise stated]).</p>

Universities	Higher Education Statistics Agency (HESA) data, enrolled student numbers, postcode	https://www.hesa.ac.uk [Data published on the HESA website is free to copy, use, share, and adapt for any purpose. HESA open data is published under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence].	UG/PGT assigned clerically based on University of Southampton Geography average contact time and calendar PGR profile based on office-based workplace profile
Health care facilities	Attendance figures at NHS Trusts, Primary Care Trusts, Treatment Centres, Independent Sector Healthcare Provider sites, distances travelled, postcode	https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics [Copyright and database rights in the NHS Digital Content are released free-of-charge under the current version of the Open Government Licence ("OGL"), except where specified, either in these terms and conditions, elsewhere on this website or in the OGL terms. This means that you can use NHS Digital Content, including copying it, adapting it, and using it for any purpose, including commercially, provided you follow these terms and conditions and the terms of the OGL].	Inpatients allocated clerically by historical bed occupancy data (no longer available); outpatients clerically by estimated appointment times; A&E by mean arrivals over 24 hours (https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/weekly-ae-sitreps-2011-12/) [You may use and re-use the information featured on this website (not including logos or photographs) unless identified as another party's copyright free of charge in any format or medium, under the terms of the Open Government Licence].
Background			
Road network	Ordnance Survey (OS) OpenMap - Local road network, Department for Transport (DfT) annual average daily flows (AADF) by area, road and vehicle type for 19 time periods from National Transport Model	https://osdatahub.os.uk/downloads/open [Contains OS data © Crown copyright and database right [2019] under the Open Government Licence]; http://www.dft.gov.uk/traffic-counts/ [All content is available under the Open Government Licence v3.0 , except where otherwise stated].	Time periods referenced to National Transport Model

Coastline	ONS 2015 Regions, clipped to Mean High Water mark	http://geoportal.statistics.gov.uk/datasets/regi-ions-december-2015-full-clipped-boundaries-in-england [Source: Office for National Statistics licensed under the Open Government Licence v.3.0 . Contains OS data © Crown copyright and database right [2019]].	N/A
Inland water	OS OpenMap - Local Surface water	https://osdatahub.os.uk/downloads/open [Contains OS data © Crown copyright and database right [2019] under the Open Government Licence].	N/A
Lookup files			
Grid references and lookup files	ONS Postcode Directory November 2012 version; Output Area and Workplace Zone 2011 population-weighted centroids	https://geoportal.statistics.gov.uk/ [Contains OS data © Crown copyright and database right [2019]; Contains Royal Mail data © Royal Mail copyright and database right [2019]; Source: Office for National Statistics licensed under the Open Government Licence v3.0].	N/A

Table 2 Pop247 England 2011 Standard Industrial Classification (SIC) groups

SIC Group: Workplace destinations filename	SIC Group: Time Profile name	SIC section name	SIC section code
Agri+Fish	1AGRI	Agriculture, forestry and fishing	A
Service	2SERV	Real estate activities Administrative and support service activities Activities of households as employers	L N T
Accom	3ACC	Accommodation and food service activities	I
Mine	4MINE	Mining and quarrying Transportation and storage Human health and social work	B H Q
Retail+Arts	5RETL	Wholesale and retail trade and repair of motor vehicles and motorcycles Arts, entertainment and recreation Other service activities	G R S
Public+Office	6OFFB	Financial and insurance activities Construction Electricity, gas, steam and air conditioning supply Water supply, sewerage, waste management and remediation activities Manufacturing Information and communication Professional, scientific and technical activities Education Public administration and defence; compulsory social security	K F D E C J M P O

Table 3 Temporal granularity for destination time profiles

Destination type	Time intervals (minutes)
Education	
Primary	15
Secondary	15
Further Education Colleges	15
Universities – UG/PGT	15
Universities - PGR	10
Health	
Inpatient	15
Outpatient	60
A&E	60
Workplaces	
Agriculture, forestry, fishing	10
Services	10
Accommodation, Food services	10
Mining, transportation	10
Retail, Arts	10
Office-based, Public sector	10

Table 4 Department for Transport National Transport Model day of week and time periods

Number	Day of week	Time period	Description
1	Mon-Fri	00:00-06:00	Off peak
2	Mon-Fri	06:00-07:00	AM peak
3	Mon-Fri	07:00-08:00	AM peak
4	Mon-Fri	08:00-09:00	AM peak
5	Mon-Fri	09:00-10:00	AM peak
6	Mon-Fri	10:00-16:00	Inter peak
7	Mon-Fri	16:00-17:00	PM peak
8	Mon-Fri	17:00-18:00	PM peak
9	Mon-Fri	18:00-19:00	PM peak
10	Mon-Fri	19:00-22:00	Shoulders
11	Mon-Fri	22:00-00:00	Off peak
12	Saturday	00:00-09:00	Saturday night
13	Saturday	09:00-14:00	Saturday Day
14	Saturday	14:00-20:00	Saturday Day
15	Saturday	20:00-00:00	Saturday night
16	Sunday	00:00-10:00	Sunday Night
17	Sunday	10:00-15:00	Sunday Day
18	Sunday	15:00-20:00	Sunday Day
19	Sunday	20:00-00:00	Sunday Night

Session parameters

The data library also contains four SessionParas files, containing the session parameters used to create the four sets of sample output files deposited here. When running SurfaceBuilder247 (in either batch or interactive mode), these files tell the program important information such as the spatial dimensions and extent of the output grid to be created, the buffer size and the input files (background, time series, origins, destinations) to be used in the run. These session parameters files will be useful for any user who wishes to run SurfaceBuilder247 for an alternative target time, and can be readily adapted.

The data library also contains empty folders for data logs, results, run logs and validation files. These are populated with log and output files when SurfaceBuilder247 is run.

2. Sample outputs

Sample outputs for England 2011 (Version 2-0) are provided as part of this data deposition. These have been created using the SurfaceBuilder247 software with the data library files described above. The output files contain the estimated count of population present in a cell at the specified target time, for a specific population sub-group.

Outputs are for 200m British National Grid (BNG) grid cells, with spatial extent 0,0 (SW corner) to 700000,600000 (NE corner). Files are GeoTIFF format to enable easy import into Geographical Information Systems (GIS) software. Note that cells with genuine No Data values (e.g. because they are in the sea or inland water) are not differentiated from (land-based) cells with zero population in the output files: both are given a value of zero.

There are four sets of sample outputs:

- 02:00 (2am) out of school and university term-time
- 02:00 (2am) during school and university term-time
- 14:00 (2pm) out of school and university term-time
- 14:00 (2pm) during school and university term-time

These days are intended to represent a typical weekday during term- and out of term-times. The term-time runs are notionally for 15/06/11 and the out of term-time for 17/08/11.

There are seven population sub-groups for the term-time outputs, plus total population which is the sum of all sub-groups (see Table 5).

Table 5 Population sub-groups: term-time

Sub-group	Details
0-4	0-4 year olds
5-9	5-9 year olds
10-15	10-15 year olds
16-17	16-17 year olds
18-64HE	18-64 year old students
18-64NSTU	18-64 year old non-students
OV65	65 years old or over
AllAges	Total population i.e. sum of all sub-groups

There are six population sub-groups for the out of term runs, plus the total population category (see Table 6). This is because it was not feasible to subdivide the 18-64 category into student and non-student categories based on the available data.

Table 6 Population sub-groups: out of term-time

Sub-group	Details
0-4	0-4 year olds
5-9	5-9 year olds
10-15	10-15 year olds
16-17	16-17 year olds
18-64	18-64 year olds
OV65	65 years old or over
AllAges	Total population i.e. sum of all sub-groups

The term-time runs include students at primary and secondary schools, students at Further Education colleges, postgraduate researchers (PGRs) and Taught (undergraduates and postgraduates) at Universities. The out of term-time runs include postgraduate researchers (PGRs) at Universities but exclude students at primary and secondary schools, students at Further Education colleges, and Taught (undergraduates and postgraduates) at Universities.

3. Batch files

Four batch files are provided. These were used to create the four sets of sample outputs deposited here. The batch files run SurfaceBuilder247 in batch mode, providing the program with details such as the location of the data library, the name of the relevant session parameters file, and the runs to be undertaken. Note that each individual SurfaceBuilder247 run is for a unique combination of population sub-group, date and time, but runs can be batched together for ease of processing, as shown in the batch files here. They can also be run in parallel to aid efficiency. Users wishing to run their own target time or different spatial extents can adapt these batch files as required.

4. References, citations and downloads

Full details of the **Pop247 framework and methodology** can be found in the following article. Please cite this article when using Pop247/Pop247NRT data and/or the SurfaceBuilder247 software in publications.

Martin D, Cockings S, Leung Y. Developing a flexible framework for spatiotemporal population modelling. *Annals of the Association of American Geographers*, 105(4), 2015, 754-772
<http://dx.doi.org/10.1080/00045608.2015.1022089>

SurfaceBuilder247 is available for free download and use from ESRC's Gateway to Research (<https://gtr.ukri.org/projects?ref=ES%2FP010768%2F1>) or from the authors' project website at <http://pop247nrt.geodata.soton.ac.uk/>