

Data Descriptor Table

<i>Dataset</i>	<i>Metadata</i>	<i>Source</i>	<i>License</i>	<i>Data Cleaning</i>
General health scores	A csv containing general health scores across London in 2011.	InFuse 2011 UK Census: https://infuse2011gf.ukdataservice.ac.uk/infusewizgeo.aspx	Open	<ul style="list-style-type: none"> Selected the MSOAs in London as the geographic area, and then the 'General health' topic, followed by selecting all the categories before downloading the dataset. In Excel, removed columns with duplicated information – GEO_TYP2, and GEO_TYPE. Changed the column names so they were easier to type in R. Deleted empty rows. Converted to csv.
Ethnicities	A csv containing data on distribution of ethnicities across London in 2011.	InFuse 2011 UK Census: https://infuse2011gf.ukdataservice.ac.uk/infusewizgeo.aspx	Open	<ul style="list-style-type: none"> Selected the MSOAs in London as the geographic area, and then the 'Ethnic group' topic, followed by selecting all the categories before downloading the dataset. In Excel, removed columns with duplicated information – GEO_TYP2, and GEO_TYPE. Changed the column names so they were easier to type in R. Deleted empty rows. Converted to csv.
Migration levels (2011)	A csv containing data on countries of birth across London in 2011.	InFuse 2011 UK Census: https://infuse2011gf.ukdataservice.ac.uk/infusewizgeo.aspx	Open	<ul style="list-style-type: none"> Selected the MSOAs in London as the geographic area, and then the 'Country of birth' topic, followed by selecting all the categories before downloading the dataset. In Excel, removed columns with duplicated information – GEO_TYP2, and GEO_TYPE. Changed the column names so they were easier to type in R. Deleted empty rows. Converted to csv.

Migration levels (2001)	A csv containing data on countries of birth across London in 2001.	InFuse 2011 UK Census: https://infuse2011gf.ukdataservice.ac.uk/infusewizgeo.aspx	Open	<ul style="list-style-type: none"> Selected the MSOAs in London as the geographic area, and then the 'Country of birth' topic, followed by selecting all the categories before downloading the dataset. In Excel, removed columns with duplicated information – GEO_TYP2, and GEO_TYPE. Also created an extra column that added up the number of people born in different countries. Changed the column names so they were easier to type in R. Deleted empty rows. Converted to csv.
Migration levels (difference)	A csv containing data on countries of birth across London in 2001 and 2011, and the difference between them.	InFuse 2001 UK Census: https://infuse.ukdataservice.ac.uk/infusewiz.aspx?cookie=openaccess InFuse 2011 UK Census: https://infuse2011gf.ukdataservice.ac.uk/infusewizgeo.aspx	Open	<ul style="list-style-type: none"> In Excel, combined the data from the migration levels datasets (both 2001 and 2011). Created a new column that calculated the difference in persons. Changed the column names so they were easier to type in R. Deleted empty rows. Converted to csv.
Main languages	A csv containing data on the number of people having non-English as their main language in 2011.	InFuse 2011 UK Census: https://infuse2011gf.ukdataservice.ac.uk/infusewizgeo.aspx	Open	<ul style="list-style-type: none"> Selected the MSOAs in London as the geographic area, and then the 'Language, main language used' topic, followed by selecting the total and English categories so that I could calculate the number of non-English speakers in R. In Excel, removed columns with duplicated information – GEO_TYP2, and GEO_TYPE. Changed the column names so they were easier to type in R. Deleted empty rows. Converted to csv.
Life expectancy	A csv containing data on life expectancies	Office for National Statistics: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/lifeexpectancyatbirthandage65bylocal	Open	<ul style="list-style-type: none"> Using excel spreadsheet to extract the relevant information – specifically Inner and Outer London from the sheet titled 'E&W LAs at birth' for both

	across London from 2011 to 2013.	areasinenglandandwalesreferencetable1		<p>male and females in 2011 to 2013 into a separate spreadsheet.</p> <ul style="list-style-type: none"> • Changed the column names so they were easier to type in R. • Deleted empty rows. • Converted to csv.
London MSOAs and Boroughs	A shp shapefile containing boundaries (vector) of London's MSOAs and Boroughs in 2011.	London Datastore: https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london	Open	<ul style="list-style-type: none"> • Downloaded the zip folder and extracted specifically the 'MSOA_2011_London_gen_MHW' and 'London_Borough_Excluding_MHW' files.
Wellbeing scores	A csv containing data on life satisfaction and worthwhile scores across London in 2011/12.	London Datastore: https://data.london.gov.uk/dataset/subjective-personal-well-being-borough	Open	<ul style="list-style-type: none"> • Downloaded the 'Personal Well-being by Borough' data from the website. In Excel, extracted the life satisfaction and worthwhile scores into a separate spreadsheet. • Changed the column names so they were easier to type in R. • Deleted empty rows. • Converted to csv.
Hospitals, GP practices and pharmacy locations	A csv containing data and coordinates of the locations of hospitals, GP practices and pharmacies across London.	NHS website dataset: https://www.nhs.uk/about-us/nhs-website-datasets/	Open	<ul style="list-style-type: none"> • Downloaded the 'Hospitals', 'GP practices' and 'Pharmacies' Excel documents. • Data was separated by a different delimiter, so used the 'text to column' feature to separate the data into columns. • Changed the column names so they were easier to type in R. • Deleted empty rows. • Converted to csv.