**Inner London Dwelling Dimensions Dataset, 2017-2018: Additional Metadata**

For each of the 5,278 floor plans collected, Archilyse provided room-level data including 1) the net floor area, 2) the dimensions of the minimum bounding rectangle (width and length), 3) the circumference, 4) the total window length, 5) the number of doors (including the IDs of rooms the doors open on to) and 6) the number of kitchens, bathroom elements and staircases. A ‘room’ is defined as a space bounded and separated from others by walls and connected to others by doors. Therefore, rooms that are separated from each other by openings other than doors were counted as only one room. For instance, connected living and dining areas or entrance halls partially separated from living rooms were counted as single rooms. At the same time, built-in storage, which meets the criteria of a ‘room’ (enclosed by walls and separated by a door), were counted as separate rooms in the raw dataset.

In modelling the data, first, a number of other data points were calculated. These include 1) compactness that indicates the shape of the room, 2) adjacent and connected rooms that indicate layout, 3) room function, 4) the net floor area calculated together with built-in storage, 5) width and length of the minimum bounding rectangle calculated together with built-in storage. Table 4.03 (see Related Resources) describes the given and calculated data points and how they were calculated.

A key calculation at room level was the room function. To label the room functions, I developed a step-by-step decision tree (See Figure 4.03 see Related Resources). First, any room with bathroom elements (shower, bathtub, WC) was given the label ‘bathroom’ and any room with kitchen elements (kitchen counter with sink) was given the label ‘kitchen’. Second, any room without a window was labelled either as ‘storage’ or ‘circulation’. To distinguish between these two labels, criteria such as the number of doors, room shape (width/length ratio, compactness), and net floor area were used. Any remaining rooms with windows were given the label ‘room’, with the exception of rooms smaller than 4 m² (labelled as ‘storage’) and with the compactness of less than 0.4 (labelled as ‘circulation’).

Storage areas with a floor area smaller than 1.5 m² that open onto one of the rooms were considered as built-in storage, and they were merged with the room they open onto. After this, a composite net floor area and composite width and length of the minimum bounding rectangle were calculated for every room. The rooms were further differentiated into categories based on the composite net floor areas of the rooms and their widths (the width of the minimum bounding rectangle). The rooms with a floor area larger than 12 m² and a width larger than 2.5 m were labelled as double rooms, the remaining rooms with a floor area larger than 8 m² and a width larger than 2.15 were labelled as a single room, and all other remaining rooms were labelled as small rooms.

To distinguish kitchens that are also used as a living room from those used solely as a kitchen, in the second step a set of criteria, including the number of rooms and the net floor area of kitchens, were used. Accordingly, any ‘kitchen’ in dwellings with no rooms was labelled as ‘studio’, any ‘kitchen’ with a net floor area between 14 and 18 m² in dwellings with one room was labelled as ‘small living-in kitchen’, and any ‘kitchen’ with a net floor area larger than 18 m² in dwellings with more than one room was labelled as ‘living-in kitchen’. In multiple-storey dwellings, the living areas with open kitchens that occupy the whole floor, regardless of their size, were labelled as ‘kitchen floor’. The ‘kitchens’ that did not satisfy these requirements were labelled as ‘working kitchens’. In dwellings with ‘working kitchens’, a potential living room was identified among the rooms. In one-storey dwellings the largest room, and in multiple-storey dwellings the largest room on the ground floor, or the room that is directly connected to the kitchen, were identified as living rooms.

Based on the data at the room level, further data points at the dwelling level were calculated. These were iteratively compared to the original floor plans to verify their validity. These include 1) the number of habitable rooms, 2) the type of rooms, 3) the kitchen type, 4) the entrance, 5) the circulation area, and 6) the grouping of functions. The number of habitable rooms was calculated as the total number of rooms including habitable kitchens, i.e. studio, small living-in kitchens, living-in kitchens, and kitchen floors. The entrance was defined as the room with a door opening to the outside, the circulation area was defined as the room that has the largest number of connected rooms, and the grouping of functions was defined as whether the bedrooms were grouped together.

Assessing the data, some of the floor plans were eliminated from the dataset. First, the required data points could not be fully extracted from all of the floor plans. These were mostly related to the different drawing conventions used in the original floor plans. These were excluded from the analysis. In the case of multi-storey dwellings, whenever a plan was eliminated, the whole set of floor plans that belonged to the same dwelling was excluded. Second, there were a few cases with exceptionally large or small dwelling sizes (total net floor area) compared to other housing in the same LSOA and with the same number of habitable rooms. Comparing the data against the floor plans showed that these anomalies resulted from either an incorrect scaling of floor plans or the inclusion of balconies, gardens, and terraces in the total floor area calculation. In the case of the former, the plans were omitted from further analysis, and in the case of the latter the balconies, gardens, and terraces were removed, and the internal floor areas were recalculated. While these cast doubt on the accuracy of the floor plans, randomly selected 100 plans showed that the dimensions were accurate within a 5 m² margin.

In its final form, a total of 1,840 plans were eliminated from the initial dataset of 5,278 floor plans. It is important to point that most of the eliminations were in houses and maisonettes of the older housing stock (Appendix A, see Related Sources).