

1

Results

This chapter reports the key results of the research presented in this dissertation. First, the different segmentations introduced in the previous chapter are reported, and the relationship of these to patterns of urban morphology is qualitatively examined. Second, the results of the quantitative evaluation of the relationship between each segmentation and house prices are reported.

Further supplementary results are reported in the Appendix.

1.1 Relationship to urban morphology

Perhaps the most intuitive way to compare the seven segmentations presented in this dissertation is graphically. Figures 1.1 through 1.7 map the way each segmentation partitions the city of Barcelona. While the scale of these city-level maps limits the amount of detail which each can show, sites of particular interest in certain segmentations are highlighted throughout the Discussion.

It should be noted that clusters have been coloured with the same colours where a clear equivalence is perceptible (for example when identified, the cluster most closely corresponding to the Ciutat Vella is pink), but these are so coloured for ease of comparison only and should not be taken to indicate any formal connection between the clusters produced by different segmentations.

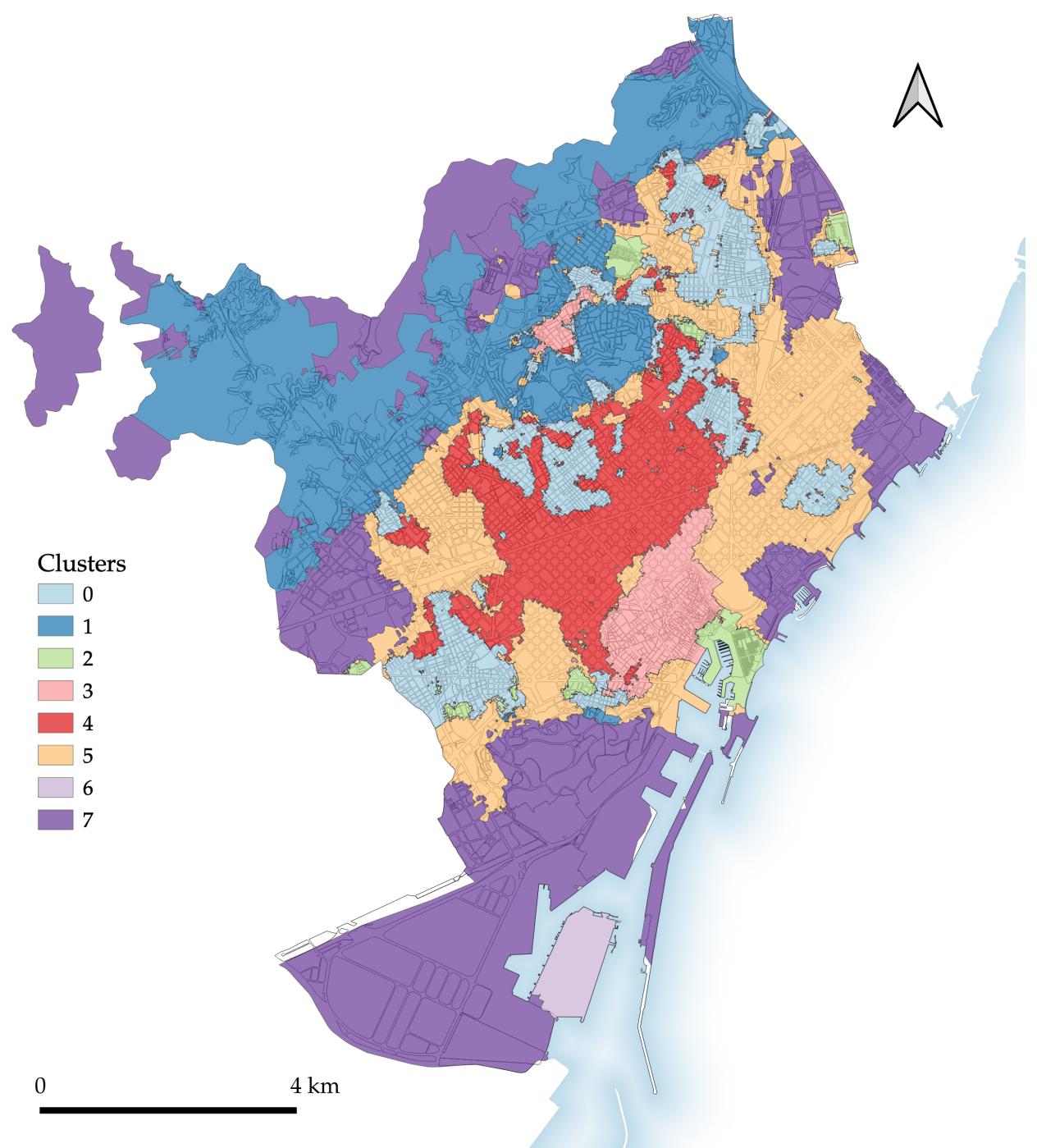


Figure 1.1: Segmentation 1: Morphological tessellation.

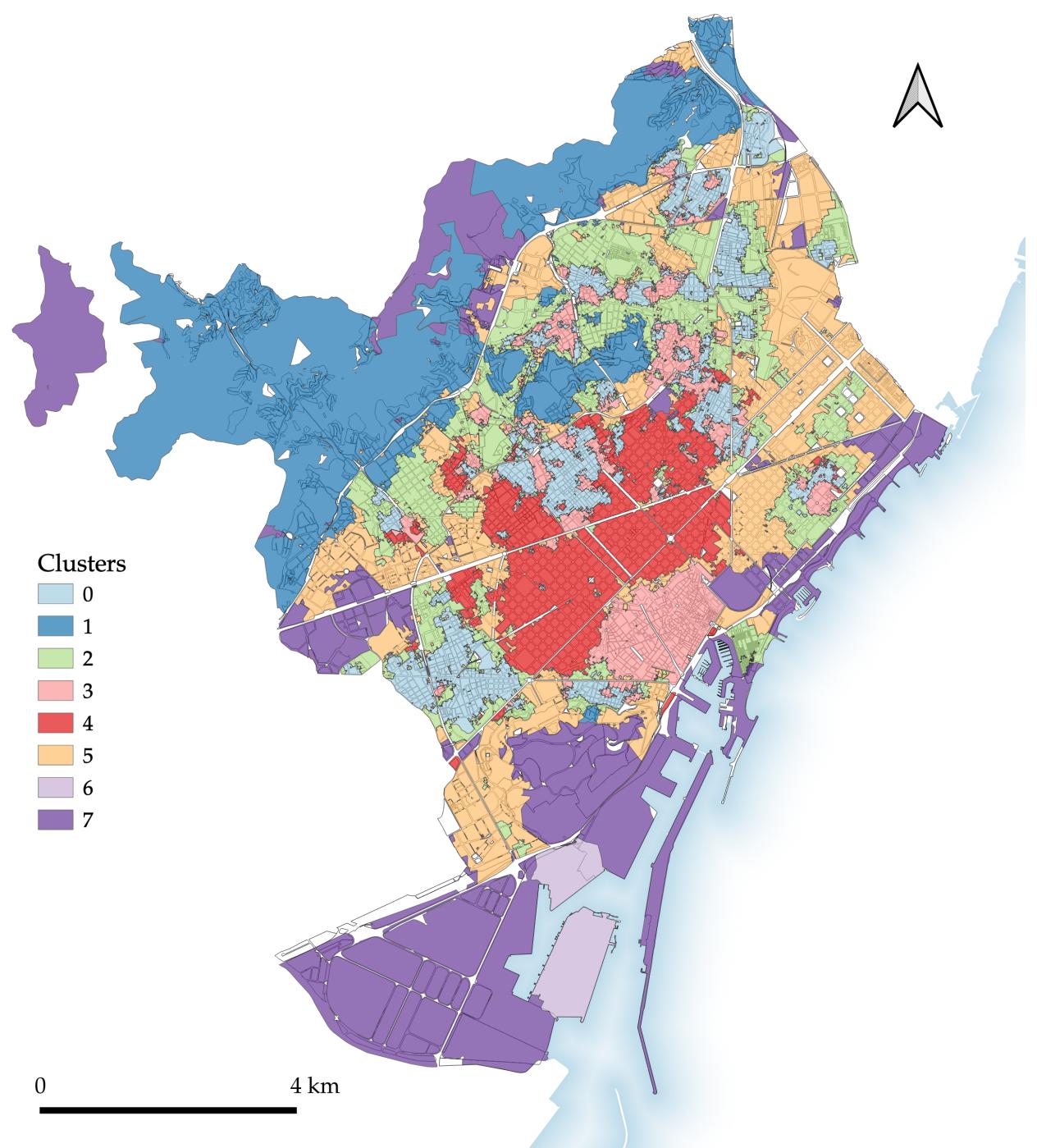


Figure 1.2: Segmentation 2: Enclosed tessellation.

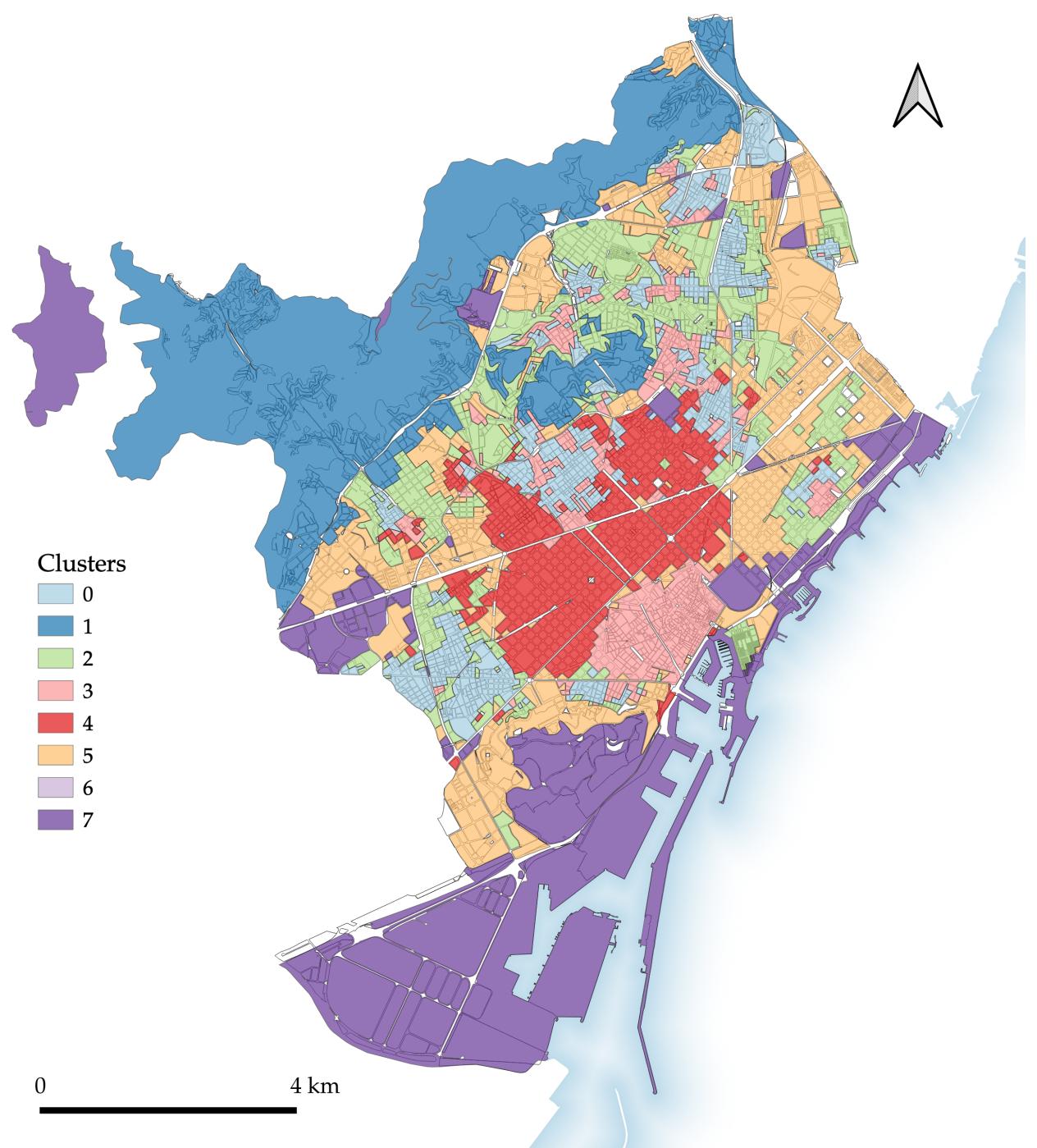


Figure 1.3: Segmentation 3: Enclosed tessellation transposed to block geometry.

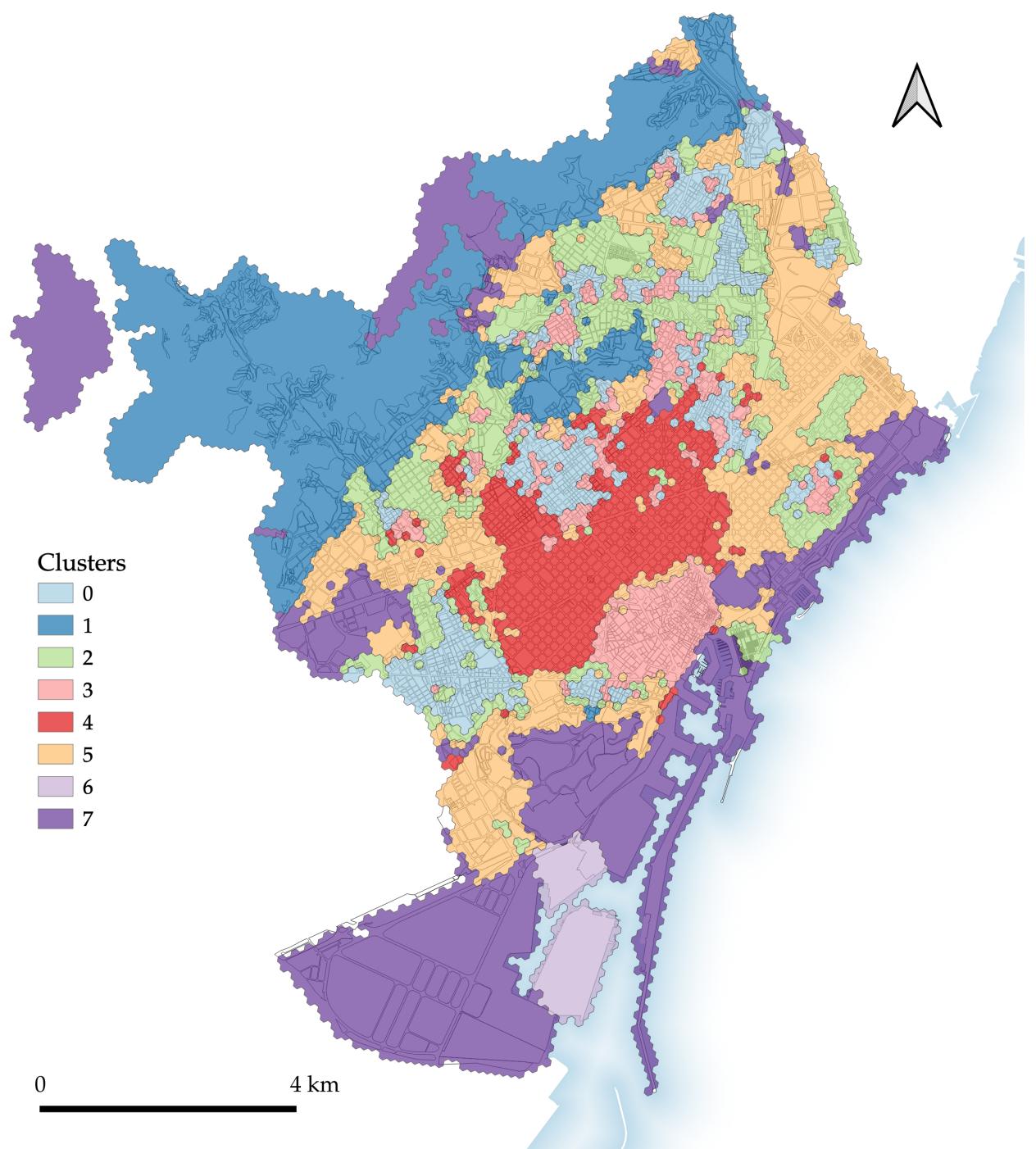


Figure 1.4: Segmentation 4: Enclosed tessellation transposed to H3 geometry.

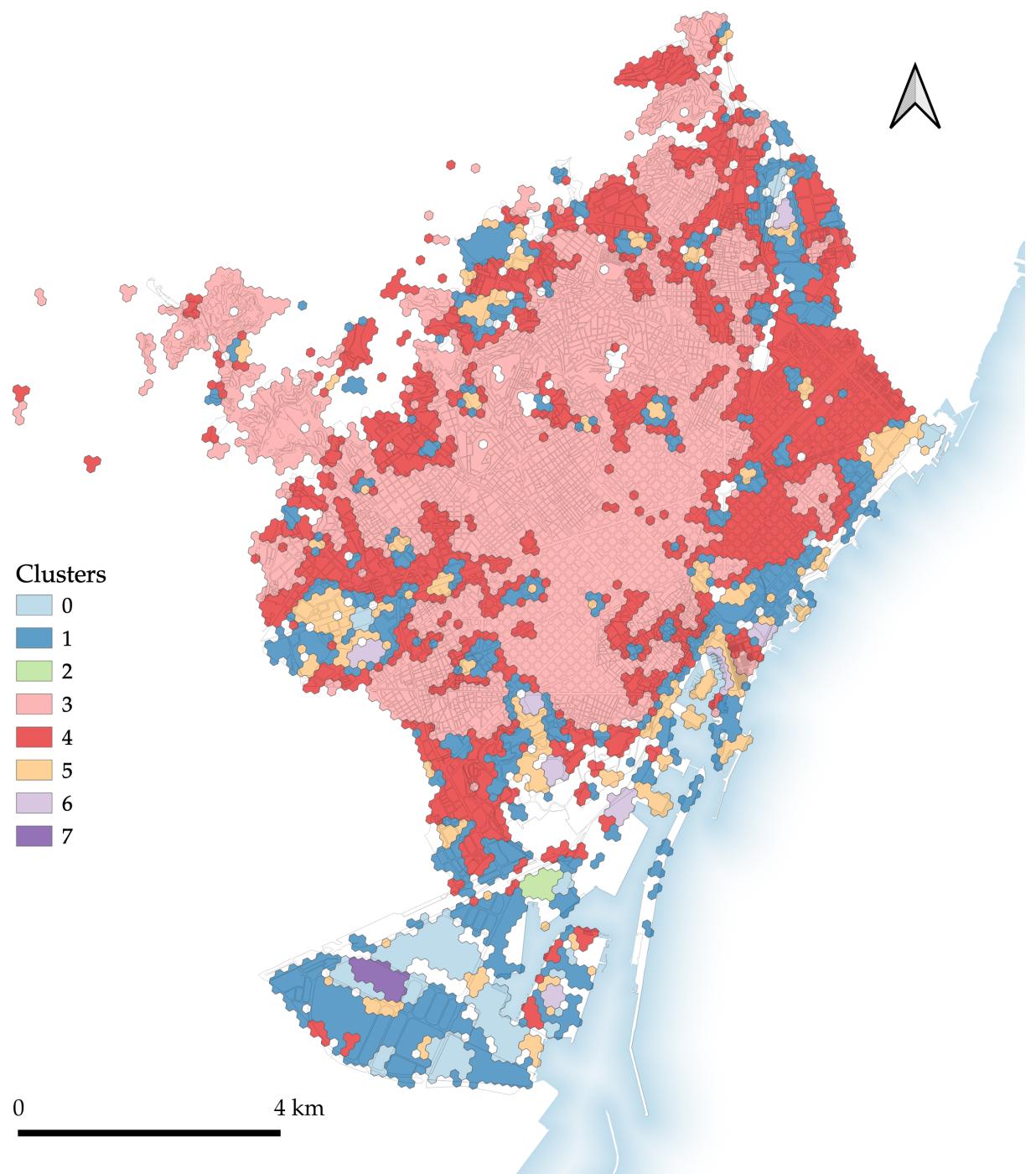


Figure 1.5: Segmentation 5: H3 'basic'.

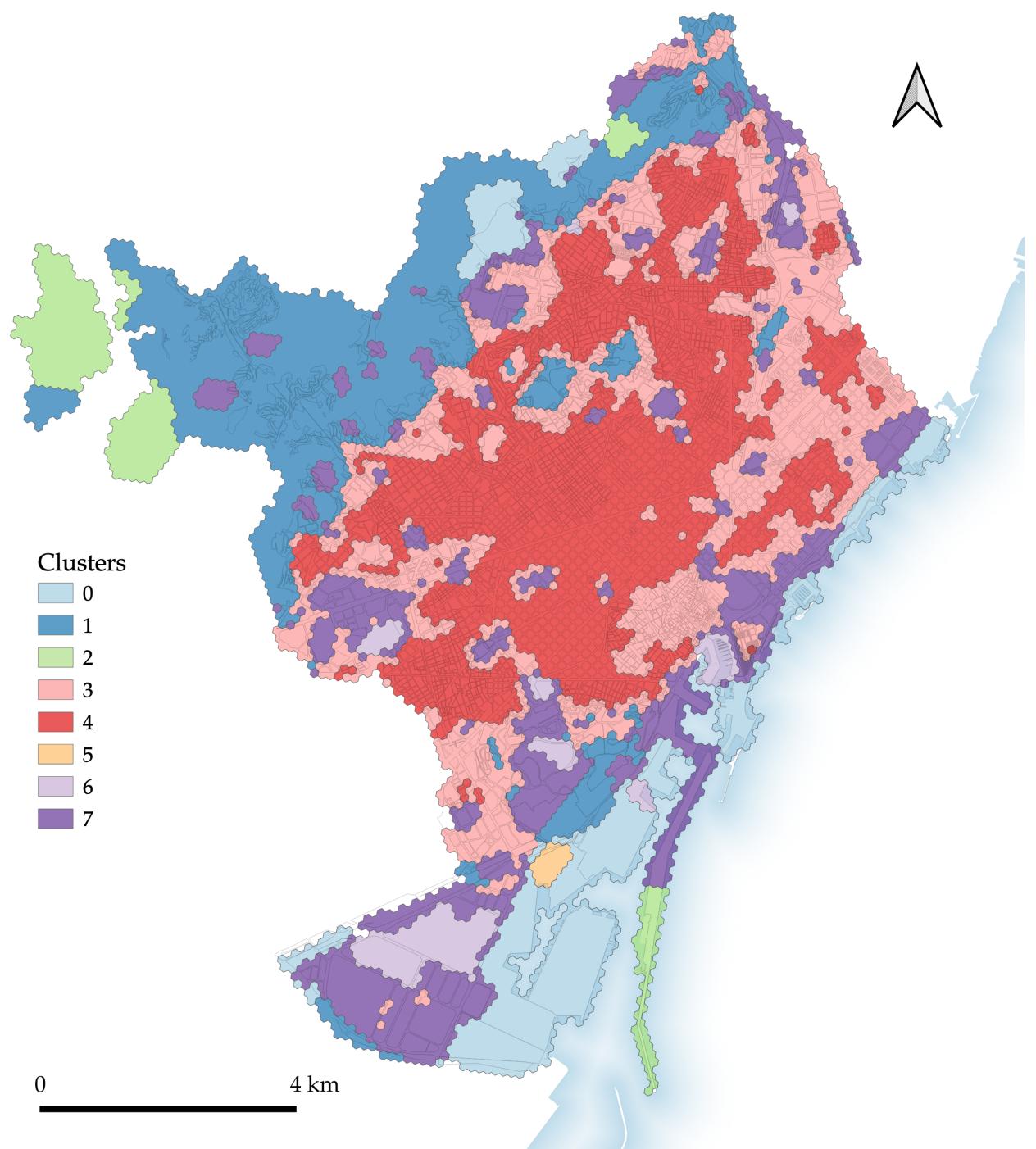


Figure 1.6: Segmentation 6: H3 clustering using characters from enclosed tessellation.

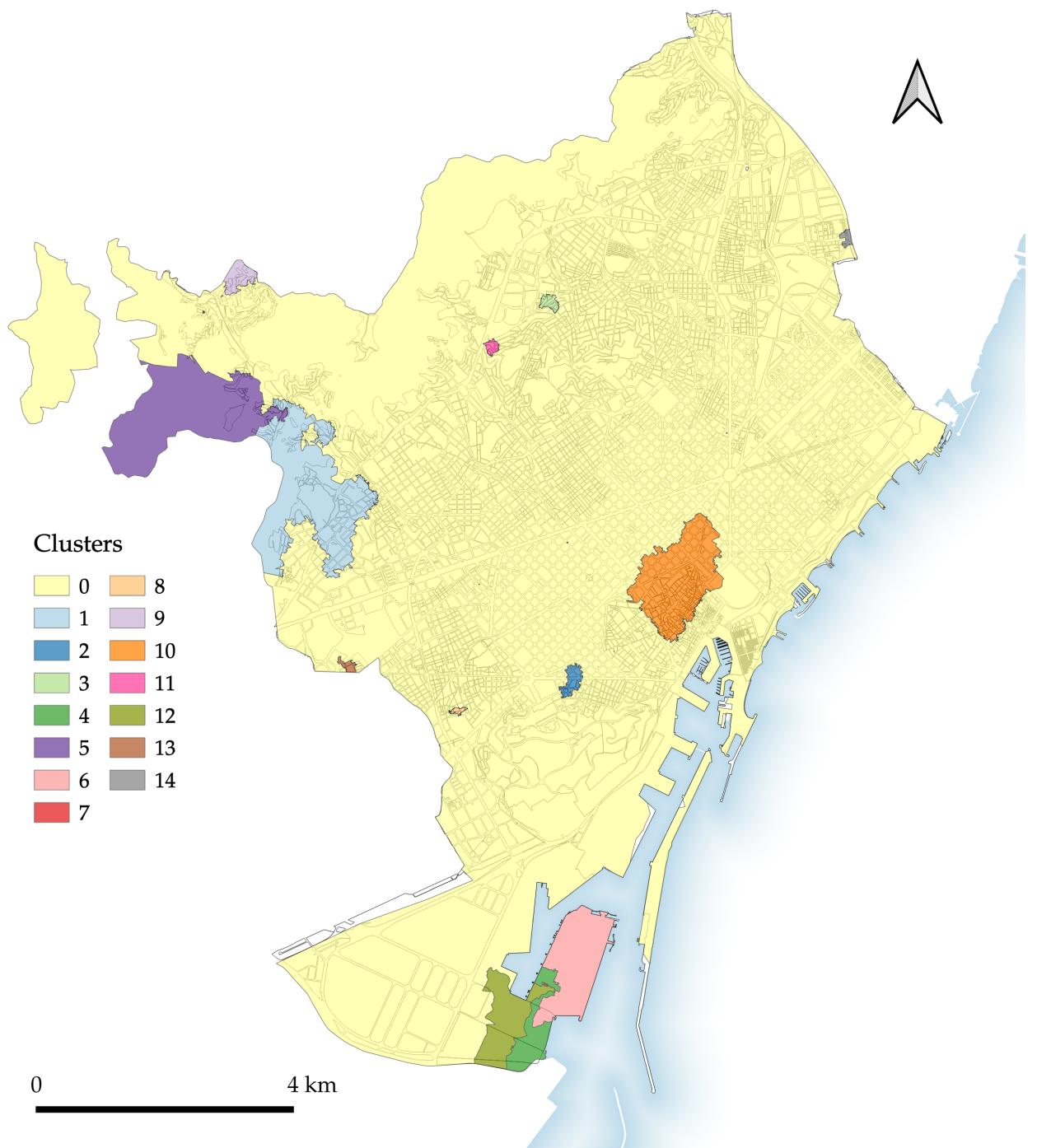


Figure 1.7: Segmentation 7: Spatially constrained clustering on morphological tessellation cells.

Although discussed at greater length in the next chapter, a simple initial description of each of the segmentation maps makes clear certain key results. The **morphological tessellation segmentation** (Figure 1.1) relatively successfully picks out certain key urban tissues, including those of the Ciutat Vella and of the Eixample, although both have clear room for improvement. For example, the Ciutat Vella type also includes a section of the Eixample immediately to the north of the old city, while the Eixample type excludes certain parts of the Eixample ((mis)classified as part of the pink and orange types) while also incorporating other areas beyond the Eixample grid.

The **enclosed tessellation segmentation** (Figure 1.2) identifies similar overall morphological trends to the preceding MT segmentation, but also comprises key differences. Immediately conspicuous are the areas omitted from classification – those enclosures not containing any buildings. The segmentation is also visibly more ‘fragmented’ than its MT counterpart, particularly in the northern part of the city.

This fragmented nature is a motivating factor in the development of the next two segmentations, which transpose the ET segmentation to **block** (Figure 1.3) and **H3** (Figure 1.4) geometries.

The **H3 ‘basic’ segmentation** (Figure 1.5) tests the performance of a segmentation produced without using ET or MT cells at any point. Its results leave a lot to be desired: despite the use of contextual characters (defined with neighbouring H3 cells), few if any of the segmentation’s types or polygons could be said to clearly constitute distinct urban tissues.

The consequent **H3 clustering using ET characters** (Figure 1.6) somewhat improves on the preceding ‘basic’ segmentation. While morphologically distinct areas are delineated to some extent, the differentiation of different types is inferior to that found in the ET or MT segmentations.

Finally, a visual inspection of the **spatially constrained segmentation** (Figure 1.7) immediately makes obvious the imbalance in the size of its clusters: 90.3% of the total area is assigned to one cluster.

1.2 Relationship to house price indices

For each segmentation, multiple metrics are calculated to measure the dispersion within/between the clusters generated. As discussed above, both type and polygon¹ level metrics are reported below.

In order to allow a comparison between the novel spatial segmentations mapped above and the existing spatial segmentations which may be used to represent spatial housing submarkets, three of the latter are included within the comparison below. The neighbourhoods, districts, and idealista polygons are mapped in Figure 1.8. This makes clear the ways in which the idealista polygons are largely coterminous with the city's administrative neighbourhoods, but sometimes merge these neighbourhoods, as with those which form the two orange idealista polygons in Nou Barris²; and sometimes separate these neighbourhoods in two, as with la Marina del Prat Vermell at the very South of the city. In other cases the geometry of the original neighbourhoods has been simplified or otherwise altered to create the idealista polygons.

1.2.1 Type metrics

Table 1.1 reports the average Quartile Coefficient of Dispersion of the average residential property sale price for all types, the mean of the areas of types in each segmentation, and the number of types included in each segmentation. Also included for each segmentation is a boxplot showing the distribution of these type areas.

Figure 1.9 plots each type from each segmentation, comparing its area (on the x -axis) with the QCoD of house prices within the type (on the y -axis). It shows a definite relationship between these two variables, and for this reason the average QCoD alone (provided in Table 1.1) is insufficient to make a fair comparison

¹As set out in the previous chapter, *type* describes all cells in the city with a certain cluster label, whereas *polygons* refer to the separate geographies of contiguous cells with a certain cluster label.

²Ciutat Meridiana, Vallbona, and Torre Baró in the North of the district; and Can Peguera and el Turó de la Peira in the South.

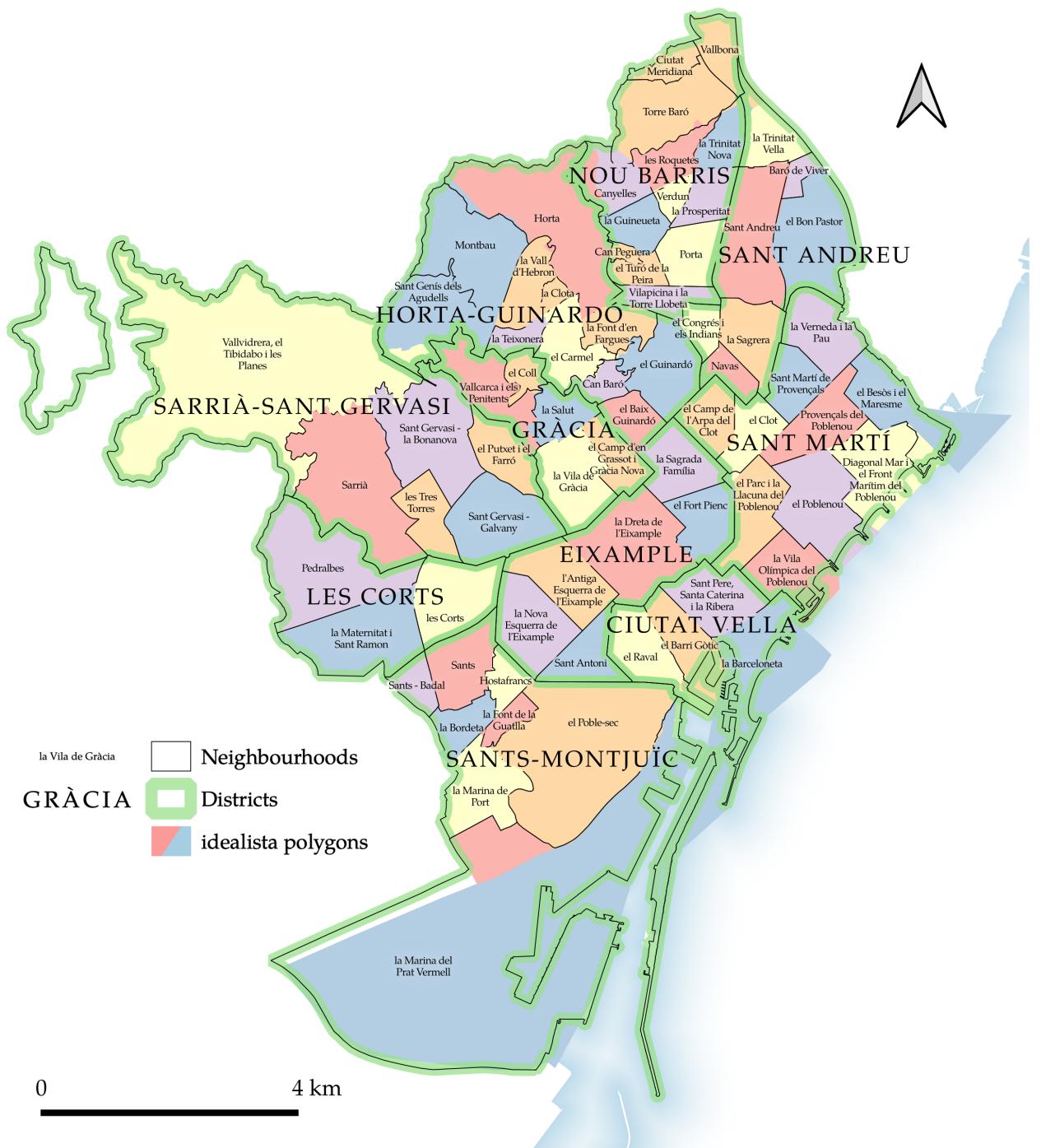


Figure 1.8: Existing spatial units: neighbourhoods, districts, and idealista polygons in Barcelona.

Table 1.1: Average type values for each segmentation.

Segmentation	Mean QCoD	Type area (km^2)		Number of units
		Mean	Distribution	
Morphological tessellation	0.206	12.62	■■■■■■■■	8
Enclosed tessellation	0.201	11.92	■■■■■■■	8
ET transposed to block	0.202	14.53	■■■■■■■	6
ET transposed to H3	0.202	13.47	■■■■■■■	8
H3 basic	0.140	10.13	■■■■■■■■	8
H3 with ET characters	0.222	13.47	■■■■■■■■	8
Spatially constrained MT	0.080	6.73	■■■■■■■■	15
Existing neighbourhoods	0.071	1.39	■■■■■■■■	73
Existing districts	0.115	10.17	■■■■■■■■	10
idealista polygons	0.075	1.42	■■■■■■■■	69

of how well segmentations capture variations in property prices: all else being equal, larger areas will tend to have larger dispersions. By plotting both the QCoD and the area of the spatial unit whose internal house price dispersion the QCoD records, a better judgement can be made about how well a given segmentation captures variation in house prices *given its area*. Note that in Figure 1.9, the dot for one type (the largest type in the spatially constrained segmentation) is not displayed, as its area of 91.2km^2 makes it an extreme outlier.

Because Figure 1.9 plots every type separately, it is difficult to make a clear judgement about how well a segmentation performs overall. This is evident by the spread of dots of the same colour in different areas of the chart: while one type within a segmentation may have a low QCoD, this may be offset by other types within the same segmentation having much greater dispersion. In order to allow better judgements of the overall performance of segmentations, Figure 1.10 therefore plots the average areas of the types in each segmentation against the corresponding average house price QCoDs, allowing the direct comparison of each segmentation's average values for these two metrics. The grey line plots a simple linear regression model fit to the points on the graph. Although clearly a poor predictor of the average QCoD of a given segmentation, the line serves as a visual aid in understanding how well a segmentation can be seen to capture house price dispersion, given the size of the units into which it partitions space (in this case, the

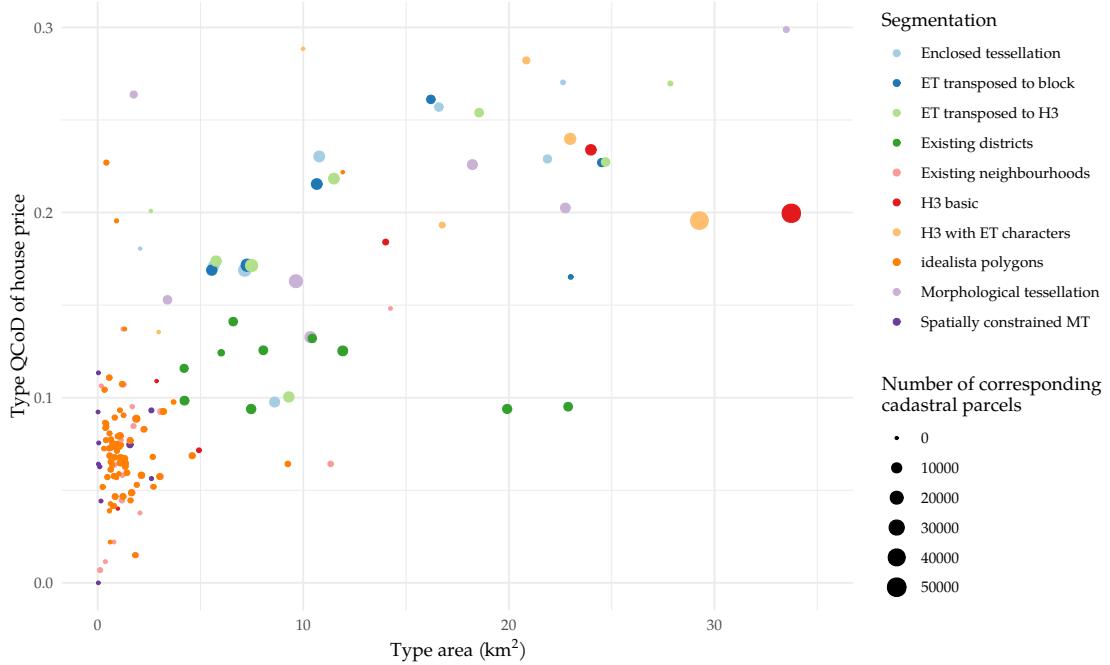


Figure 1.9: Every type plot by house price QCoD and area.

different types). Segmentations plot below the line can be seen to have lower levels of dispersion given the average areas of their types, and therefore better capture variation in house prices. Conversely, segmentations plot above the line can be seen to have higher QCoD values than might be expected given the average areas of their types, and therefore perform worse as a delineation of housing submarkets.

It should be noted that this system of interpretation is not derived from any particular empirical base, but rather the plots provide a useful heuristic for comparing the performance of segmentations which divide space into different numbers of units of different sizes.

1.2.2 Polygon metrics

Treating each type—each colour on the segmentation map—as the ultimate spatial unit generated by the segmentation process is one way of assessing the segmentations. Alternatively, each *polygon* produced can be seen as a separate spatial unit: in this conceptualisation, areas which are assigned the same cluster

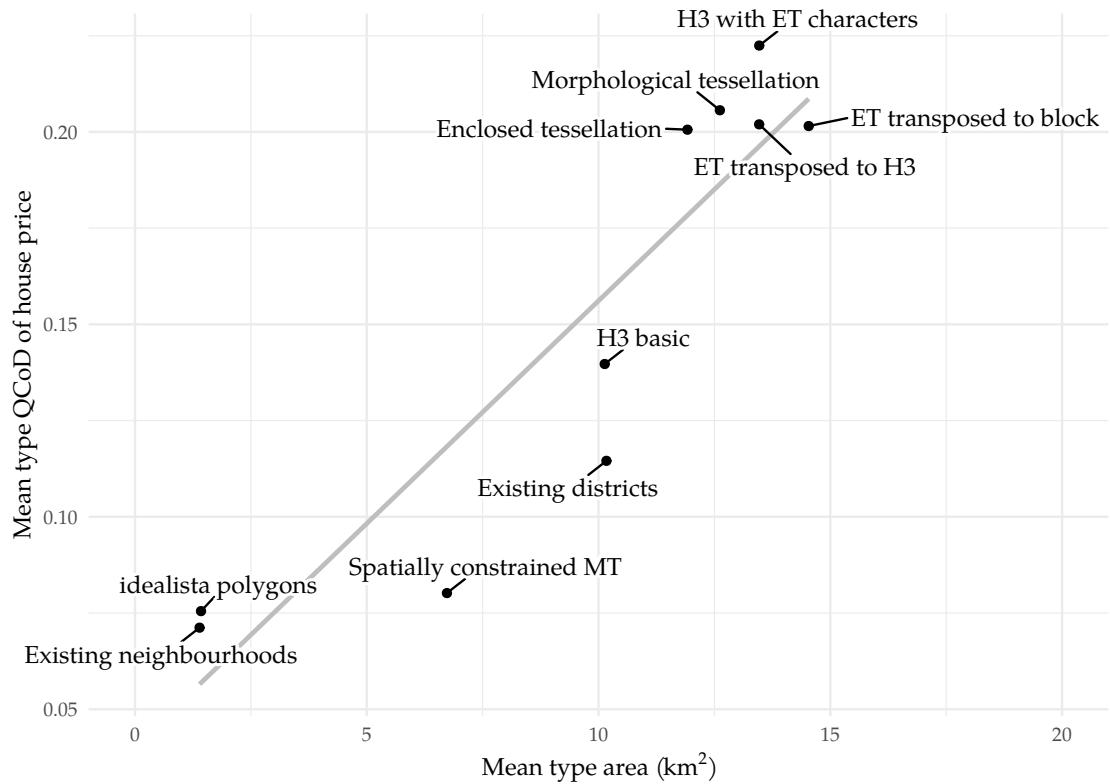


Figure 1.10: Segmentation averages of typologies plot by house price QCoD and area.

label—the same colour on the map—but are located in different parts of the city will be counted as separate units.

Table 1.2 shows the average results of the same statistics as Table 1.1 for the same segmentations, but calculated at the polygon level rather than the type level.

As a generality, the smaller a unit is, the fewer house price data points it is likely to contain: larger units are therefore generally more robust when calculating the QCoD, and types more robust than polygons. Because many of the polygons have very small areas (notably those composed of a single or very few MT or ET cells), many contain few cadastral parcels with attached house price information: of the 3,721 polygons into which the ten segmentations examined here are divided³, 1,732 (46.5%) correspond to fewer than ten house price data points, including 632 (17%) with no corresponding cadastral parcels. Figure ?? demonstrates that there is also a geographical pattern to this validation data, meaning that certain

³Seven novel and three existing.

kinds of types and polygons are more likely to have few or no relevant data from which to calculate the QCoD.

This makes a calculation of the QCoD for house prices within these polygons at best not robust and at worst impossible. For this reason, the averages shown in Table 1.2 and plot in Figure 1.12 exclude polygons containing fewer than ten house price data points. As these polygons only contain a few cadastral parcels which are geographically proximate and therefore likely also numerically proximate in terms of average house price, their QCoD is likely to be low: there is likely to be minimal dispersion in an area containing only a few properties. Including these polygons would therefore have weighted the average QCoD values to suggest lower levels of dispersion, but only on the grounds of including many small polygons, which would likely not be seen as constituting distinct housing submarkets. A version of Figure 1.12 which does not exclude these polygons in its calculations is provided in Appendix A as Figure ??.

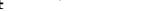
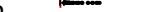
The polygon area distribution boxplots are again shown alongside the average area value, but the range for these has been artificially truncated: values (all outliers) larger than 18km^2 are beyond the range of the plot. Because the types and polygons for the spatially constrained segmentation are identical (since spatial contiguity is a condition of the clustering process, so the types it generates cannot be multipart geometries⁴), the 91.2km^2 type discussed previously is also counted as a polygon⁵. If plotted with a range inclusive of this polygon, the majority of the other boxplots would be rendered illegible, so for this reason the range is limited.

The different units represented in the two tables is evident when comparing the ‘Number of units’ column in each table: in Table 1.1 this reports the number of types (eight for the enclosed tessellation segmentation; the same when it is transposed to the H3 geometry), while in Table 1.2 this reports the number of polygons (1,666 for the enclosed tessellation segmentation; reducing to 249 when

⁴Except in the few cases where the city boundary itself contains separate geometries, such as islands.

⁵Or to be more precise, the large majority of the type is counted as a polygon: as can be seen in Figure 1.7, it also encompasses two islands (one literal and one figurative), which are counted as separate polygons.

Table 1.2: Average polygon values for each segmentation.

Segmentation	Mean QCoD	Type area (km ²)				Number of units
		Mean	Distribution			
Morphological tessellation	0.076	0.79		..	.	119
Enclosed tessellation	0.067	0.24		.	.	311
ET transposed to block	0.066	0.26		.	.	322
ET transposed to H3	0.074	0.45		206
H3 basic	0.071	0.53		..	.	116
H3 with ET characters	0.075	0.97		81
Spatially constrained MT	0.080	8.71		.	.	11
Existing neighbourhoods	0.071	1.36		73
Existing districts	0.115	9.90			.	10
idealista polygons	0.073	1.26		68

this is transposed to the H3 geometry). This difference is also reflected in the areas of the units: polygons tend to be much smaller than the types to which they belong.

Because the types and polygons of the existing spatial units (neighbourhoods, districts, and idealista polygons) are spatially coterminous (their types include few or no multipart geometries), they show the fewest changes when comparing types with polygons.

Figure 1.11 replicates Figure 1.9, but reporting the values for polygons and not types. In order to more clearly show the spread of values, the area of polygons is mapped to the x -axis logarithmically: as shown in the axis label, each axis tick multiplies by a factor of ten (0.1 km^2 , 1 km^2 , 10 km^2 , etc). This is made necessary by the distribution of areas among the polygons being examined: while there are a few notable polygons with large areas⁶, there are many with very small areas. 53.4% of polygons are smaller than 0.01 km^2 ($10,000 \text{ m}^2$, or about the size of Liverpool's Abercromby Square) and 34.2% of polygons are smaller than 0.001 km^2 ($1,000 \text{ m}^2$, or about a twelfth the size of a block in the Eixample).

Figure 1.12 replicates Figure 1.10, but again reporting the values for polygons and not types.

⁶Of the 3,721 polygons into which the ten segmentations examined here are divided, 185 (5%) have an area of more than 1 km²; 32 (0.86%) have an area of more than 5 km²; and 16 (0.43%) have an area of more than 10 km².

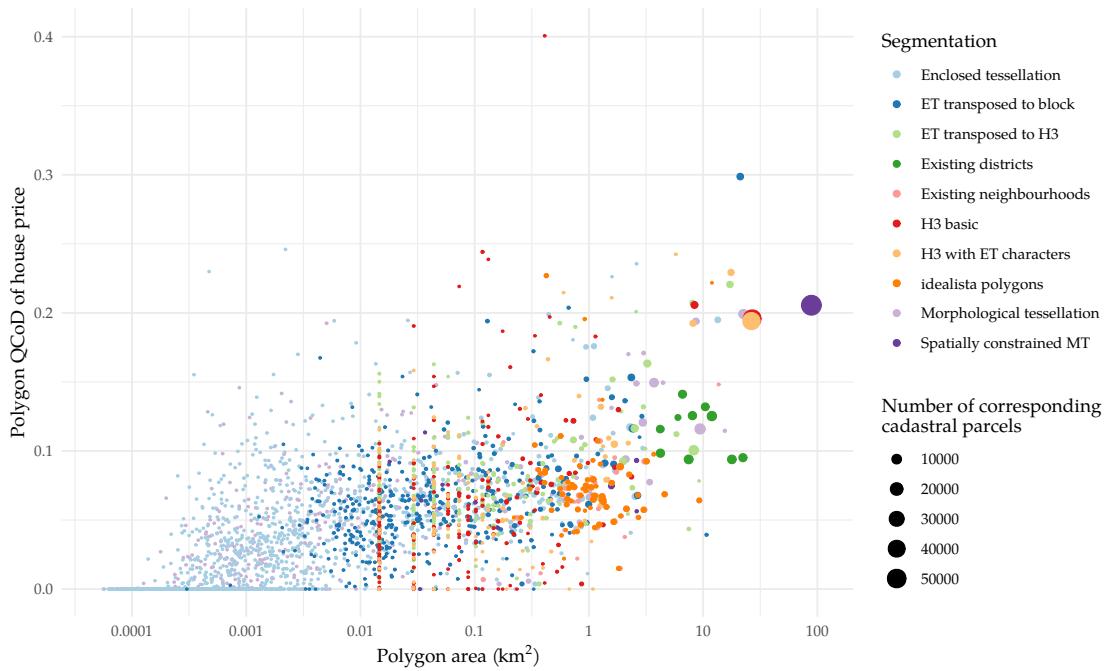


Figure 1.11: Every polygon plot by house price QCoD and area.

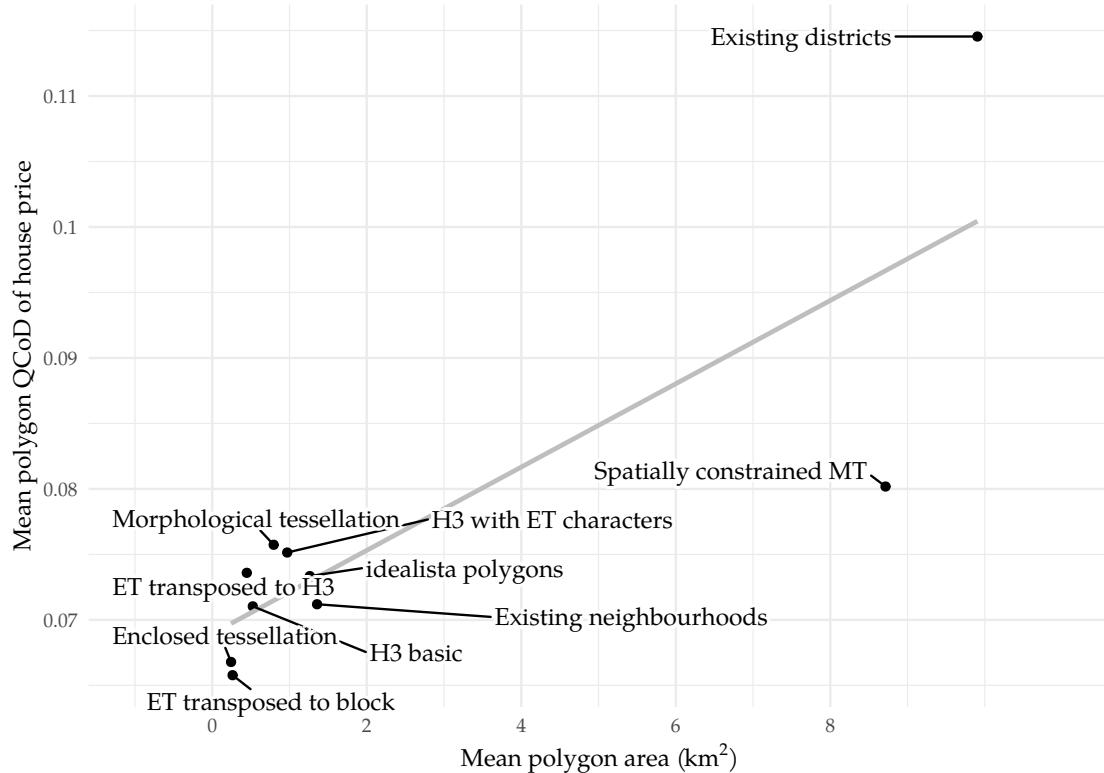


Figure 1.12: Segmentation averages of polygons plot by house price QCoD and area.