architecture

MDPI

Article

Quantifying Typological Repetition, Compactness, and Domestic Spatial Quality in Finnish Apartment Layouts: Evidence from Aurinkolahti, Helsinki

Vera Bijelić [1] and Yasmany García-Ramírez [2,*]

[1] School of Architecture, Universidad Internacional del Ecuador, Loja 110101, Ecuador; bvera@uide.edu.ec
[2] Departamento de Ingeniería Civil, Arquitectura y Geociencias, Universidad Técnica Particular de Loja, Loja 110101, Ecuador
* Correspondence: ydgarcia1@utpl.edu.ec

Abstract

Apartment floor plans have traditionally been interpreted through qualitative typological readings, but coded floor-plan datasets now enable reproducible quantitative analysis. This article examines architectural metadata from apartment units in Aurinkolahti, Helsinki, Finland, to explore typological repetition, spatial compactness, and selected indicators of domestic spatial quality. Rather than reconstructing full apartment geometries, the study treats each mapped layout as an architectural metadata record and distinguishes it from the number of built apartments represented by that layout. This stock-sensitive distinction allows repeated floor plans to be analyzed as components of housing production, not merely as individual cases. The workflow includes data cleaning, classification-coverage assessment, weighted and unweighted summaries, temporal grouping, compactness metrics, repetition indicators, kitchen-related quality markers, exploratory clustering, and statistical modeling. Results show typological concentration around a limited number of layout families and identify the 2021–2023 period as combining smaller weighted mean floor areas, higher repetition intensity, and lower kitchen natural-light shares. Coded floor-plan metadata are therefore positioned as a limited but useful intermediate form of architectural evidence: more structured than visual typological description, but not a substitute for direct plan analysis or comprehensive housing-quality assessment.

Keywords: apartment layouts; housing typology; spatial compactness; domestic spatial quality; floor-plan metadata

Check for updates

Academic Editors: Tilanka Chandrasekera and Mallikarachchige Nishan Rasanga Wijetunge

Received: 1 June 2026
Revised: 30 June 2026
Accepted: 2 July 2026
Published: 5 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Apartment floor plans are more than technical representations of individual dwellings. They also reflect broader patterns of housing production, domestic organization, spatial standardization, and architectural priorities [1]. For decades, architectural research has relied on floor plans to compare housing typologies, circulation systems, spatial hierarchies, and changing models of domestic life. Manuals and case-based repositories such as Floor Plan Manual Housing have contributed to this tradition by documenting recurring apartment configurations and offering systematic ways of interpreting housing morphology and layout organization [2]. However, recent typomorphological research has also shown that systematic quantitative methods can complement historically qualitative approaches by enabling comparison across larger urban and architectural datasets [3].

The growing availability of coded architectural datasets has created new opportunities to connect typological interpretation with computational and data-driven analysis. Rather than replacing the architectural reading of plans, these datasets make it possible to identify recurring spatial patterns that are difficult to detect through isolated case studies alone. Recent research has shown that floor-plan analysis increasingly relies on computational methods capable of extracting geometric and semantic information from architectural drawings, although standardization, dataset diversity, and generalizability remain persistent challenges [4,5]. In this context, apartment layouts become not only drawings to be interpreted individually, but also data structures capable of revealing broader tendencies in housing production.

This distinction is especially important in apartment housing, where the same layout may be repeated across several floors, stairwells, or buildings. A single floor-plan record may therefore represent dozens of built apartments. Consequently, quantitative analysis must distinguish between the number of mapped layout records and the number of dwellings represented by those records. This shift from plan-by-plan observation toward stock-sensitive interpretation expands the analytical possibilities of architectural typology research. This stock-sensitive perspective is consistent with previous Finnish housing-stock research showing that apparently diverse apartment buildings may still contain recurring flat types that structure a substantial proportion of the dwelling stock [6].

Previous studies have demonstrated the value of typological analysis for understanding both existing housing stocks and their long-term transformation. For example, Kaasalainen and Huuhka [6] examined 8745 Finnish flats built between the 1960s and 1980s and identified 18 recurring apartment types based on 10 basic layouts. Their findings showed that substantial typological repetition can emerge even in housing stocks that were not formally standardized. Similarly, quantitative apartment-layout studies have used clustering and building-drawing information to identify typical layouts and standard design patterns across large housing datasets [7]. Together, these studies frame housing not as a collection of isolated buildings, but as a system of repeated spatial solutions shaped by broader production logics.

Quantitative floor-plan research has also contributed to debates on adaptability, dwelling morphology, and domestic quality. For instance, Femenias and Geromel [8] analyzed 313 contemporary apartments modified by owner-occupiers and found that characteristics such as living-space size and floor-plan fragmentation were associated with later spatial rearrangements. Their work suggests that apartment layouts are not neutral configurations, but spatial structures that can support or constrain long-term adaptability and everyday domestic life. Recent studies on spatial quality have similarly argued that housing evaluation requires more than floor-area measurements alone. Connectivity, openness, integration, daylight access, and room organization all contribute to how residential layouts may support usability and adaptation. In particular, floor-plan openness has been explicitly operationalized through quantitative indicators such as openness score, weighted openness score, and openness potential [9,10].

Within this broader discussion, the Finnish context offers a particularly relevant case for examining the relationship between apartment typology, compactness, and domestic spatial quality. Earlier research has already shown that housing layouts in Finland carry implications for adaptability, usability, and long-term stock management [6]. More recently, Meriläinen and Tervo [11] documented significant changes in apartment building morphology in Aurinkolahti between 2000 and 2023, including increasing building depths, a growing number of units, and more complex typological structures, with implications for natural light, adequate space, and usability.

The present study builds on this context by analyzing the dataset titled Architectural floor plan data from apartment units in Aurinkolahti, Helsinki, Finland, 2000–2023 [12]. The dataset was compiled using open GIS information from kartta.hel.fi and floor-plan drawings obtained from public building-permit archives. It contains coded metadata from apartment units completed between 2000 and March 2023 and was originally developed to classify recurring layout types and examine spatial organization in Finnish apartment housing. As such, it provides a rare opportunity to study apartment layouts at a scale that extends beyond conventional architectural case studies while remaining grounded in specific floor-plan evidence.

Rather than redrawing or vectorizing floor plans, this study approaches the dataset as a source of coded architectural metadata. This position differs from automatic floor-plan recognition approaches that attempt to extract walls, doors, rooms, and other geometric objects from rasterized plan images [4]. The analysis focuses on three interconnected dimensions. The first is typological repetition, understood as the extent to which individual layout records represent repeated built apartments. The second is spatial compactness, examined through floor area, unit depth, and area-per-room relationships rather than floor area alone. The third is domestic spatial quality, approached through coded indicators such as kitchen size, kitchen natural light, and bedroom-access arrangements. Although these variables do not provide a complete assessment of housing quality or environmental performance, they offer architecturally meaningful evidence that can support comparative analysis.

Methodologically, the study also responds to ongoing discussions about the use of semantically enriched floor-plan datasets in architectural research. Recent literature has emphasized that coded datasets become significantly more valuable when geometric and semantic attributes are analyzed together, particularly through graph-based and clustering approaches capable of identifying recurring spatial patterns and typological families [13,14]. At the same time, researchers have noted important limitations related to benchmarking, dataset diversity, and the risk of reducing housing quality to narrow quantitative indicators [15]. This study acknowledges those limitations and treats quantitative classification as a complement to architectural interpretation rather than a replacement for the close reading of floor plans.

This study therefore uses coded apartment-layout metadata to examine three related tendencies in the Aurinkolahti housing stock: typological repetition, spatial compactness, and selected domestic spatial-organization markers. The aim is not to provide a complete assessment of housing quality, nor to reconstruct full floor-plan geometry, but to show how repeated apartment layouts can be analyzed as part of a represented housing stock. The analysis combines weighted summaries, typological profiles, schematic plan diagrams, exploratory clustering and statistical checks to connect quantitative patterns with architectural interpretation.

The article is guided by four research questions:

1. How can coded apartment floor-plan metadata be organized into a reproducible workflow for analyzing typological repetition, compactness, and domestic spatial-quality markers?
2. How does the interpretation of apartment typology change when mapped layout records are weighted by the number of represented apartments?
3. What temporal patterns can be identified in layout repetition, compactness, and selected domestic-quality indicators across the Aurinkolahti apartment stock?
4. To what extent can exploratory data-driven classification complement architectural typological interpretation without replacing the close reading of floor plans?

Unlike previous analyses of Finnish apartment housing, which either relied on fixed typological categories without weighting [6] or examined building-level morphology without

unit-level metadata analysis [11], the present study combines apartment-count weighting, multi-indicator compactness assessment, and exploratory data-driven classification within a single reproducible workflow applied to unit-level coded data.

2. Materials and Methods

2.1. Research Design

This study applies a quantitative and interpretive workflow to coded apartment-layout metadata. Each mapped layout is treated as one architectural record, while the number of identical apartments represented by that layout is used as a weighting variable. This distinction allows the analysis to describe both the diversity of mapped layouts and their actual representation in the built housing stock (see Figure 1).

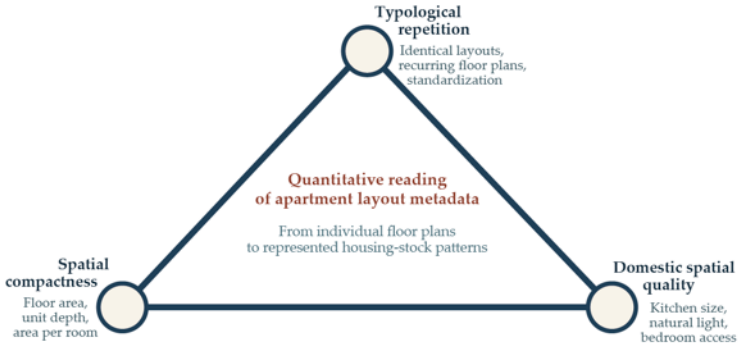


Figure 1. Conceptual framework for analyzing apartment layout metadata through typological repetition, spatial compactness, and domestic spatial quality.

To strengthen the architectural readability of the quantitative workflow, the analysis also includes representative schematic plans for the main typological families and data-driven clusters. These diagrams provide a spatial counterpart to the tabular and statistical summaries and clarify how compactness, circulation, depth, and kitchen-light conditions are expressed at the level of apartment organization.

The methodological sequence consisted of six stages: dataset import and cleaning; classification-coverage assessment; construction of weighted; calculation of compactness, repetition, and domestic-quality indicators; and exploratory data-driven classification.

2.2. Data Source

The dataset was obtained from Zenodo record 19894248, titled Architectural floor plan data from apartment units in Aurinkolahti, Helsinki, Finland, 2000–2023. The dataset presents a mapping of apartment units from the housing stock of Aurinkolahti, Helsinki, completed between 2000 and March 2023 [12]. Data were collected using open GIS data available from kartta.hel.fi and floor-plan drawings from public building-permit archives in February and March 2023. The dataset was created to classify recurring layout types and to study spatial organization in urban Finnish apartments and its development during the 2000s. The spreadsheet includes information such as unit identifier, building name, permit year, completion year, number of floors, building type, studied floor plan, number of identical floor plans in the building, number of units in each floor plan, total number of identical units, unit floor area, unit depth, room count, initial and final layout classification, bedroom-entrance position, kitchen size, natural light in the kitchen, and analysis notes.

2.3. Analytical Dimensions

The study was organized around three analytical dimensions: typological repetition, spatial compactness, and domestic spatial quality. These dimensions were selected because

they allow the dataset to be read not only as a list of floor-plan categories, but as a structured representation of recurring layout production, dimensional organization, and selected domestic-quality attributes.

Typological repetition refers to the extent to which a mapped apartment layout represents multiple identical apartments. Spatial compactness refers to dimensional relationships among floor area, unit depth, and room structure. Domestic spatial quality refers to coded attributes associated with the kitchen and bedroom-access arrangement. These quality-related markers are not treated as complete measures of habitability, but as partial indicators available within the dataset. The analytical dimensions and derived indicators are summarized in Table 1.

Table 1. Analytical dimensions and derived indicators.

Analytical Dimension	Description	Derived Indicators
Typological repetition	Measures how strongly apartment layouts recur across the housing stock	Repetition intensity; represented apartment count; building-level repetition
Spatial compactness	Describes dimensional relationships within apartment layouts	Floor area; unit depth; area per room; depth-area relationships
Domestic spatial quality	Captures selected coded attributes related to everyday interior quality	Kitchen larger than 7 m^2; kitchen with natural light; bedroom access from entry/corridor
Temporal development	Allows comparison across construction periods	Completion-period grouping
Data-driven layout families	Explores whether coded attributes form interpretable groups	Gower distance; PAM clustering; silhouette-based cluster selection

Note. Indicators are reported descriptively in the Section 3. Formula-based definitions were simplified in the main text, and full calculations are provided in the reproducible R workflow.

2.4. Data Preparation and Cleaning

The spreadsheet was imported into R 4.4.3 (version 4.3.1; R Foundation for Statistical Computing, Vienna, Austria) [16] using RStudio (version 2023.09.1; Posit Software, PBC, Boston, MA, USA) [17]. Column names were standardized, numeric fields were converted into numeric format, and text-based categorical fields were cleaned for consistency. The workflow retained the original logic of the dataset while creating derived variables for analysis. The number of represented apartments associated with each mapped layout was used as the weighting variable.

Binary indicators were derived from the spreadsheet annotations for kitchen size, kitchen natural light, and bedroom access. Marked cells were coded as presence (1) and empty or missing cells as absence (0). While this approach is consistent with the original dataset structure, the resulting indicators remain dependent on the completeness and consistency of the source annotations. The R code used for data processing and analysis is publicly available in the accompanying Zenodo repository to ensure transparency and reproducibility.

2.5. Operational Definition of Typological Families

The labels Type A, Type B, Type C, Type D, Type E, and Type F refer to grouped layout families defined in the source dataset. They are used here as operational categories specific to the Aurinkolahti dataset, not as universal apartment typologies. Room-group labels indicate the number of bedrooms: ST refers to studio apartments, 1BR to one-bedroom apartments, 2BR to two-bedroom apartments, and 3BR+ to apartments with three or more bedrooms.

Each typological family was profiled using the architectural variables available in the dataset, including mapped records, represented apartment count, room group, floor area, unit depth, kitchen-size marker, kitchen natural-light marker, and bedroom-access marker. This allowed the labels to be interpreted through measurable spatial attributes rather than as purely nominal categories.

Figure 2 visualizes these operational definitions through representative schematic plans. The schemes have been reordered alphabetically in rows, from Type A to Type F, to improve readability and consistency. They synthesize dominant spatial-organization patterns inferred from the coded dataset and are intended as comparative interpretive diagrams, not as exact reproductions of individual permit plans.

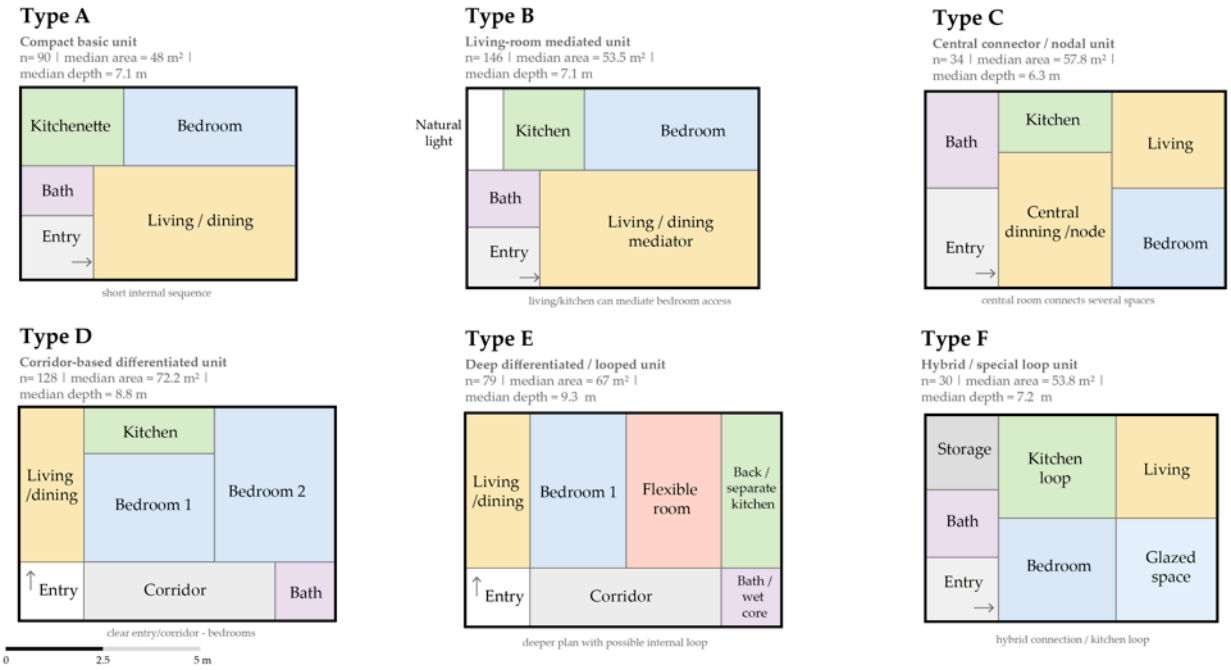


Figure 2. Schematic spatial organization of the six operational typological families in the Aurinkolahti apartment dataset. The drawings are schematic interpretive diagrams derived from coded dataset variables and are not exact reproductions of individual permit plans.

2.6. Classification Coverage and Weighting Logic

Weighted summaries were used to account for the fact that one mapped layout record may represent several identical apartments. For each record, the weighting variable was the total number of identical apartments represented by that layout. Weighted means and weighted shares were then calculated using this apartment-count weight. In practical terms, layouts repeated many times in the built stock contributed more strongly to weighted results than layouts occurring only once. The full calculation procedure is provided in the reproducible R script accompanying the article.

2.7. Compactness Indicators

Spatial compactness was examined through floor area, unit depth, and area per room. Area per room was calculated by dividing apartment floor area by the coded room count. Studio units were treated as one-room dwellings for this calculation. Two additional proportional descriptors were calculated in the R workflow—the depth-area ratio and the depth-square-root-area ratio—but these were used only as exploratory checks and are not emphasized in the main result tables.

2.8. Repetition and Standardization Analysis

Repetition intensity was defined as the number of identical apartments represented by each mapped layout record. At the period or typological-family level, repetition was summarized using weighted means so that highly repeated layouts had proportionally greater influence on the result. This approach allowed repetition to be interpreted as a stock-level indicator of standardization rather than as a simple count of unique layout records.

2.9. Domestic Spatial-Quality Indicators

This study does not attempt to measure the overall quality of individual apartments. Apartment quality depends on multiple spatial, environmental, technical, and experiential factors, many of which are not available in the dataset. These include, among others, façade orientation, the number of cardinal directions served by windows, detailed window dimensions, daylight depth, acoustic conditions, accessibility, ventilation performance, the number of apartments served by a staircase or elevator, and post-occupancy experience. The analysis therefore focuses only on selected coded spatial-quality markers available in the source dataset: kitchens larger than 7 m^2, kitchens with natural light, and bedroom access from the entry or corridor.

These variables are interpreted as partial indicators of spatial organization and domestic functionality, not as comprehensive measures of housing quality. Kitchen size and kitchen natural light provide limited evidence about the position and environmental exposure of a key domestic space, while bedroom access from the entry or corridor offers a simplified indication of circulation and privacy structure. The area-per-room indicator complements these variables by relating apartment size to room count, but it also remains a dimensional descriptor rather than a direct measure of quality.

The three domestic spatial-organization markers were coded as binary variables: presence or absence of a kitchen larger than 7 m^2, presence or absence of kitchen natural light, and presence or absence of bedroom access from the entry or corridor. Weighted shares were then calculated for typological families and completion periods using the apartment-count weight. The corresponding R code is provided in the reproducible workflow.

2.10. Crosswalk Between Detailed and Grouped Classifications

The dataset contains two levels of typological classification: detailed layout labels describing specific apartment configurations, and broader grouped classifications representing more general typological families. Examining the relationship between these two levels is important for understanding how individual layout variations contribute to wider patterns within the housing stock.

To support this analysis, a crosswalk was created, linking detailed layout labels to their corresponding grouped classifications. This procedure provides a transparent overview of the classification structure and helps clarify how numerous fine-grained apartment layouts are consolidated into a smaller number of broader typological categories. By making these relationships explicit, the analysis facilitates the interpretation of both typological diversity and typological concentration within the dataset.

2.11. Exploratory Data-Driven Classification

In addition to the predefined typological classifications contained in the dataset, an exploratory clustering analysis was conducted to investigate whether recurring apartment layout families could be identified from combinations of spatial, typological, and domestic-quality attributes. The objective was not to replace the existing architectural classification

system, but rather to examine whether data-driven methods could reveal interpretable patterns that complement conventional typological analysis.

Because the dataset contains a mixture of numeric and categorical variables, similarities between apartment layout records were calculated using Gower distance, which is specifically designed for mixed-data environments. The resulting dissimilarity matrix was then analyzed using Partitioning Around Medoids (PAM) clustering, a robust clustering method that identifies representative observations while reducing sensitivity to extreme values.

To determine an appropriate number of clusters, several alternative solutions were evaluated using average silhouette width. The selected solution was chosen on the basis of its relative interpretability and internal consistency rather than as evidence of definitive apartment categories. Consequently, the resulting clusters should be understood as exploratory layout families that highlight recurring combinations of characteristics within the dataset.

2.12. Exploratory Statistical Modelling

To complement the descriptive analyses, two exploratory statistical models were fitted to examine associations among the variables contained in the coded floor-plan dataset. These models were intended as analytical tools for identifying patterns and relationships within the data rather than for establishing causal effects.

The first model examined floor area as a function of selected temporal, dimensional, typological, and domestic-quality attributes:

$$A_i = \beta_0 + \beta_1 Y_{c,i} + \beta_2 D_i + \beta_3 T_i + \beta_4 K_i + \varepsilon_i \quad (1)$$

where A_i is the floor area, $Y_{c,i}$ is the completion year centered around the mean completion year of the analytical subset, D_i is the unit depth, T_i represents typological-family effects, K_i is the kitchen natural-light marker, and ε_i is the error term.

The second model examined kitchen natural light as a binary outcome. The probability p_i that apartment layout i had coded kitchen natural light was modeled using a logistic specification:

$$log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 Y_{c,i} + \beta_2 A_i + \beta_3 D_i + \beta_4 T_i + \beta_5 L_i \quad (2)$$

where A_i is the floor area, D_i is the unit depth, T_i represents the typological-family effects, and L_i indicates whether the kitchen is larger than 7 m^2. Both models were interpreted as exploratory summaries of association within the dataset.

2.13. Software and Reproducibility

All analyses were implemented in R 4.4.3 [16] using RStudio [17] as the integrated development environment. Spreadsheet import, data cleaning, transformation, statistical analysis, clustering, visualization, and output generation were performed through a reproducible R workflow.

Spreadsheet data were imported with readxl 1.4.5, and column-name cleaning was performed with janitor 2.2.1. Data manipulation relied on the tidyverse 2.0.0 ecosystem, including dplyr 1.2.1, readr 2.2.0, tidyr 1.3.1, tibble 3.3.1, purrr 1.2.2, forcats 1.0.0, stringr 1.6.0, and lubridate 1.9.4. Excel outputs were exported with writexl 1.5.4.

Exploratory model outputs were extracted with broom 1.0.9. Mixed-variable similarity and clustering were implemented with cluster 2.1.8.1, while multivariate and clustering visualization support was provided by FactoMineR 2.12 and factoextra 1.0.7.

3. Results

3.1. Dataset Structure and Classification Coverage

The cleaned local dataset contained 624 mapped apartment-unit records representing 3789 apartments across 116 buildings completed between 2000 and 2023. Although the Zenodo record describes 622 unique units representing 3787 apartments, the local cleaned file contained two additional mapped records and two additional represented apartments. This difference was retained transparently and did not require manual removal.

Classification coverage was assessed before typological interpretation. As shown in Table 2, 507 mapped records had a valid grouped typological classification, representing 2776 apartments or 73.3% of the weighted stock. The remaining 117 records represented 1013 apartments, equivalent to 26.7% of the weighted stock. These unclassified observations were retained for dataset-level summaries but excluded from typological dominance claims.

Table 2. Classification coverage of the grouped apartment typology.

Classification Status	Records	Weighted Apartments	Share of Records	Share of Weighted Apartments
Classified typology	507	2776	81.2%	73.3%
Unclassified/missing typology	117	1013	18.8%	26.7%

3.2. Weighting Sensitivity

The comparison between mapped records and represented apartments confirmed that weighting is necessary. As seen in Figure 3, the relative importance of some typological groups changes when repeated units are considered. The largest typological weighting shift occurred for Type B, whose weighted share was 4.0 percentage points lower than its unweighted record share. At the temporal level, the largest shift occurred in the 2021–2023 period, where the weighted share was 3.7 percentage points higher than the unweighted record share. These differences indicate that raw row counts do not fully represent the built apartment stock and confirm the need to account for layout repetition in subsequent analyses.

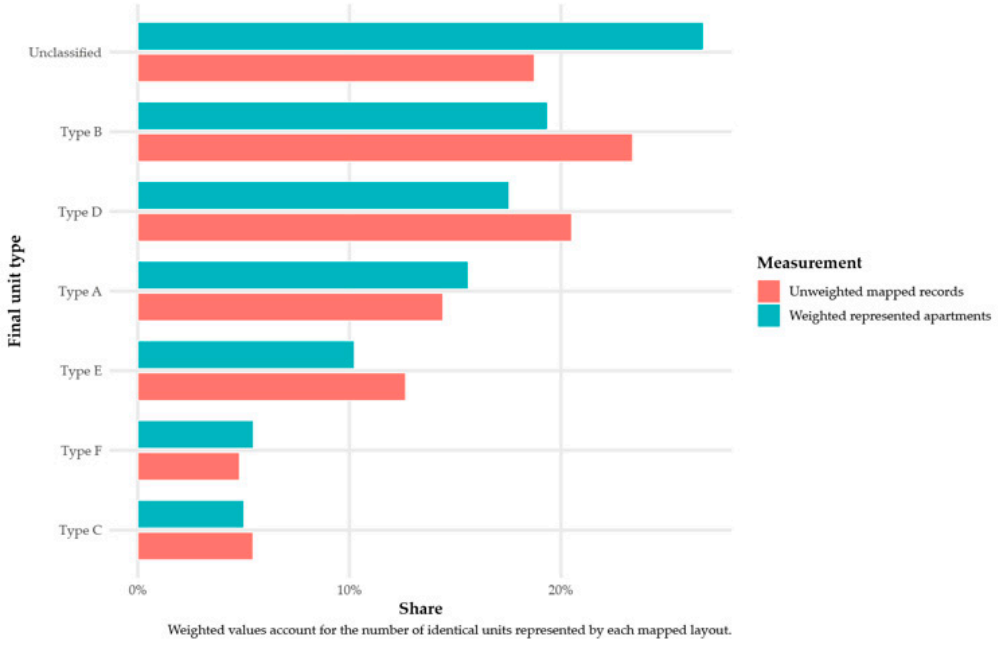


Figure 3. Effect of weighting on typological composition: mapped records versus represented apartments.

3.3. Characterization of Grouped Typological Families

Before interpreting the distribution of typological families, the grouped labels were empirically profiled using the coded architectural metadata available in the dataset. This step is necessary because labels such as Type A, Type B, or Type D are dataset-specific grouped classifications rather than self-explanatory architectural types. Table 3 summarizes each family in terms of mapped records, represented apartments, share of the classified weighted stock, dominant room group, dominant building type, floor area, unit depth, kitchen-related indicators, bedroom-access marker, and repetition intensity.

Table 3. Empirical profile of grouped typological families.

Typological Family	Dominant Room Group	Dominant Building Type	Mean Floor Area (m^2)	Median Floor Area (m^2)	Mean Unit Depth (m)	Median Unit Depth (m)	Kitchen > 7 m^2 (%)	Kitchen Natural Light (%)	Bedroom Access from Entry/ Corridor (%)	Mean Repetition
Type A	1BR	kytketty point block and row houses (9)	48.6	45.5	7.1	7	10.00%	5.20%	5.70%	9
Type B	1BR	corner lamella (3 + 4 + 4)	55	51	7.3	7.1	35.10%	32.20%	14.40%	7.7
Type C	1BR	point block (6)	63.6	58	7.2	6.4	15.80%	22.10%	27.90%	7.3
Type D	2BR	point block (6)	69.3	68.5	8.9	9.1	64.50%	24.10%	66.60%	6.5
Type E	1BR	point block (5)	63.9	64.5	8.7	8	38.90%	40.20%	87.90%	6.4
Type F	1BR	point block (6)	55.2	52	7.8	8.2	40.60%	24.20%	11.60%	9.2

Note. 1BR = one-bedroom apartment; 2BR = two-bedroom apartment. Types A–F are grouped typological-family labels from the source dataset and should be interpreted as operational categories specific to this analysis.

Type B was the most represented grouped family, with 146 mapped records and 734 represented apartments, accounting for 26.4% of the classified weighted stock. It was dominated by 1BR layouts and was empirically characterized as a mid-sized, shallower family with a comparatively high kitchen natural-light share. Type D was the second most represented family, with 128 records and 665 represented apartments, accounting for 24.0% of the classified weighted stock. In contrast to Type B, Type D was dominated by 2BR layouts and showed a larger average floor area and greater unit depth. Type A followed with 90 records and 592 represented apartments, accounting for 21.3% of the classified weighted stock; this was dominated by 1BR layouts and profiled as a smaller, shallower family with a low kitchen natural-light share.

The remaining grouped families represented smaller shares of the classified stock. Type E accounted for 14.0% and was characterized as a mid-sized, moderate-depth 1BR family with a high kitchen natural-light share and a high bedroom-corridor access share. Type F accounted for 7.5% and was profiled as a mid-sized, moderate-depth 1BR family with moderate kitchen natural-light share. Type C represented 6.8% of the classified weighted stock and was profiled as a mid-sized, shallower 1BR family with moderate kitchen natural-light share.

These profiles do not define the complete geometric configuration of each type. Rather, they provide an empirical basis for interpreting the typological labels used in the subsequent analyses. Architectural claims about corridor organization, kitchen integration, room adjacency, façade exposure, or orientation would require direct inspection or vectorized reconstruction of the original floor plans.

3.4. Typological Composition and Temporal Structure

Among the classified apartments, Type B was the most represented typological family, accounting for 734 represented apartments. Type D followed with 665 apartments, and Type A with 592 apartments. Together, these three types represented 71.7% of the classified weighted stock.

The temporal distribution showed that the 2000–2005 period represented the largest weighted share of apartments, with 1521 represented units. The most recent period, 2021–2023, was smaller in absolute terms but became analytically important because of its compactness and repetition patterns. Figure 4 presents the typological composition of apartment layouts by completion period. The figure highlights shifts in the relative prevalence of different typological families and provides context for the subsequent analyses of repetition and compactness.

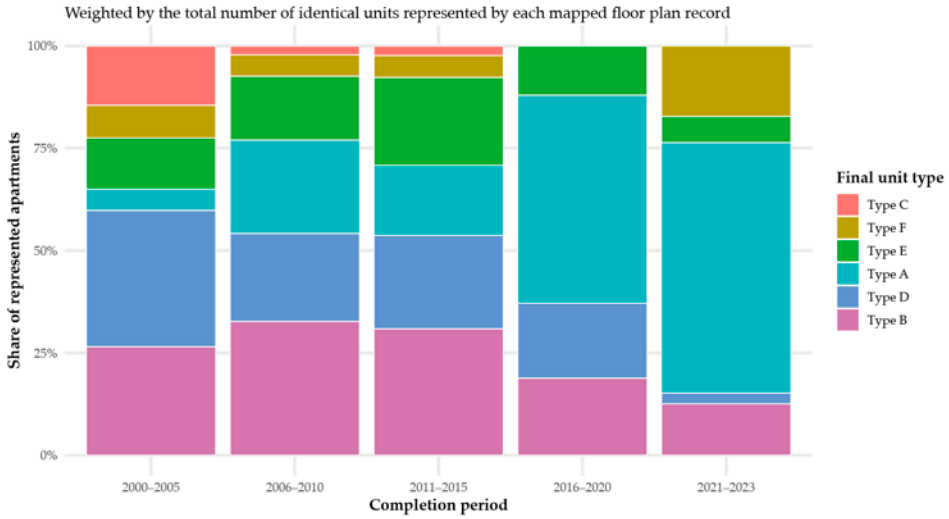


Figure 4. Typological composition of apartment layouts by completion period.

3.5. Spatial Compactness

The weighted mean floor area was 58.91 m^2, and the approximate weighted median was 56.00 m^2. The weighted mean unit depth was 7.87 m, with an approximate weighted median of 7.40 m. The weighted mean area per room was 43.80 m^2/room.

The temporal comparison showed a relevant contrast. The lowest weighted mean floor area occurred in 2021–2023, with 52.11 m^2, while the highest occurred in 2006–2010, with 61.63 m^2. This suggests that the most recent period is associated with more compact apartment production. Figure 5 illustrates the temporal variation in apartment floor area across the five periods. This visualization offers an initial indication of changes in dwelling size over time and serves as a basis for the more detailed compactness analysis presented later in the study.

To complement the temporal analysis, the relationship between apartment floor area and unit depth was examined directly (Figure 6). The scatterplot shows how dimensional compactness varies across mapped layouts, while point size represents repetition intensity. This visualization makes it possible to distinguish isolated layout records from floor plans that represent larger numbers of identical apartments.

Figure 6 shows that compactness is not captured by floor area alone. Layouts with similar floor areas may differ in unit depth, and highly repeated layouts appear across different size-depth combinations. This supports the interpretation of compactness as a relationship between apartment size, spatial proportion, and repetition rather than as a single-dimensional measure. The schematic plans in Figure 2 provide a complementary

architectural reading of these differences by contrasting shallower compact layouts with deeper and more differentiated apartment organizations.

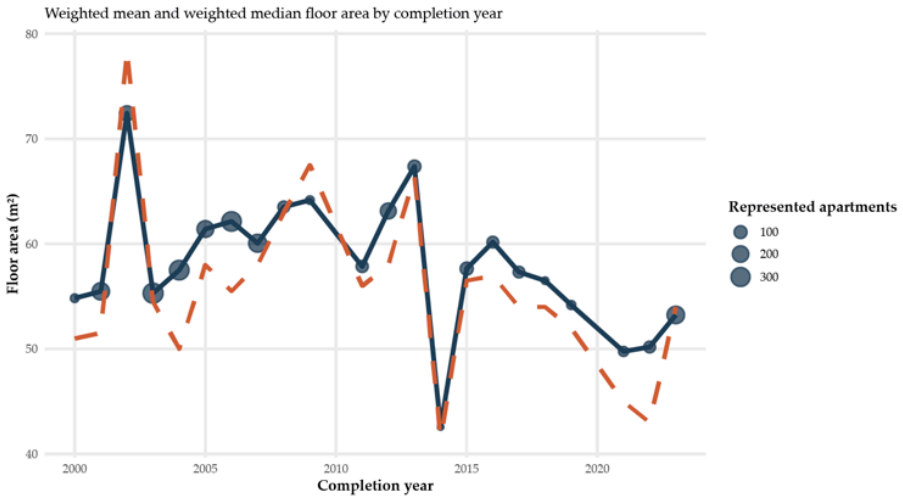


Figure 5. Temporal variation in apartment floor area.

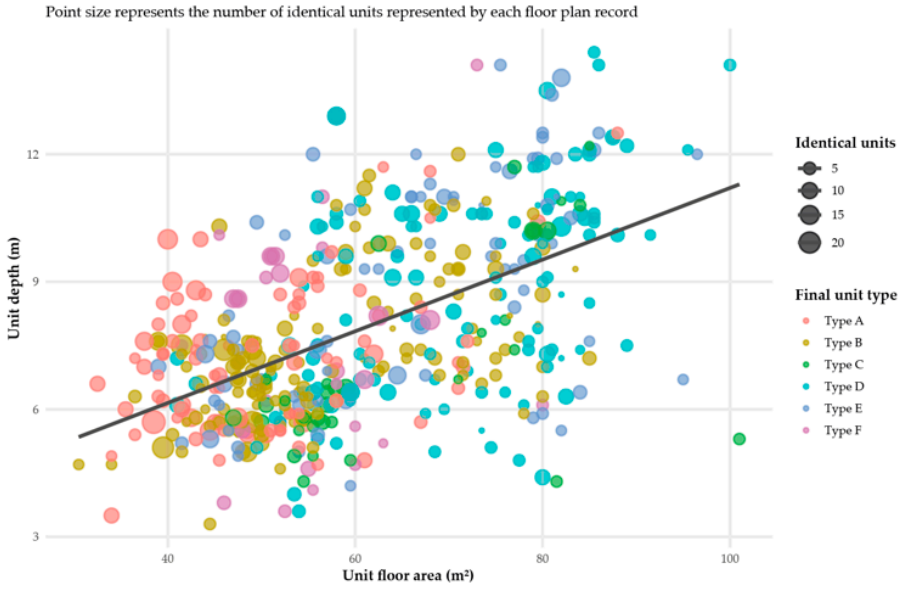


Figure 6. Relationship between apartment floor area and unit depth.

3.6. Domestic Spatial Quality: Kitchen Size and Natural Light

The weighted share of apartments with kitchens larger than 7 m^2 was 26.7%. The weighted share of apartments with kitchen natural light was 17.8%, while the share with bedroom entrance from the entry/corridor was 26.4%.

The temporal contrast was especially marked for kitchen natural light. The highest period-level share occurred in 2000–2005, with 24.3%, whereas the lowest occurred in 2021–2023, with only 5.7%. This suggests that recent compactness may coincide with a reduction in this partial marker of domestic spatial quality. This pattern is shown in Figure 7.

Figure 7 shows that the decline in kitchen natural light is not evenly distributed across typological families or completion periods. The 2021–2023 period combines a low share of kitchens with natural light with the compactness and repetition patterns reported above. In architectural terms, this indicates a changing relationship between kitchen space, façade access, and apartment depth. However, kitchen natural light is treated here only as a partial spatial-organization marker, not as a comprehensive measure of housing quality.

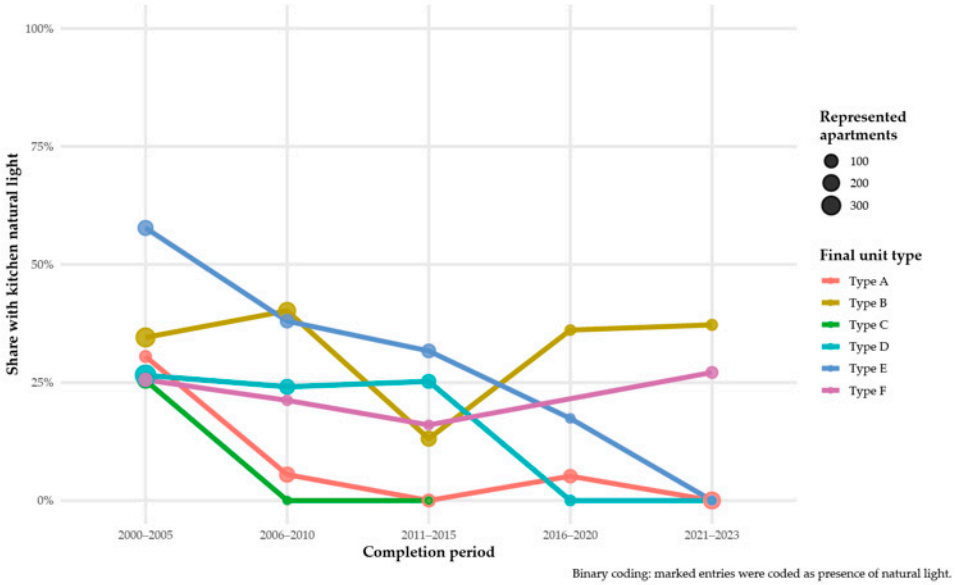


Figure 7. Share of apartments with kitchen natural light by completion period and typological family.

3.7. Typological Repetition and Standardization

The strongest layout repetition appeared in As Oy Helsingin Hyperion, completed in 2023, where one mapped layout represented 45 identical apartments. At the period level, the highest mean repetition occurred in 2021–2023, with a mean repetition intensity of 14.87.

This finding indicates that the most recent period combines three relevant characteristics: the lowest weighted mean floor area, the lowest share of kitchen natural light, and the highest mean repetition intensity. This combination suggests a shift toward more compact and standardized apartment production in the most recent years of the dataset. Repetition patterns by period are shown in Figure 8.

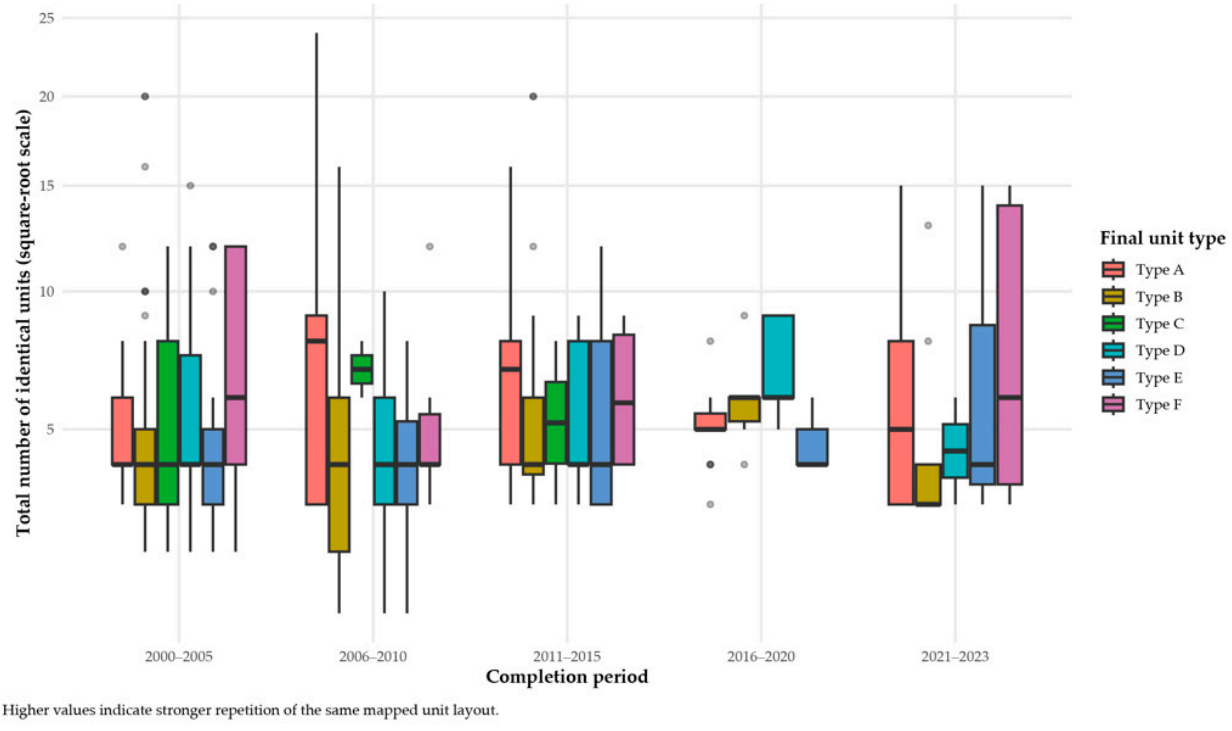


Figure 8. Typological repetition across completion periods.

3.8. Detailed-to-Grouped Typological Structure

The detailed-to-grouped typological heatmap showed how fine-grained layout labels were consolidated into broader typological families. This visualization supports the interpretation that the dataset contains detailed variation, but that much of the classified stock is organized around a smaller number of broader typological groups. This structure is shown in Figure 9.

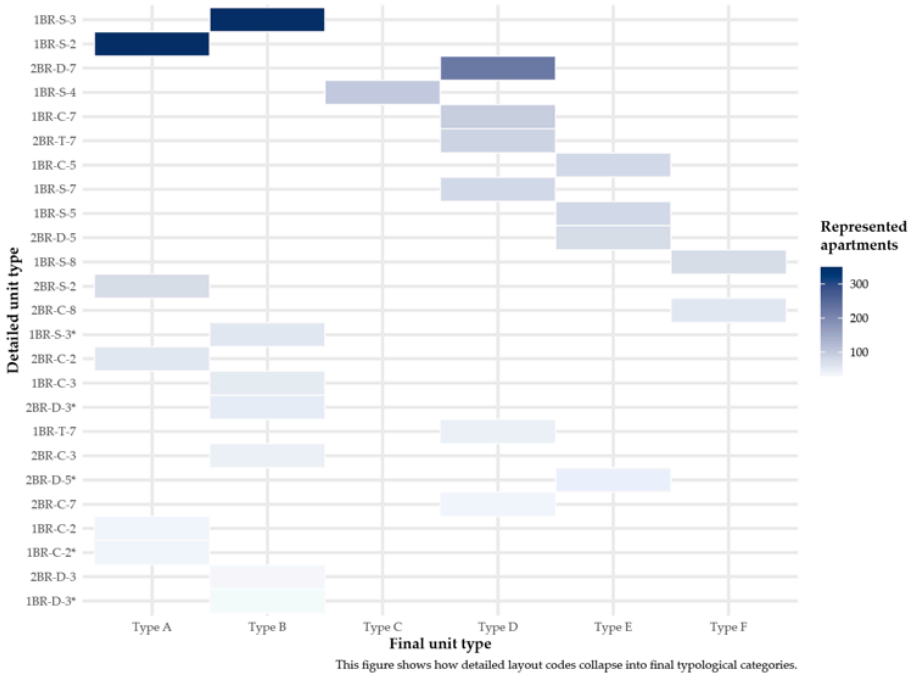


Figure 9. Crosswalk between detailed layout labels and broader typological families.

3.9. Data-Driven Layout Families

The exploratory clustering selected a two-cluster solution based on average silhouette width. The silhouette value was 0.31, indicating moderate rather than strong separation. The clusters should therefore be interpreted as exploratory layout tendencies rather than definitive architectural typologies.

Cluster 1 represented 1779 apartments and was dominated by Type B and 1BR layouts. Its mean floor area was 50.74 m^2, mean depth was 6.96 m, and kitchen natural light was present in 15.3% of represented apartments. This cluster can be interpreted as a compact and relatively shallow layout tendency.

Cluster 2 represented 997 apartments and was dominated by Type D and 2BR layouts. Its mean floor area was 73.51 m^2, mean depth was 9.50 m, and kitchen natural light was present in 40.4% of represented apartments. This cluster corresponds to a larger and deeper layout tendency.

The cluster profiles are summarized in Table 4, while Figure 10 shows the ordination of the two data-driven layout families. Figure 2 provides the corresponding schematic reference for interpreting these clusters spatially.

Table 4. Spatial profiles of the exploratory data-driven layout clusters.

Cluster	Weighted Apartments	Dominant Family	Dominant Room Group	Mean Floor Area	Mean Depth
1	1779	Type B	1BR	50.74 m^2	6.96 m
2	997	Type D	2BR	73.51 m^2	9.50 m



Figure 10. Data-driven classification of apartment layout families.

3.10. Exploratory Statistical Models

The two exploratory statistical models were fitted using the classified analytical subset with complete model variables. Both models used 507 mapped apartment-layout records, representing 2776 apartments after weighting by the number of identical units associated with each record. Completion year was centered around the mean year of the modeling subset, which was 2008.507. Therefore, the model intercepts refer to the centered temporal condition rather than to an uninterpretable year-zero baseline.

The first model was a weighted linear regression with apartment floor area as the dependent variable (see Table 5). The model explained a moderate share of floor-area variation, with $R^2 = 0.485$ and adjusted $R^2 = 0.476$. Unit depth was positively associated with floor area, indicating that deeper apartment layouts tended to have larger floor areas. Completion year also showed a positive association with floor area, although the magnitude of this effect was small compared with the typological and dimensional predictors. All reported typological-family coefficients were positive and statistically significant relative to the reference family. The largest estimated differences were observed for Type D and Type C. Kitchen natural light was also positively associated with floor area, suggesting that layouts with coded kitchen natural light tended to be larger after accounting for centered completion year, unit depth, and typological family.

The second model was a weighted logistic regression with kitchen natural light as the binary outcome (see Table 6). The model indicated that completion year was negatively associated with the odds of kitchen natural light. In other words, more recent completion years were associated with lower odds of a kitchen being coded as naturally lit, after accounting for floor area, unit depth, typological family, and kitchen size. Floor area showed a small positive association, whereas unit depth showed a negative association. Thus, larger apartments had slightly higher odds of kitchen natural light, while deeper units had lower odds, all else equal.

Table 5. Exploratory weighted linear model for apartment floor area.

Predictor	Estimate	95% CI	*p*-Value
Intercept	28.40	24.70 to 32.10	<0.001
Centered completion year	0.210	0.058 to 0.362	0.0068
Unit depth	2.615	2.186 to 3.043	<0.001
Type B	5.607	2.784 to 8.430	<0.001
Type C	16.000	11.700 to 20.200	<0.001
Type D	16.400	13.300 to 19.500	<0.001
Type E	10.000	6.680 to 13.400	<0.001
Type F	4.580	0.675 to 8.480	0.0216
Kitchen natural light	6.420	4.230 to 8.610	<0.001

Model fit: $R^2 = 0.485$, adjusted $R^2 = 0.476$, $n = 507$, $p < 0.001$. Completion year was centered around the mean completion year of the modeling subset (2008.507). The model was weighted by the number of identical apartments represented by each mapped layout record. Coefficients should be interpreted as exploratory associations within the coded dataset, not as evidence of causal effects.

Table 6. Exploratory weighted logistic model for kitchen natural light.

Predictor	Odds Ratio	95% CI	*p*-Value
Intercept	0.055	0.029 to 0.102	<0.001
Centered completion year	0.871	0.852 to 0.891	<0.001
Floor area	1.018	1.007 to 1.030	0.0018
Unit depth	0.870	0.822 to 0.921	<0.001
Type B	2.591	1.700 to 4.060	<0.001
Type C	1.239	0.692 to 2.230	0.471
Type D	0.546	0.337 to 0.901	0.0157
Type E	4.362	2.760 to 7.050	<0.001
Type F	1.846	1.070 to 3.200	0.0271
Kitchen larger than 7 m^2	12.400	9.340 to 16.600	<0.001

Model fit: null deviance = 3079.64; residual deviance = 2238.46; n = 507. Completion year was centered around the mean completion year of the modeling subset (2008.507). The model was weighted by the number of identical apartments represented by each mapped layout record. Odds ratios should be interpreted as exploratory associations within the coded dataset, not as evidence of causal effects.

The strongest association in the logistic model was observed for kitchens larger than 7 m^2. Layouts with kitchens larger than 7 m^2 had markedly higher odds of kitchen natural light. Typological-family effects were also visible: Type B, Type E, and Type F showed higher odds of kitchen natural light relative to the reference family, whereas Type D showed lower odds. Type C did not show a statistically clear difference from the reference family.

The exploratory models support the descriptive interpretation developed in the previous subsections. Floor area was associated with unit depth, typological family, and kitchen natural light, while kitchen natural light was associated with completion year, floor area, unit depth, typological family, and kitchen size. The logistic model is especially relevant because it reinforces the temporal pattern observed descriptively: more recent completion years were associated with lower odds of kitchen natural light even after controlling for apartment size, depth, typological family, and kitchen size. Nevertheless, these models should be interpreted as exploratory associations within the coded and weighted dataset, not as causal explanations of apartment design quality.

4. Discussion

4.1. From Individual Floor Plans to a Stock-Sensitive Interpretation of Apartment Layouts

This study demonstrates that coded floor-plan metadata can support a more stock-sensitive interpretation of apartment production than conventional record-by-record typological analysis. While the Aurinkolahti dataset was originally created to classify recurring apartment layouts and document spatial organization in Finnish housing between 2000 and

2023 [12], the present analysis extends that objective by treating each mapped layout not only as a typological observation, but also as a representation of multiple built apartments.

The distinction matters because it changes the scale of interpretation. When every floor-plan record is treated equally, the analysis primarily reflects the internal structure of the dataset. However, when layouts are weighted according to the number of apartments they represent, the interpretation shifts closer to the actual built housing stock. The weighting sensitivity analysis did not completely transform the overall results, but it did alter the relative prominence of several typological families and time periods. This suggests that housing research benefits from considering apartment layouts not simply as isolated design cases, but as recurring spatial products embedded within broader systems of housing production.

The findings align with earlier Finnish housing research by Kaasalainen and Huuhka [6], who identified substantial typological repetition within a large and apparently diverse housing stock. Similar observations have also emerged in recent computational studies using graph-based and clustering approaches to analyze large floor-plan datasets [7,18,19]. Together, these studies reinforce the idea that housing diversity at the level of individual plans can coexist with strong structural repetition at the scale of the housing stock.

At a broader methodological level, the study also supports recent arguments that coded architectural datasets can expand typological analysis beyond descriptive visual comparison. Research on graph-based floor-plan representations has shown that connectivity structures and semantic annotations can reveal relationships between layouts that are difficult to identify manually [14]. In this sense, the present workflow contributes to a growing body of research seeking to combine architectural interpretation with reproducible quantitative analysis.

4.2. Typological Concentration and Measurable Repetition

One of the clearest findings is the concentration of the housing stock around a relatively small number of broader typological families. Although the dataset contained substantial variation at the detailed layout level, Type B, Type D, and Type A together accounted for more than 70% of the classified weighted stock. This does not imply a lack of architectural variation. Rather, it suggests that much of the observed diversity is organized around recurring spatial structures that become visible once the data are interpreted at a broader typological scale.

This relationship between apparent diversity and underlying concentration is central to the study's argument. Apartment production may appear varied when individual layout labels are examined separately, while still relying heavily on repeated housing models. A similar phenomenon was observed in Finnish apartment housing, where recurring flat types represented a substantial share of the housing stock despite formal variation between buildings [6]. The present findings reinforce that interpretation by showing how repetition becomes measurable once apartment layouts are analyzed through weighted representation rather than isolated records.

The strongest example of this phenomenon was observed in As Oy Helsingin Hyperion, where a single mapped layout represented 45 identical apartments. In this case, repetition is no longer simply a visual similarity between drawings; it becomes a quantifiable property of housing production. The temporal results further strengthen this interpretation. The highest average repetition intensity was observed during the 2021–2023 period, suggesting that recent apartment production in the dataset is characterized not only by compactness, but also by increasing standardization and replication of layouts.

These results resonate with recent research on automated floor-plan analysis and typological clustering, which has emphasized the growing importance of scalable methods capable of identifying repeated spatial patterns across large housing datasets [13,15]. At the same time, they also raise broader architectural questions about how repetition interacts with housing quality, adaptability, and domestic experience.

4.3. Compactness as a Spatial and Typological Condition

The compactness findings require careful interpretation. The lowest weighted mean apartment area was observed in the 2021–2023 period, whereas the highest occurred between 2006 and 2010. This pattern supports the interpretation that recent apartment production within the dataset is associated with smaller dwelling sizes. However, compactness cannot be reduced to floor area alone. Apartments with similar surface areas may differ substantially in depth, room configuration, circulation, and access to natural light.

For this reason, the analysis incorporated additional indicators such as unit depth and area-per-room relationships. Recent studies on spatial compactness have similarly argued that compactness should be understood as a configurational condition rather than merely a geometric one [20]. Graph-based and space-syntax research has shown that connectivity, integration, and spatial permeability often provide a more meaningful understanding of housing usability than area measurements alone [21].

The exploratory weighted linear model reinforces this interpretation. Floor area was not only associated with typological family, but also with unit depth and kitchen natural light. In particular, unit depth showed a positive association with floor area, confirming that larger dwellings in the dataset also tend to be deeper. Kitchen natural light was likewise positively associated with floor area, suggesting that naturally lit kitchens are more likely to appear in larger layouts once year, depth, and typological-family effects are considered. These results do not redefine compactness as a statistical outcome, but they support the methodological decision to interpret compactness as a combined dimensional and typological condition rather than as floor area alone.

The Aurinkolahti findings are particularly relevant in relation to the increasing building depths and greater typological complexity documented in recent apartment developments within the same neighborhood [11]. Although the present study did not analyze building morphology directly, it complements their work from the apartment-unit perspective by showing how compactness, repetition, and selected domestic-quality indicators interact within the housing stock.

Importantly, compactness should not automatically be interpreted negatively. Compact apartments may support affordability, spatial efficiency, and urban density goals. The concern emerges when compactness coincides with strong repetition and weaker domestic-quality indicators. This concern is consistent with daylight research in compact residential forms, where the provision of daylight to interior or deeper spaces has been identified as a significant design challenge [22]. In the present dataset, the 2021–2023 period combines the smallest weighted average floor areas, the highest repetition intensity, and the lowest proportion of kitchens with natural light. While these findings do not demonstrate a definitive decline in housing quality, they do identify a pattern that deserves closer architectural attention.

The interpretation of the 2021–2023 period requires caution because it overlaps with the COVID-19 pandemic and its aftermath. The present dataset does not allow the pandemic-related effects to be isolated from longer-term housing-production trends. Therefore, the compactness and repetition patterns observed in this period should be read as empirical signals within the dataset rather than as conclusive evidence of a permanent post-pandemic typological shift.

4.4. Kitchen Natural Light and Domestic Spatial Quality

The kitchen-related findings are among the most architecturally significant results of the study. Only 17.8% of represented apartments included coded kitchen natural light, and the proportion declined substantially over time, from 24.3% in 2000–2005 to 5.7% in 2021–2023. These figures should be interpreted cautiously. The dataset does not include detailed information on window dimensions, façade orientation, daylight performance, or measured illuminance. Nevertheless, kitchen natural light remains a meaningful proxy when analyzed alongside compactness and repetition indicators.

The exploratory logistic model adds an important multivariable check to this descriptive pattern. After accounting for floor area, unit depth, typological family, and whether the kitchen was larger than 7 m^2, centered completion year remained negatively associated with the odds of kitchen natural light. This suggests that the lower prevalence of naturally lit kitchens in more recent layouts is not simply an artifact of apartment size, depth, or typological composition. At the same time, the model should not be interpreted as demonstrating a causal temporal decline in design quality. This indicates an association within the coded and weighted dataset, and therefore strengthens the descriptive interpretation without replacing architectural judgement.

The decline in coded kitchen natural light should not be interpreted as a direct measure of declining housing quality. Contemporary apartments may rely on mechanical ventilation, local exhaust systems, and integrated living–dining–kitchen arrangements, meaning that a kitchen without a direct façade opening can still be technically serviceable and spatially functional. For this reason, kitchen natural light is interpreted here as a partial spatial-quality marker. Its decline remains architecturally relevant because it signals a changing relationship between kitchen space, façade access, and apartment depth, but it must be read together with ventilation, room usability, open-plan organization, dwelling size, and resident preferences.

The importance of this finding lies less in the individual variable itself than in its convergence with broader spatial trends. Kitchens occupy a central role in everyday domestic routines, and recent research on small urban homes in Helsinki has shown that kitchen and dining areas can present deficiencies in functionality, furnishability, daylight incidence, and adaptability when compact layouts are used to reduce the apartment floor area [23]. Previous studies have emphasized that residents often prioritize factors such as room organization, window orientation, and spatial openness when evaluating housing quality [24,25]. In this context, the declining presence of naturally lit kitchens raises important questions about how compact apartment layouts are being spatially organized.

The findings also connect with broader discussions about domestic adaptability and spatial adequacy. A previous study demonstrated that characteristics such as fragmentation and living-space organization can influence how apartments are modified by occupants over time [8]. Similarly, the kitchen design and spatial flexibility play an important role in the long-term adaptability of multiresidential housing [26]. Although the present study did not examine post-occupancy transformations, it supports the broader argument that floor-plan characteristics should be understood as indicators of potential domestic performance rather than neutral geometric descriptors.

At the same time, the study acknowledges the limitations of using isolated spatial indicators to evaluate housing quality. This is also consistent with research showing that in small homes, floor area alone is not the defining factor in improving housing design quality; the shape and orientation of spaces also matter [23]. Recent literature has emphasized that meaningful assessments of domestic quality require multi-metric approaches that combine spatial, functional, environmental, and social dimensions [19,27]. The present analysis

therefore treats kitchen natural light as a partial but architecturally relevant indicator rather than a definitive measure of housing quality.

4.5. Exploratory Clustering and Emerging Layout Families

The clustering analysis identified two exploratory layout families with distinct spatial tendencies. The first cluster was characterized by smaller and shallower predominantly one-bedroom apartments with lower shares of kitchen natural light. The second cluster contained larger and deeper apartments, mainly with two bedrooms and higher proportions of naturally lit kitchens. Although these differences are architecturally interpretable, the moderate silhouette value indicates that the separation between groups is not especially strong.

The addition of schematic archetypal plans strengthens this interpretation by showing that Cluster 1 and Cluster 2 differ not only in mean area and depth, but also in their likely organizational logic: a compact living-mediated layout in Cluster 1 and a deeper corridor-differentiated layout in Cluster 2.

For this reason, the clustering results should be understood as exploratory rather than definitive. Their value lies not in producing rigid apartment categories, but in identifying recurring combinations of spatial characteristics within the dataset. This distinction is important because computational classification methods can sometimes create a misleading impression of analytical certainty. In practice, apartment typologies often overlap, hybridize, and evolve gradually, rather than forming sharply bounded categories.

Recent apartment-building research has similarly used hierarchical clustering to identify typical layouts while still linking the resulting types to architectural interpretation and building-scale classifications [7]. Hybrid approaches that combine semantic, geometric, and topological attributes have proven useful for detecting recurring spatial structures while still requiring architectural interpretation to contextualize the results [28,29]. The present findings support this perspective. The clusters are readable and meaningful, but they should be interpreted as analytical tendencies rather than fixed housing types.

More broadly, the clustering workflow demonstrates how coded apartment metadata can bridge descriptive statistics and architectural reasoning. Rather than replacing typological interpretation, exploratory clustering can help identify spatial relationships that warrant closer architectural examination.

4.6. Contributions and Limitations of Coded Floor-Plan Metadata

This study demonstrates how publicly available architectural datasets can be transformed into reproducible workflows for analyzing repetition, compactness, and selected domestic spatial-quality indicators. Recent literature has emphasized that semantically enriched floor-plan datasets create new possibilities for scalable and transparent housing analysis, particularly when combined with graph-based methods and structured annotations [14,30]. The contribution of this study lies in showing how relatively simple spreadsheet metadata can still support meaningful architectural interpretation when analyzed systematically.

The dataset also imposes clear limits on what can be claimed. The analysis is based on coded metadata rather than vectorized geometric models. This limitation is important because automatic floor-plan analysis typically aims to recover geometric and semantic elements such as walls, doors, rooms, and vectorized spatial structures, which are not directly available in the present spreadsheet-based dataset [4]. As a result, it cannot directly evaluate spatial syntax, façade exposure, circulation geometry, window orientation, daylight depth, or lived residential experience. In this respect, the study differs from

more advanced graph-based or space-syntax analyses that rely on detailed geometric representations of apartment layouts [21].

The study therefore avoids claiming an objective assessment of apartment quality. Its contribution is more limited: it identifies how selected measurable markers of compactness, repetition, and domestic spatial organization vary across the Aurinkolahti apartment stock. Several relevant quality-related variables could not be assessed, including façade orientation, window configuration, daylight performance, staircase and elevator organization, ventilation performance, acoustic conditions, and resident experience.

The 2021–2023 findings should also be interpreted cautiously because this period overlaps with the COVID-19 pandemic and its aftermath. The analysis cannot determine whether the observed compactness and repetition patterns represent a stable long-term shift, a pandemic-related acceleration of existing tendencies, or a temporary response to exceptional market and production conditions.

4.7. Implications for Housing Research and Design Practice

Taken together, the findings suggest that compactness, repetition, and domestic spatial-quality indicators should be interpreted relationally rather than independently. A decrease in apartment area alone would provide a limited interpretation of recent housing production. Likewise, repetition by itself could simply reflect standard construction practices. The significance emerges through the convergence of multiple tendencies: the most recent period combines smaller apartments, stronger layout repetition, and lower shares of kitchens with natural light. The exploratory models reinforce this convergence by showing that the association between more recent completion years and lower odds of kitchen natural light persists even when other dimensional and typological variables are considered.

For architectural practice, this does not necessarily imply that contemporary apartment layouts are inherently deficient. However, the kitchen and dining-area evidence from recent Helsinki housing suggests that compact design strategies require evaluation through multiple quality criteria rather than through area efficiency alone [23]. Questions of usability, daylight access, adaptability, and everyday domestic organization remain central, particularly in increasingly dense urban contexts.

The findings also reinforce recent calls for multi-dimensional approaches to housing-quality assessment. Research on residential spatial quality has shown that occupants evaluate apartments through a combination of configurational, environmental, and social characteristics rather than through size alone [24,27]. In this sense, coded floor-plan datasets can help expand housing research beyond descriptive typologies toward more integrated interpretations of housing performance and domestic experience.

Overall, the findings should be read as evidence of measurable tendencies in the Aurinkolahti dataset rather than as a complete evaluation of residential quality. Coded metadata can reveal relevant stock-level patterns in compactness, repetition, and selected domestic spatial markers, but these patterns still require architectural judgement, and where possible, direct plan-based analysis.

5. Conclusions

The main contribution of the study is methodological. It demonstrates how coded floor-plan metadata can operate as an intermediate evidence layer between visual typological interpretation and full geometric floor-plan analysis. By combining apartment-count weighting, compactness indicators, repetition measures, exploratory clustering, and selected spatial-organization markers, the workflow provides a reproducible way to describe stock-level tendencies in apartment layouts. The addition of schematic typological and cluster-based plans further helps translate coded metadata back into architectural form,

making the relationship between compactness, circulation, depth, and domestic spatial markers more legible. However, the findings should be interpreted as measurable tendencies rather than complete evaluations of residential quality. Future research should combine coded metadata with direct plan analysis, façade orientation, window configuration, staircase and elevator organization, daylight simulation, ventilation performance, and resident experience.

Author Contributions: Conceptualization, Y.G.-R. and V.B.; methodology, V.B.; software, Y.G.-R.; validation, Y.G.-R. and V.B.; formal analysis, Y.G.-R.; investigation, V.B.; resources, V.B.; data curation, Y.G.-R.; writing—original draft preparation, Y.G.-R.; writing—review and editing, Y.G.-R. and V.B.; visualization, Y.G.-R.; supervision, Y.G.-R. and V.B.; project administration, V.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available on Zenodo at https://doi.org/10.5281/zenodo.20489313 (Accessed on 1 June 2026).

Acknowledgments: The authors acknowledge the use of GitHub Copilot (OpenAI Codex-based assistance) with the GPT-5 model for supporting code debugging and error revision in the R scripts used in this study, as well as for minor English language editing and proofreading during manuscript preparation. The authors reviewed, edited, and approved all final content of the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Baker, J.; Oppewal, H. The Effects of Floor Plan Representations on Preferences for Apartments. *J. Hous. Built Environ.* **2023**, *38*, 727–752. [CrossRef] [PubMed]
2. Urban Housing Lab. *Floor Plan Manual Housing*; Heckmann, O., Schneider, F., Zapel, E., Eds.; Birkhäuser Publishers: Basel, Switzerland, 2017.
3. Berghauser Pont, M.; Stavroulaki, G.; Bobkova, E.; Gil, J.; Marcus, L.; Olsson, J.; Sun, K.; Serra, M.; Hausleitner, B.; Dhanani, A.; et al. The Spatial Distribution and Frequency of Street, Plot and Building Types across Five European Cities. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 1226–1242. [CrossRef]
4. Pizarro, P.N.; Hitschfeld, N.; Sipiran, I.; Saavedra, J.M. Automatic Floor Plan Analysis and Recognition. *Autom. Constr.* **2022**, *140*, 104348. [CrossRef]
5. Lu, Z.; Wang, T.; Guo, J.; Meng, W.; Xiao, J.; Zhang, W.; Zhang, X. Data-Driven Floor Plan Understanding in Rural Residential Buildings via Deep Recognition. *Inf. Sci.* **2021**, *567*, 58–74. [CrossRef]
6. Kaasalainen, T.; Huuhka, S. Homogenous Homes of Finland: 'Standard' Flats in Non-Standardized Blocks. *Build. Res. Inf.* **2016**, *44*, 229–247. [CrossRef]
7. Prasetya, A.; Ho, A.D.; Kubota, T. Analysis of Typical Layout of Apartment Buildings in Indonesia. *Buildings* **2023**, *13*, 1387. [CrossRef]
8. Femenias, P.; Geromel, F. Adaptable Housing? A Quantitative Study of Contemporary Apartment Layouts That Have Been Rearranged by End-Users. *J. Hous. Built Environ.* **2020**, *35*, 481–505. [CrossRef]
9. Rockow, Z.R.; Ross, B.E. An Areal Openness Model (AOM) for Quantifying the "Openness" of Floor Plans. *Int. J. Build. Pathol. Adapt.* **2021**, *39*, 470–489. [CrossRef]
10. Park, K.; Ergan, S.; Feng, C. Quality Assessment of Residential Layout Designs Generated by Relational Generative Adversarial Networks (GANs). *Autom. Constr.* **2024**, *158*, 105243. [CrossRef]
11. Meriläinen, S.; Tervo, A. The Changing Typology of Urban Apartment Buildings in Aurinkolahti. *Build. Cities* **2025**, *6*, 50–69. [CrossRef]
12. Meriläinen, S. Architectural Floor Plan Data from Apartment Units in Aurinkolahti, Helsinki, Finland, 2000–2023. Available online: https://zenodo.org/records/19894248 (accessed on 28 May 2026).
13. Massafra, A.; Al-Harasis, D.H.; Stefanini, L.; Jabi, W. Semi-Automated Dataset Generation for Residential Buildings Using Graph-Based Topological Modelling. *Buildings* **2025**, *15*, 1283. [CrossRef]

14. Ziaee, A.; Suter, G. Multi-Unit Space Function and Space Access Element Classification in Apartment Buildings Using Machine Learning and Graph Deep Learning. *J. Build. Eng.* **2025**, *30*, 113472. [CrossRef]
15. Xu, Z.; Jha, N.; Mehadi, S.; Mandal, M. Automatic Floor Plan Analysis: Datasets, Methods, and Applications (2000–2025). *Autom. Constr.* **2025**, *178*, 106378. [CrossRef]
16. R Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2026.
17. R Studio Team. *RStudio: Integrated Development Environment for R*; R Studio Team: Boston, MA, USA, 2023.
18. Martínez-Rocamora, A.; Díaz-Cuevas, P.; Camarillo-Naranjo, J.; González-Vallejo, P. Identification of Residential Building Typologies by Applying Clustering Techniques to Cadastral Data. *J. Build. Eng.* **2024**, *86*, 108912. [CrossRef]
19. Wang, Y.; Zhu, Y.; Wang, K.; Li, X. A Hybrid Deep Learning Approach to Investigating Architectural Morphology: A Workflow Combining Graph and Image Data to Classify High-Rise Residential Building Floorplans. *J. Build. Eng.* **2025**, *111*, 113255. [CrossRef]
20. Van Nes, A. Spatial Configurations and Walkability Potentials: Measuring Urban Compactness with Space Syntax. *Sustainability* **2021**, *13*, 5785. [CrossRef]
21. Dawes, M.J.; Ostwald, M.J. Space Syntax: Mathematics and the Social Logic of Architecture. In *Handbook of the Mathematics of the Arts and Sciences*; Springer: Cham, Switzerland, 2021.
22. Freewan, A.A.Y.; Gharaibeh, A.A.; Jamhawi, M.M. Improving Daylight Performance of Light Wells in Residential Buildings: Nourishing Compact Sustainable Urban Form. *Sustain. Cities Soc.* **2014**, *13*, 32–40. [CrossRef]
23. Tervo, A.; Meriläinen, S.; Lilius, J. Floor Plan Analysis of Kitchens and Dining Areas in Small Urban Homes. *Nord. J. Archit. Res.* **2025**, *37*, 71–100.
24. Bao, X.; Zhang, T.; Dewancker, B.J.; Liu, S. Exploring the Unit Spatial Layout Preference for Urban Multi-Unit Residential Buildings: A Survey in Beijing, China. *Sustainability* **2023**, *15*, 12013. [CrossRef]
25. Brkanić Mihić, I.; Koški, D. Residents' and Architects' Perceptions of Apartments' Spatial Characteristics—Identifying Differences in Opinion on a Case Study of Osijek, Croatia. *Buildings* **2024**, *14*, 1496. [CrossRef]
26. Ollár, A.; Granath, K.; Femenías, P.; Rahe, U. Is There a Need for New Kitchen Design? Assessing the Adaptative Capacity of Space to Enable Circularity in Multiresidential Buildings. *Front. Archit. Res.* **2022**, *11*, 891–916. [CrossRef]
27. Brkanić Mihić, I. Housing Quality Assessment Model Based on the Spatial Characteristics of an Apartment. *Buildings* **2023**, *13*, 2181. [CrossRef]
28. Azizi, V.; Usman, M.; Patel, S.; Schaumann, D.; Zhou, H.; Faloutsos, P.; Kapadia, M. Floorplan Embedding with Latent Semantics and Human Behavior Annotations. In *SimAUD 2020: Proceedings of the 11th Annual Symposium on Simulation for Architecture and Urban Design*; Association for Computing Machinery, Inc.: New York, NY, USA, 2020.
29. Yang, L.; Huang, W. Representation and Assessment of Spatial Design Using a Hierarchical Graph Neural Network: Classification of Shopping Center Types. *Autom. Constr.* **2023**, *145*, 104727. [CrossRef]
30. Mostafavi, F.; van Engelenburg, C.; Khademi, S.; Vrachliotis, G. Floor Plan Generation: The Interplay among Data, Machine, and Designer. *Int. J. Archit. Comput.* **2025**, *23*, 370–386. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.