Mapping the Spatial Orientation of Local Development Policy: An AI-Based Textual Analysis of Townscape Image Handbooks (TAK) in Budapest

Gergő Bendegúz Békési [1] – Zsolt Kasza [2] – István Tózsa [3]

Abstract

While the local policies of municipalities are conventionally evaluated through fiscal budgets, their underlying narrative inclinations often remain latent. This paper maps the text-driven spatial strategies within a metropolitan governance structure by applying computational linguistics to Budapest, Hungary. Specifically, we conduct an empirical experiment using a dictionary-based Natural Language Processing (NLP) pipeline – the Label to Vector (Schopf et al., 2022) framework – to analyse the official Townscape Image Handbooks (Településképi Arculati Kézikönyv – TAK) of the 23 administrative districts of Budapest. Building upon the urban marketing foundations established by Ashworth and Voogd, this study quantifies the intensity of place-marketing narratives across three critical socioeconomic target groups: tourism, residential development, and business premises (capital attraction). By deploying choropleth mapping based on decile distribution and calculating structural deviations from the municipal baseline, this paper uncovers hidden spatial-social anomalies embedded within urban planning documents. Ultimately, the study demonstrates the predictive power of automated text analysis in regional planning and highlights the emerging role of artificial intelligence in auditing institutional narratives.

Keywords: Townscape Image Handbooks (TAK), urban marketing, Natural Language Processing (NLP), text vectorization, spatial orientation, Budapest districts

1. Introduction

1.1. Institutional and Theoretical Framework of the Townscape Image Handbook (TAK) System

As of 2026, the Townscape Image Handbooks (Településképi Arculati Kézikönyv – TAK[4]) constitute the most extensive and voluminous architectural and urban development document

[1] PhD student, Department of Automation and Applied Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary; e-mail: bekesigergobendeguz@gmail.com

[2] Former PhD student at the National University of Public Service, Member of the Hungarian Marketing Association's Town-Marketing Expert Group; e-mail: kaszazsolt@gmail.com

[3] Professor emeritus, National University of Public Service, Head of the Economic Geography and Urban Marketing Centre, John von Neumann University; e-mail: tozsa.istvan@nje.hu

[4] Településképi Arculati Kézikönyv (TAK) translates to Townscape Image Handbook in English. To ensure readability and structural consistency with regional planning jargon, the Hungarian acronym TAK is retained throughout this paper.

system ever implemented in Hungary.[5]5 The analytical focus of the present study is directed toward the implicit place-marketing potential embedded within these documents. We argue that the ongoing statutory updating of the TAK system in 2026 provides a critical opportunity to integrate place-marketing strategies, which we consider the most effective tool for holistic municipal development. Ensuring a conceptual alignment between the marketing narratives of the handbooks and newly drafted urban development documents is paramount. This integration is explicitly mandated by contemporary legislation; specifically, Section 22 of Act C of 2023 on Hungarian Architecture[6] decrees a tight institutional and functional coupling between local handbooks, zoning ordinances, and broader municipal development frameworks within a multi-level governance structure.

The empirical baseline for this research stems from a comprehensive study conducted in 2024, which analysed a representative sample of 164 handbooks drawn from the total universe of 3,177 Hungarian municipalities (Kasza, 2024). The preliminary findings of this computational, marketing-focused corpus analysis were published in May 2025 in *Területi Statisztika* (Békési, Kasza, and Tózsa, 2025). Utilizing a dictionary-based Natural Language Processing (NLP) framework – specifically, the Label to Vector text vectorization method – that national study demonstrated that nearly two-thirds (66.04%) of the evaluated text corpus contained distinct place-marketing elements, despite the fact that the TAK documents were originally commissioned for architectural, preservationist, and zoning purposes rather than promotional ones. Furthermore, the national sample revealed a structural bias toward specific target audiences within the classic Ashworth – Voogd (1993) urban marketing framework: 83.7% of the marketing narratives targeted the residential component (population retention and attraction), while 16.3% focused on business premises (capital attraction), leaving tourism largely unaddressed at the rural and national scale.

The immediate empirical relevance of this paper is underscored by the regulatory landscape of 2026, which legally obligates Hungarian municipal governments – regardless of the political outcomes of the general elections held on April 12, 2026[7] – to initiate the comprehensive overhaul and drafting of their new-generation urban development plans.[8] Within this nationwide transition, the Municipality of Budapest has assumed a pioneering role. The capital's urban planning apparatus finalized the regulatory drafts by the end of 2025,[9]

[5] The statutory obligation to draft and maintain a Townscape Image Handbook was legally mandated for all 3,177 Hungarian municipalities, including the capital, under Act LXXIV of 2016 on the Protection of Townscapes (2016. évi LXXIV. törvény a településkép védelméről).

[6] Act C of 2023 on Hungarian Architecture (2023. évi C. törvény a magyar építészetről) was adopted by the Hungarian Parliament on December 12, 2023. Subsections (3) and (4) of Section 22 explicitly regulate the division of planning labour and regulatory hierarchy between the metropolitan Municipality of Budapest and its constituent administrative districts, establishing that district-level townscape handbooks and regulatory plans must operate within the structural guidelines set by the capital's master assembly. Available at: https://mkogy.jogtar.hu/jogszabaly?docid=A2300100.TV (Accessed: March 7, 2026).

[7] Regardless of the outcomes of the Hungarian general elections held on April 12, 2026, the structural and temporal parameters of municipal urban planning remain strictly bound by the prior statutory schedule.

[8] See the relevant harmonizing provisions of Act XXXIX of 2021 on the Amendment of Act XXI of 1996 on Regional Development and Regional Planning (2021. évi XXXIX. törvény a területfejlesztésről és a területrendezésről szóló 1996. évi XXI. törvény módosításáról).

[9] Based on institutional communications provided by the Urban Planning Department (Urbanisztikai Osztály, Városfejlesztési Csoport) within the Mayor's Office of the Municipality of Budapest (Budapest Főváros Főpolgármesteri Hivatal Várostervezési Főosztály).

culminating in the official adoption of the new unified Metropolitan Development and Zoning Master Plan by the General Assembly on February 28, 2026.[10]

This legislative milestone effectively opened the regulatory window for the 23 autonomous administrative districts of Budapest to launch their respective localized planning phases. This process is driven by a strict statutory sunset clause: after July 1, 2027, traditional, older-generation development and planning instruments can no longer be legally applied.[11] Crucially, the administrative districts do not possess decentralized autonomy regarding the selection of the external consultants tasked with drafting these new master plans. Instead, they are legally integrated into a centralized public procurement framework managed exclusively by a professional entity 100% owned by the Municipality of Budapest.[12] The final allocation of these planning contracts took place throughout March 2026; for instance, in the 14th district (Zugló), the winning planning consortium tasked with designing the district's new master plan was officially presented to the public on June 1, 2026.

Consequently, this study proposes a replicable, artificial intelligence-driven methodology designed to audit and synchronize the implicit marketing content of the existing TAK corpus – which remains legally valid after July 1, 2027, despite pending structural updates – with the newly mandated statutory development plans.

The nationwide baseline assessment published by Békési, Kasza, and Tózsa (2025) underscored an institutional paradox: from a place-marketing perspective, the primary beneficiaries of the mandatory TAK framework were the smallest micro-settlements (those with fewer than 500 inhabitants). In 2017, these micro-villages and tiny settlements accounted for over 35% of the total 3,177 Hungarian municipalities,[13] with the vast majority receiving a 1 million HUF non-repayable central government grant to finance the drafting of their respective handbooks. Furthermore, since October 1, 2019, these local handbooks have been structurally integrated into a centralized, publicly accessible digital platform.[14] Prior to this mandatory digitization, micro-villages (under 200 residents) and small villages (200 - 499 residents) possessed highly restricted or non-existent capabilities to project their institutional identity into the digital space. The central repository provided an unprecedented geopolitical and promotional window for

[10] Proposal for the Adoption of the Budapest Municipal Plan and its Environmental Assessment (Javaslat a Fővárosi Településterv és a környezeti értékelése elfogadására). Under Act C of 2023, this newly enacted Municipal Plan structurally replaces the former 'Budapest 2030' Urban Development Concept, the 'Home in Budapest' Integrated Urban Development Strategy (ITS), and the historical Budapest Spatial Structure Plan (TSZT). The core adopted document, titled 'From a City to a Home' (Városból Otthon), establishes the mandatory operational boundaries for all constituent districts. Available at: https://budapest.hu/hirek/2026/01/27/a-januar-28-i-fovarosi-kozgyules-eloterjeszteseinek-osszefoglaloja (Accessed: March 7, 2026).

[11] These legally phasing-out documents primarily include the traditional Integrated Settlement Development Strategies (Integrált Településfejlesztési Stratégia – ITS) and the standalone Settlement Development Concepts (Településfejlesztési Koncepció – TFK).

[12] The public procurement process for the newly mandated district plans was structurally centralized and managed via BFVT Kft. (Budapesti Fővárosi Városépítési Tervező Kft. / Budapest Metropolitan Urban Planning and Design Ltd.). This entity, 100% owned by the Municipality of Budapest, effectively streamlined the bidding consortia for the administrative districts, with tender conclusions finalized in March 2026. Source: https://www.bfvt.hu/kapcsolat (Accessed: March 7, 2026).

[13] Quantitatively, this demographic threshold encompassed more than 1,100 individual rural municipalities across Hungary in 2017.

[14] The official national depository is hosted via the centralized architectural portal: https://tak.e-epites.hu/# (Accessed: March 7, 2026).

these isolated communities to broadcast their local history, prominent historical figures, and tangible architectural or natural endowments.

The statutory obligation to include a socio-historical narrative and a comprehensive catalogue of local assets was enforced by Government Decree 314/2012 (XI. 8.).[15] Although this regulatory mandate did not inherently guarantee that minor municipalities would strategically utilize the handbooks as promotional vehicles, the institutional pressure for self-identification and place branding triggered a grassroots search for unique socio-cultural assets (Unique Selling Propositions – USP). For instance, the handbook of Várda[16] highlighted that the internationally acclaimed Hungarian graphic artist Endre Szász spent his final creative years in the village. This structural drive for self-definition proved critical for the approximately 1,100 micro-settlements that lacked significant historical architecture or landmark monuments. Unable to leverage monolithic heritage assets like the Basilica of Esztergom or the Castle of Eger, these under-500-inhabitant villages strategically pivoted toward environmental and ecological branding within their TAK documents; examples include Mogyoróska (80 inhabitants), which branded itself through the presence of a rare botanical specimen (*Daphne mezereum*), or Szaknyér (43 inhabitants), which anchored its spatial identity around a historically commemorative linden tree, planted in memory of the coronation of Maria Theresa and distinguished by a trunk circumference of 3.6 metres.

1.2. The Metropolitan Contrast: Budapest vs. the Rural Baseline

In stark contrast to the aforementioned micro-settlements, the empirical core of the present study shifts the analytical lens toward the largest, demographically densest, and financially most autonomous administrative cluster in Hungary: the 23 administrative districts of the capital city, Budapest. To illustrate the structural scale of this urban-rural divide within the total universe of 3,177 Hungarian municipalities, Budapest represents a concentrated population of approximately 1.7 million inhabitants. Conversely, the entire cohort of micro-settlements with fewer than 500 residents aggregated a mere 0.3 million occupants during the 2017 baseline period according to data from the Hungarian Central Statistical Office (KSH)[17].

This demographic polarization fundamentally altered the institutional challenges and narrative motivations during the initial drafting of the handbooks. For the micro-settlements, the primary planning dilemma was content scarcity – specifically, how to generate a mandatory 20-to-30-page townscape volume for municipalities characterized by extreme spatial minimalism. A poignant example is the village of Gagyapáti, which in 2017 registered a total of merely 14 physical structures and 19 permanent residents. However, having received central government funding, these minor localities recognized a rational self-interest in producing high-quality, substantive documents. Driven by inter-municipal competition, they understood that to retain current residents or attract external visitors, they needed to articulate a compelling spatial identity. This baseline competition catalysed the implicit place-marketing transformation of the

[15] Government Decree 314/2012. (XI. 8.) on Settlement Development Concepts, Integrated Settlement Development Strategies, and Settlement Planning Instruments (314/2012. (XI. 8.) Korm. rendelet a településfejlesztési koncepcióról, az integrált településfejlesztési stratégiáról és a településrendezési eszközökről az építészeti-műszaki tervtanácsokról és a településkép érvényesítési eszközökről).

[16] Várda was registered as a 505-inhabitant municipality during the 2017 baseline period, subsequently dipping below the 500-resident threshold. Notably, as of 2026, Várda has secured the national 'City Marketing Diamond' award for three consecutive years, representing a unique achievement within the micro-settlement category.

[17] www.ksh.hu

TAK documents in the rural segment, characterized by a heavy reliance on residential positioning (retaining and attracting people) rather than business premises or capital attraction.

The 23 administrative districts of Budapest operated under diametrically opposed institutional and financial conditions. Receiving no central budgetary subsidies due to their size and fiscal strength,[18] the metropolitan planning apparatuses did not struggle with a scarcity of narrative material, but rather with information overload. The core challenge for district planners was a complex exercise in comprehensive inventory listing: determining how to systematically catalogue an immense density of historic architectural monuments, rapidly changing townscape character areas, and urban green spaces without omitting critical local endowments.

1.3. Literature Review and International Relevance

This study represents the first comprehensive endeavour to report the place-marketing dimensions of the Hungarian TAK framework to the international academic community in English, expanding upon localized, pilot-scale university working papers within the region (e.g., Kasza, 2020). The justification for broadcasting these Central and Eastern European (CEE) empirical findings on a global editorial platform is threefold:

- First, the conceptual blueprint of the Hungarian TAK system is heavily derived from international planning innovations, specifically the Design Guide framework implemented in County Cork, Ireland, in 2003 (Shanahan, 2003). The structural framework of the TAK is thus built upon internationally imported planning knowledge. Irish planning expert Mike Shanahan played a foundational role in delineating the operational parameters of the Hungarian handbook system, which suggests that the place-marketing methodologies generated within the Hungarian context possess high structural exportability to other transitioning CEE jurisdictions.

- Second, the strategic outreach of the TAK corpus explicitly transcends Hungary's formal administrative borders (encompassing 9.5 million residents) and is conceptually calibrated toward the broader sociocultural ecosystem of the Carpathian Basin, which populates approximately 13 million individuals.[19] The framework was designed to project institutional place identity across geopolitical boundaries so that Hungarian diasporas residing in Romania, Slovakia, Serbia, and Ukraine could frictionlessly engage with localized architectural and socioeconomic identities, thereby mitigating administrative and cultural barriers in potential cases of remigration.

- Third, targeting a high-impact international publication platform facilitates the integration of this research into the broader paradigm of global metropolis studies. As the primary metropolis[20] in Hungary, Budapest exerts a profound spatial and economic gravity. Examining its district-level planning configurations yields vital comparative

[18] Under the 2017 statutory allocation rules, only Hungarian municipalities with a permanent population of fewer than 10,000 inhabitants were eligible to apply for the one-time, non-repayable central government subsidy of 1 million HUF to cover the consulting fees associated with the TAK compilation.

[19] The institutional coordination of the initial handbook rollouts was managed by the State Secretariat for Architecture and Construction (Építészeti és Építésügyi Helyettes Államtitkárság), which operated until 2022 under the leadership of State Secretary Zsolt Füleky. The initial mandate of the secretariat explicitly included the spatial and architectural heritage protection of cross-border Hungarian communities.

[20] In contemporary regional science and urban demography, a metropolis is structurally defined as an urban agglomeration maintaining a permanent population threshold between 1 and 10 million inhabitants.

insights for metropolitan governance and spatial path-dependencies across competing CEE capitals, such as Prague, Warsaw, Bucharest, Belgrade, Bratislava, Zagreb, and Vienna.

Within the nascent body of literature, Bíró and Láposi (2018) provided a comparative ex-ante analysis of how Scottish urban planners utilize townscape manuals relative to the newly introduced Hungarian handbooks. In terms of empirical evaluation, the first representative assessment of the initial TAK output was published by Makovényi and Simon (2021). Notably, these authors indicated that townscape preferences and aesthetic guidelines are structurally non-neutral, demonstrating statistically significant correlations with the political orientation (left-wing versus right-wing affiliation) of the presiding municipal mayors. Between 2017 and 2025, the overarching characteristic of Hungarian townscape literature was its strict containment within architectural, historical, and preservationist paradigms, primarily championed by the Chamber of Hungarian Architects (Magyar Építész Kamara – MÉK).

In contrast to these purely physical planning approaches, this paper positions place marketing as the missing institutional bridge required to synthesize the TAK corpus with incoming municipal development and zoning ordinances. Place marketing operates as a holistic, multi-disciplinary matrix that can dynamically align localized townscape identities with economic development vectors (Piskóti, 2023). Consequently, this study offers a computational methodology to enable municipal assemblies and planning consortia to exploit the latent marketing profiles of their handbooks, structurally transforming fragmented planning documents into a synergistic, market-oriented urban strategy.

2. Research Objectives

The initiation of this empirical investigation is grounded in the structural assumption that municipal governments, as early as the 2017 baseline period, actively sought to exploit the latent place-marketing opportunities embedded within the mandatory statutory frameworks of townscape planning. Furthermore, the methodological innovations developed within this paper are designed to be highly scalable; they can be frictionlessly extended to next-generation municipal and regional development documents that – in accordance with transitioning statutory and professional requirements – will complement or replace the legacy handbook system in the upcoming planning cycle.

The overarching objective of this research is to demonstrate that the conceptual link between Townscape Image Handbooks (TAK) and place-marketing parameters can be quantified and verified through primary, data-driven empirical research, moving beyond the traditionally speculative or purely aesthetic narratives found in mainstream architectural publications. In this framework, a qualitative, theory-driven conceptualization based on external place-marketing scholarship serves as an indispensable diagnostic tool, defining the relevant analytical categories and precisely calibrating the specific keywords required for the subsequent large-scale quantitative text vectorization (Békési, Kasza, and Tózsa, 2025).

Consequently, the core research question driving this study is formulated as follows: To what structural extent, and through what spatial characteristics, does place marketing manifest within metropolitan townscape handbooks? To systematically address this inquiry, this study deploys a rigorous mixed-methods approach. The qualitative phase derives and calibrates the place-marketing categories and keyword sets from established external theoretical frameworks, while the quantitative phase applies computational linguistics to the TAK corpus to statistically verify

the intensity, variance, and frequency of these place-marketing trajectories across administrative boundaries. This dual lens provides a holistic typological mapping of institutional narratives.

The broader institutional significance of this study lies in its capacity to demonstrate that if the place-marketing dimensions of the handbooks can be empirically verified, urban planning and promotional strategy can be structurally unified into a single, synergistic development paradigm for municipal governance. Contemporary urban development cannot sustainably operate without an integrated place-marketing philosophy. The mandatory introduction of the TAK system in 2017 forced municipal assemblies to make a strategic, long-term decision: whether to path-dependently preserve or dynamically develop their townscape identities. For municipalities opting for preservationist pathways, these established narrative parameters cannot be legally or strategically ignored in the post-July 1, 2027 planning landscape.

To ensure strict methodological replicability, the operational architecture of this research was structurally modelled after the classic research process continuum formulated by Earl Babbie (2008). This linear methodological pipeline systematically executes four core stages: (1) conceptualization of place-marketing dimensions within public planning, (2) operationalization of text corpuses into measurable units, (3) selection of the advanced Natural Language Processing (NLP) framework, and (4) formulation of the total metropolitan sampling frame (Babbie, 2008, p. 126).

Finally, this study directly contributes to expanding the analytical effectiveness of algorithmic integration within spatial sciences. By introducing a novel scientific perspective regarding the regional application and automated auditing of AI-based innovations, this research builds upon contemporary computational baselines in regional analytics, particularly mirroring the state-of-the-art manifestations of automated language evaluation and narrative processing recently established within the regional planning discourse by Lülök, Dobos, and Sebestyén (2026).

3. Data and Methodology

3.1. Label to Vector (Lbl2Vec) Pipeline and Operationalisation

Because the Townscape Image Handbook (TAK) is a completely novel planning instrument introduced post-2017, the initial phase of this research required the establishment of a standardized nomenclature to precisely calibrate core concepts, keywords, and definitions against shifting statutory frameworks.

The second phase involved operationalization—specifically, selecting an objective metric to quantify the latent place-marketing profiles embedded within the text corpus. Detecting place-marketing narratives through conventional human intelligence is highly constrained by volume; the average length of a single TAK document ranges from 30 to 300 pages. The cognitive limits of manual content analysis, combined with the inherent subjectivity of human coders evaluating narrative styles, necessitated a quantitative, data-driven approach applied to the 1+23 metropolitan sample (comprising the central Municipality of Budapest and its 23 administrative districts). To filter out human bias and objectively determine the presence, intensity, and strategic orientation of the institutional links between the handbooks and place-marketing vectors, this study deploys the unsupervised Label to Vector (Lbl2Vec) computational framework (Schopf et al., 2022).

The third phase dictated the formal selection of the analytical research design. In alignment with the classic methodological continuum proposed by Malhotra (2001, p. 93), all three dimensions of research modelling were systematically executed: (1) a verbal/textual model, (2) a graphical model, and (3) a mathematical model. The verbal content analysis and subsequent text vectorization were processed across the 1+23 sample via the Lbl2Vec architecture. The graphical outputs, demonstrating the localized spatial distribution of these narratives, are compiled in the Appendix and visual annexes of this paper.

Building upon the diagnostic baseline established in the qualitative phase, the quantitative stage applies rigorous mathematical-textual classification to verify whether place-marketing characteristics manifest significantly within the official planning corpus. Crucially, the pipeline calculates how these embedded narratives function as strategic supply-side offers directed toward three distinct socioeconomic target groups: **local inhabitants** (residential orientation), external **visitors** (tourism orientation), and potential **investors** looking to acquire commercial real estate or establish industrial facilities (business premises orientation).

Methodologically, this framework is rooted in automated text classification, where the objective is to categorize unstructured textual data into predefined thematic domains. The underlying Natural Language Processing (NLP) pipeline operates via text vectorization: textual elements are transformed into dense, multi-dimensional numerical vectors that mathematically capture semantic relationships and contextual similarities between words. This enables a robust numerical representation of the documents' core concepts. Leveraging semantic similarity metrics, the algorithm identifies latent themes and contexts, categorizing documents even in the total absence of explicit target labels or manually annotated training data.

These algorithmic methods are capable of processing vast amounts of textual data rapidly and efficiently, drastically reducing the temporal and financial resources required for manual content analysis. This computational efficiency is particularly advantageous for the current research design, which required the systematic evaluation and classification of an institutional corpus encompassing 24 standalone master documents totalling 3,626 pages. Furthermore, the underlying deep-learning architecture facilitates a profound structural analysis of the text, enabling the recognition of latent conceptual structures and implicit institutional paths. Consequently, the classification does not merely flag isolated, explicit keywords but dynamically maps the broader underlying themes and strategic concepts that define localized municipal and regional marketing functions (Chang et al., 2024).

However, unlike conventional supervised text classification tasks where a machine learning model is trained on pre-labelled data, this study faces an unsupervised classification environment. The planning documents contain no prior, explicit metadata tags indicating marketing target groups; instead, we possess only a calibrated set of core keywords defining each strategic category. In this specific context, the Lbl2Vec approach formulated by Schopf et al. (2022) emerges as a highly relevant and methodologically superior framework. Lbl2Vec is an innovative technique engineered specifically to execute text classification in unsupervised settings by exploiting word-to-document co-occurrences and the local context provided by specific keyword seeds.

In this context, the text-vectorization methodology utilized by the Lbl2Vec framework becomes uniquely relevant and effective (Schopf et al., 2022). Lbl2Vec is an innovative technique engineered specifically to execute text classification in unsupervised settings by exploiting word-to-document co-occurrences and the specific local context provided by pre-calibrated

keyword seeds. This framework fundamentally relies on the core principles of dense text vectorization and semantic similarity to automatically identify, and group documents based on their latent thematic architecture, bypassing the need for manually annotated training sets.

In the initial algorithmic step, Lbl2Vec operationalizes the document corpus into a joint semantic space by leveraging pre-trained word embeddings that mathematically represent the deep conceptual meanings of words and phrases. This numeric transformation allows machine learning architectures to systematically process unstructured textual data. During the vectorization of a given document D, each individual word w_i within the document is represented as a V-dimensional dense vector v_i, where V denotes the predefined embedding dimension. The overarching structural vector representation d of the complete document corpus is mathematically computed as the centroid (mean) of all its constituent word vectors:

$$d = \frac{1}{N}\sum_{i=1}^{N} v_i \tag{1}$$

where N represents the total word count contained within the evaluated document, and v_i denotes the corresponding vector of the i-th word.

Furthermore, the pipeline dynamically vectorizes the user-defined keyword seeds, operating under the same mathematical principles applied to the broader text corpus. Through this mechanism, each input keyword is transformed into a dense vector that numerically quantifies its specific semantic and contextual meaning. The vectorization of these benchmark keywords mirrors the exact pipeline of the general text processing: every input keyword k is mapped as a discrete V-dimensional vector in the shared semantic space.

Subsequently, the Lbl2Vec architecture executes a comparative analysis between the generated document vectors and the keyword vectors to rigorously evaluate the semantic similarity across the joint vector space. This spatial pipeline enables the model to identify which specific planning documents exhibit a tight conceptual proximity to the respective keyword seeds, thereby implicitly classifying the unstructured text based on the underlying themes. To precisely quantify this semantic similarity between a document vector d and a keyword vector k, the framework deploys the cosine similarity metric, which computes the cosine of the angle between the two dense vectors, systematically indicating their proximity within the multi-dimensional vector space.

The cosine similarity between the document vector d and the keyword vector k is mathematically formulated as follows:

$$\mathrm{sim}(d,k) = \frac{d \cdot k}{|d||k|} \tag{2}$$

where $\cdot$ denotes the standard scalar (dot) product of the vectors, and $|d|$ and $|k|$ represent their respective Euclidean norms. Consequently, leveraging this strict semantic similarity measurement, the Lbl2Vec framework dynamically identifies and aggregates the text corpuses according to the localized themes and contexts associated with each strategic keyword, achieving robust classification even in environments entirely devoid of explicit metadata tags or prior classification parameters (Schopf et al., 2022).

3.2. Analysing the Place-Marketing Function via Monte Carlo Simulation

The computational framework delineated in the previous section allows for a rigorous empirical validation of whether, and to what spatial extent, the evaluated planning documents exhibit an intrinsic place- or regional-marketing function. The operational pipeline first requires the calibration of two distinct keyword sets: a specific target set encapsulating core place-marketing concepts, and a baseline control set representing words highly prevalent in general Hungarian text corpuses. Utilizing the Lbl2Vec algorithm, a definitive place-marketing function can be robustly verified within the Townscape Image Handbooks (TAK) if the document classification yields a statistically significant bias toward the target marketing set compared to the general linguistic baseline.

By leveraging highly frequent, general Hungarian nouns as a counterweight to place-marketing vocabulary, the research design subjects the underlying hypothesis to a conservative statistical stress test. General baseline words possess an inherent statistical advantage due to their high baseline probability of occurrence across any standard text corpus, outperforming purely randomized word selections. To eliminate arbitrariness, we utilize the official linguistic frequency repository compiled by the Hungarian Research Centre for Linguistics (2014).

To ensure the statistical robustness and replicability of the empirical findings, the classification experiment is executed iteratively through a computational simulation. In the research design presented here, the target set of 10 place-marketing benchmark keywords is systematically benchmarked against a control group of 10 general nouns randomly sampled from the top 1,000 most frequent Hungarian nouns. This randomized sampling pipeline is repeated over $N = 100$ iterations. The calibrated place-marketing seed keywords selected for this baseline test are identity (identitás), local (helyi), place (hely), spirit [of place] (szelleme), image (imázs), brand (márka), townscape-related (településképi), identity/visual (arculati), character-driven (karakterű), and townscape (településkép).

Consequently, the control group in each iteration serves as an un-biased linguistic baseline drawn from the broader Hungarian vocabulary. To eliminate stochastic anomalies and random noise, the parameters of the Monte Carlo simulation are formally structured and operationalized through the following mathematical sequence:

$$\mathrm{W} = \{w_1, w_2, \dots, w_{10}\} \tag{3}$$

$$T = \{t_1, t_2, \dots, t_{10}\} \tag{4}$$

where w_i denotes the i-th word randomly selected from the top 1,000 most frequent Hungarian nouns, W represents the aggregated control vector set for a given simulation, t_i signifies the i-th pre-calibrated place-marketing keyword, and T constitutes the structural target set of marketing vectors. Following the algorithmic pipeline specified in Equation (1), the comprehensive dense vector representations of sets W and T are calculated as the mean centroids of their respective constituent word embeddings. These structural document-level representations are mathematically denoted as w and t

$$W \xrightarrow{embedding} w \tag{5}$$

$$T \xrightarrow{embedding} t \tag{6}$$

Subsequently, deploying the Lbl2Vec protocol, the semantic proximity of each digitized TAK document vector (where d_i represents the dense vector of the i-th handbook corpus) relative to the baseline centroid w and the marketing centroid t is systematically computed via the cosine similarity metric established in Equation (2):

$$\text{sim}(d_i, w) = S_{wi} \tag{7}$$

$$\text{sim}(d_i, t) = S_{ti} \tag{8}$$

For each individual handbook document i within the metropolitan corpus, a deterministic classification rule is applied: if $S_{wi} < S_{ti}$, the i-th document is structurally assigned to the place-marketing category; conversely, if $S_{wi} \geq S_{ti}$, the document is categorized as non-marketing, exhibiting a tighter proximity to the general linguistic baseline.

Following this step, the specific proportion of documents classified into the place-marketing cluster *(p_t)* and the proportion assigned to the randomized control baseline *(p_w)* are calculated for each discrete simulation run. Because these metrics represent complementary probability rates, they are constrained by definition:

$$p_t = 1 - p_w \tag{9}$$

This computational pipeline is replicated for $N = 100$ autonomous simulation iterations, where a completely new, randomized control word set W is generated for each consecutive run.

The execution of the Monte Carlo simulation yields a distribution of 100 discrete data points for both p_t and p_w. These descriptive proportions capture the specific structural density of the TAK corpus that exhibits a superior semantic affinity toward place-marketing orientations rather than randomized baseline linguistics across varying iterations.

We can thus formalize the percentage of documents aligning with the place-marketing narrative in any given simulation as a continuous random variable X. Over the course of the 100 simulated runs, the empirical mean of the individual observations X_i serves as an unbiased estimator for the mathematical expected value of the distribution:

$$\bar{X} = \frac{1}{100} \sum_{i=1}^{100} X_i \tag{10}$$

The mathematical expected value and the corresponding variance of the sample mean estimator $\bar{X}$ are structurally defined according to probability theory as follows:

$$E[\bar{X}] = E[X] \tag{11}$$

$$VAR[\bar{X}] = \frac{VAR[X]}{100} \tag{12}$$

Operating under this 100-fold sampling matrix, the empirical evaluation of the metropolitan Budapest text corpus was executed, establishing the mathematical foundation to reveal the numerical density and spatial distribution of local development inclinations.

3.2.1. Empirical Simulation Results and Density Function Analysis

The numerical execution of the 100-fold Monte Carlo simulation yielded a robust empirical probability distribution, which structurally validates the presence of localized development and promotional narratives across the metropolitan planning corpus. The resulting probability density configurations are visualized in **Figure 1**, illustrating the comparative distribution of the target place-marketing framework against the randomized linguistic baseline.

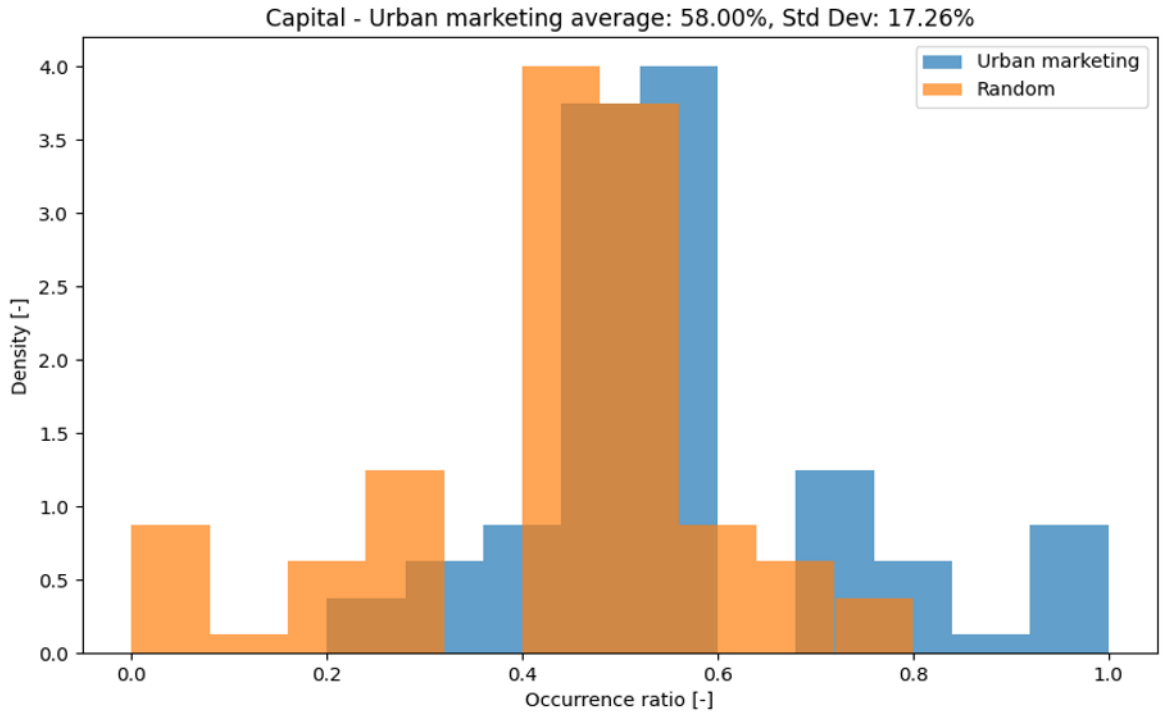


Figure 1. Empirical density distribution of urban marketing occurrence ratios within the Budapest TAK corpus

Source: Authors' own editing

Quantitatively, the computational processing reveals that the mean occurrence ratio of place-marketing narratives across the entire 1+23 Budapest corpus stands at **58.00%**, maintaining a sample standard deviation (σ) of **17.26%**.

As structurally demonstrated by the empirical density functions in **Figure 1**, the target Urban marketing distribution (represented by the blue vector cluster) exhibits a pronounced rightward shift along the horizontal axis relative to the Random control baseline distribution (represented by the orange cluster). This spatial separation in the joint semantic vector space confirms that despite the mathematical advantages inherent to high-frequency common nouns, the semantic profile of the Townscape Image Handbooks is systematically saturated with place-marketing vectors. The significant variance ($\sigma = 17.26\%$) concurrently underscores that while the baseline marketing inclination is omnipresent, individual administrative districts display high heterogeneity regarding their promotional density, reflecting distinct local planning micro-strategies. With the overarching place-marketing function of the metropolitan corpus mathematically verified, the analytical framework transitions to decoding the specific target-group typologies embedded within these institutional narratives.

3.3. Target Group Typology and Keyword Operationalization

Following the verification of the general place-marketing function of the Budapest Townscape Image Handbook (TAK) corpus, the next analytical step focused on identifying the dominant socioeconomic target group orientation embedded within the individual textual segments of the documents. In this phase, the study extended the previously introduced Label to Vector methodological logic to the typological classification of place-marketing narratives.

The analytical premise was that the strategic orientation of each document segment can be inferred from its semantic proximity to theoretically predefined keyword groups. Accordingly, the corpus was first divided into smaller textual units, and each document segment was transformed into a dense vector representation within the same semantic space used for the keyword-based classification. In parallel, the three place-marketing target groups derived from the Ashworth and Voogd framework – tourism, residential orientation, and business premises – were operationalized through separate, pre-calibrated keyword groups. These keyword groups were also vectorized, thereby creating three benchmark semantic centroids against which the document segments could be compared.

The classification was therefore based on the semantic similarity between the vectorized document segments and the vectorized keyword groups. For each textual unit, cosine similarity was calculated in relation to all three target-group vectors. The segment was then assigned to the category with which it exhibited the highest semantic proximity. In other words, the method determined whether a given part of a TAK document was most strongly aligned with tourism-oriented, residential-oriented, or business-premises-oriented place-marketing narratives.

This procedure preserved the unsupervised character of the Label to Vector framework. No manually annotated training dataset was required, and the classification did not depend on explicit labels embedded in the original documents. Instead, the method relied on theoretically grounded keyword groups and their semantic relationship to the textual content of the documents. This approach made it possible to identify latent strategic orientations even when the handbooks did not explicitly refer to tourism, population attraction, or capital attraction as formal planning objectives.

To operationalize the place-marketing framework established by Ashworth and Voogd (1993), three discrete target categories were defined, each represented by a calibrated set of ten seed keywords. These keyword groups served as the semantic reference points of the Label to Vector classification procedure. The computational parameters are summarized in **Table 1.**

Table 1. Calibrated keyword seeds for the tri-partite urban marketing typology within the NLP pipeline

Target Group / Category	Seed Keywords: English translation (Hungarian original)
Tourism	architectural (építészeti), protected (védett), local (helyi), historical (történelmi), natural (természeti), cultural (kulturális), heritage (örökség), world heritage (világörökség), world heritage-related (világörökségi), water (víz)
Residential	buildings (épületek), residential house (lakóház), settlement (település), city/town (város), village (falu), house (ház), residential area (lakóterület), community-related (közösségi), traffic/transportation (közlekedési), residential structure (lakóépület)
Business Premises	industrial (ipari), park (park), agricultural (mezőgazdasági), economic (gazdasági), source/resource (forrás), technical (műszaki), regulatory (szabályozási), industry (ipar), economy (gazdaság), administrative (közigazgatási)

Source: Authors' own compilation.

Formally, let d_j denote the dense vector representation of the j-th document segment, and let c_T, c_R, and c_B denote the centroid vectors of the tourism, residential, and business-premises keyword groups, respectively. Each centroid was calculated as the mean vector of the seed keywords belonging to the given category. The semantic affinity of a document segment to each target group was then computed using cosine similarity:

$$sim(d_j, c_m) = \frac{d_j \cdot c_m}{|d_j||c_m|} \tag{13}$$

where c_m denotes one of the three target-group centroids. The final category assignment of the document segment was determined by the highest similarity value:

$$class(d_j) = \arg \max_{m \in T,R,B} sim(d_j, c_m) \tag{14}$$

where T refers to tourism, R to residential orientation, and B to business premises.

The empirical execution of this pipeline on the macro-baseline document representing the central Municipality of Budapest (Budapest Főváros TAK) revealed a distinct hierarchy in the capital's macro-narrative: **66.80%** of the core text corpus is dominantly aligned with the Tourism dimension, while **18.80%** is dedicated to the Residential paradigm, and **14.40%** targets Business Premises (capital attraction).

3.3.1. Baseline Distribution Analysis

The empirical output of the topic modelling configuration is systematically illustrated in **Figure 2**, which maps the continuous probability density functions (PDF) generated for the Tourism, Residential, and Business premises marketing dimensions across the metropolitan corpus.

As structurally demonstrated by the mathematical density distributions in **Figure 2**, the Tourism function (represented by the grey density cluster) exhibits a pronounced rightward skewness, maintaining a dominant modal peak at higher probability thresholds. This spatial pattern statistically validates its structural hegemony within the metropolitan narrative. Conversely, the Residential trajectory (red cluster) and the Business premises orientation (green cluster) are tightly constrained within lower, well-defined occurrence intervals, with mathematical mean baselines established at 18.80% and 14.40%, respectively.

Consequently, the deployment of this model yields a highly nuanced, non-binary typological classification framework that successfully captures the multi-thematic composition of individual planning documents. Rather than forcing a discrete handbook into a single rigid category, the algorithm models localized policy cross-contamination by calculating continuous probability vectors. This analytical resolution serves as the baseline matrix required to evaluate the explicit strategic orientations and spatial anomalies across the 23 administrative districts of Budapest, as cross-referenced in the comprehensive dataset detailed in **Table 2**.

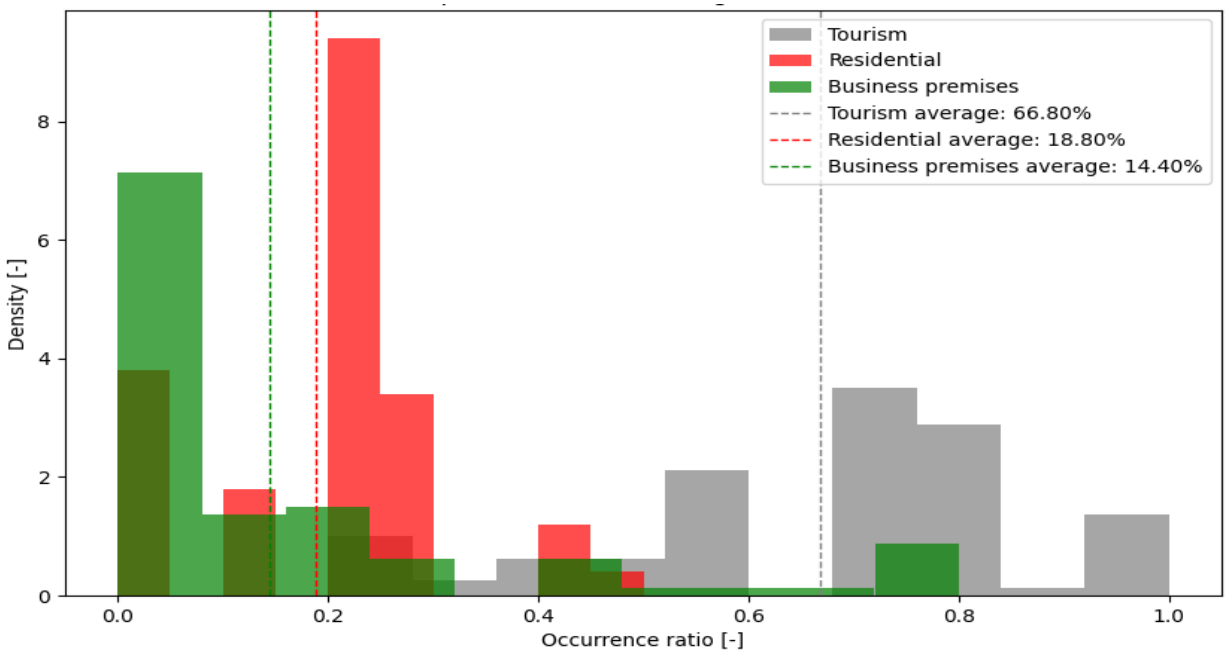


Figure 2: Probability density distribution of the tri-partite urban marketing functions within the Budapest TAK corpus

Source: Authors' own editing.

3.3.2. Empirical Results and Spatial Analysis

To systematically decode the spatial distribution and strategic heterogeneity of place-marketing configurations across the metropolitan fabric, the computational pipeline was executed across all 23 administrative districts of Budapest. The foundational empirical matrix generated by the NLP algorithm, harmonized with localized structural indicators, is structurally compiled in **Table 2**.

Table 2. Comprehensive place-marketing profiles, target-group distributions, and institutional deviations across the 1+23 Budapest TAK sample

District name	TAK Length (pages)	TAK Year of issue	Population[21] 2025.01.01.	**Deviation** from Capital Baseline (rounded percentage points) Marketing (Ma) Tourism (Tu) Residential (Re) Business premise (Bp)	**Marketing** Average (Avg) Standard deviation (SD)	Breakdown by **Target groups** Tourism (Tu) Residential (Re) Business premise (Bp)
BUDAPEST	279	2017	1 700 000		Avg: 58.00%, SD: 17.26%	Tu. Avg.: 66.80% Re. Avg.: 18.80% Bp. Avg.: 14.40%
1st District (Budavár)[22] Várkerület (The Castle)	120	2017	26 011	Ma.: + 7 Tu.: - 15 Re.: - 6 Bp.: + 21	Avg.: 64.50%, SD.: 30.18%	Tu. Avg.: 51.70% Re. Avg.:12.60% Bp. Avg.: 35.70%

[21] Data source: https://nepesseginfo.hu/budapest-xviii-nepessege-lakossaga

[22] The official administrative nomenclature of the 23 districts of Budapest and their respective municipal neighbourhoods is harmonized in accordance with the national urban repository. See: Pénzcentrum (2022) Budapest districts 2022: Here are the names of the Budapest map districts. Available at: https://www.penzcentrum.hu/utazas/20220313/budapest-keruletek-2022-ime-a-budapest-terkep-keruletek-nevei-1122842 (Downloaded: March 7, 2026).

2nd District (Rózsadomb)	210	2017	87 275	Ma.: + 4 Tu.: + 7 Re.: - 2 Bp.: - 5	Avg.: 62.30% SD.: 33.10%	Tu. Avg.: 73.50% Re. Avg.: 17.20% Bp. Avg.: 9.30%
3rd District Óbuda - Békásmegyer	376	2017	125 154	Ma.: + 14 Tu.: + 3 Re.: + 2 Bp.: - 5	Avg.: 71.90% SD.: 32.27%	Tu. Avg.: 69.90% Re. Avg.: 20.40% Bp. Avg.: 9.70%
4th District Újpest	116	2017	95 971	Ma.: + 5 Tu.: + 2 Re.: + 2 Bp.: - 4	Avg.: 62.60% SD.: 26.93%	Tu. Avg.: 68.30% Re. Avg.: 20.80% Bp. Avg.: 10.90%
5th District Belváros – Lipótváros The City	136	2017	25 848	Ma.: + 4 Tu.: + 12 Re.: - 4 Bp.: - 8	Avg.: 62.30%, SD.: 23.70%	Tu. Avg.: 78.50% Re. Avg.: 15.10% Bp. Avg.: 6.40%
6th District Terézváros	199	N/A	36 874	Ma.: + 10 Tu.: - 27 Re.: + 19 Bp.: + 8	Avg.: 67.80%, SD.: 23.22%	Tu. Avg.: 40.10% Re. Avg.: 37.40% Bp. Avg.: 22.50%
7th District Erzsébetváros	171	2017	50 730	Ma.: + 6 Tu.: - 20 Re.: + 6 Bp.: + 14	Avg.: 64.00% SD.: 22.00%	Tu. Avg.: 46.90% Re. Avg.: 24.80% Bp. Avg.: 28.30%
8th District Józsefváros	81	2022	70 809	Ma.: + 2 Tu.: - 25 Re.: + 29 Bp.: - 4	Avg.: 60.30% SD.: 20.52%	Tu. Avg.: 41.40% Re. Avg.: 47.90% Bp. Avg.: 10.70%
9th District Ferencváros	128	N/A	54 316	Ma.: + 11 Tu.: - 22 Re.: + 22 Bp.: - 1	Avg.: 68.70% SD.: 20.13%	Tu. Avg.: 45.20% Re. Avg.: 41.20% Bp. Avg.: 13.60%
10th District Kőbánya	239	2017	72 082	Ma.: + 4 Tu.: + 12 Re.: + 2 Bp.: - 14	Avg.: 62.07% SD.: 33.81%	Tu. Avg.: 78.40% Re. Avg.: 20.80% Bp. Avg.: 0.80%
11th District Újbuda (Kelenföld)	94	2017	132 615	Ma.: +1 Tu.: - 7 Re.: + 7 Bp.: +1	Avg.: 58.90% SD.: 25.02%	Tu. Avg.: 59.90% Re. Avg.: 25.90% Bp. Avg.: 14.20%
12th District Hegyvidék	166	2017	57 784	Ma.: - 15 Tu.: - 26 Re.: - 18 Bp.: + 44	Avg.: 43.40% SD.: 18.18%	Tu. Avg.: 40.50% Re. Avg.: 0.90% Bp. Avg.: 58.60%
13th District (Angyalföld)	181	2021	115 525	Ma.: - 4 Tu.: - 31 Re.: + 16 Bp.: + 14	Avg.: 54.10% SD.: 17.09%	Tu. Avg.: 36.30% Re. Avg.: 35.10% Bp. Avg.: 28.60%
14th District Zugló	192	2017	113 418	Ma.: + 18 Tu.: - 28 Re.: + 9 Bp.: + 19	Avg.: 76.30% SD.: 26.82%	Tu. Avg.: 38.80% Re. Avg.: 28.20% Bp. Avg.: 33.00%
15th District (Újpalota)	156	2017	77 824	Ma.: + 22 Tu.: - 47 Re.: + 31 Bp.: + 16	Avg.: 80.20% SD.: 10.39%	Tu. Avg.: 19.80% Re. Avg.: 49.60% Bp. Avg.: 30.60%
16th District (Mátyásföld)	135	2017	73 045	Ma.: + 23 Tu.: - 4 Re.: + 12 Bp.: - 8	Avg.: 81.00% SD.: 27.07%	Tu. Avg.: 63.10% Re. Avg.: 30.50% Bp. Avg.: 6.40%
17th District Rákosmente	117	2017	87 625	Ma.: + 18 Tu.: - 6 Re.: + 6 Bp.: 0	Avg.: 75.60% SD.: 15.83%	Tu. Avg.: 61.20% Re. Avg.: 24.30% Bp. Avg.: 14.50%
18th District Pestszentlőrinc - Pestszentimre	139	2017	100 363	Ma.: + 3 Tu.: - 24 Re.: + 19 Bp.: + 5	Avg.: 61.00% SD.: 18.47%	Tu. Avg.: 42.80% Re. Avg.: 37.80% Bp. Avg.: 19.40%

19th District Kispest	69	2017	57 799	Ma.: + 18 Tu.: - 4 Re.: - 8 Bp.: +13	Avg.: 75.90% SD.: 18.34%	Tu. Avg.: 62.40% Re. Avg.: 10.50% Bp. Avg.: 27.10%
20th District Pesterzsébet	54	2017	63 117	Ma.: + 2 Tu.: - 38 Re.: +36 Bp.: +1	Avg.: 59.90% SD.: 21.05%	Tu. Avg.: 29.20% Re. Avg.: 55.00% Bp. Avg.: 15.80%
21st District Csepel	65	2017	73 148	Ma.: +3 Tu.: - 10 Re.: + 10 Bp.: - 1	Avg.: 61.00% SD.: 26.85%	Tu. Avg.: 57.10% Re. Avg.: 29.20% Bp. Avg.: 13.70%
22nd District Budafok - Tétény	144	N/A	55 589	Ma.: + 11 Tu.: -1 Re.: - 16 Bp.: + 17	Avg.: 68.80% SD.: 18.18%	Tu. Avg.: 66.20% Re. Avg.: 2.60% Bp. Avg.: 31.20%
23rd District Soroksár	59	2021	22 448	Ma.: + 17 Tu.: - 44 Re.: + 26 Bp.: + 18	Avg.: 75.00% SD.: 18.36%	Tu. Avg.: 22.60% Re. Avg.: 44.90% Bp. Avg.: 32.50%

Source: Authors' own editing

3.3.3. Spatial Disaggregation and Choropleth Mapping of Strategic Orientations

To provide a precise reference matrix for the subsequent spatial and target-group analyses, the administrative and geographical configuration of the study area must be established. Budapest operates under a unique, two-tiered municipal governance structure composed of the central Capital Municipality and 23 autonomous administrative districts, topographically bifurcated by the Danube River into the western Buda side and the eastern Pest side. The spatial distribution and administrative indexing of these 23 districts are systematically mapped in **Figure 3**, serving as the baseline geographic key for the international reader to locate and contextualize the localized planning anomalies discussed in the empirical sections.

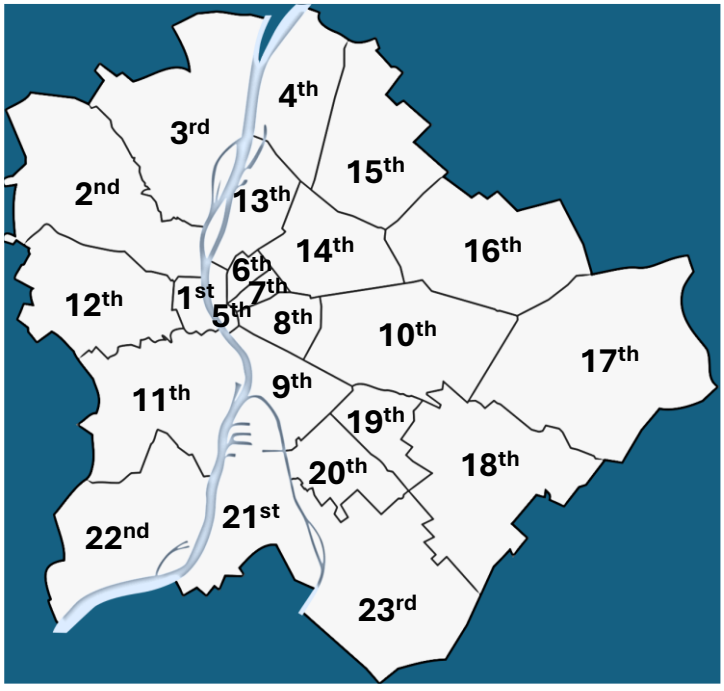


Figure 3. Geographic and administrative mapping of the 23 constituent districts of Budapest

Source: Authors' own editing.

Building upon this administrative geometry, the computational linguistic outputs were spatially projected onto the metropolitan tissue. The localized variance regarding overall marketing density and specific target-group inclinations is detailed in the following thematic sub-sections.

4. Empirical Results and Spatial Analysis

4.1. Macro-Comparative Assessment: National Baseline vs. Metropolitan Budapest

To structurally contextualize the strategic specificities of the metropolitan planning landscape, the baseline results derived from the 1+23 Budapest corpus were systematically benchmarked against the national representative sample previously established by Békési, Kasza, and Tózsa (2025). This macro-comparative analysis reveals a profound structural polarization between the rural/national planning consensus and the capital's narrative strategies.

Quantitatively, while the national representative sample exhibited a higher overall place-marketing density – with nearly two-thirds (66.04%) of the corpus saturated with promotional vectors—metropolitan Budapest demonstrates a more conservative mean marketing density of 58.00%. However, the most drastic structural divergence manifests within the target-group preferences. The overwhelming residential dominance observed at the national level, where 83.70% of all marketing narratives targeted the retention or attraction of inhabitants, collapses to a mere 18.80% within the capital. Conversely, the narrative capacity dedicated to capital attraction and business premises remains remarkably stable across both spatial scales, indexing at 16.30% nationally and 14.40% in Budapest.

The definitive structural anomaly is localized within the Tourism target group. While tourist-oriented narratives remained entirely negligible or fell below statistical significance thresholds across the rural national sample, they surge to an absolute structural hegemony within the Budapest matrix, averaging an unprecedented 66.80% of the marketing profile. This extreme concentration directly mirrors the geopolitical and macroeconomic hegemony of Budapest within the Hungarian tourism sector. The formal statistical parameters of this urban-rural divide are compiled in Table 3.

Table 3: Macro-comparative matrix of place-marketing trajectories between the national representative sample and the metropolitan Budapest corpus

No.	Analytical Metric / Core Parameter	National Representative Sample (*Kasza, 2024; N = 164*)	Metropolitan Budapest Corpus (*Present Study; N = 24*)
1	Share of place-marketing density within the cohort (%)	66,037%	58,00 %
2	Residential-oriented narrative preference (%)	83,7%	18,80 %
3	Business premises-oriented (capital attraction) preference (%)	16,3%	14,40 %
4	Tourism-oriented narrative preference (%)	Statistically negligible	66,80 % *(Exceeds 50% in 13 out of 24 cases)*

Source: Authors' own editing based on Békési, Kasza, and Tózsa (2025).

The detailed spatial disaggregation of these metropolitan marketing profiles across the 23 administrative districts of Budapest is operationally visualized via Datawrapper choropleth mapping in **Figures 4** through **7.** These automated GIS outputs deliver a multi-layered

diagnostic reading of local development commitments, illustrating not only the absolute marketing intensity of individual district-level handbooks but also mapping their structural inclinations – or lack thereof – toward tourists, residential populations, and corporate enterprises across the urban tissue.

4.2. Spatial Analysis of General Urban Marketing Density

The absolute density of implicit urban marketing vectors calculated across the district-level text corpuses is structurally disaggregated through a decile-based choropleth configuration in **Figure 4.** The empirical values range from a minimum threshold of 43.40% to a maximum processing ceiling of 81.00% of the total narrative density.

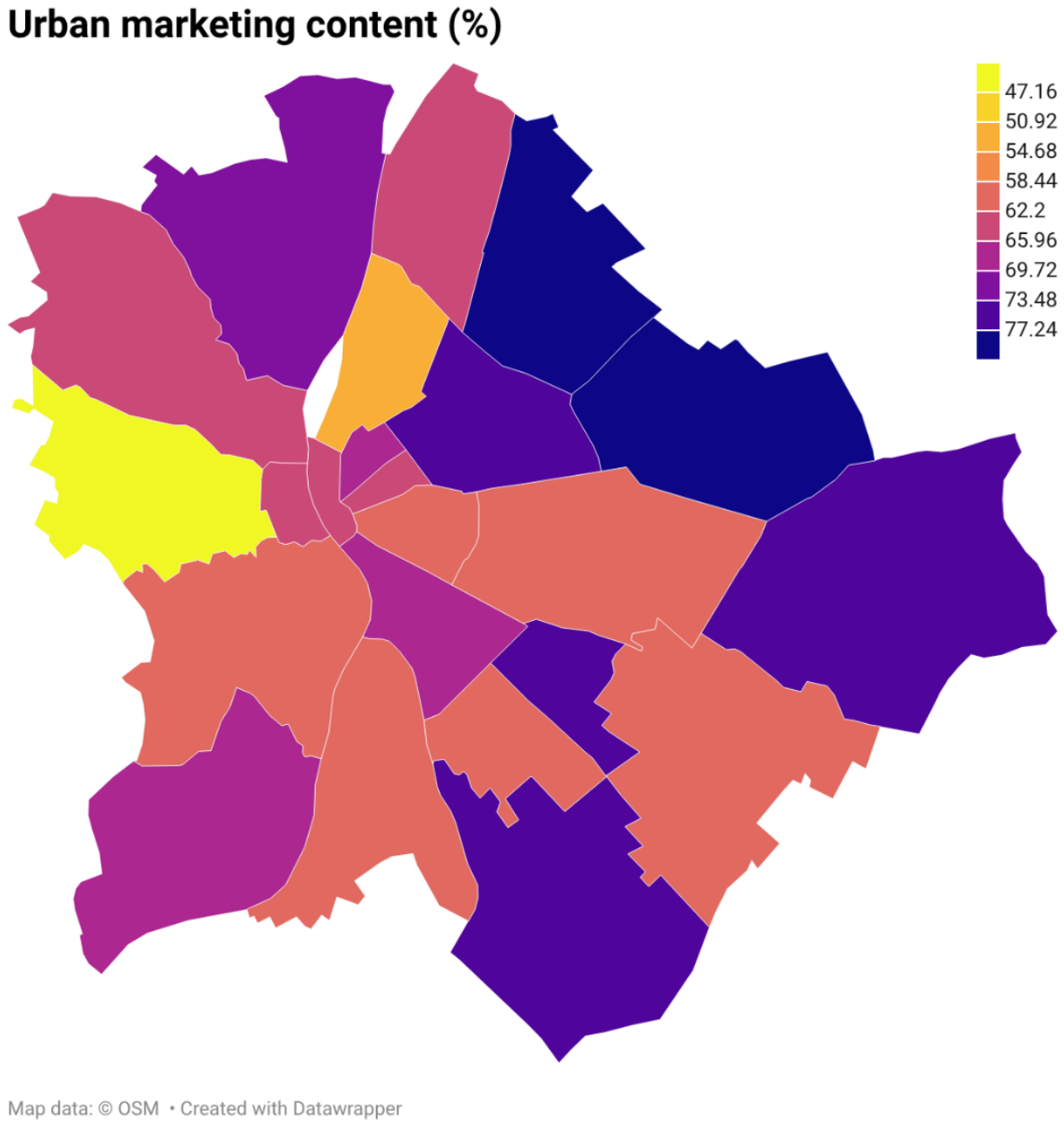


Figure 4. Decile distribution of overall urban marketing content density within the Budapest district TAK corpuses

Source: Authors' own editing, created with Datawrapper

The spatial distribution mapped in **Figure 4** reveals a distinct concentric and structural polarization across the metropolitan fabric, contradicting intuitive assumptions regarding place-promotion intensity. The core analytical takeaways are structured as follows:

- **The Periphery - Centrum Paradigm**: The highest density of structural place-marketing narratives (represented by the dark blue and deep purple deciles exceeding 70%) is systematically concentrated along the outer administrative periphery of the capital. This pattern is particularly dominant across the eastern and south-eastern Pest-Side industrial and transition zones – such as the 15th, 16th, 17th, and 23rd Districts. Methodologically, this spatial cluster signals high institutional pressure; these peri-urban districts lack monolithic historical attractions, forcing local planners to strategically weaponize their handbooks as active text-driven promotional assets to attract capital and stabilize local demographics.

- **The Transitional Belt:** The mid-tier deciles (ranging from 58.44% to 65.96%, mapped in coral and magenta) predominantly partition the traditional transition and dynamic residential zones of the city, including the 4th and 10th Districts. Here, place-marketing elements operate in structural equilibrium with standard building-code regulations, balancing preservationist guidelines with targeted residential supply-side offers.

- **The Elite Buda Anomaly:** Conversely, the lowest relative place-marketing density is localized within the affluent, high-status residential and greenfield zones of the Buda hills, most notably within the 12th District (indexing at the absolute cohort minimum of 43.40%, marked in bright yellow). This constitutes a major spatial anomaly. From an economic geographical perspective, this low density reflects high positional security and institutional complacency; the 12th district possesses a firmly established, self-sustaining high-prestige brand, meaning that local planning authorities face no structural necessity to actively utilize the TAK as an aggressive marketing vehicle.

4.3. Spatial Patterns of Tourism-Oriented Supply-Side Offers

The probabilistic density of tourist-oriented development and branding strategies calculated via the NLP pipeline is spatially mapped across the metropolitan tissue in **Figure 5**. The continuous occurrence parameters range from a localized minimum cohort baseline of 19.80% to a maximum concentration peak of 78.50% of the total strategic text configuration.

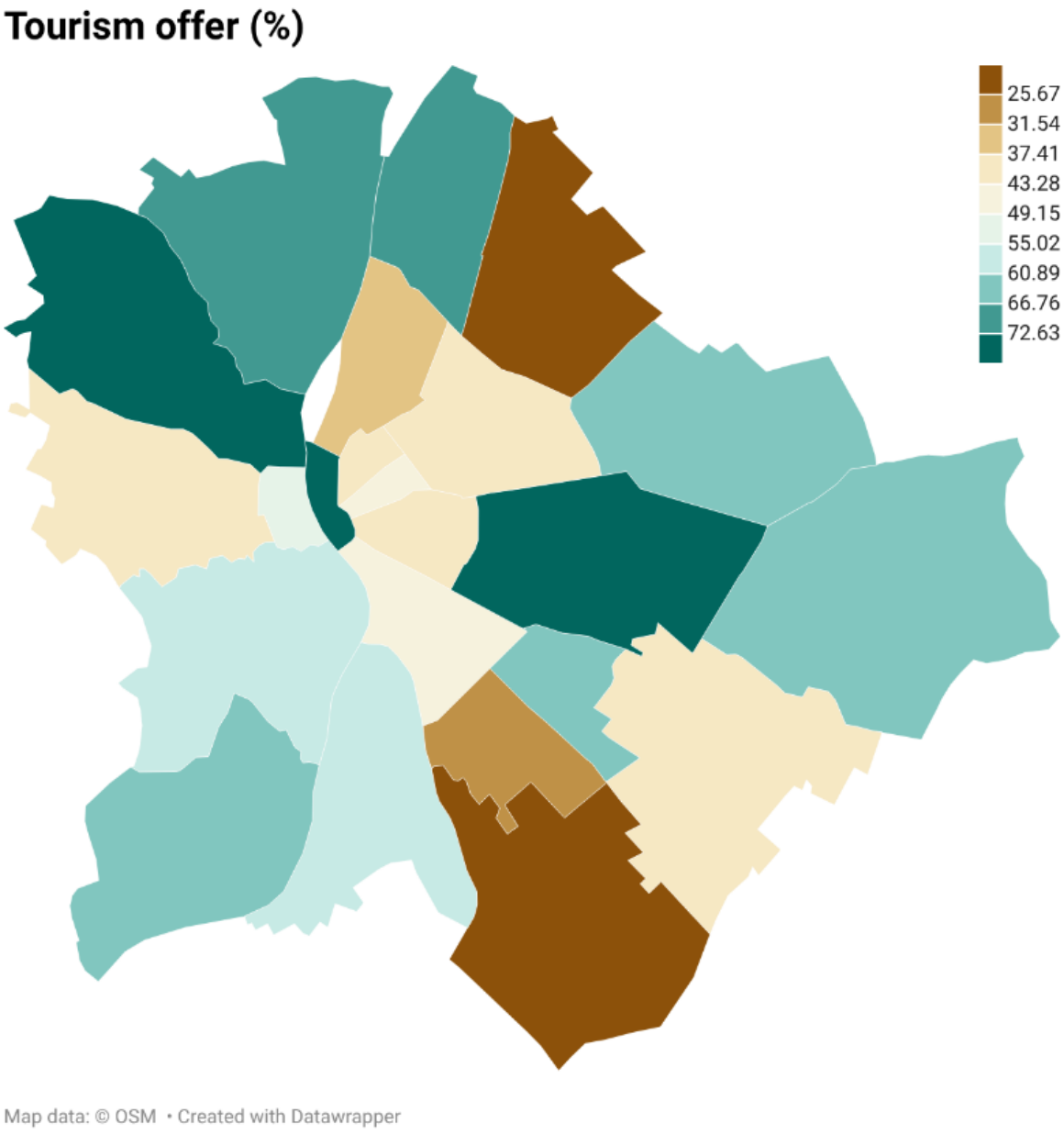


Figure 5. Decile distribution of implicit tourism-oriented urban marketing density within the Budapest district TAK corpuses

Source: Authors' own editing, created with Datawrapper.

The spatial architecture visualized in **Figure 5** reveals a compelling, highly polarized urban layout that shifts the traditional understanding of metropolitan tourism gravity away from a singular historic core toward distinct functional axes:

- **The Northwestern Historical and Recreational Axis:** The highest density interval (exceeding 66.76%, mapped in deep teal) forms a continuous, high-intensity geographic block across the northwestern segment of the capital. This pattern is particularly visible in the historical and topographically rich northwestern urban fabric, where the 2nd and 3rd Districts, together with the 4th and 5th Districts in the core metropolitan zone, serve as key examples of high tourism-oriented textual density. In these spaces, the TAK narrative is structurally saturated with cultural heritage, historic preservation, and environmental

tourism vectors. Local planning authorities in these prestigious districts heavily prioritize the projection of a high-end cultural and recreational identity.

- **The Kőbánya (10th District) Theoretical Anomaly:** A major, highly unexpected structural anomaly is localized within the eastern industrial heartland of the Pest side, specifically the 10th District (Kőbánya), which indexes in the higher dark-teal decile. While intuitively classified as a brownfield industrial transition zone, the NLP pipeline successfully uncovers a massive hidden structural orientation toward cultural/industrial tourism within its handbook. This is explained by the district's extensive, unique subterranean limestone quarry systems, the historical Brewery ecosystem, and emerging post-industrial heritage revitalizations. Local planners aggressively leveraged the TAK to structurally rebrand Kőbánya from a legacy industrial zone into an avant-garde industrial heritage destination.

- **The South-Budapest Non-Tourism Belt:** Conversely, the absolute minimum density thresholds (ranging from 22.60% to 37.41%, mapped in deep brown and beige) are concentrated across the southern and south-eastern agrarian and heavy-industrial peripheries, such as the 23rd (Soroksár) District. In these peri-urban territories, the local planning apparatus completely discards tourism as a viable economic or spatial trajectory, leaving the handbook narrative optimized for alternative developmental target groups.

4.4. Spatial Patterns of Residential-Oriented Supply-Side Offers

The spatial allocation of residential-oriented marketing density and population-retention narratives across the 23 administrative units is structurally compiled in **Figure 6**. The empirical values display high local volatility, ranging from an absolute minimum threshold of 0.90% to a maximum strategic concentration peak of 55.00% of the localized handbook text.

The computational map visualized in **Figure 6** reveals a clear spatial layout that reflects the structural realities of suburbanization, housing development capacities, and local demographic targets:

- **The Outer Residential and Suburbanization Belts:** The absolute highest density cluster (exceeding 44.18%, mapped in dark green) is localized within two distinct metropolitan growth vectors. On the Pest side, this pattern is particularly visible in selected northern, inner-transition, and south-eastern residential environments, with the 15th and 20th Districts serving as prominent examples of high residential-oriented textual density. On the Buda side, this high-intensity narrative is not dominant. In these specific environments, local planning authorities utilize the TAK as an active residential marketing asset, highlighting housing expansion zones, family-friendly green infrastructures, and residential community spaces to actively attract middle-class demographics and urban-suburban migrators.

- **The Compact Inner Core and Prestige Anomaly:** Conversely, the absolute minimum density thresholds (ranging from 0.90% to 17.13%, marked in light beige) are heavily concentrated in selected central and high-status saturated urban environments, including the historic central business district of the 5th District and, most notably, the hyper-affluent residential enclaves of the Buda hills in the 12th District. From an urban marketing perspective, this spatial pattern is highly logical: these central and elite zones are structurally saturated, offering zero territorial capacity for large-scale greenfield or brownfield housing developments. Consequently, local planners completely omit expansionist residential

recruitment from their handbooks, focusing instead on high-end preservation or commercial spatial functions.

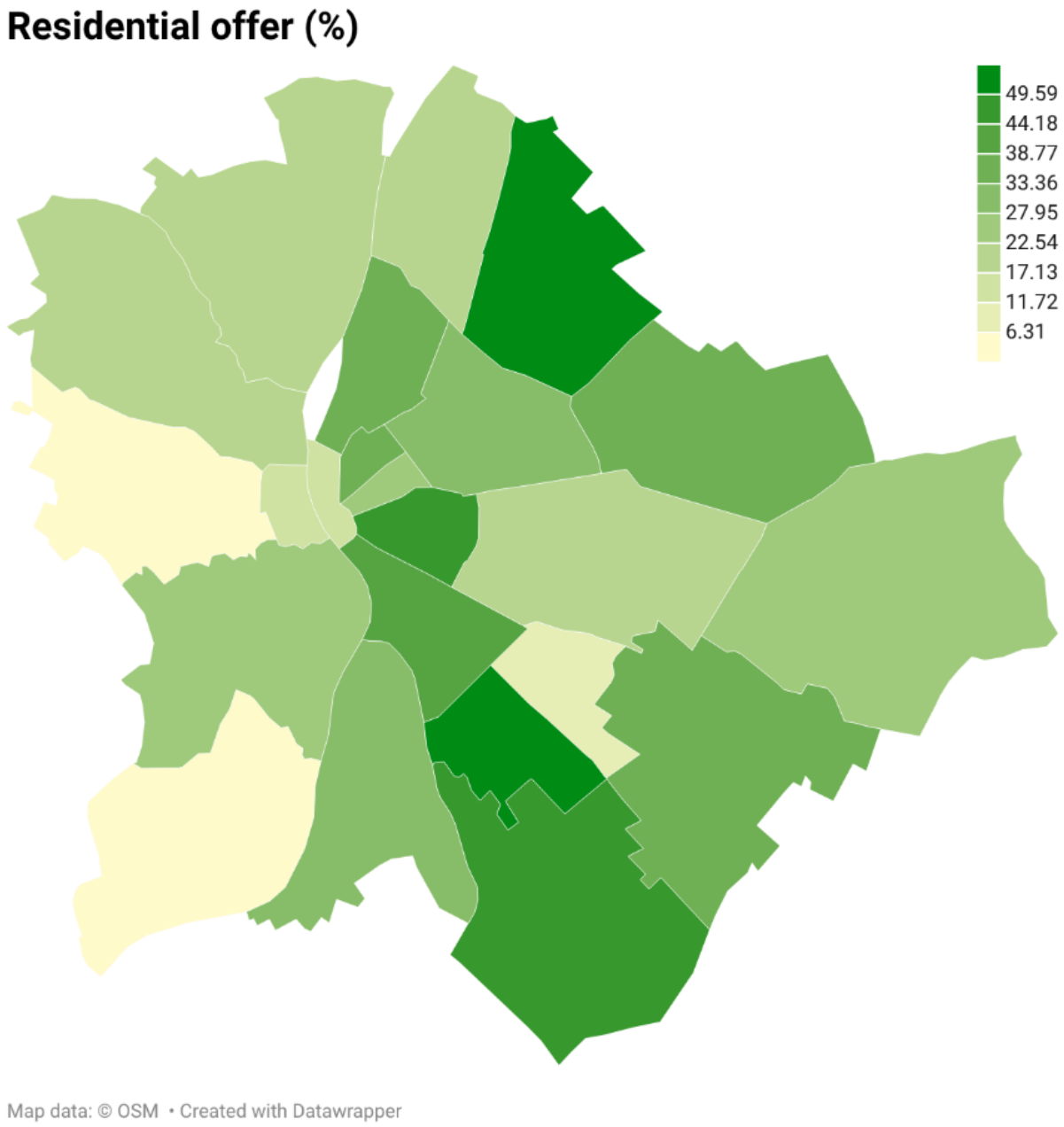


Figure 6. Decile distribution of implicit residential-oriented urban marketing density within the Budapest district TAK corpuses

Source: Authors' own editing, created with Datawrapper

4.5. Spatial Patterns of Business Premises-Oriented Supply-Side Offers

The probabilistic spatial density of business premises-oriented narratives and capital-attraction vectors calculated across the metropolitan corpus is systematically disaggregated in **Figure 7.**

The continuous text-metric thresholds display extreme localized polarization, ranging from a baseline minimum of 0.80% to an absolute cohort peak of 58.60% of the localized planning text.

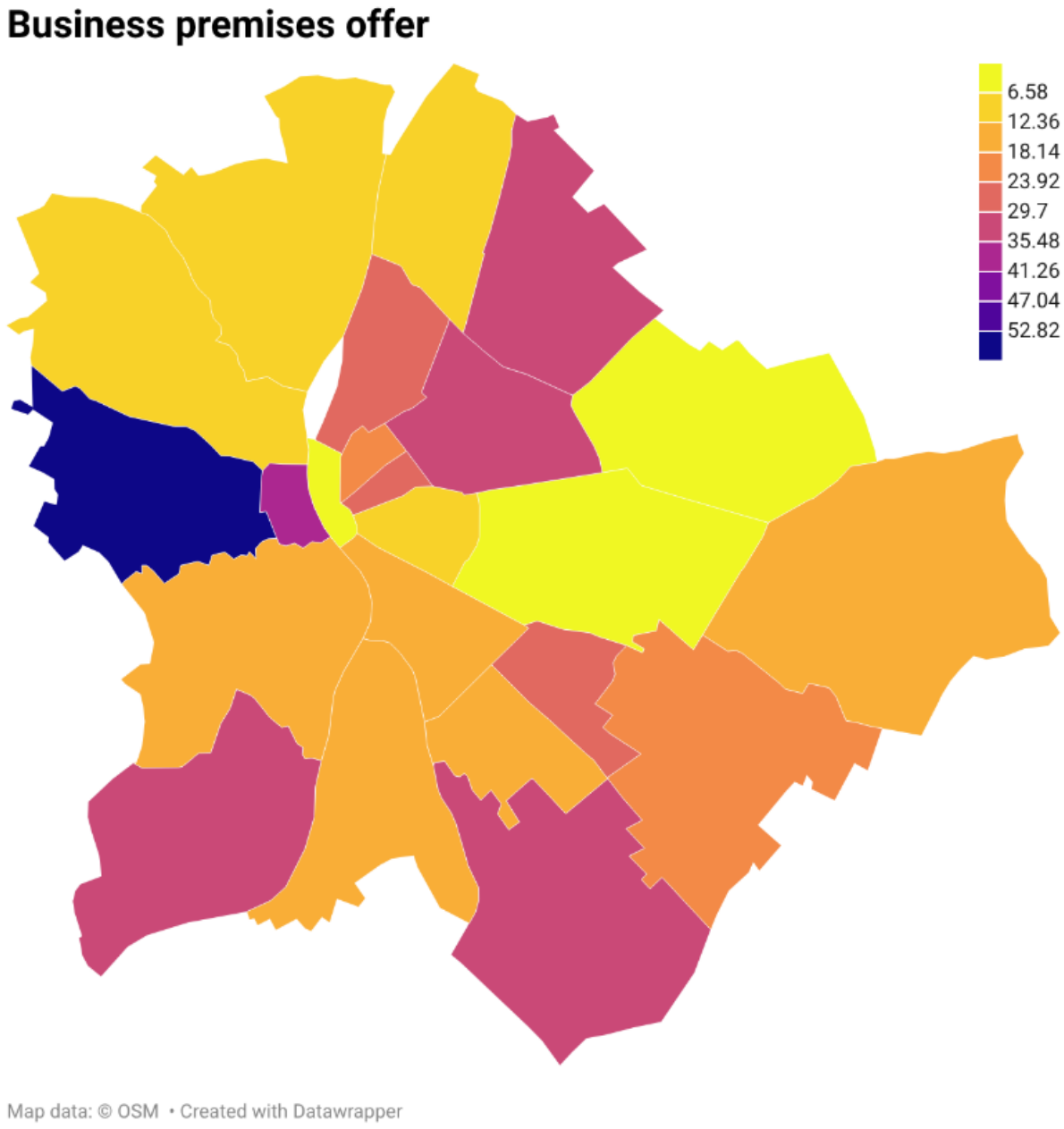


Figure 7: Distribution of business premises-oriented content by deciles across city districts in Budapest

Source: Authors' own editing, created with Datawrapper.

The spatial patterns visualized in **Figure 7** reveal a highly unconventional, counter-intuitive macroeconomic configuration across the capital's administrative geography:

- **The High-Value Buda Corporate Enclave (The 12th District Paradox):** The absolute structural peak of capital-attraction and business-premises density (exceeding 52.82%,

mapped in deep purple and dark blue) is heavily concentrated within the 12th District (Hegyvidék). This represents a profound planning paradox. While topographically and socially established as a low-density, greenfield elite residential zone, the NLP pipeline successfully demonstrates that its handbook narrative is aggressively optimized for business-premises promotion. Lacking heavy industrial or agricultural capacities, local planning authorities strategically weaponized the TAK to attract high-end, clean commercial assets—such as corporate headquarters, private medical clinics, high-tech R&D facilities, and premium diplomatic/business office spaces.

- **The Post-Industrial Pest Heartland:** The mid-to-high deciles (ranging from 29.70% to 41.26%, mapped in deep magenta and coral) form a belt across the southern and eastern periphery of the Pest side, including the 15th and 23rd Districts. In these traditional industrial, brownfield, and transport-logistics zones, the business-premises offer represents a consistent, path-dependent structural baseline. Local planners systematically use the handbooks to signal available territorial capacities for light manufacturing, commercial logistics, and innovative post-industrial brownfield redevelopments.

- **The Saturated Northern and Central Zones:** Conversely, the absolute minimum density thresholds (ranging from 0.80% to 12.36%, marked in bright yellow) span across a large geographic block in northern Buda (2nd and 3rd Districts) and eastern Pest (10th and 16th Districts). In these sectors, the local planning frameworks heavily deprioritize business and industrial capital recruitment within their handbooks, actively sheltering their territories from corporate spatial saturation in favour of residential or heritage tourism development paths.

5. Discussion and Policy Implications

5.1. Evaluation of General Place-Marketing Density within the TAK Corpus

The empirical findings generated across the metropolitan Budapest planning framework demonstrate that although the Townscape Image Handbooks (TAK) were originally mandated under strict architectural, aesthetic, and non-promotional statutory guidelines, both the 2024 national baseline sample (***N = 164***) and the 2026 metropolitan sample (***N = 24***) exhibit a profound, statistically significant saturation of implicit place-marketing content. The analytical auditing of the capital's planning text was executed via a two-tiered protocol: the initial phase quantified the absolute distribution of promotional density, while the subsequent phase structurally decoded the target-group orientations embedded within those identified marketing clusters.

Methodologically, the computational processing transformed a comprehensive text corpus totalling **3,626 pages** derived from the 24 standalone master handbooks (***d = documents***) alongside the calibrated dictionary arrays (***k = keywords***) into dense vector trajectories via the Label to Vector neural architecture.

The quantitative distribution of absolute place-marketing density across the administrative boundaries displays a significant spatial spread, ranging from a cohort minimum of 43.40% in the 12th District to an absolute ceiling of 81.00% in the 16th District. Crucially, the 12th District constitutes the singular administrative anomaly across the entire metropolis where the calculated place-marketing density fails to cross the 50.00% absolute equilibrium threshold. In

all other 22 administrative units, as well as within the macro-level capital handbook (Budapest Főváros TAK), promotional vectors consistently dominate the planning text.

When benchmarked against the national baseline, the metropolitan corpus exhibits high internal heterogeneity: 10 district handbooks surpass the national place-marketing average, while 13 index below it. The high-intensity cluster that significantly outpaces the national rural baseline of **66.04**% is geographically constrained within the following administrative units:

(1) 3rd District 71.90%, (2) 6th District 67.80%, (3) 9th District 68.70%, (4) 14th District 76.30%, (5) 15th District 80.20%. (6) 16th District **81.00**% (Cohort Maximum), (7) 17th District 75.60%, (8) 19th District 75.90%. (9) 22nd District 68.80%, (10) 23rd District 75.00%.

The granular graphical disaggregation of these individual district profiles, tracking both absolute density and target-group probability shares, is structurally compiled for all 23 units within the Appendix of this paper.

5.2. Target-Group Typology and Institutional Planning Orientations

The secondary phase of the computational text matrix successfully confirms that the identified place-marketing density is not structurally amorphous but is precisely calibrated toward the three key socio-economic target groups defined within the Ashworth–Voogd framework. The empirical processing tracks extreme localized variance across all three strategic tracks:

- **Tourism probability scores** range from a localized minimum of 19.80% (15th District) to an absolute structural peak of **78.50**% (5th District).
- **Residential probability scores** range from an absolute minimum of 0.90% (12th District) to a structural maximum of **55.00**% (20th District).
- **Business premises probability scores** range from a minimum baseline of 0.80% (10th District) to an absolute cohort peak of **58.60**% (12th District).

The structural configurations of the 12th and 10th Districts represent the most severe spatial anomalies detected across the metropolitan fabric. As noted, the 12th District stands as the unique planning entity where the baseline promotional inclination drops below the 50.00% threshold. Concurrently, its internal target-group distribution reveals a highly asymmetric composition: it commands the absolute maximum score for Business premises (58.60%) while recording the absolute cohort minimum for Residential narratives (0.90%). Diametrically opposed to this layout is the 10th District (Kőbánya), which compresses its capital-attraction capacity to the cohort minimum of 0.80% while aggressively reallocating its narrative assets to establish a commanding near-maximum in Tourism positioning (78.40%), however, the absolute maximum within this tourism-oriented dimension is still recorded by the inner-city 5th District.

From a strategic management perspective, these empirical benchmarks serve a dual function. First, they establish a standardized reference matrix allowing local urban planners and municipal assemblies to precisely position, evaluate, and calibrate their localized developmental priorities against competing districts, the capital municipality baseline, and national trends.

Second, from a macro-governance perspective, possessing a quantitative, empirical map of local narrative inclinations is of paramount importance for central state secretariats and

metropolitan planning authorities. Urban development cannot operate sustainably in a data vacuum. Having an explicit, text-derived indicator of district-level planning inclinations allows macro-planners to accurately predict where spatial developments will encounter the least structural friction. It reveals precisely which districts maintain high institutional readiness for greenfield or brownfield capital investments (Business premises), where local governance is optimized to absorb suburban demographic expansion (Residential), and where the institutional matrix is structurally configured to prioritize heritage and ecological conservation (Tourism).

This metropolitan distribution underscores a profound structural transformation relative to the national baseline. At the national scale, place-marketing was restricted to a singular, path-dependent track: **83.70**% of all rural narratives exclusively targeted Residential recruitment, with Business premises capturing a minor 16.30% and Tourism failing to achieve statistical significance. Within the metropolitan ecosystem of Budapest, this monochromatic model is replaced by a diversified, highly active multi-thematic architecture. The absolute dominance of Tourism is empirically verified by the fact that in 13 out of the 24 evaluated handbooks – representing more than 54% of the total metropolitan corpus – the tourist-oriented supply-side offer systematically exceeds the 50.00% absolute majority threshold.

5.2.1. Institutional Dissonance and Structural Deviation Analysis

To rigorously map the alignment – or lack thereof – between the overarching metropolitan planning directives and localized district-level execution, a comparative delta (Δ) analysis was conducted. This model mathematically subtracts the master baseline values of the central Capital Municipality (Budapest Főváros TAK) from the individual probability metrics of the 23 administrative districts across all four parameters (General Marketing, Tourism, Residential, and Business Premises). The resulting spatial pattern of institutional harmony and independent local path-dependencies is systematically visualized in **Figure 8**.

The strategic indicators compiled in **Figure 8** serve as an empirical diagnostic map, exposing that systemic institutional synchronization is highly fragmented across the metropolis. The spatial variations reveal a deeply decentralized planning landscape, characterized by the following specific local anomalies:

- **The Cohort Baselines (Harmony *vs.* Polarization):** The highest degree of structural compliance and the smallest overall deviation from the central capital master plan are localized within the 4th (Újpest) and the 11th (Újbuda) Districts, which closely mirror the metropolitan planning equilibrium. Conversely, some of the strongest structural polarizations and spatial distances from the baseline are visible in the 15th and 20th Districts, which exemplify highly divergent local planning profiles.
- **Target-Group Extremes and Local Resistance:** The most severe planning divergence manifests within the tourism sector, where peripheral and transition units display aggressive local resistance to mass-tourist branding. The absolute negative deltas are recorded by the 15th and 23rd Districts, which reject the capital's tourism heavy narrative by -47.0 and -44.2 percentage points, respectively. Regarding population recruitment, the 20th District (Pesterzsébet) marks the absolute metropolitan peak in proactive community promotion, exhibiting a massive +36.2 percentage points positive deviation in residential narratives relative to the capital framework.

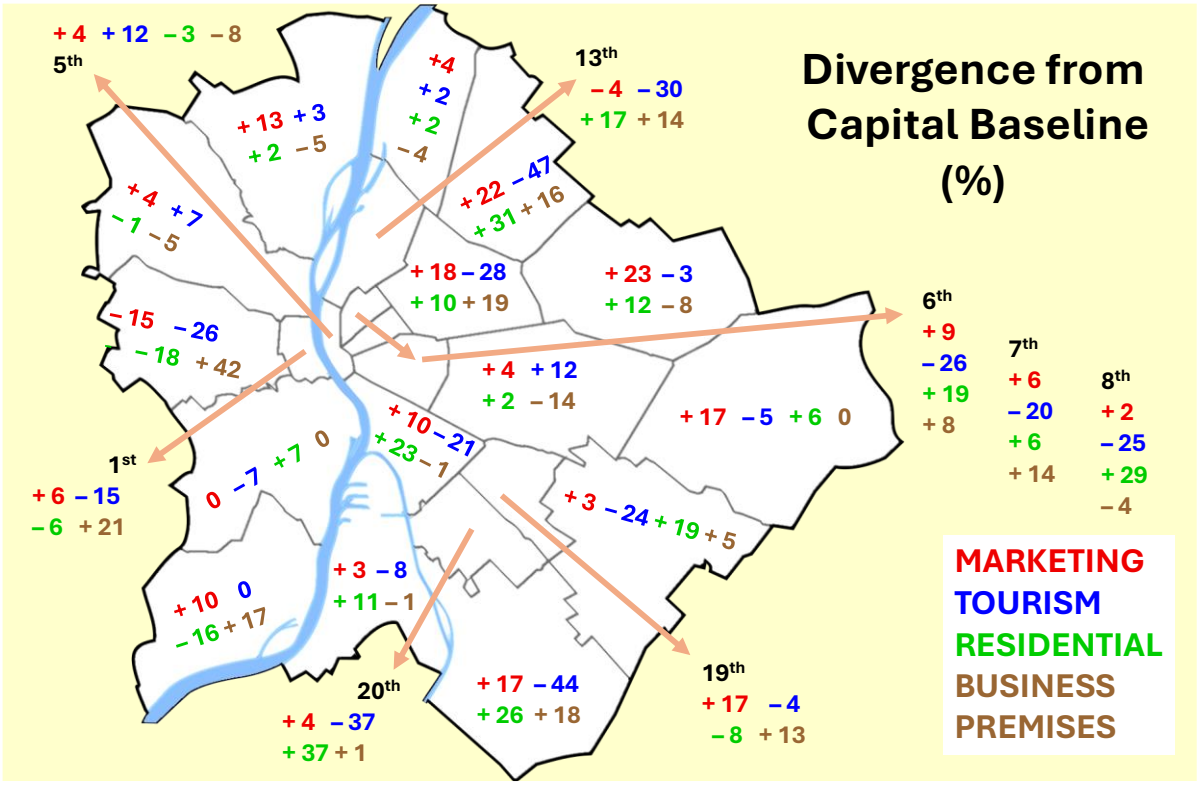


Figure 8: The divergence of district-level TAK contents from the capital's baseline, having used Label-2-Vector analysis to show percentage shifts in marketing, tourism, residential and business premises content types.

Source: Authors' own editing, created with Datawrapper.

- **The Capital Attraction and Business Premise Volatility:** In the corporate and capital attraction sector, the 12th District (Hegyvidék) commands a dominant, unique position, pacing +44.2 percentage points higher than the central capital baseline. A compelling structural anomaly is similarly detected within the historical heartland of the 1st District (Várnegyed); while intuitively classified as a tourism hotspot, its handbook surprisingly indices -15.1 percentage points below the capital baseline for tourism, while jumping +21.3 percentage points higher in business premises and high-value localized commercial infrastructure recruitment. Conversely, the absolute strongest narrative rejection of corporate spatial saturation is localized in the 10th District (Kőbánya), which drops -13.6 percentage points below the capital baseline for business-premises promotion.
- **The Volumetric and Lifecycle Footprint of the Corpus:** These local anomalies are strongly correlated with the structural properties of the individual text corpuses. The total metropolitan macro-sample displays high structural variance in document length and update frequencies. The absolute volumetric minimum is localized within the southern transition zone, where the 20th District operates with a highly condensed handbook of merely 54 pages, followed closely by the 23rd District at 59 pages. Conversely, the absolute volumetric maximums are concentrated in the northern and eastern zones; the 3rd District (Óbuda-Békásmegyer) commands the longest text corpus at 376 pages, while the 10th District (Kőbánya) represents the secondary ceiling at 239 pages. Crucially, as of the 2026 planning cycle, the vast majority of these handbooks still reflect the historical baseline parameters established in 2017. Out of the 24 evaluated documents, a mere cohort of three units has successfully undergone complete statutory updates or comprehensive rewrites: the 13th and 23rd Districts, alongside the

8th District (Józsefváros), which represents the most modern updated text corpus, finalized in 2022. The comprehensive updating of the central capital handbook is slated for conclusion later in 2026.

Ultimately, this multi-criteria delta matrix underscores that the vast majority of Budapest's administrative districts do not structurally synchronize their townscape handbooks with the central municipality's tourism-heavy master narrative. Instead, they execute highly localized, distinct planning paths tailored to their unique socio-geographic challenges. The automated NLP pipeline functions as an effective institutional audit tool, exposing these hidden tensions between centralized metropolitan branding and localized, district-level realities.

6. The Future of AI in Regional and Urban Planning

The complex, mixed-methods computational framework introduced in this study – demonstrated through the intra-metropolitan empirical lens of Budapest – offers a replicable methodology to extract latent institutional insights, which are critical for aligning place-marketing parameters with strategic urban development. Rather than treating textual planning documents as passive regulatory scripts, this paper positions automated text auditing as an active tool for predictive planning governance. To provide a standardized, future-oriented operational matrix for regional scientists and municipal planning authorities, the conceptual and computational toolkit driving this AI-based urban planning revolution is systematically synthesized in **Table 4**.

Table 4. Automated conceptual and computational architecture for AI-driven place-marketing and spatial planning audits

No.	Advanced Planning Concept / NLP Method	Definition Formal Methodological Specification and Regional Planning Utility
1	Pre-2023 baseline paradigm (Human intelligence baseline)	Prior to the widespread democratization of deep-learning LLM architectures in 2023, analysing the promotional dimensions of the TAK corpus relied strictly on manual, qualitative content analysis. In the 2017–2022 baseline cycle, research was restricted to testing qualitative, localized hypotheses regarding whether marketing intensity correlated with: (i) public engagement metrics at municipal town halls, (ii) the prior acquisition of national 'City Marketing Diamond' prestige awards, (iii) the collaborative or isolated drafting of handbooks by unified municipal joints, or (iv) functional integration with regional digital planning portals (ÉTAK).[23]
2	Machine Learning scaling algorithms	Represents a paradigm shift in text processing capacity. While unable to substitute human planning intuition, it bypasses cognitive constraints by successfully processing vast text corpuses (potentially exceeding 10,000 pages). This facilitated a methodological leap, enabling representative national testing (***N=164***) and the integration of rigorous mathematical-statistical stress tests alongside traditional qualitative and visual townscape assessments.
3	Unsupervised text classification	Refers to a non-traditional text categorization pipeline implemented when pre-annotated, hand-labelled training metadata is unavailable. Because historical qualitative audits of the TAK corpus were incapable of generating discrete, scalable training labels[24], supervised learning was mathematically unfeasible. In this unsupervised environment, the neural network relies exclusively on target keyword vectors to discover latent semantic structures.

[23] Architectural Landscape Image Handbook

[24] Prior to the deployment of automated computational pipelines, several structural hypotheses were evaluated utilizing manual human intelligence. However, due to cognitive and resource constraints, none of these qualitative approaches could be scaled to meet the minimum representative sampling threshold – formally defined as 5% of the total 3,177 Hungarian TAK universe, equating to a statistical baseline of at least 159 documents. Consequently, manual text pre-annotation (labelling) remained technically unfeasible, structurally necessitating the transition to unsupervised, dictionary-based keyword vectorization seeds to process the representative national and metropolitan datasets.

4	Label to Vector (Lbl2Vec) framework	A specialized, unsupervised neural embedding technique engineered to classify unstructured text based on spatial proximities in a joint semantic vector space. By performing dense text vectorization and computing multi-dimensional cosine similarity profiles across 100 simulation iterations, the pipeline establishes whether a target document aligns with a place-marketing matrix or a randomized linguistic control baseline.
5	Dense word embeddings	The foundational mathematical matrix that maps unstructured linguistic syntax into dense, low-dimensional numerical arrays readable by deep-learning architectures. Vectorizing the complete planning text ($\boldsymbol{D}$) alongside target marketing vectors ($\vec{\boldsymbol{k}}_1$) and baseline control vectors ($\vec{\boldsymbol{k}}_2$) allows algorithms to determine the absolute strategic inclination of a city's development policy by measuring the exact cosine angles between their vector trajectories.

Source: Authors' own compilation

7. Conclusion

The empirical trajectory of this study establishes a novel methodological milestone within spatial sciences by demonstrating how advanced computational linguistics can structurally redefine our understanding of local development policies. While contemporary regional research conventionally relies on rigid fiscal indicators or explicit zoning maps to evaluate municipal trajectories, this paper proves that the text-driven auditing of local planning corpuses uncovers hidden strategic inclinations that would otherwise remain latent. By successfully executing an unsupervised Label to Vector (Lbl2Vec) and class-based NLP pipeline across a massive 3,626-page text corpus spanning the 23 administrative districts of Budapest, this research marks a profound paradigm shift from manual, speculative content analysis to automated, multi-dimensional narrative mapping.

The overarching significance of this methodology is directly tied to the explosive, transformative evolution of artificial intelligence frameworks post-2023. This technological leap has equipped regional scientists and public administration experts with unprecedented computational capacities to filter out human coder bias. The intra-metropolitan empirical lens of Budapest successfully validates this tool, exposing a clear, structural polarization between the centralized, tourism-heavy master identity enforced by the Capital Municipality (66.80% tourism baseline) and the highly fragmented, "renitent" functional path-dependencies executed by individual administrative districts.

Looking forward, the academic and practical implications of these findings transcend traditional urban geography. If these innovative, AI-based textual analyses structurally diffuse across the social sciences, they can function as primary catalysts for improving overall economic competitiveness within the European Union. Specifically, as systematically demonstrated by Csontos, Éltető, and Sass (2026), such advanced text-auditing frameworks will directly contribute to optimizing Foreign Direct Investment (FDI) inflows. By precision-mapping the latent, narrative inclinations of local societies embedded within documents like the TAK, macro-planners and multinational enterprises gain an empirical tool to anticipate localized community friction, thereby discovering proactive, structural pathways to mitigate or entirely avoid the disruptive NIMBY (Not In My Back Yard) phenomenon in attracting high-value global capital assets.

Ultimately, the successful deployment of these algorithmic innovations relies heavily on human institutional readiness and cultural adoption. In this context, we firmly trust in the socio-technological paradigm established by Fási, Marton, and Szűcs (2025), which proves that younger demographic cohorts and individuals with higher educational attainments exhibit significantly less friction and fear regarding rapid technological disruptions, demonstrating an intrinsic openness toward artificial intelligence (AI). Regional development – regardless of

whether it operates at the microscopic settlement level, the meso-scale county structure, or the macro-metropolitan matrix – will be inevitably and profoundly transformed by AI-driven diagnostics, such as the neural pipeline presented within this study. Far from substituting human planning intuition, these computational innovations empower future generations of urban architects to ensure that a city's strategically projected brand identity is flawlessly, mathematically synchronized with its evolving socioeconomic and regulatory spatial realities.

If the innovative AI-based textual analyses – such as the tools presented in this study – spread in social sciences it can result in improving competitiveness in the EU, they will contribute to FDI investments as proved by Csontos-Éltető-Sass (2026), because the detection of the "hidden" inclination of local societies, manifested in the TAK texts e.g. will find ways to avoid the NIMBY phenomena in attracting FDI.

We trust in the statement of Fási-Marton-Szűcs (2025) namely that younger age groups and those with higher education – in local and regional development, too – are less fearful of technological development, primarily artificial intelligence (AI). Regional development, regardless of whether it is settlement-, county- or regional-level development, are also significantly influenced in the future by AI-based devices, like the ones presented in our study now.

Future research may extend the present Budapest-based corpus after the completion of the updated Budapest Townscape Image Handbook in 2026, enabling a direct comparison between the 2017 and 2026 capital-level TAK documents. This would also make it possible to examine the alignment between the updated Budapest TAK and the newly adopted metropolitan development and zoning master plan, as well as the extent to which district-level TAK updates follow or diverge from the capital's strategic framework. Finally, an enlarged Budapest sample could provide a basis for comparing the 25 capital-level and district-level TAK documents with the 25 Townscape Image Handbooks of Hungarian cities with county rights, thereby broadening the national relevance of the AI-based textual audit.

Acknowledgements

The authors acknowledge the use of Gemini (Google) for language polishing and stylistic support during the preparation of the manuscript. The authors remain fully responsible for the content, analysis, and conclusions of the study.

References

Ashworth, G. J. – Voogd, H. 1993. Selling the City -Marketing Approaches in Public Sector Urban Planning – Wiley 226 p. ISBN: 047194470X, 9780471944706

Babbie, E. 2008. A társadalomkutatások módszertana (Methodology of Social Research – in Hungarian). Balassi Kiadó, Bp. ISBN: 978-963-506-764-0, 715 p.

Bíró H. – Láposi R. 2018. A skót meló – Városarculat Skóciában (The Scottish job – Townscape image in Scotland in Hungarian) = Építészfórum 2018.01.10. https://epiteszforum.hu/print/a-skot-melo-varosarculat-skociaban1

Békesi G. B. – Kasza Zs. – Tózsa I. 2024. A Településképi Arculati Kézikönyvek és a településmarketing kapcsolatának vizsgálata – (The Relationship between Townscape Image Handbooks and Urban Marketing – in Hungarian) = ***Területi Statisztika***, 65.3. pp 390-413. DOI: https://doi.org/10.15196/TS650305 Download: https://www.ksh.hu/statszemle_archive/terstat/2025/2025_03/ts650305.pdf

Chang, Y. – Wang, X. – Wang, J. – Wu, Y. – Yang, L. – Zhu, K. – Xie, X. 2024. A survey on evaluation of large language models. ACM Transactions on Intelligent Systems and Technology, 15(3): 1–45. DOI: https://doi.org/10.1145/3641289

Csontos T. – Éltető A. – Sass M. 2026. Foreign direct investment and trade patterns in the European Union: implications for competitiveness = ***Regional Statistics***, Vol. 16. No. 2. 2026: 388–420; DOI: https://doi.org/10.15196/RS160208 Download: https://www.ksh.hu/statszemle_archive/regstat/2026/2026_02/rs160208.pdf

Fási Cs. – Marton Á. Szűcs P. 2025. On the way to Industry 5.0? Analysing the relationship between attitudes regarding AI and material deprivation in Hungary = ***Regional Statistics***, Vol. 15. No. 6. 2025: 1124–1149; DOI: https://doi.org/10.15196/RS150604 Download: https://www.ksh.hu/statszemle_archive/regstat/2025/2025_06/rs150604.pdf

Kasza, Zs. 2020. Visual Appearance of the Architectural Image in the Kaposvár Image Manuals. Corvinus Regional Studies, 5 (7): pp 1–9. Download: http://web.uni-corvinus.hu/~ljeney/CRS_2020_7_Kasza.pdf

Kasza Zs. 2024. Arculati kézikönyvek marketingje (Marketing in Townscape Image Handbooks – in Hungarian) Doctoral theses manuscript. 244 p. Download: https://en.uni-nke.hu/document/en-uni-nke-hu/M%C5%B1helyvita%20KZS.pdf

Lülök G. – Dobos I. – Sebestyén Z. 2026. Statistics on the use of AI technologies in the member states of the EU = ***Regional Statistics***, Vol. 16. No. 1. 2026: 3-32; DOI: https://doi.org/10.15196/RS160101 Download: https://www.ksh.hu/statszemle_archive/regstat/2026/2026_01/rs160101.pdf

Makovényi, F. – Simon, M. 2021. A településképi arculati kézikönyvek és a településképi rendeletek vizsgálata a nyilvánosság számára való elérhetőség, a minőség és a koherencia szempontjából. (Examination of Townscape Image Handbooks and Urban Image Regulations in Terms of Public Accessibility, Quality and Coherence – in Hungarian.) In: Dér, A. – Kovács, Zs. (szerk.): Az építészeti kultúra terjesztésének hazai évtizedei – Magyar Építész Kamara, Budapest, pp. 29–48.

Malhotra, N. K. 2001. Marketingkutatás (Marketing Research – in Hungarian) Akadémiai Kiadó, Bp. ISBN: 963-16-3084-6, 904 p.

Nyelvtudományi Kutatóközpont, Nyelvtechnológiai Kutatócsoport 2014. MNSZ gyakorisági listák (Hungarian National Corpus: Frequency Lists – in Hungarian). Download: http://corpus.nytud.hu/cgi-bin/mnszgyak

Piskóti I. 2023. Településmarketing stratégiai kézikönyv (Urban Marketing Strategic Handbook – in Hungarian) – Economic Geography and Urban Marketing Centre, John von Neumann University, Geo Publications. Download: https://geo.nje.hu/storage/content/4605/files/01-Piskoti.pdf?r=1b73066bbae14beae45bdf6975f6d8fa

Schopf, T. – Braun, D. – Matthes, F. 2022. Lbl2Vec: an embedding-based approach for unsupervised document retrieval on predefined topics. arXiv preprint arXiv:2210.06023. https://arxiv.org/abs/2210.0602

Shanahan, M. 2003. Cork Rural Design Guide Building a New House in the Countryside – Colin Buchanan and Partners Ltd. Ireland 128 p. ISBN No. 0 9525 8694 0 Download: http://corkcocoplans.ie/wp-content/uploads/bsk-pdf-manager/2016/07/Rural-Design-Guide-2nd-Edition-2010.pdf

Legal References

(1) Government Decree No. 314/2012 (XI. 8.) on the settlement development concept, the integrated settlement development strategy, settlement planning instruments, and certain specific legal institutions of settlement planning. Note: this decree is no longer in force; it has been replaced by Government Decree No. 419/2021 (VII. 15.).

(2) Act CXLIII of 2015 on Public Procurement.
In force: 1 January 2026 – 31 August 2026.
https://net.jogtar.hu/jogszabaly?docid=a1500143.tv

(3) Act LXXIV of 2016 on the Protection of Townscape.
https://mkogy.jogtar.hu/jogszabaly?docid=A1600074.TV

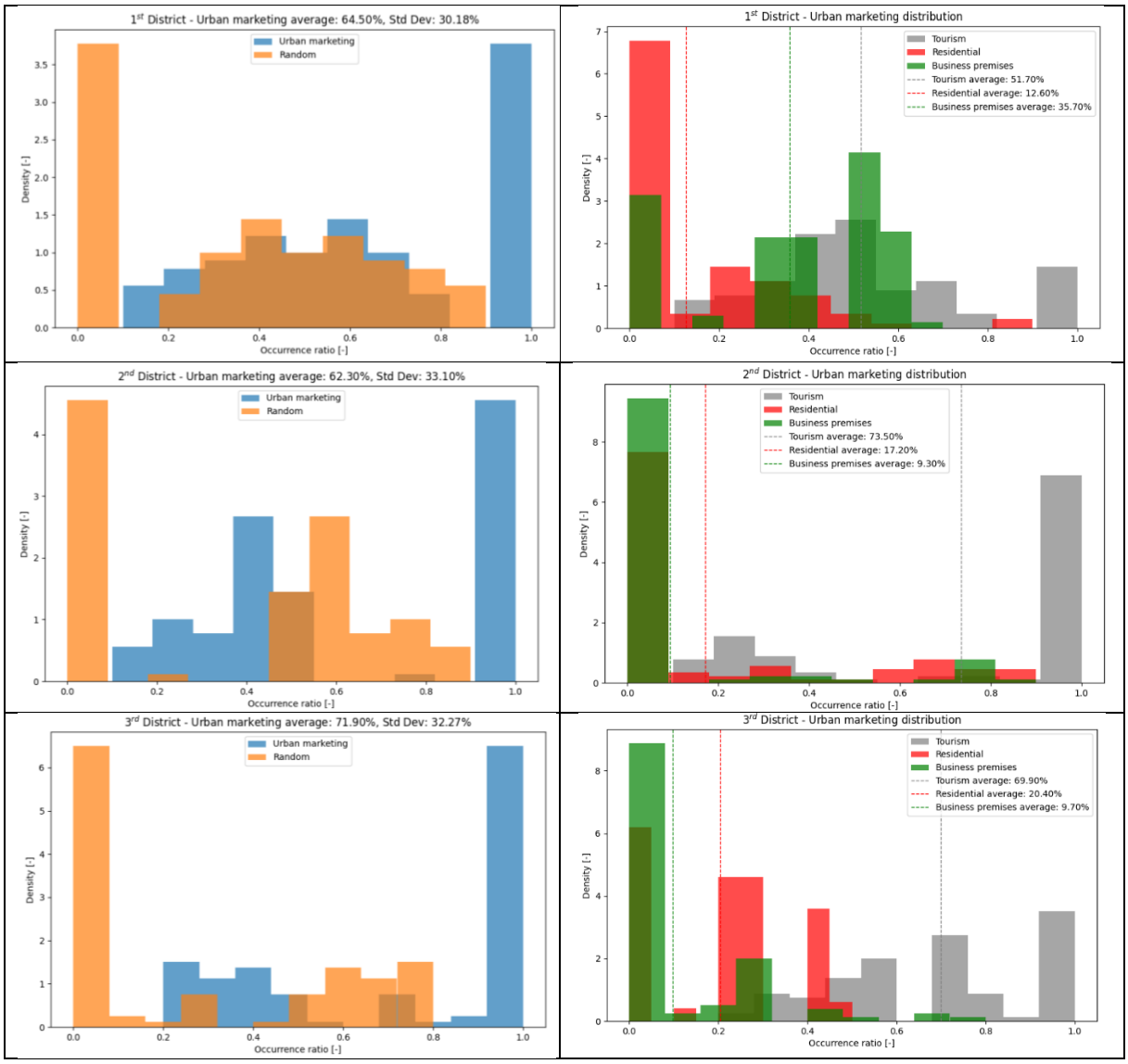
(4) Act XXXIX of 2021 on the Amendment of Certain Acts Related to Settlement Planning.
https://mkogy.jogtar.hu/jogszabaly?docid=A2100039.TV

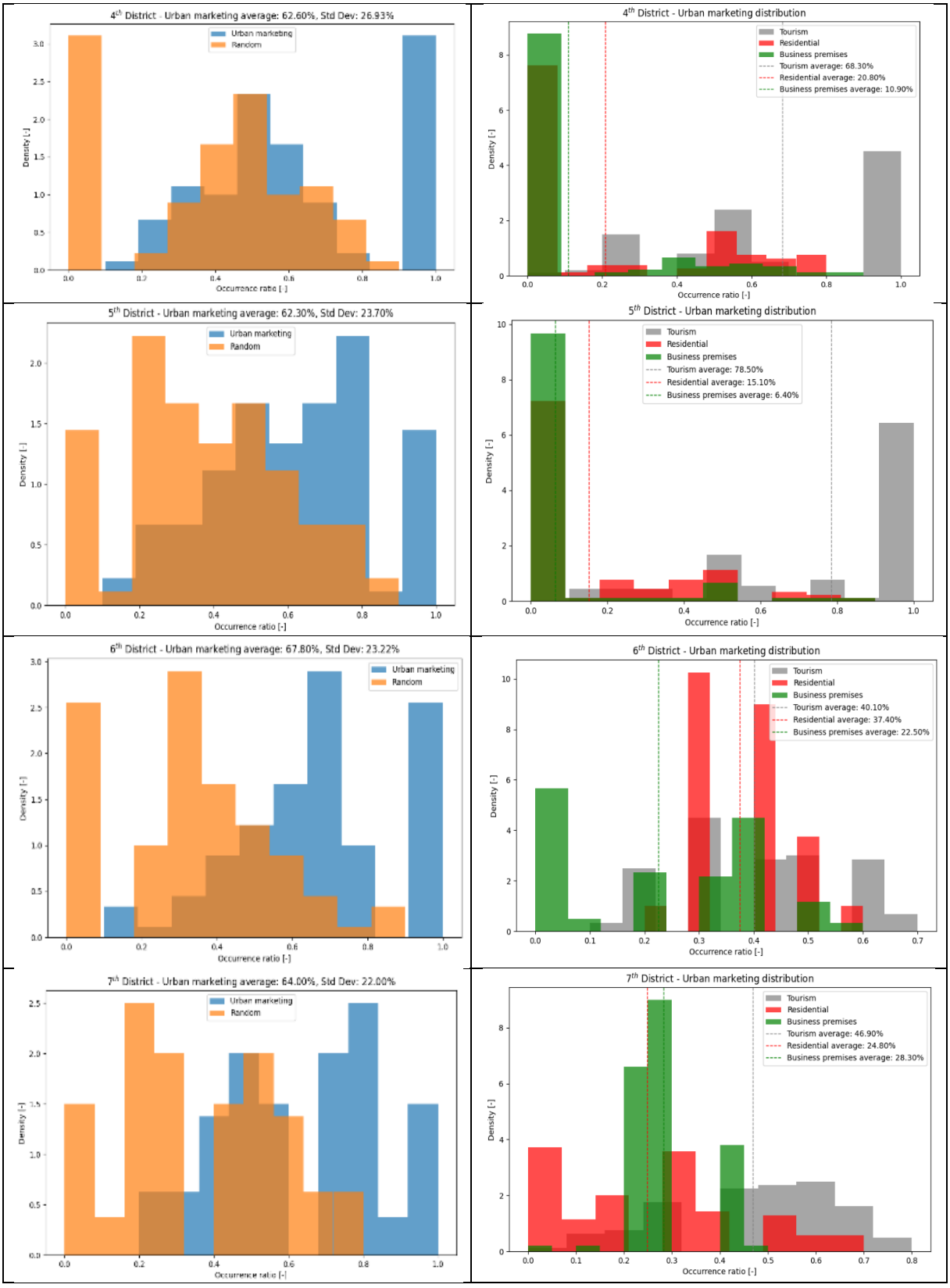
(5) Act C of 2023 on Hungarian Architecture.
In force: 14 May 2026 – 30 June 2027.
https://net.jogtar.hu/jogszabaly?docid=a2300100.tv

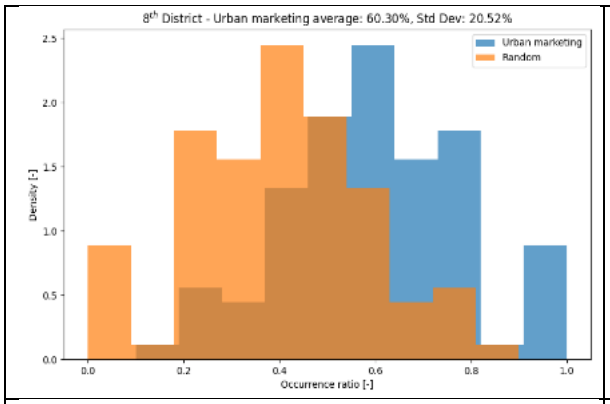
(6) Government Decree No. 419/2021 (VII. 15.) on the Content, Preparation and Adoption Procedure of Settlement Plans, and on Certain Specific Legal Institutions of Settlement Planning.
https://net.jogtar.hu/jogszabaly?docid=a2100419.kor

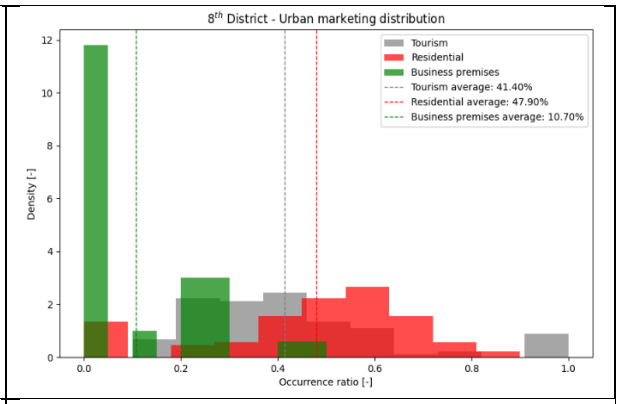
Appendix

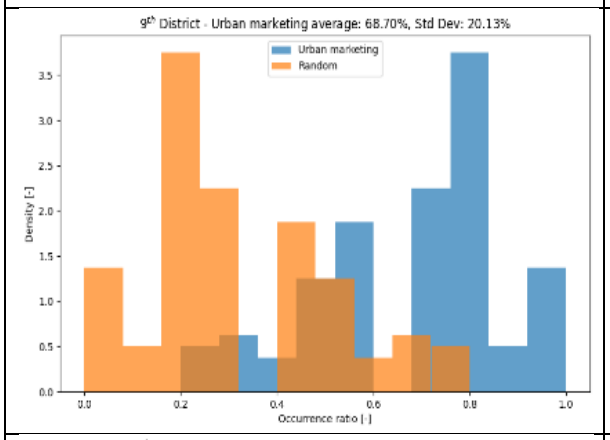
Diagrams showing the urban marketing contents of the Townscape Image Handbooks of the 23 Budapest districts, alongside the tourism-, residential-, business premises-oriented distributions of their marketing elements.

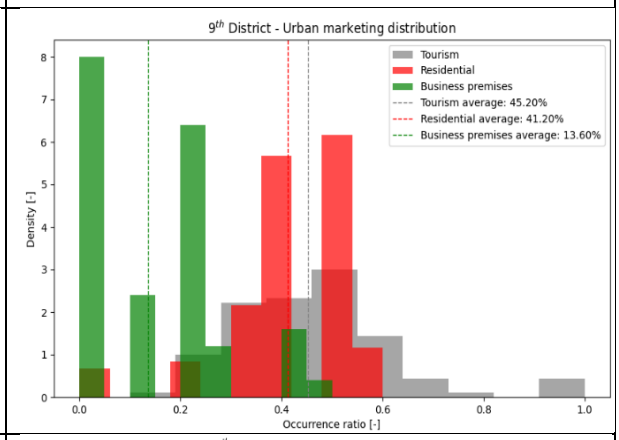


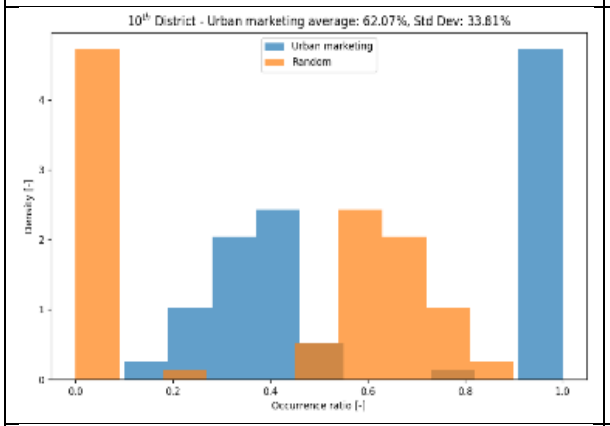

4th District - Urban marketing average: 62.60%, Std Dev: 26.93%
Urban marketing
Random
Density [-]
Occurrence ratio [-]
4th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 68.30%
Residential average: 20.80%
Business premises average: 10.90%
5th District - Urban marketing average: 62.30%, Std Dev: 23.70%
5th District - Urban marketing distribution
Tourism average: 78.50%
Residential average: 15.10%
Business premises average: 6.40%
6th District - Urban marketing average: 67.80%, Std Dev: 23.22%
6th District - Urban marketing distribution
Tourism average: 40.10%
Residential average: 37.40%
Business premises average: 22.50%
7th District - Urban marketing average: 64.00%, Std Dev: 22.00%
7th District - Urban marketing distribution
Tourism average: 46.90%
Residential average: 24.80%
Business premises average: 28.30%

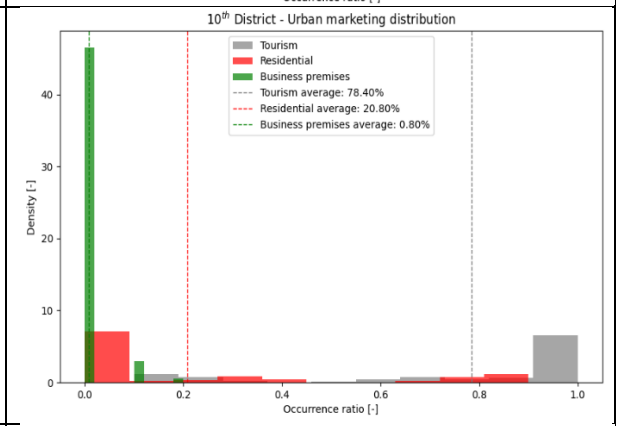

8th District - Urban marketing average: 60.30%, Std Dev: 20.52%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

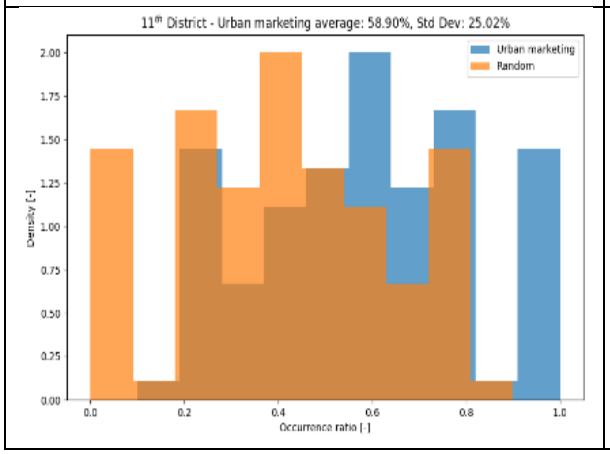

8th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 41.40%
Residential average: 47.90%
Business premises average: 10.70%
Density [-]
Occurrence ratio [-]

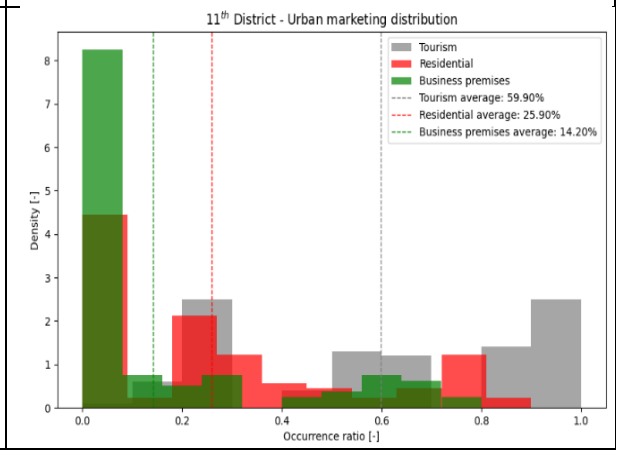

9th District - Urban marketing average: 68.70%, Std Dev: 20.13%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

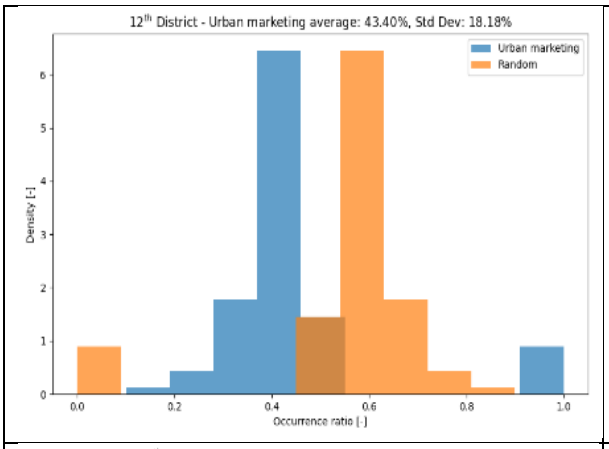

9th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 45.20%
Residential average: 41.20%
Business premises average: 13.60%
Density [-]
Occurrence ratio [-]

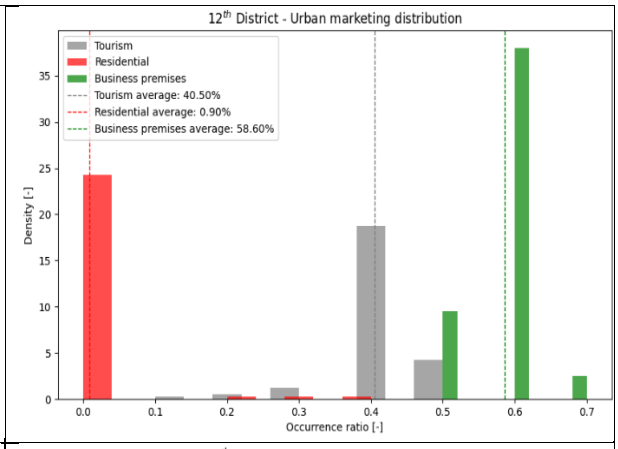

10th District - Urban marketing average: 62.07%, Std Dev: 33.81%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

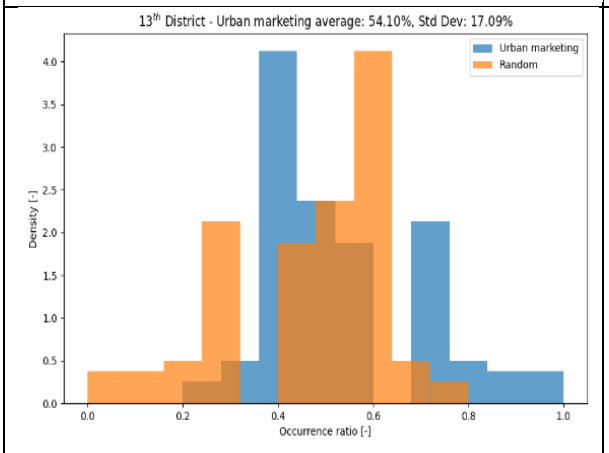

10th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 78.40%
Residential average: 20.80%
Business premises average: 0.80%
Density [-]
Occurrence ratio [-]

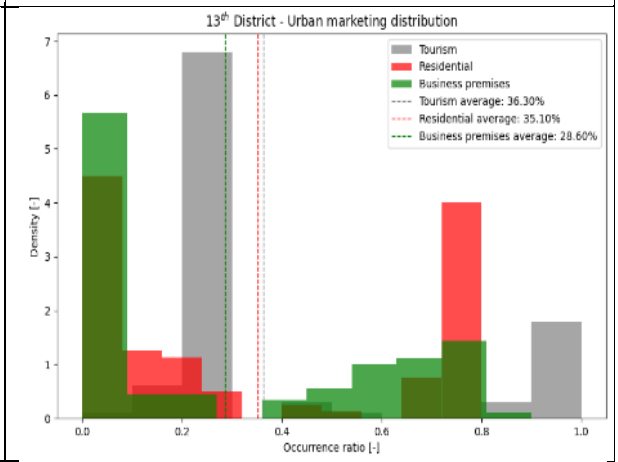

11th District - Urban marketing average: 58.90%, Std Dev: 25.02%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

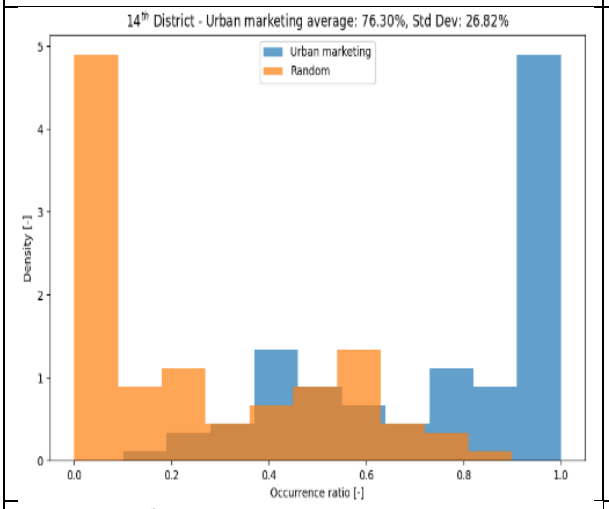

11th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 59.90%
Residential average: 25.90%
Business premises average: 14.20%
Density [-]
Occurrence ratio [-]

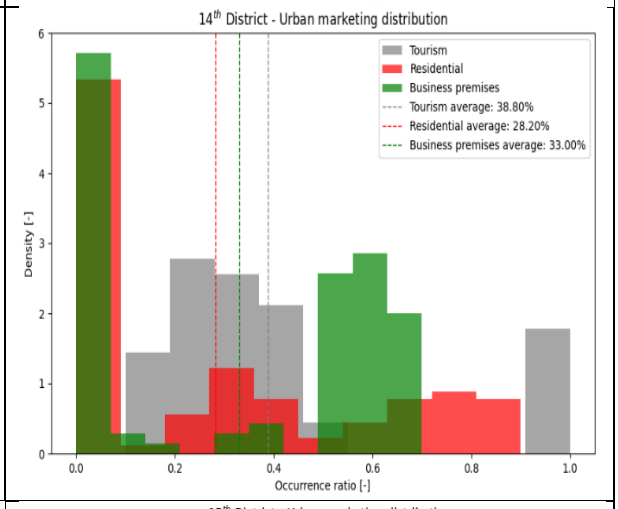

12th District - Urban marketing average: 43.40%, Std Dev: 18.18%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

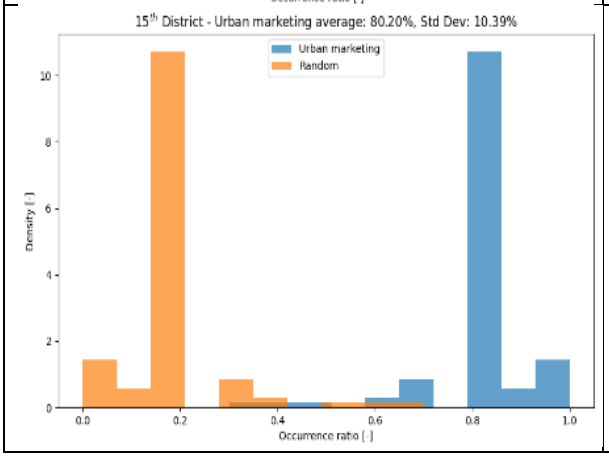

12th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 40.50%
Residential average: 0.90%
Business premises average: 58.60%
Density [-]
Occurrence ratio [-]

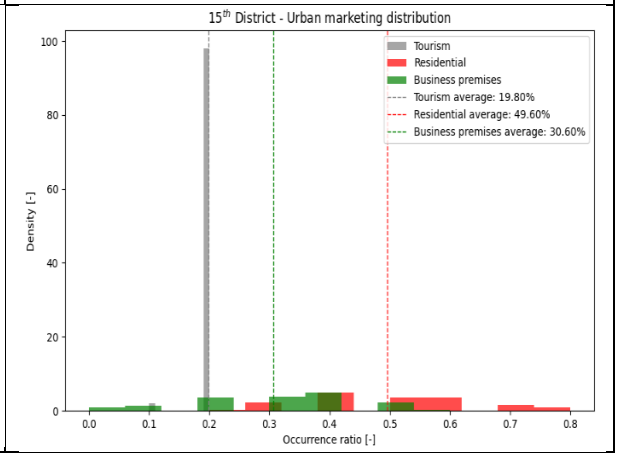

13th District - Urban marketing average: 54.10%, Std Dev: 17.09%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

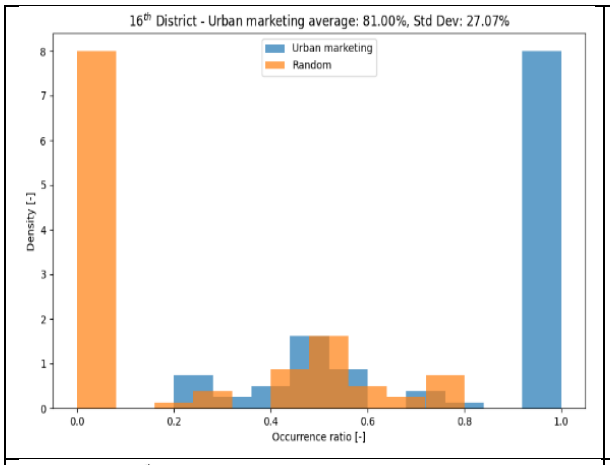

13th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 36.30%
Residential average: 35.10%
Business premises average: 28.60%
Density [-]
Occurrence ratio [-]

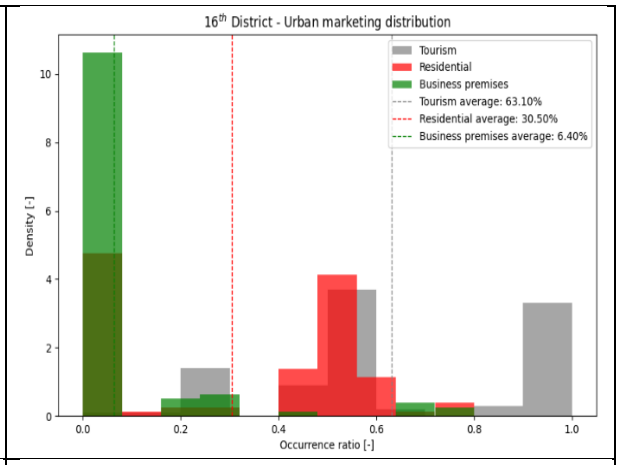

14th District - Urban marketing average: 76.30%, Std Dev: 26.82%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

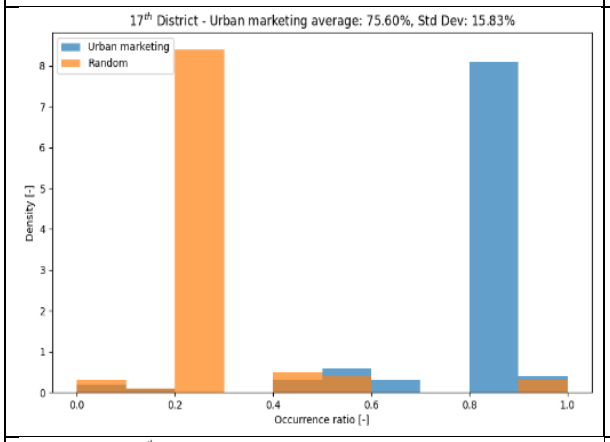

14th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 38.80%
Residential average: 28.20%
Business premises average: 33.00%
Density [-]
Occurrence ratio [-]

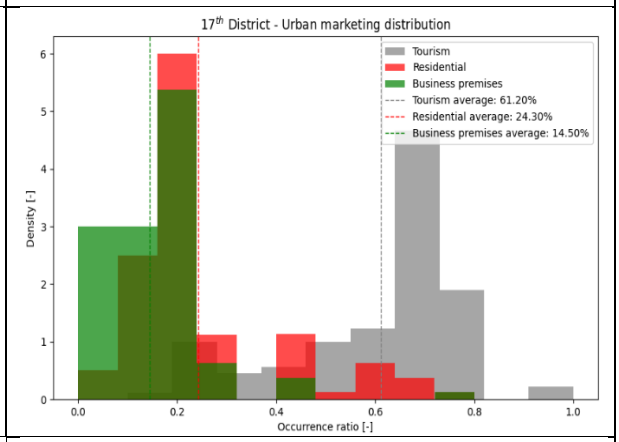

15th District - Urban marketing average: 80.20%, Std Dev: 10.39%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

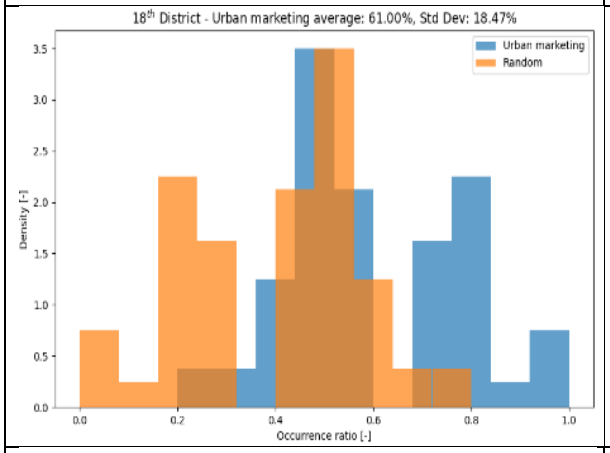

15th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 19.80%
Residential average: 49.60%
Business premises average: 30.60%
Density [-]
Occurrence ratio [-]

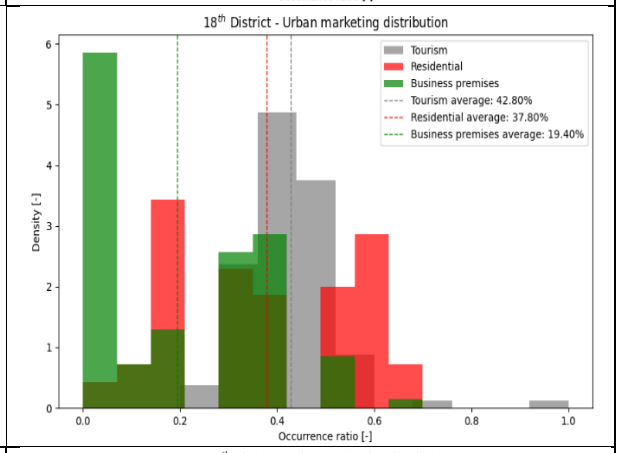

16th District - Urban marketing average: 81.00%, Std Dev: 27.07%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

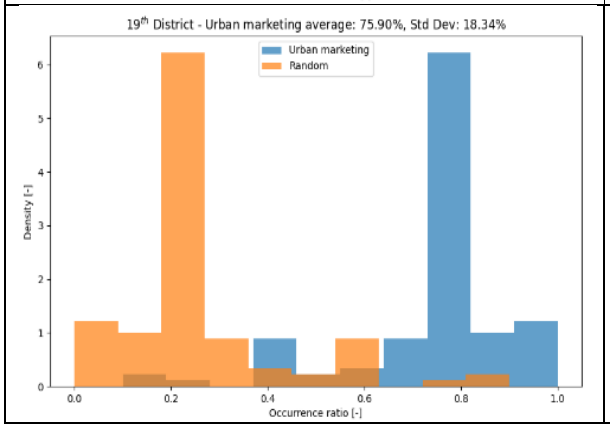

16th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 63.10%
Residential average: 30.50%
Business premises average: 6.40%
Density [-]
Occurrence ratio [-]

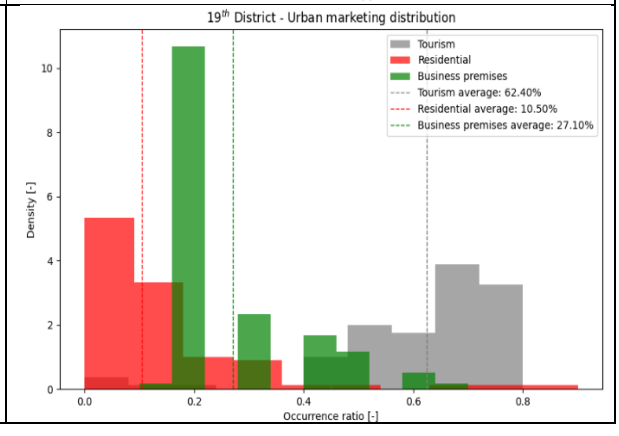

17th District - Urban marketing average: 75.60%, Std Dev: 15.83%
Urban marketing
Random
Density [-]
Occurrence ratio [-]

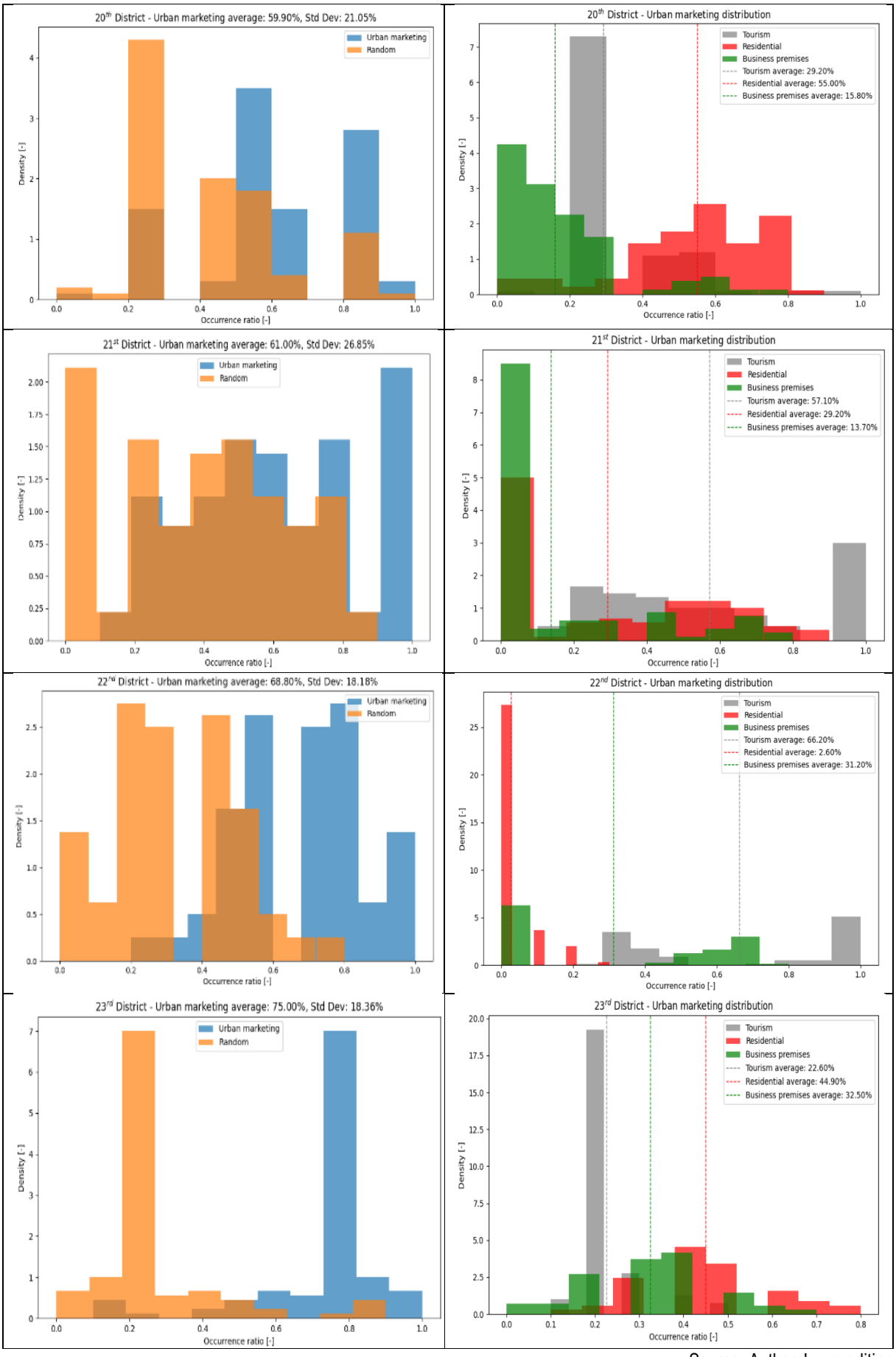

17th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 61.20%
Residential average: 24.30%
Business premises average: 14.50%
Density [-]
Occurrence ratio [-]


18th District - Urban marketing average: 61.00%, Std Dev: 18.47%
Urban marketing
Random
Density [-]
Occurrence ratio [-]


18th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 42.80%
Residential average: 37.80%
Business premises average: 19.40%
Density [-]
Occurrence ratio [-]


19th District - Urban marketing average: 75.90%, Std Dev: 18.34%
Urban marketing
Random
Density [-]
Occurrence ratio [-]


19th District - Urban marketing distribution
Tourism
Residential
Business premises
Tourism average: 62.40%
Residential average: 10.50%
Business premises average: 27.10%
Density [-]
Occurrence ratio [-]



Source: Authors' own editing