EXECUTIVE SUMMARY OF THE THESIS

Methodological design for an Urban Inequality Indicator (UII) based on subjective bias

TESI DI LAUREA MAGISTRALE IN MANAGEMENT ENGINEERING INGEGNERIA GESTIONALE

AUTHOR: ADRIAN CANTERA

ADVISOR: FEDERICO BARTOLOMUCCI

ACADEMIC YEAR: 2025-2026

---

1. Introduction

Traditional spatial data processes often reinforce urban inequalities because they lack a subjective dimension. Despite the rise of smart city data, a gap remains in how urban injustice is measured and visualized. This research addresses the disconnection between objective data and the cognitive biases influencing its interpretation, emphasizing how design choices like color and interactivity impact political decision-making.

To solve this, this master's thesis develops the Urban Inequality Indicator (UII) for Milan's 88 NILs. The methodology integrates citizen surveys to incorporate subjective bias into the indicator's design. A comparative analysis shows that geometric aggregation more accurately identifies critical deprivation than arithmetic methods by preventing compensability between dimensions.

Key findings reveal that urban injustice is a mutable variable across age cohorts rather than a static metric. Through spatial autocorrelation (Local Moran's I), clusters of inequality were identified. Ultimately, the thesis concludes that addressing cognitive bias in data representation supports inclusive, focused governance. The UII is proposed as a tool for municipalities to reshape cities based on real citizens' opinions and as a potential new standard for alleviating inequality.

2. Theoretical Framework & Methodological Design

2.1. Research Questions

This study seeks to answer two core questions:

RQ1: To what extent do perceptions of urban inequality and spatial justice diverge across different demographic age groups?

RQ2: How can dynamic data visualization strategies effectively represent these age-dependent disparities to facilitate inclusive urban policymaking?

3. Literature Review

3.1. Foundations of urban data and statistical modeling

The ultimate goal of extracting and visualizing urban data is to reveal hidden societal dynamics. However, meaningful insights require a deep understanding of mathematical concepts. This technical foundation must be combined with subjective human perception to construct realistic and inclusive composite indicators.

To generate real outcomes, analysis must account for spatial dependencies, specifically spatial correlation (how similar or diverse neighboring zones are) and spatial heterogeneity (the uneven distribution of urban resources). While simple models (Gibbs, 1997) are useful, they risk hiding marginalized voices. Making a mixed-method approach came as essential. Looking at the technical metrics:

- **Global indicators** (Moran's I): These measure overarching patterns to determine if clustering or dispersion exists across a whole society. However, they cannot pinpoint specific locations.

- **Local indicators** (LISA): These identify localized heterogeneity, revealing specific neighborhood "hotspots" (high-high) and "coldspots" (low-low).

Shifting from a global to a local perspective is crucial for urban planning. Local indicators allow city councils to target resources exactly where needed (like, affordable housing in low-income hotspots). Furthermore, this approach helps capture varying perceptions of inequality across age groups, such as younger populations (18-30) prioritizing digital access versus older populations (60+) focusing on healthcare.

3.1.1. Why Data Extraction Matters for Cities

In the current neoliberal "smart city" paradigm, citizens are often reduced to mere consumers. To counter this, spatial data points must be understood within their social context to reveal uneven opportunities.

While data extraction drives predictive analytics (such as modeling critical infrastructure decay), raw data is not neutral. Transforming data into insights requires a human cognitive layer. Recognizing this social dimension, such as how environmental benefits correlate with income, is far more important than the extraction itself to ensure equitable urban policymaking.

3.1.1. Statistical Frameworks: Composite Indicators, Normalization, and Aggregation

Composite Indicators (CIs) aggregate multidimensional data to improve interpretability. Building them requires following strict steps (like the OECD framework), with a critical focus on data normalization (for comparability) and aggregation logic. There are two options for aggregating:

- **Additive methods** (Arithmetic Mean): These are compensatory. A high value in one category can offset a low value in another, giving a false sense of balance.

- **Multiplicative methods (**Geometric Mean): These are non-compensatory. A low value (representing severe deprivation) is not canceled out by a high value elsewhere. This makes it the superior choice for analyzing urban social data without falsifying reality.

Finally, rather than using purely mathematical weighting like Principal Component Analysis (PCA), this study derives its weights from subjective surveys of Milan's residents. This ensures the resulting indicator reflects actual lived experiences rather than just statistical abstractions.

3.2. Smart cities, data ethics, and inequality

The rise of "smart cities" relies heavily on Information and Communication Technologies (ICTs) and sensor networks to optimize urban resources and establish evidence-based "smart governance" (Nastjuk et al., 2022). However, this technocratic approach often reduces citizens to passive "consumer users" rather than active participants in civic planning ((Kitchin et al., 2018). By flattening complex urban dynamics into mere data points on digital platforms (Bibri, 2022), city planners risk misclassifying neighborhoods and ignoring the social realities of their inhabitants, making rigorous technical and ethical oversight essential (Lim & Taeihagh, 2019).

Relying strictly on these automated data platforms introduces significant ethical risks, primarily because different demographic groups experience and perceive urban inequalities in fundamentally different ways (Beaunoyer et al., 2020). For example, while younger populations might view digital infrastructure as a critical urban need, older adults typically prioritize healthcare access and walkability (Schuster-Olbrich et al., 2024). Because human bias ultimately dictates what data is measured and displayed, failing to account for these diverging priorities can lead to policies that unintentionally disadvantage vulnerable groups while overrepresenting others (Wang et al., 2025).

Furthermore, the ubiquitous surveillance required by smart cities raises profound privacy concerns (Ziosi et al., 2024) and frequently exacerbates algorithmic bias. While algorithms themselves are not inherently malicious, they are trained on historical data that reflects past societal prejudices, leading to disparate treatment of marginalized communities. An example is predictive policing, where historical data causes algorithms to falsely correlate certain neighborhoods with inherent criminality, resulting in disproportionate police presence rather than the much-needed reallocation of social resources (Madison Cutler, 2024). To prevent smart city tools from exclusively serving wealthier classes, urban data analysis must actively prioritize equity and continuously question the neutrality of its outcomes.

3.3. Mapping injustice

To effectively drive inclusive urban policy, data representations must make the distinct spatial realities of different demographic groups visible. Dynamic spatial tools, particularly Geographical Information Systems (GIS), allow researchers to overlay multiple dimensions of inequality, such as highlighting gaps in digital connectivity for younger populations or mapping healthcare and mobility accessibility for seniors. However, mapping is not a neutral technical exercise, relying on non-uniform or overly generic data risks underrepresenting vulnerable communities or stigmatizing specific neighborhoods. For example, spatial studies have revealed how uneven urban growth and a lack of access to public green spaces left African American communities disproportionately vulnerable during the COVID-19 pandemic (Pallathadka et al., 2022) . Therefore, spatial representations must be grounded in robust, ethically considered datasets to accurately reflect socio-economic disparities without introducing unwanted interpretative bias.

Communicating these spatial inequalities effectively requires tailoring visual outputs to the diverse literacy levels and needs of various stakeholders (Jeremy W. Crampton, 2010) . While policymakers often seek high-level, multi-factor analyses, everyday citizens generally prioritize clear, actionable solutions to localized problems. Among the various mapping typologies, including point, heat, grid, and choropleth maps, the ultimate goal must always be maximum interpretability. Interestingly, literature indicates that highly interactive web tools do not necessarily improve understanding. In fact, for certain demographics like the elderly, the cognitive burden of navigating complex interactive platforms can actually create a barrier to comprehension. Consequently, well-designed static representations frequently perform just as well, proving that clarity and the democratization of information should always take precedence over technological attractiveness.

The ultimate effectiveness of data visualization heavily depends on deliberate design choices, such as color palettes, dimensionality, and data granularity, which inherently carry profound social impacts. Misguided design choices, like using a progressive color gradient to represent different racial groups or employing "rainbow scales" for socioeconomic data, can confuse users, create false statistical boundaries, and unintentionally marginalize communities. Best practices dictate that darker colors and higher opacities should intuitively represent larger quantities to align with human cognitive biases, and standard 2D models are consistently more comprehensible than complex 3D representations (Brandie M Stewart; Jessica Cipolla, 2009). Ultimately, achieving the appropriate data granularity and employing empathetic, transparent visual design are essential to ensure that urban mapping serves as a reliable tool for social equity rather than a source of misinterpretation.

3.4 Answering the research questions

The perception of urban inequality is highly subjective and varies significantly across different socio-demographic realities. For instance, studies examining political, racial, and economic backgrounds reveal that communities with lower per capita income systematically report lower satisfaction levels regarding urban opportunities (Smith et al., 2023). However, age emerges as a particularly critical factor in shaping how urban injustices are experienced.

While younger residents typically evaluate a city based on affordable rent, socio-cultural offerings, and digital connectivity, elderly populations prioritize practical everyday needs like street walkability and proximity to healthcare services. Because policies that solely boost digital infrastructure might completely bypass the needs of older citizens, building a truly inclusive city requires a multifactorial analysis that integrates these generational differences to ensure fairer municipal decision-making (Li et al., 2025).

To translate these age-dependent disparities into actionable urban policymaking, cities require broad, tailored composite indicators that integrate subjective age bias with empirical data. Such an indicator can accurately pinpoint whether a specific neighborhood requires increased investment in broadband networks for the youth or enhanced medical accessibility for seniors. However, possessing this accurate metric is only half the challenge; the data must be presented through accessible platforms, such as interactive maps and control panels, that promote clear understanding among all stakeholders. Ensuring these platforms are effective requires rigorous evaluation through qualitative surveys and audited tests to gauge user comprehension and ensure the tools match the diverse literacy levels of the public.

Ultimately, the design of these data visualization strategies must carefully balance analytical complexity with visual accessibility to avoid alienating specific groups. Older adults often face technological barriers and demonstrate reluctance toward complex software infrastructures, dictating that graphical interfaces must remain highly intuitive. In fact, research demonstrates that well-designed static representations frequently outperform complex interactive tools for older users, whereas younger demographics prefer and benefit from high interactivity (Xexakis & Trutnevyte, 2019). Therefore, inclusive urban policymaking relies on finding a vital design midpoint, employing inclusive communication techniques that successfully present complex, multifactorial variables without overwhelming the observer. By focusing these visual representations directly on the distinct needs of citizens, municipalities can successfully leverage technology to unblock urban realities and drive equitable development.

4. Methodological Design of the UII

A mixed-method approach was employed. The UII is constructed upon five dimensions of inequality for Milan:

Economical, Educational, Sociodemographical, Health, and Living Environment.

Each dimension comprises four sub-categories rental costs, proximity to universities, pollution levels, green space index…). Data underwent min-max normalization to preserve value distances and was aggregated using a geometric mean to avoid masking low performance. Subjective weights were then derived from a survey administered to Milan's inhabitants, categorized by age (18-30, 30-60, 60+).

Data was gathered at the granular level of the 88 NILs encompassing the central city of Milan, utilizing open datasets from sources like ISTAT and the Comune di Milano. To translate these calculations into visual insights, an interactive web mapping platform was developed using Python libraries (Geopandas, Folium) and GIS tools. The platform outputs choropleth maps apply tailored colorimetry and scaling to highlight stress zones across the different sub-indicators and the final composite index.

The survey results proved that "urban justice" is a perceived reality that shifts according to the user's life stage. The baseline weights assigned to dimensions changed drastically across age cohorts. The economic category (particularly rent prices) was highly critical for younger groups (18-30 and 30-60), while the 60+ cohort distributed importance more heavily toward health and socio-demographic indices.

5. Results

To accurately represent the distribution of urban resources in Milan, this study first tested the mathematical weighting of the Urban Inequality Indicator (UII) by comparing compensatory (arithmetic mean) and non-compensatory (geometric mean) aggregation methods. While an arithmetic approach, similar to the HDI, allows high scores in one dimension to obscure severe deficiencies in another. The geometric mean effectively highlights localized disparities, such as those found in the Padova-Turro-Crescenzago or Porta Magenta zones. Based purely on objective data, Milan exhibits high overall levels of inequity, though specific spatial analyses using global and local indicators (Moran's I and Geary's C) confirm that these inequalities are clustered rather than distribute randomly. Notably, this unbiased weighted baseline identifies in the city center, specifically a "Low-Low" cold spot around Brera, as an area of relatively low inequality due to abundant leisure and public transit options.

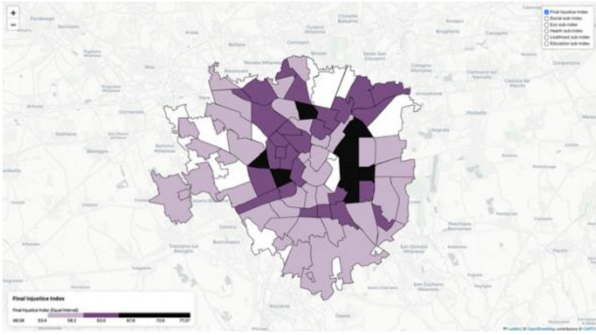

Figure 1: General representation of the final outcome representation

However, when subjective, age-based survey data is integrated into the model, this supposedly objective urban reality fractures into diverging demographic experiences. The objective spatial clusters vanish, proving that human perception of injustice is not strictly bound by geographic proximity, but by distinct life stages. For citizens under the age of sixty (both the 18-30 and 30-60 cohorts), the city center transforms from a privileged "cold spot" into a massive epicenter of urban inequality. This dramatic shift is primarily driven by the crippling cost of living, gentrification, and high rent prices, which completely overshadow the objective benefits of central infrastructure. Younger demographics also heavily penalize zones lacking digital connectivity, emphasizing that what a machine algorithm calculates as a well-resourced zone may actually feel deeply unjust and inaccessible to youth.

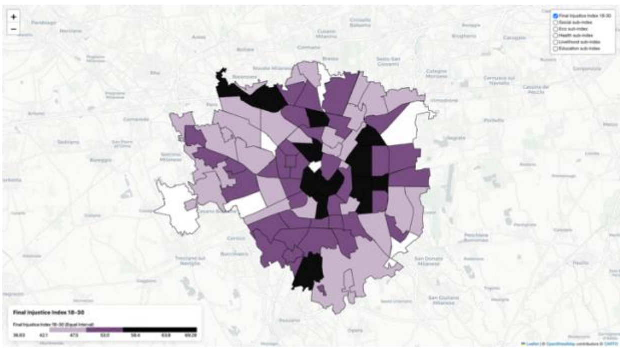

Figure 2: Representation of the Milan's urban reality by the 18-30 age gap

Conversely, the 60+ demographic presents a fundamentally different map of injustice, characterized by a much lower concentration of perceived inequality in the city center. This senior cohort naturally deprioritizes rental costs, focusing instead on population density and the proximity of healthcare infrastructure.

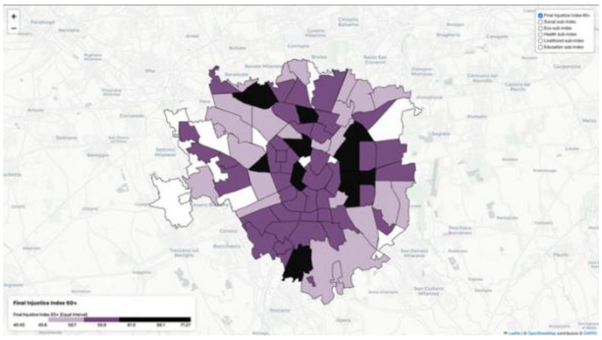

Figure 3: Representation of the Milan's urban reality by the 60+ age gap

These clear differences underscore the danger of relying solely on objective, unbiased weighted data for municipal governance. If urban planners aggregate divergent demographic needs into a single, mathematically uniform calculation, they actively mask the specific socioeconomic burdens felt by different groups. To truly enhance citizens' lives, policymakers must abandon the reductionist view of residents as generic data points and instead use subjective, age-differenciated indicators to direct political willingness and resources exactly where each demographic needs them most.

6. Conclusions

This thesis successfully breaks from traditional, socially blind paradigms of urban data analysis by proving that spatial measurements in cities like Milan are deeply influenced by demographic biases. A core achievement of this work is demonstrating that a "one-size-fits-all" measurement for citizens does not exist. To address this, the study introduces the novel Urban Inequality Indicator (UII), which achieves mathematical balance. By applying a subjective weighting system derived from age-specific demographic targets, the UII bridges raw data with sociological reality to answer how people actually interpret urban space (RQ1). This data is then translated into actionable insights for policymakers through an interactive, Python- and GIS-based web platform. By generating age-tailored choropleth maps across Milan's 88 NILs, the platform prioritizes ethical data treatment and interpretability to effectively communicate complex spatial injustices to stakeholders (RQ2).

Despite the robust framework provided by the UII, the study encountered specific limitations that contextualize its findings. First, gathering a sufficiently large and representative sample size for the subjective questionnaire proved challenging, particularly within the senior demographic, although both online and in-person methods were utilized to secure truthful responses. Second, the availability of updated, granular open data at the specific NIL level was highly constrained. For instance, critical dimensions like neighborhood-level internet connectivity, a primary concern for the under-thirty cohort, could not be included due to a lack of freely accessible municipal data. Finally, while the visualization platform's design choices were heavily grounded in established academic literature, its actual impact on interpretability was not empirically tested with real policymakers in a controlled experimental environment, leaving that practical validation outside the current scope.

6.1. Future Research

Looking forward, the highly scalable nature of the UII methodology opens clear, actionable avenues for future research. The immediate next step should involve applying this cross-demographic spatial justice framework to other major European cities experiencing similar urban pressures, such as Barcelona, Paris, or Amsterdam. Additionally, to move beyond static historical datasets, future iterations must integrate real-time data feeds via APIs from urban IoT devices, coupled with recurrent citizen surveys to dynamically adjust demographic weights over time. Lastly, future studies should track more diverse dimensions of the surveyed individuals. Specifically, cross-referencing citizens' age groups with their exact residential locations will provide a much deeper, hyper-localized understanding of how geographic placement directly shapes an individual's subjective perception of their urban ecosystem.

7. Bibliography

Beaunoyer, E., Dupéré, S., & Guitton, M. J. (2020). COVID-19 and digital inequalities: Reciprocal impacts and mitigation strategies. Computers in Human Behavior, 111, 106424. https://doi.org/10.1016/J.CHB.2020.106424

Bibri, S. E. (2022). The Social Shaping of the Metaverse as an Alternative to the Imaginaries of Data-Driven Smart Cities: A Study in Science, Technology, and Society. https://doi.org/10.3390/smartcities5030043

Brandie M Stewart; Jessica Cipolla. (2009). (PDF) Extraneous information and graph comprehension: Implications for effective design choices. https://www.researchgate.net/publication/235295298_Extraneous_information_and_graph_comprehension_Implications_for_effective_design_choices

Gibbs, P. E. (1997). What is Occam's Razor? https://www.researchgate.net/publication/330171618

Jeremy W. Crampton. (2010). (PDF) An Introduction to Critical Cartography. https://www.researchgate.net/publication/241435510_An_Introduction_to_Critical_Cartography

Kitchin, R., Cardullo, P., & Di Feliciantonio, C. (2018). Citizenship, Justice and the Right to the Smart City. https://doi.org/10.31235/osf.io/b8aq5

Li, J., Li, X., & Zhou, X. (2025). FAP-CD: Fairness-Driven Age-Friendly Community Planning via Conditional Diffusion Generation. Proceedings of the AAAI Conference on Artificial Intelligence, 39(27), 28168–28176. https://doi.org/10.1609/AAAI.V39I27.35036

Lim, H. S. M., & Taeihagh, A. (2019). Algorithmic decision-making in AVs: Understanding ethical and technical concerns for smart cities. Sustainability (Switzerland), 11(20). https://doi.org/10.3390/su11205791

Madison Cutler. (2024). DEVELOPMENT OF PREDICTIVE POLICING TECHNOLOGY AND JUSTIFICATION FOR USE.

Nastjuk, I., Trang, S., & Papageorgiou, E. I. (2022). Smart cities and smart governance models for future cities: Current research and future directions. In Electronic Markets (Vol. 32, Number 4, pp. 1917–1924). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s12525-022-00609-0

Pallathadka, A., Pallathadka, L., Rao, S., Chang, H., & Van Dommelen, D. (2022). Using GIS-based spatial analysis to determine urban greenspace accessibility for different racial groups in the backdrop of COVID-19: a case study of four US cities. GeoJournal, 87(6), 4879–4899. https://doi.org/10.1007/s10708-021-10538-8

Schuster-Olbrich, J. P., Marquet, O., Miralles-Guasch, C., & Fuentes Arce, L. (2024). Spatial patterns and drivers of urban expansion: An exploratory spatial analysis of the Metropolitan Region of Santiago, Chile, from 1997 to 2013. Cities, 153. https://doi.org/10.1016/j.cities.2024.105305

Smith, R. M., Deb, D., Blizard, Z., & Midgett, R. (2023). A Planner's quest for identifying spatial (in)justice in local communities: A case study of urban census tracts in North Carolina, USA. Applied Geography, 158. https://doi.org/10.1016/j.apgeog.2023.103030

Wang, Z., Cheng, Q., Gu, Z., Liu, C., Cun, D., Shi, X., Wu, X., Zhou, Z., He, X., Vlacic, L., & Liu, Z. (2025). Analyzing Mega-mobility Systems in Smart Cities: A Macro–Micro Integration with Feedback Paradigm Empowered by Artificial Intelligence. Research, 8. https://doi.org/10.34133/RESEARCH.0982

Xexakis, G., & Trutnevyte, E. (2019). Are interactive web-tools for environmental scenario visualization worth the effort? An experimental study on the Swiss electricity supply scenarios 2035. Environmental Modelling and Software, 119, 124–134. https://doi.org/10.1016/j.envsoft.2019.05.014

Ziosi, M., Hewitt, B., Juneja, P., Taddeo, M., & Floridi, L. (2024). Smart cities: reviewing the debate about their ethical implications. AI and Society, 39(3), 1185–1200. https://doi.org/10.1007/s00146-022-01558-0