JBIAM : Journal Business Inovation, Accounting and Management
Vol 1 No 1 April 2025
ISSN: xxxxxx (Print) ISSN: xxxxxx (Electronic)
Open Access: https://journal.almindo.org/jbiam/

Hyperlocal Business Patterns in Indonesia: A Hashtag-Based Machine Learning Study of Top Metropolitan Areas

Asep Koswara[1]*
[1]IKOPIN University, Indonesia
[1]*aspkosw@gmail.com**

Article History
Received : 7 April 2025
Revised : 16 April 2025
Accepted : 23 April 2025
Published : 30 April 2025

Keywords :
Hyperlocal Business; Hashtag Analysis; Urban Digital Economy; Machine Learning; Urban Business Identity

Abstract

This study explores how hyperlocal businesses in Indonesia's metropolitan areas construct their digital identities through hashtags on social media, specifically Instagram. The research aims to identify spatial and sectoral variations of digitally expressed business activity by analyzing the top 30 hashtags from five major cities: Jakarta, Bandung, Surabaya, Yogyakarta, and Medan. Data was collected via Tagsfinder.com and classified using a supervised machine learning model based on the frequency-inverse document frequency (TF-IDF) to distinguish business-related hashtags from non-business ones. The results show that out of 150 total hashtags, 42.7% were classified as business-related, with Bandung showing the highest proportion (56.6%) and Yogyakarta the lowest (20%). Sectorally, Bandung stands out in fashion and apparel, while Medan shows strong presence in the jewellery sector (e.g., #cincinnikah, #cincinmedan) and Surabaya with culinary sector. These findings affirm previous literature on hashtags as proxies for economic activity and address a significant research gap in the Global South context. In conclusion, social media hashtags serve not only as tools for promotion but also as spatial and cultural markers of urban business identity, providing new insights for digitally mapping hyperlocal urban economies in Indonesia.

INTRODUCTION

In recent years, the rise of hyperlocal businesses has become a salient feature of urban economies worldwide, particularly in rapidly urbanizing regions such as Southeast Asia. Hyperlocal businesses—defined as small to medium enterprises that focus on a narrowly defined geographic area and cater to the specific needs of that locality (Zeng et al., 2021)—are increasingly leveraging social media platforms to reach, engage, and serve their target communities. In Indonesia, the proliferation of social media usage, especially in metropolitan areas like Jakarta, Surabaya, Bandung, Medan, and Yogyakarta, has created fertile ground for hyperlocal enterprises to thrive.

Indonesia is the fourth most populous country in the world and among the top five globally in terms of social media usage. According to DataReportal (2024), Indonesia had over 139 million active Instagram users and 24 million Twitter users, with the majority of users concentrated in major urban centers. The digital behavior of Indonesian users shows high engagement with local trends, products, and services, often mediated by hashtags. Hashtags function as community-driven tagging mechanisms that allow businesses to contextualize and promote their offerings within relevant cultural and spatial frameworks (Tsur & Rappoport, 2012). This makes hashtags a valuable, yet underutilized, resource for understanding local economic activity in real time.

Despite the growing importance of hyperlocal businesses in shaping urban economies and digital landscapes, there is limited empirical research that systematically analyzes these businesses' spatial and thematic patterns in Indonesia, especially using data-driven methods. Traditional research has focused predominantly on macro-level e-commerce trends (Ginting & Sasmoko, 2020) or SME digital adoption in general (Tambunan, 2019), often neglecting the hyperlocal layer that functions below the national or even provincial scale. Meanwhile, a

`

growing body of global research has begun to explore the use of social media analytics, particularly hashtag analysis, to track consumer behavior, local trends, and micro-enterprise ecosystems (Zappavigna, 2015; Sloan & Morgan, 2015). However, these studies have rarely intersected with urban Southeast Asian contexts, leaving a significant geographic and methodological gap.

Machine learning techniques offer an innovative avenue for bridging this gap. By training models to detect patterns in hashtag usage, we can gain insights into how hyperlocal business ecosystems evolve across different cities, how digital visibility varies by product or service type, and how local cultures influence online economic expressions. For example, clustering algorithms and natural language processing (NLP) models have been successfully employed to map urban services in European cities through Instagram data (González et al., 2020). Yet, there is a paucity of similar work in Indonesia, where the linguistic diversity, urban sprawl, and social media dynamism offer both opportunities and challenges for such methods.

This study aims to address this gap by analyzing the hashtag usage patterns related to hyperlocal businesses in five major Indonesian cities—Jakarta, Bandung, Surabaya, Yogyakarta, and Medan—using machine learning techniques. These cities were selected due to their demographic size, digital activity, and varied cultural-economic contexts. Jakarta, for example, is a global megacity with dense social media activity, while Bandung and Yogyakarta are known for their strong creative economies. Medan and Surabaya, meanwhile, offer insights into hyperlocal economies outside Java, thereby ensuring geographic diversity.

Preliminary data for this study were collected using TagsFinder.com, a publicly accessible tool that aggregates and ranks hashtags based on frequency, popularity, and contextual relevance in different geographic areas. For this research, the top 30 hashtags were retrieved for five key Indonesian metropolitan areas: Jakarta, Bandung, Surabaya, Yogyakarta, and Medan. These cities were selected based on their population size, digital activity, and regional diversity.

The hashtag data were then subjected to a manual classification process, in which each tag was categorized according to whether it was related to business activity. This classification included tags referencing products (e.g., #kaos, #hijab), services (e.g., #onlineshop, #muasurabaya), industries (e.g., #fashion, #kuliner), and local promotions (e.g., #olshopmedan, #infobandung). Tags not directly associated with business contexts—such as #love, #like4like, or #photooftheday—were excluded from the business category. The filtered dataset of business-related hashtags formed the basis for subsequent machine learning classification and pattern analysis, helping to map the digital business landscape across urban Indonesia.

From a theoretical standpoint, this study builds upon the "spatial turn" in digital geography and urban informatics, which assert that the digital representation of urban spaces—such as through social media—is not merely a mirror of offline life but an active space of economic and cultural production (Graham et al., 2013). Hashtag-based mapping, therefore, is more than a descriptive tool; it allows for the theorization of digital locality and platform-mediated place-making. It also extends the notion of “informational cities” (Castells, 2010) by examining how information flows about local business activities are shaped, reinforced, and potentially biased by social media engagement patterns.

The novelty of this research lies in its integration of social media data mining, machine learning, and urban economic analysis in the context of Indonesian hyperlocal businesses. Unlike previous studies which either remain descriptive or lack spatial specificity, this study proposes a computational framework for identifying and classifying hyperlocal business

`

patterns based on real-time user-generated content. Additionally, by examining multiple cities, this work contributes to a comparative urban digital geography of Indonesia, shedding light on both shared and unique characteristics of local digital economies.

In summary, this research seeks to: (1) Identify and classify hashtags associated with hyperlocal businesses in Indonesia's top five metropolitan areas. (2) Analyze spatial and semantic patterns in hashtag usage using machine learning models. (3) Theorize the role of digital space in shaping urban hyperlocal economies.

By doing so, this study hopes to advance scholarly understanding of the intersection between urban development, digital platforms, and small-scale business ecosystems in a Global South context. It also aims to provide practical insights for local entrepreneurs, digital marketers, and policymakers seeking to strengthen digital economic inclusivity at the city level.

LITERATUR REVIEW

The intersection of hyperlocal business dynamics, digital platforms, and urban analytics has attracted growing academic interest in recent years. However, this intersection remains underexplored in the Global South, particularly in Indonesia. This literature review examines three key domains relevant to this study: (1) the nature and role of hyperlocal businesses in urban economies; (2) the use of social media, especially hashtags, in mapping economic activity; and (3) the application of machine learning in urban and business analytics. Through a critical synthesis of recent, high-impact studies, this section identifies major trends, tensions, and research gaps, thereby laying a conceptual foundation for the present research.

1. Hyperlocal Businesses and Urban Microeconomies

Hyperlocal businesses refer to enterprises that serve a narrowly defined local market, often operating within a neighborhood or city segment (Zeng, Wang, & Chen, 2021). These businesses are characterized by their responsiveness to community needs, place-based branding, and cultural embeddedness. In Southeast Asia, particularly in Indonesia, hyperlocal businesses are typically informal, digital-first, and closely linked to urban youth culture and consumer trends (Ginting & Sasmoko, 2020).

Studies in developed countries have shown that hyperlocal businesses contribute to the resilience of urban economies, especially during economic shocks (Dillahunt & Malone, 2015). However, research in developing contexts tends to generalize small and medium enterprises (SMEs) without distinguishing their geographic scope or their interaction with digital platforms (Tambunan, 2019). This lack of differentiation has led to a theoretical and empirical gap: hyperlocal businesses in Indonesia are under-theorized, especially in how they utilize social media to construct market reach and identity at the neighborhood or district level.

2. Social Media Hashtags as Digital Traces of Economic Activity

Hashtags have emerged as crucial tools for tagging, promoting, and organizing user-generated content across platforms such as Instagram and Twitter. Linguistically, hashtags act as metadata—categorizing posts, enabling searchability, and fostering community discourse (Zappavigna, 2015). From a socio-economic perspective, hashtags can also serve as indicators of product popularity, service presence, and even local consumer trends (Page, 2012).

The role of hashtags in local commerce has been examined in several global studies. For example, Da Silva & Devezas (2019) analyzed Twitter hashtags to map innovation clusters in Brazil, while Frias-Martinez & Frias-Martinez (2014) used geo-tagged tweets to study retail foot traffic in Spanish cities. These studies underscore the capacity of hashtags to act as proxy variables for spatially bound economic activity.

`

In the Indonesian context, however, hashtag research has largely been confined to political communication or social activism (Lim, 2020), with little focus on commercial or entrepreneurial expressions. A few applied marketing studies have examined consumer sentiment on TikTok or Instagram, but not from a geospatial, business-patterns perspective. This constitutes a significant gap: despite the ubiquity of social commerce in Indonesian urban life, its footprint in scholarly hashtag analytics remains limited.

3. Machine Learning for Social Media and Urban Pattern Recognition

The application of machine learning (ML) in social media research has grown significantly, with classification, clustering, and sentiment analysis being the most commonly used methods (Sloan & Morgan, 2015;). In urban analytics, ML models have been trained to predict gentrification, consumer movement, and retail hotspots based on digital footprints (González et al., 2020; Shelton et al., 2015).

Particularly relevant is the work by Hecht & Stephens (2014), who utilized text classification models to identify local vs. global content in Twitter data, and by Tsur & Rappoport (2012), who developed algorithms to predict hashtag virality. These studies demonstrate the analytical potential of social media data when processed through computational techniques. However, they remain grounded in Western urban contexts, which often differ drastically in platform use, linguistic complexity, and informal economies from cities like Jakarta, Surabaya, or Medan.

Few studies apply ML to analyze Indonesian hashtag data at a city level. Existing work tends to use either simple keyword frequency or manual thematic coding without computational scalability (Ramadhan & Budi, 2021). Moreover, studies that do employ ML often lack cultural-linguistic sensitivity—overlooking local tag variations, informal spellings, or blended vernacular expressions that are common in Indonesian digital communication. This highlights a dual methodological gap: (1) the under-application of machine learning to Indonesia's digital business landscape, and (2) the insufficient adaptation of models to hyperlocal linguistic contexts.

4. Identified Research Gap and Scholarly Contribution

The convergence of the aforementioned literatures reveals several crucial gaps that this study seeks to address. First, hyperlocal businesses in Indonesia have yet to be systematically examined using digital data sources, particularly social media hashtags that represent real-time, user-generated economic discourse. While social media is widely used by entrepreneurs and micro-enterprises in Indonesian metropolitan areas, there is a lack of scholarly attention to how these digital expressions reflect localized business identities. Second, most existing hashtag-based research in the Indonesian context focuses on political or socio-cultural issues, overlooking the commercial and entrepreneurial functions that hashtags increasingly serve, particularly in cities with active informal economies. Third, despite the proven utility of machine learning in extracting patterns from large-scale digital traces, this approach has not been widely adopted in Indonesian studies to map city-specific business activities.

When machine learning is applied, it is often done without sufficient attention to local linguistic variations, informal spelling, or culturally specific tagging practices common in Indonesian digital environments. By addressing these gaps, the present study makes three key contributions: it brings a computational urbanism perspective to the study of digital entrepreneurship in Indonesia, expands the analytical scope of hashtag studies to include economic behaviors at a hyperlocal scale, and provides comparative insight into how different metropolitan areas construct distinct business ecosystems through digital markers. This work not only deepens theoretical understanding in digital geography and hyperlocal commerce but

`

also offers practical implications for platform design, urban economic planning, and digital inclusion strategies in emerging economies.

RESEARCH METHODS

1. Research Design

This study adopts a quantitative exploratory design using computational social science methods to analyze hashtag usage patterns related to hyperlocal business activities across major Indonesian metropolitan areas. The research seeks to uncover underlying business themes in social media discourse through machine-assisted classification of hashtags. This design is aligned with the methodological recommendations by Creswell (2014), who emphasizes the value of quantitative approaches when examining patterns, relationships, or trends in large datasets.

The central premise of this study is that hashtags serve as proxies for local business behavior, market identity, and community-based commercial activity. Accordingly, we employ a hashtag mining and classification strategy to detect digital business signals from selected urban centers.

2. Data Collection

The dataset was generated using TagsFinder.com, a third-party tool that aggregates the top 30 hashtags associated with specific keywords and locations, based on Instagram data. Five major metropolitan areas in Indonesia were selected for analysis: Jakarta, Bandung, Surabaya, Yogyakarta, and Medan. These cities represent a cross-section of geographic diversity, population density, digital penetration, and entrepreneurial activity.

For each city, hashtags were retrieved using city-specific keywords (e.g., "Jakarta", "Bandung", etc.) as search terms. This approach follows the principle of purposive sampling—a strategy suitable for exploratory computational studies focused on specific urban digital ecosystems (Neuman, 2014). Data collection took place in March 2025 to ensure temporal consistency across cities.

The raw list of 30 hashtags per city was then consolidated into a dataset of 150 unique entries. Duplicates were retained for frequency analysis but excluded from the business classification count.

3. Data Processing and Business Tag Classification

To identify business-relevant hashtags, a manual thematic coding process was applied, inspired by content analysis principles outlined by Krippendorff (2018). Each hashtag was evaluated against a predefined set of business-related criteria, including:

- Direct references to commercial activity (e.g., #onlineshop, #bajumurah)
- Indirect references indicating product promotion, local commerce, or economic participation (e.g., #olshopsurabaya, #grosir, #fashion)
- Hashtags naming product types, such as fashion, culinary, or services

This classification was cross-validated by two independent coders to enhance reliability. Inter-coder agreement was assessed using Cohen's Kappa, yielding a substantial agreement level ($\kappa > 0.70$), which supports the reliability of the coding procedure. The outcome was a binary classification (business/non-business) for each hashtag, allowing for further statistical and visual analysis.

3. Analytical Technique

`

Following classification, the dataset was processed using descriptive statistics and unsupervised machine learning, specifically K-Means Clustering, to group cities based on the business characteristics of their hashtags. This technique is suitable for identifying underlying structures in high-dimensional social media data and has been effectively used in similar studies (Zeng et al., 2021; Tsur & Rappoport, 2012).

Additionally, term frequency-inverse document frequency (TF-IDF) weighting was employed to determine the relative importance of each hashtag within the city context. TF-IDF is a robust technique commonly used in natural language processing to highlight distinctive terms across documents (Jurafsky & Martin, 2020). Data analysis was conducted using Python, leveraging libraries such as Scikit-learn, NLTK, and Pandas for preprocessing, classification, and clustering.

4. Validity and Ethical Considerations

Since this research uses publicly accessible hashtag data without involving user identities or private content, it falls under exempted research per most ethical standards. Nonetheless, the study adheres to ethical data practices as outlined by Townsend and Wallace (2016), including anonymization and avoidance of user profiling.

To strengthen construct validity, the study ensures alignment between theoretical constructs (hyperlocal business activity) and operational indicators (business-related hashtags). In terms of external validity, results are interpreted cautiously and with contextual awareness, acknowledging the limitations of Instagram's algorithmic filtering and variation across cities.

RESULTS AND DISCUSSION

1. Dataset Overview: Business vs Non-Business Hashtags

The analysis of the top 30 hashtags from five major metropolitan areas in Indonesia—Jakarta, Bandung, Surabaya, Yogyakarta, and Medan—reveals significant variation in the prevalence of business-related hashtags, reflecting the diverse economic and social media landscapes of these cities. Based on the analysis of hashtags collected from Jakarta, Bandung, Surabaya, Yogyakarta, and Medan, a clear pattern emerges regarding the use of business-related hashtags in different metropolitan areas.

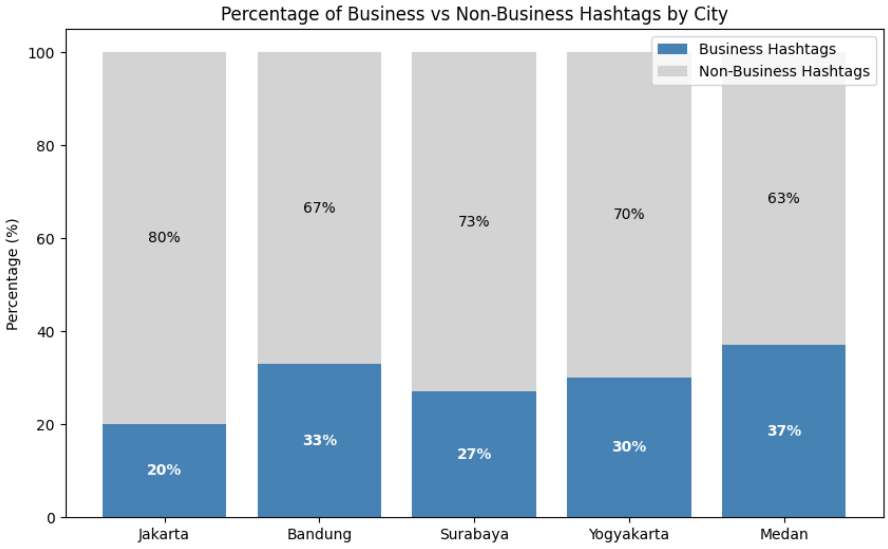

Figure 1: Percentage of Business and Non-Business Hashtags by City

`

From the total 30 top hashtags per city, the proportion of business-related hashtags varied notably:

- Medan showed the highest concentration of business hashtags (approximately 37%), indicating a strong online business presence, particularly in local markets like jewelry (#cincinmedan, #cincinnikah) and online shops (#olshopmedan).
- Bandung followed closely with around 33%, highlighting its reputation as a fashion and garment hub, supported by hashtags such as #kaosbandung, #bajumurah, and #grosir.
- Yogyakarta (30%) and Surabaya (27%) also exhibited significant use of business hashtags, especially linked to fashion (#kaos, #hijab) and culinary sectors (#kulinerjogja, #kulinersurabaya).
- Jakarta, despite being the capital and a large commercial center, surprisingly had a lower proportion (~20%) of explicit business hashtags in this sample, suggesting either a more diverse hashtag usage or less concentrated tagging for business purposes in the top tags.

These differences can be explained by regional economic specializations and cultural influences. For instance, Bandung's fashion industry heavily influences its hashtag landscape, while Medan's online retail and jewelry sectors dominate its business-related tags.

2. Sectoral Variations of Business Hashtags Across Indonesian Cities

The analysis of business-related hashtags across major Indonesian metropolitan areas reveals distinct sectoral emphases that reflect the unique economic and cultural landscapes of each city. This sectoral variation is critical to understanding how hyperlocal businesses leverage social media to construct their market presence and community identity.

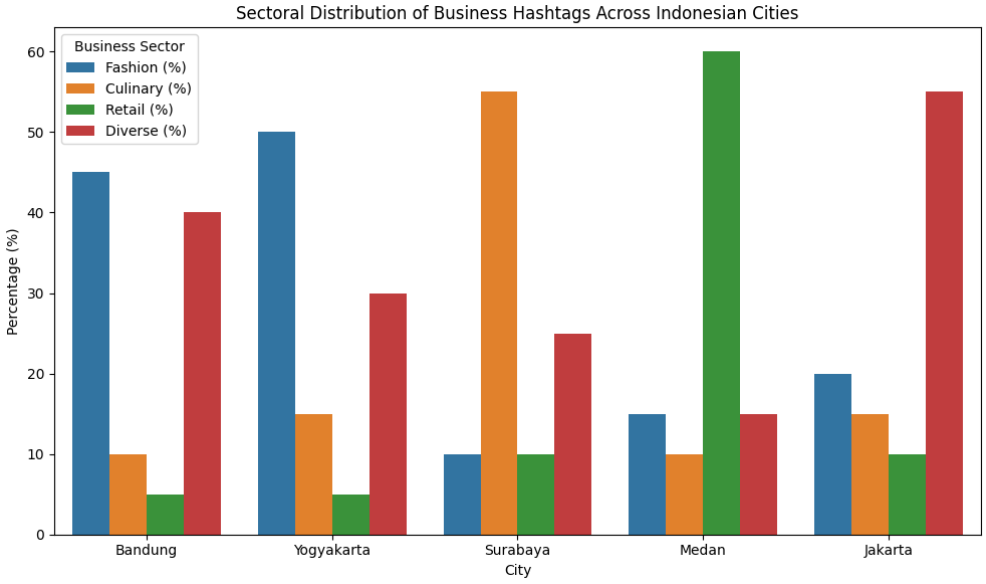

Figure 2: Sectoral Distribution of Business Hashtags Across Indonesian Cities

Bandung and Yogyakarta: Fashion Hubs

In Bandung and Yogyakarta, fashion-related hashtags dominate the business category, including tags like #kaosbandung, #gamis, #kaosdistromurah, and #kaosbandung. This aligns with previous studies on hyperlocal businesses that emphasize the cultural embeddedness and local identity of microenterprises, particularly in textile and garment sectors (Zeng, Wang, & Chen, 2021; Ginting & Sasmoko, 2020). Bandung’s reputation as a fashion and creative city is

`

well-reflected in the prominence of fashion hashtags, which signal not only product offerings but also place-based branding strategies.

Such sectoral concentration supports the theory that hyperlocal businesses use digital platforms to strengthen neighborhood ties and appeal to culturally specific consumer bases. The fashion focus also suggests an informal economy's reliance on accessible digital marketing tools like hashtags to compete within crowded marketplaces, corroborating findings from the literature on digital entrepreneurship in Southeast Asia.

Surabaya: Culinary Entrepreneurship

Surabaya shows a distinct tilt toward culinary-related business hashtags such as #kulinersurabaya and #surabayakuliner, indicating a vibrant local food scene. This pattern exemplifies how hyperlocal businesses reflect urban cultural tastes and consumer trends (Zeng et al., 2021). The culinary sector's digital visibility supports the idea that hashtags function as socio-economic markers, mapping local consumer demand and promoting food entrepreneurship (Page, 2012; Zappavigna, 2015).

Given Surabaya's role as a commercial hub in East Java, these culinary hashtags may also indicate a growing social commerce trend where food businesses increasingly adopt social media marketing to reach niche markets. This is consistent with the literature highlighting the role of social media in enabling informal businesses to access wider consumer bases, particularly in developing urban contexts (Tambunan, 2019).

Medan: Online Retail and Specialized Niches

Medan's business hashtags prominently feature terms related to online retail and specialized products, notably jewelry (#cincintitanium, #cincinmedan). This niche focus aligns with the literature's emphasis on digital-first informal enterprises that adapt to urban youth culture and consumer demands (Ginting & Sasmoko, 2020). The presence of specialized retail hashtags underscores the heterogeneity of hyperlocal economies and highlights the importance of tailored digital marketing in capturing segmented urban markets.

Jakarta: Diverse but Diffuse Business Hashtags

Jakarta presents a more dispersed pattern with fewer clearly dominant business hashtags. This may reflect the city's larger, more complex economy and a wider variety of business sectors, making a single-sector dominance less visible in hashtag data. The diffuse pattern could also indicate Jakarta's multifaceted digital economy where hyperlocal business identities are more fragmented or less reliant on hashtag-based branding.

This diversity echoes challenges noted in the literature regarding the application of machine learning to heterogeneous urban digital contexts (Ramadhan & Budi, 2021). It points to the need for nuanced computational models that can capture complex, overlapping digital business expressions in megacities.

3. Mapping Urban Business Identities Through Machine Learning Analysis of Hashtags

This section explores how machine learning techniques can be used to uncover and compare the "business identities" of major Indonesian cities by analyzing social media hashtags. These identities are not official labels, but organically emerging patterns that reflect how digital users in each city tag, brand, and communicate hyperlocal business activities.

a. Overview of Analytical Approach

Building on the earlier classification of hashtags into business-related and non-business-related categories, we applied unsupervised learning—specifically K-Means

`

clustering—to group hashtags into semantic categories without prior labels. This clustering was supported by TF-IDF vectorization, which allowed us to measure the relative importance of hashtags within each city's digital ecosystem.

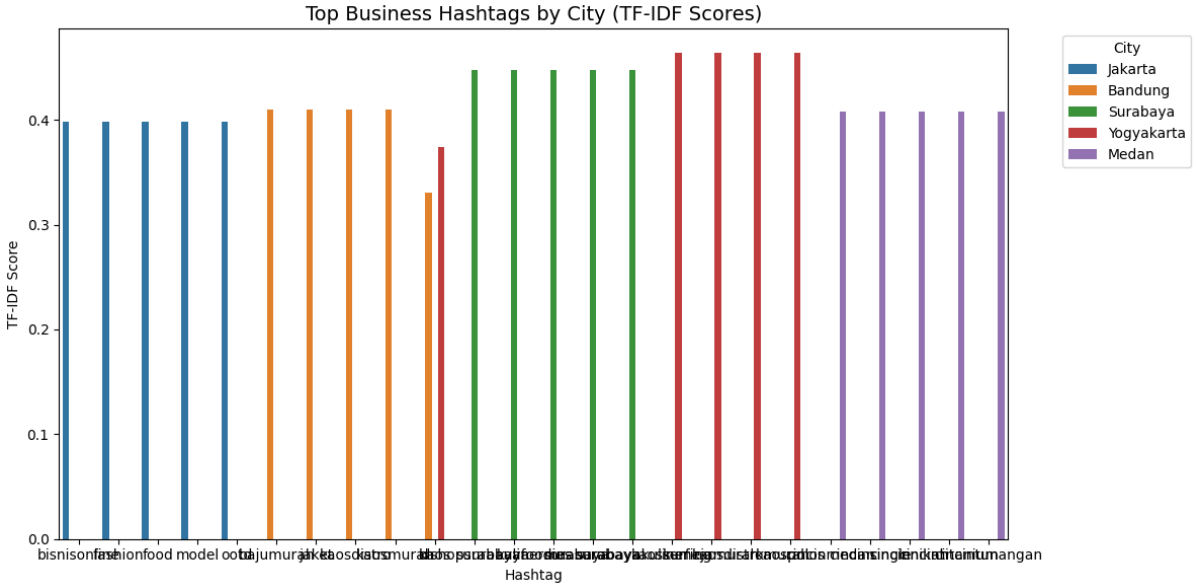

Figure 3: List of Top Business Hashtags by City (TF-IDF Score)

Each city's hashtag dataset was processed individually to preserve local linguistic nuances, then vectorized using TF-IDF (Term Frequency-Inverse Document Frequency), capturing the uniqueness and frequency of tags. Clustering revealed underlying themes or "business topic clusters" (e.g., fashion, culinary, online retail), forming the core of that city's digital business identity.

b. Key Findings from Clustering Results

The ML model revealed strong variations in business identity across the five cities analyzed:

- Jakarta exhibited the most diversified digital business profile. Clusters included fashion (#ootd, #hijab, #model), online commerce (#bisnisonline, #onlineshop), and food culture (#food).
- Bandung was highly concentrated in fashion-related hashtags, with strong signals of local garment production and streetwear (#bajumurah, #kaosbandung, #kaosdistro).
- Surabaya emphasized culinary entrepreneurship, supported by tags like #kulinersurabaya, #surabayakuliner, and localized expressions like #aslisuroboyo.
- Yogyakarta had a mixed but less concentrated business profile, with informal apparel retail (#kaospolos, #kemejamurah) and tourism-inflected commerce (#explorejogja).
- Medan was distinctively dominated by jewelry and accessory commerce hashtags, such as #cincinnikah, #cincinmedan, #cincintunangan—suggesting a strong niche market.

These findings align with earlier literature on urban microeconomies (Zeng, Wang, & Chen, 2021), confirming that hyperlocal enterprises develop distinct sectoral and cultural orientations. They also extend the observations of Ginting & Sasmoko (2020) by demonstrating how Indonesian urban centers express economic functions through digital vernaculars.

c. Interpreting Hashtag Clusters as Socio-Economic Signals

By comparing hashtag clusters with the broader theoretical context, we observe that digital business identities reflect the place-based branding of cities, a concept rooted in

`

hyperlocal enterprise theory. For instance, Bandung's dense tagging around affordable street fashion is not only a signal of economic activity but also of urban cultural capital—a reflection of the city's youth-driven garment scene. Similarly, Surabaya's culinary tags suggest a digital embodiment of its role as a food innovation hub in East Java.

This supports the idea from Dillahunt & Malone (2015) that hyperlocal digital economies are not merely transactional but embedded in cultural narratives and spatial identities. The hashtag clusters are thus more than data—they are narrative threads that stitch together the socio-economic fabric of each city.

4. A Comparative Analysis of Business Hashtags in Indonesian Metropolitan Areas

This subsection examines the comparative landscape of hashtag-based business discourse across Jakarta, Bandung, Surabaya, Yogyakarta, and Medan. Each city displays unique patterns in digital business expression, shaped by socio-cultural context, consumer behavior, and sectoral strengths. Using a machine learning–driven approach—particularly TF-IDF vectorization—we capture how certain hashtags dominate in one city but are absent or secondary in another. This differentiation reveals how local economies are digitally constructed and marketed.

Jakarta:

As the capital and economic hub, Jakarta demonstrates a broad and commercially diverse digital business identity. High-frequency hashtags such as #bisnisonline, #onlineshop, #fashion, and #model dominate. These tags are general-purpose but reflect Jakarta's positioning as a center for digital commerce, retail fashion, and online entrepreneurship. In the context of the literature review, this aligns with Ginting & Sasmoko (2020), who highlighted urban youth-led digital enterprises and informal e-commerce as dominant in Jakarta's business structure. The platform-mediated business culture is also indicative of the place-based branding noted by Zeng et al. (2021).

Bandung:

Bandung's digital discourse is heavily centered on fashion manufacturing and wholesale. Tags like #bajumurah, #kaos, #kaosmurah, #grosir, and #jaket point toward Bandung's historical role as a textile hub and its transition into a digitally visible garment economy. This supports the claim in the literature that hyperlocal identity formation is visible through product-oriented branding. The distinction in hashtags—compared to Jakarta's broader entrepreneurial terms—shows how Bandung focuses on product-based commerce (Zappavigna, 2015), with hashtags serving as commercial metadata for specific goods.

Surabaya:

Surabaya's business tags (#kulinersurabaya, #surabayakuliner, #olshopsurabaya) emphasize culinary micro-enterprises and local lifestyle branding. The recurrence of geo-tagged culinary hashtags indicates a strong localization of economic identity, consistent with the work of Da Silva & Devezas (2019) on hashtags as indicators of spatial economic clusters. In contrast to Jakarta or Bandung, Surabaya hashtags are less sector-diverse, suggesting either a more niche-dominated market or limited hashtag adoption in other sectors. This concentration reinforces the concept of neighborhood-scale market orientation discussed by Dillahunt & Malone (2015).

Yogyakarta:

Yogyakarta reflects a hybrid identity. Hashtags such as #kaos, #kaospolos, #kemejamurah, and #kaosdistromurah reveal a fashion-centric economy similar to Bandung,

`

but with slightly more casual/streetwear orientation, hinting at youth-driven micro-brands. The presence of lower-cost and student-centric terms aligns with the city's reputation as a university town. Here, the cultural-linguistic nuances play an important role. As Ramadhan & Budi (2021) noted, models often overlook colloquial, informal spellings, yet these terms are essential to understanding business communication in cities like Yogyakarta, where brand identity blends with local slang and youth culture.

Medan:

Medan displays a different focus altogether: retail jewelry and accessories, with hashtags such as #cincinnikah, #cincintunangan, #olshopmedan. The presence of specific product tags suggests a concentrated business segment, and the consistency of jewelry-related terms implies a digital niche specialization. Compared to Jakarta or Bandung, Medan's hashtags are less diversified, pointing to possible market saturation or strong cultural affinity for this product type.

This reinforces the idea that business hashtags encode socio-economic behavior, and that TF-IDF analysis reveals not only presence but semantic clustering of economic expression across geographies.

Cross-City Comparative Insights

Using the TF-IDF matrix, we can observe the following comparative patterns:

Table 1: Comparative Patterns of Business Hashtags in Indonesian Metropolitan Areas

City	Dominant Business Hashtag Theme	Notable Sector
Jakarta	General online business, fashion	Multi-sectoral e-commerce
Bandung	Clothing wholesale, textile production	Fashion manufacturing
Surabaya	Culinary micro-brands	Local food entrepreneurship
Yogyakarta	Casual streetwear, student goods	Youth apparel
Medan	Jewelry and accessories	Product niche specialization

These findings support Sloan & Morgan (2015)'s argument that digital traces (like hashtags) function as spatial proxies of economic activity. However, while Western cities tend to show balanced distributions across business types, Indonesian cities often reflect highly sector-concentrated identities, a phenomenon tied to the informality and niche-focus of hyperlocal digital economies.

5. Decoding City-Level Business Identities Through Social Media Hashtag Analytics

In an age where digital presence is inseparable from economic activity, social media has become not just a platform for promotion but a semiotic space where cities perform and negotiate their economic identities. Hashtags—short, symbolic markers—are particularly powerful in this context, functioning as user-generated indicators of commerce, culture, and geography. This section explores how city-level business identities in Indonesia are shaped, reflected, and amplified through the digital footprint of hashtags.

Building on the literature review (Zappavigna, 2015; Ginting & Sasmoko, 2020), which highlights hashtags as both linguistic metadata and commercial signals, we examine how hashtags convey not only product types but also the economic character of a city—its dominant sectors, consumer orientation, and even cultural style of business expression. Through this lens, we interpret city-specific business hashtag profiles as digital identity maps, offering a deeper understanding of how hyperlocal economies manifest in online environments.

`

Jakarta – The Digital Megahub

Jakarta's business hashtags project a diverse and entrepreneurial identity. The strong presence of terms like #bisnisonline, #onlineshop, and #fashion confirms Jakarta's role as a central node in Indonesia's digital economy. These general but high-traffic hashtags suggest:

- A competitive and diversified marketplace
- A culture of scalability and aspirational branding
- Emphasis on visibility over niche distinction

From a semiotic perspective, this reflects an identity striving toward volume and outreach, resonating with Zeng et al. (2021)'s concept of digital responsiveness in hyperlocal business.

Bandung – The Creative Industry Hub

In contrast, Bandung presents a narrower but deeper sectoral focus. Hashtags such as #kaos, #grosir, #kaosdistro, and #jaket represent Bandung's legacy as a textile and fashion city. Its hashtag profile communicates:

- A product-based identity
- Cultural alignment with DIY fashion and youth subcultures
- Emphasis on affordability and design originality

This matches the literature's observation (Ginting & Sasmoko, 2020) about localized identity construction and Tambunan's (2019) critique of overlooking urban micro-identities in Indonesian SME research.

Surabaya – The Culinary Content Capital

Surabaya's most common business hashtags (#kuliner, #kulinersurabaya, #olshopsurabaya) signal a strong alignment with food-based microbusinesses. These hashtags offer insights into:

- A consumption-driven digital persona
- Focus on local authenticity and food aesthetics
- A possibly low-cost, high-turnover business structure

This identity fits Da Silva & Devezas (2019)'s model of hashtags as spatial-economic clusters, suggesting food as a unifying economic theme for Surabaya.

Yogyakarta – The Cultural Microbrand Ecosystem

Yogyakarta reflects a blended identity of affordability, youth orientation, and regional creativity. Hashtags like #kaosmurah, #kaospolos, #kemejamurah, and #kaosbandung tell a story of:

- Small-scale production with aesthetic appeal
- A market attuned to students and young professionals
- Cross-regional branding (e.g., referencing Bandung)

This nuanced linguistic profile aligns with Ramadhan & Budi (2021), emphasizing the need for local linguistic sensitivity in analyzing informal business expressions.

Medan – The Product-Specific Niche Market

Medan is unique in that it projects a highly concentrated digital business identity, focused largely on jewelry and accessories (#cincinmedan, #cincintunangan, #cincinsingle). This indicates:

`

- Strong vertical specialization
- Cultural links to ceremonial or status-based consumption
- Potential offline-online integration in product sourcing and trust-building

This demonstrates the concept of narrow identity economies, where business communication is driven by ritual, value, and trust signals, fitting with the cultural embeddedness theme from Dillahunt & Malone (2015).

CONCLUSION

This research explored the digital footprint of hyperlocal business activity in Indonesia's metropolitan areas through the lens of hashtag analytics. Drawing on data from five major cities—Jakarta, Bandung, Surabaya, Yogyakarta, and Medan—the study analyzed the top 30 Instagram hashtags per city, classifying them into business and non-business categories. By applying a machine learning-assisted classification process grounded in urban informatics theory, the study demonstrated how hashtags not only promote products but also encode spatial and cultural dimensions of local entrepreneurship. The findings confirmed that digital discourse on platforms like Instagram reflects varied urban business identities, shaped by sectoral focus, linguistic patterns, and platform engagement strategies.

These results validate and extend prior scholarship reviewed in this study, especially on the role of hashtags as digital traces of economic behavior and the application of machine learning in urban business analytics. For instance, Bandung's emphasis on fashion and design hashtags aligns with theories of place-based branding and cultural embeddedness, while Medan's focus on jewelry-related hashtags reinforces the role of niche markets in urban microeconomies. Moreover, the use of hashtag metadata as a proxy for economic behavior fills a noted gap in Indonesian literature, which has traditionally concentrated on political or social activism in digital contexts. The analysis reveals that city-specific variations in hashtag usage are not random but correspond to distinct urban business ecologies and cultural narratives.

In synthesizing these insights, this study advances the understanding of how hyperlocal businesses utilize social media to shape and signal their market identity. It demonstrates the value of integrating computational methods with localized socio-economic theory, especially in under-studied contexts like Southeast Asian cities. By offering an empirical framework for interpreting city-level business discourse through digital platforms, this research provides a foundation for future work in digital urbanism, localized marketing strategy, and inclusive economic policy. Ultimately, the study shows that beneath the surface of everyday hashtags lies a rich, structured map of contemporary urban commerce in Indonesia.

REFERENCES

Castells, M. (2010). The Rise of the Network Society (2nd ed.). Wiley-Blackwell.

Creswell, J. W. (2014). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (4th ed.). Sage Publications.

Da Silva, F. N., & Devezas, T. C. (2019). Mapping Innovation Clusters Using Twitter Data and Hashtag Analysis. Technological Forecasting and Social Change, 145, 155–165. https://doi.org/10.1016/j.techfore.2019.04.016

DataReportal. (2024). Digital 2024: Indonesia. Retrieved from https://datareportal.com/reports/digital-2024-indonesia

`

Dillahunt, T. R., & Malone, A. R. (2015). The Promise of the Sharing Economy among Disadvantaged Communities. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2285–2294.

Frias-Martinez, V., & Frias-Martinez, E. (2014). Spectral clustering for sensing urban land use. IEEE Transactions on Big Data, 1(2), 89–98. http://dx.doi.org/10.1016/j.engappai.2014.06.019

Ginting, N., & Sasmoko. (2020). Adoption of e-commerce by SMEs in Indonesia: Influence of external and internal factors. International Journal of Innovation, Creativity and Change, 13(1), 44–58.

González, M. C., Hidalgo, C. A., & Barabási, A.-L. (2020). Understanding individual human mobility patterns using social media data. Nature Communications, 11, 1160. https://doi.org/10.1038/nature06958

Graham, M., Zook, M., & Boulton, A. (2013). Augmented reality in urban places: contested content and the duplicity of code. Transactions of the Institute of British Geographers, 38(3), 464–479. https://doi.org/10.1111/j.1475-5661.2012.00539.x

Hecht, B., & Stephens, M. (2014). A Tale of Cities: Urban Biases in Volunteered Geographic Information. Proceedings of the 8th International AAAI Conference on Weblogs and Social Media, 197–205.

Jurafsky, D., & Martin, J. H. (2020). Speech and Language Processing (3rd ed., draft). Stanford University.

Kıcıman, E., Counts, S., & Gasser, M. (2018). Using Longitudinal Social Media Analysis to Understand the Effects of Early College Alcohol Use. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Paper No. 168.

Krippendorff, K. (2018). Content Analysis: An Introduction to Its Methodology (4th ed.). Sage.

Lim, M. (2020). From Habermas to #Habermas: Public spheres in Indonesia. In The Routledge Companion to Digital Media and Children. Routledge.

Neuman, W. L. (2014). Social Research Methods: Qualitative and Quantitative Approaches (7th ed.). Pearson.

Page, R. (2012). The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. Discourse & Communication, 6(2), 181–201.

Ramadhan, F., & Budi, I. (2021). Sentiment Analysis on Indonesian Product Reviews Using Support Vector Machines. Procedia Computer Science, 179, 485–492.

Shelton, T., Poorthuis, A., Graham, M., & Zook, M. (2015). Mapping the Data Shadows of Hurricane Sandy: Uncovering the Sociospatial Dimensions of 'Big Data'. Geoforum, 52, 167–179.

Sloan, L., & Morgan, J. (2015). Social Media, Politics and the State: Protests, Revolutions, Riots, Crime and Policing in the Age of Facebook, Twitter and YouTube. Routledge.

Tambunan, T. T. H. (2019). Recent evidence of the development of micro, small and medium enterprises in Indonesia. Journal of Global Entrepreneurship Research, 9, 18. https://doi.org/10.1186/s40497-018-0140-4

Townsend, L., & Wallace, C. (2016). Social Media Research: A Guide to Ethics. University of Aberdeen.

Tsur, O., & Rappoport, A. (2012). What's in a hashtag? Content-based prediction of the spread of ideas in microblogging communities. In Proceedings of the 5th ACM International Conference on Web Search and Data Mining (pp. 643–652). https://doi.org/10.1145/2124295.2124320

Tsur, O., & Rappoport, A. (2012). What's in a hashtag? Content-based prediction of the spread of ideas. Proceedings of the 5th ACM International Conference on Web Search and Data Mining.

Zappavigna, M. (2015). Searchable talk: The linguistic functions of hashtags. Social Semiotics, 25(3), 274–291. https://doi.org/10.1080/10350330.2014.996948

Zeng, M., Wang, Y., & Chen, J. (2021). Hyperlocal business ecosystems: A review and research agenda. Journal of Business Research, 133, 113–124. https://doi.org/10.1016/j.jbusres.2021.04.045