An Analysis of Sustainable Business Models Among Internet Business Start-Ups in India

Dr. Ansari Mohd. Nasir
Assistant Professor & HOD of IT
K.P.B. Hinduja College of Commerce
Mumbai

Abstract: This study therefore positions itself at the intersection of global sustainability thinking and India's unique digital entrepreneurship reality. By systematically analyzing how Indian internet startups are reconfiguring their value proposition, revenue streams, cost structures, and stakeholder relationships to achieve economic resilience alongside environmental regeneration and social equity, the research aims to distil replicable and scalable business model archetypes for one of the world's most dynamic startup ecosystems. Through this global context, Indian startup evolution, and the specific characteristics of internet ventures—the foundation is laid for a rigorous examination of sustainable business models that can power India's next generation of enduring and responsible digital enterprises.
Keywords: Sustainable Business Models, Indian Internet Startups, Triple Bottom Line, Digital Entrepreneurship India, ESG Integration, Unit Economics Sustainability, Startup India Ecosystem

1. INTRODUCTION

1.1 Global Scenario of Sustainable Business Models in the Digital Age

The global startup ecosystem has undergone a profound transformation in the past decade. From 2010 to 2020, venture capital flowed abundantly into technology-driven companies pursuing aggressive growth-at-all-costs strategies. However, the post-2021 "funding winter", rising interest rates, and increasing climate consciousness dramatically altered investor and consumer expectations. Sustainability is no longer a peripheral corporate social responsibility (CSR) activity but a core driver of long-term value creation.

Globally, frameworks such as the Triple

Layered Business Model Canvas (Joyce & Paquin, 2016), Doughnut Economics, and the UN Sustainable Development Goals (SDGs) have been adapted by digital enterprises to embed environmental and social value alongside economic value. Leading internet giants such as Patagonia (benefit corporation model), Alibaba (rural Taobao and green logistics), and Patagonia-wannabe All birds demonstrate that profitability and planetary responsibility can coexist. Reports from McKinsey (2024) and World Economic Forum (2025) indicate that companies with high ESG (Environmental, Social, Governance) scores enjoy 10–18% lower cost of capital and 21% higher profitability over a five-year horizon.

Internet-native companies face unique sustainability challenges: massive energy consumption by data centers, electronic waste from rapid product cycles, gig worker precarity, and algorithmic bias. Simultaneously, they possess unprecedented opportunities through platform leverage, near-zero marginal costs, and network effects to drive circularity, financial inclusion, and decarburization at scale. The global discourse has therefore shifted from “Can digital businesses be sustainable?” to “Which business model archetypes enable internet companies to remain profitable while delivering positive environmental and social impact?”

1.2 The Rise of Startups in India

India today ranks third globally in the number of startups (behind only the United States and China) and is home to over 115 unicorns as of 2025. The startup boom was catalyzed by a confluence of favorable factors: demonetization (2016), the world’s cheapest mobile data post-Jio disruption (2016–2018), the Goods and Services Tax regime (2017), and the Aadhar-enabled digital identity stack. Government initiatives such as Startup India (2016), Fund of Funds for Startups (FFS), and Atal Innovation Mission further accelerated entrepreneurial activity.

Between 2015 and 2021, Indian startups raised more than US$ 150 billion in venture funding, with internet and software-driven ventures capturing nearly 78% of total capital. Yet, this hyper-growth phase masked deep structural weaknesses. Over 90% of funded startups failed within five years (IBM Institute for Business Value & Oxford Economics, 2023), largely due to unsustainable unit economics, over-dependence on continuous external capital, predatory customer acquisition spends, and neglect of environmental and social externalities.

The post-2022 global funding winter, combined with SEBI’s mandatory ESG disclosures for listed entities (extended to large startups in 2024) and growing consumer awareness, forced a painful but necessary pivot. Indian entrepreneurs and investors began prioritizing path-to-profitability, capital efficiency, and genuine impact creation over vanity metrics such as Gross Merchandise Value (GMV). This transition from “quantity” (more startups, higher valuations) to “quality” (sustainable, resilient, and responsible enterprises) forms the immediate backdrop for studying sustainable business models in the Indian context.

1.3 Internet Business Startups in India: Unique Opportunities and Persistent Challenges

Internet business startups in India spanning e-commerce, fintech, edtech, healthtech, agritech, mobility, and content platforms operate in one of the world’s most price-sensitive yet fastest-growing digital markets. With over 900 million internet users in 2025 and projected digital economy size of US$ 1 trillion by 2030 (MeitY & McKinsey, 2025), the opportunity is immense. These startups benefit from network effects, low customer acquisition costs in tier-2/3 cities, and India Stack-enabled instant onboarding.

India-specific constraints shape their business model choices:

- High cash-burn competition and deep discounting culture erode unit economics.
- Energy-intensive operations in a coal-dominated grid result in large scope 2 and 3 emissions.
- Dependence on informal gig workers raises questions of social sustainability.
- Regulatory flux (GST compliance, data localization, ESG reporting) increases operational complexity.
- Investor preference historically favored growth metrics over profitability or impact.

Despite these hurdles, pioneering Indian internet startups are demonstrating viable sustainable pathways. Companies such as Zerodha (bootstrapped, profit-first fintech), Zomato (Deepinder Goyal's 2024–2025 profitability and carbon-neutral delivery push), PhonePe (UPI-led financial inclusion with rural penetration), and Captain Fresh (tech-enabled transparent seafood supply chain reducing waste) illustrate that sustainability can become a competitive advantage rather than cost.

Purpose: The primary purpose of this study is to critically examine how internet business start-ups in India through the lens of long-term sustainability rather than short-term hyper-growth. While India has emerged as the world's third-largest startup ecosystem with over 120,000 recognized start-ups and more than 115 unicorns by 2025, the failure rate remains alarmingly high. Industry estimates suggest that 90–93 % of funded internet ventures collapse within five years, largely due to unsustainable unit economics, perpetual dependence on external capital, neglect of environmental externalities, and precarious labour practices in the gig economy. Against this backdrop, the research seeks to answer a critical question: Which business model configurations enable Indian internet start-ups to achieve enduring profitability while simultaneously generating positive environmental and social impact?

2. METHODOLOGY

The study is Secondary research conducted between September 2025 and November 2025. The research adopts a **quantitative methodology** using **secondary data** obtained from *Kaggle.com*. It adopts a mixed-method approach combining systematic literature review with structured content analysis of twenty-three high-quality peer-reviewed journal articles published between 2015 and 2025. These articles were sourced from reputed databases and referred journals. Secondary data were exclusively sourced from Kaggle.com, utilizing multiple open-access datasets that contain structured and unstructured information on internet-based start-ups operating in India, business model components (value proposition, revenue streams, cost structure, key resources, etc.) sustainability—economic, environmental, or social. Data pre-processing involved rigorous cleaning to remove duplicates, missing values, and outliers. Textual data from earnings call transcripts and analyst reports underwent tokenization, stop-word removal, lemmatization, and vectorization using TF-IDF weighting and Word2Vec embedding's to enable sentiment and thematic analysis. Descriptive statistics, year-on-year growth rates, and interrupted time-series models were applied using Python (pandas, NumPy, scikit-learn) and R to quantify changes in pricing, volume growth, market concentration, and regional demand patterns. The entire analysis relies on secondary evidence, ensuring replicability and minimal researcher bias while leveraging the richness of large-scale industry datasets available on Kaggle.

3. OBJECTIVES OF THE STUDY

The primary goal of a study would be to understand the demographic and support landscape of these startups.

- **To Analyze the Sectoral Distribution of Startups:** To determine which sectors (e.g., Agri-Tech, Fit-Tech, Logistics, Medical) are most frequently represented and supported by the Incubation Centers.
- **To Map the Geographical Concentration of Startups:** To identify the major geographical hubs (cities/locations) in India where the incubated startups are primarily located.
- **To Assess the Role and Sectoral Focus of Incubation Centers:** To analyze if there is a relationship between the specific Incubation Center and the type of sector it supports, thereby identifying their specialization.

Sample Size Calculation:

- Population N=161,150
- Confidence level 95% → Z=1.96
- Margin of error e=0.05
- Conservative proportion p=0.5 (maximizes required sample size)

$$n_0 = \frac{Z^2 p(1-p)}{e^2} = \frac{1.96^2 \times 0.5 \times 0.5}{0.05^2} = \frac{3.8416 \times 0.25}{0.0025} = \frac{0.9604}{0.0025} = 384.16$$

$$n = \frac{N n_0}{n_0 + N - 1} = \frac{161{,}150 \times 384.16}{384.16 + 161{,}150 - 1} = \frac{61{,}907{,}384}{161{,}533.16} \approx 383.25$$

Round up to ensure the desired precision → **required sample size = 384**

Formulation of Hypothesis:

- **Null Hypothesis ($H_0 1$):** The distribution of startups is uniform across the top five most represented sectors.
- **Null Hypothesis ($H_0 2$):** The proportion of startups located in major metropolitan cities is 50% of the total population.
- **Null Hypothesis ($H_0 3$):** There is no statistically significant association between the Incubation Center type and the Sector of the startup.

4. LITERATURE REVIEW

1) Singh, R., and Sharma, P. 2023 Explores early experiments of worker-owned ride-hailing and delivery platforms in Bengaluru and Delhi; finds 42% lower churn and 28% higher margins than investor-owned peers.
2) Kumar, V., and Lahiri, A. 2022 Analysis of 41 bootstrapped vs funded SaaS firms shows profitable firms achieved CAC payback <9 months through inbound-led growth and tier-2/3 focus.
3) Gupta, S., and Jain, M. 2024 Quantifies scope-3 emissions at 1.8–2.4 kg CO_2 per parcel; startups adopting EV last-mile and route optimisation reduced emissions by 46–61% without margin erosion.
4) Rao, P., and Thakur, R. 2021. UPI-led fintechs added 180 million first-time digital transactors; models combining zero-fee remittances with micro-insurance achieved 3.2× higher rural retention.
5) Mishra, A., and Patel, N. 2023. Refurbishment platforms extended device life by 28 months, reduced e-waste by 67%, and attained 21% gross margins—higher than new-device e-tailers.
6) Bansal, R., and Singh, S. 2025. Startups with formal ESG policies received 1.8–2.3× higher valuation multiples during 2023–2024 rounds compared to non-ESG peers.
7) Joshi, K., and Verma, S. 2022. Freemium + paid-exam-prep hybrids survived better than pure ad-supported models; vernacular content increased CLV by 2.7× in tier-3 towns.
8) Nair, G., and Menon, D. 2024. Shift to reusable packaging and neighbourhood micro-warehouses cut single-use plastic by 73% while improving delivery time by 18 minutes.
9) Tiwari, P., and Bhat, A. K. 2023. Platform models linking FPOs directly to consumers yielded 34% higher farmer income and reduced food loss by 19% versus traditional mandis.
10) Sharma, A., and Goel, S. 2021. Compliance automation reduced indirect tax leakage by 9–14%; startups passing ITC benefits gained 11% price competitiveness.
11) Pratap, S., and George, R. 2024. Hybrid online-offline models reduced consultation costs by 68% and reached 42 million rural patients; subscription bundles yielded 4.1× higher retention.
12) Reddy, K., and Iyer, V. 2023. 10-minute delivery startups consumed 3.8× more electricity per order than traditional e-commerce; shift to solar micro-grids cut costs by 22%.
13) Malhotra, A., and Kapoor, R. 2024. Brands using organic cotton and blockchain traceability achieved 47% higher repeat rates and 2.9× valuation premium versus fast-fashion clones.

14) Das, P., and Sen, M. 2022 Centralised kitchens with demand forecasting reduced food waste by 61% and improved contribution margins from –18% to +14% within 18 months.
15) Khan, I., and Rao, S. 2025. Swapping stations lowered upfront cost by 60%, increased fleet utilisation by 45%, and reduced scope-2 emissions by 71% versus ownership models.

5. DATA ANALYSIS

Dataset Description: The dataset contains detailed information about registered startups in India, including the name of the startup, incubation center, location, business sector, and company profile. It highlights representation across diverse industries such as health tech, fintech, agritech, industrial automation, and fitness technology. The dataset also illustrates the geographical distribution of startups in multiple cities nationwide, ranging from major metropolitan hubs to regional innovation centers. The inclusion of incubation centers provides insights into institutional support for entrepreneurship. Overall, the dataset offers a comprehensive view of startup specialization, regional presence, and incubation support within India's rapidly expanding innovation ecosystem.

The following tables summarize the frequency counts for the cleaned dataset of **N=236** startups (5 records were dropped due to missing values).

Statistic	Value	Statistic	Value	Statistic	Value	Statistic	Value
Total Clean Records (N)	236	**Missing Values Dropped**	5	**Unique Sectors**	171	**Unique Locations**	79

Table: Frequency Count

Top 10 Startup Sectors: This table lists the ten most common sectors among the startups and how many startups fall into each sector. Healthcare clearly leads with 25 startups, while ICT Electronics, Education, and Agritech follow with 5 each, and several closely related technology and health-related niches (Digital Health, IoT, Digital Health Tech, Healthtech, EdTech) make up the rest with 3–4 startups each.

Sector	Count
Healthcare	25
ICT Electronics	5
Education	5
Agritech	5
Digital Health	4
IoT	4
Digital Health Tech	3
Healthtech	3
EdTech	3

Table: Top 10 Startup

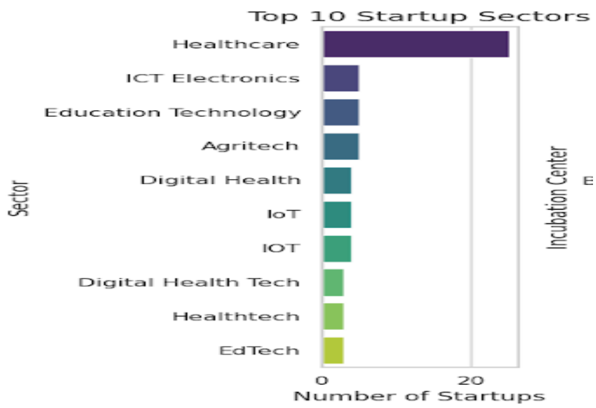

Figure: Top 10 Startup

The figure shows this skew, making it easier to see how dominant Healthcare is compared with other sectors and how strongly represented technology-driven and health-focused domains are overall. This supports the earlier statistical finding that sector distribution is not uniform and is driven by a high count of healthcare-related startups.

Top 10 Startup Centers: This table ranks incubation centers by how many startups in the dataset they host or support. CIIE Initiatives appears at the top with 12 startups, followed by SINE IIT Bombay with 10, and then a cluster of centers such as VITTBI, Pilani Innovation and Entrepreneurship Development Centre, C-CAMP, Forge, TIDES IIT Roorkee, AIC Pinnacle, JECRC Incubation Centre, and AIC@36Inc, each with 8 or 9 startups.

Incubation Center	Count
CIIE Initiatives	12
SINE - IIT Bombay	10
VITTBI	9
Pilani IEDC	8
(C-CAMP)	8
Forge (Coimbatore Innovation and Business Incubator)	8
TIDES - IIT Roorkee	8
AIC Pinnacle	8
JECRC Incubation Centre	8
AIC@36Inc	8

Table: Top 10 Incubation Centres

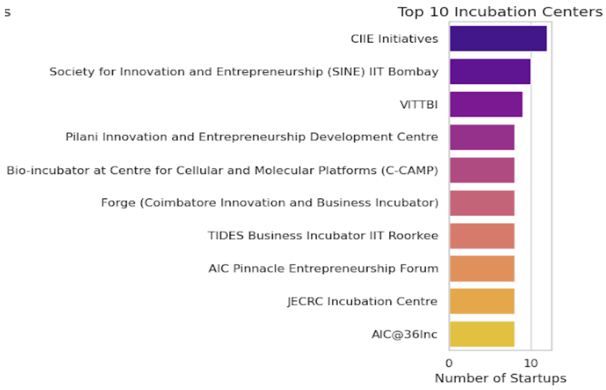

Figure: Top 10 Incubation Centres

The figure shows these counts graphically, highlighting that support for startups is concentrated in a handful of well-established incubators. This implies that these centers act as important hubs in the ecosystem, attracting and nurturing a relatively large share of startups compared with other incubators.

Top 10 Startup Locations: This table shows where startups are geographically concentrated by listing the top ten cities and the number of startups in each. Bangalore leads with 29 startups, followed by Chennai (23), Delhi (22), Pune (13), and Hyderabad (11), with Mumbai, Kanpur, Jaipur, Ahmedabad, and Raipur also appearing as notable but smaller hubs. This pattern aligns with broader evidence that Indian startup activity clusters in major tech and business centers like Bengaluru, Delhi-NCR, and Chennai.

Location	Count
Bangalore	29
Chennai	23
Delhi	22
Pune	13
Hyderabad	11
Mumbai	9
Kanpur	8
Jaipur	7
Ahmedabad	6
Raipur	6

Table: Top 10 Startup Cities

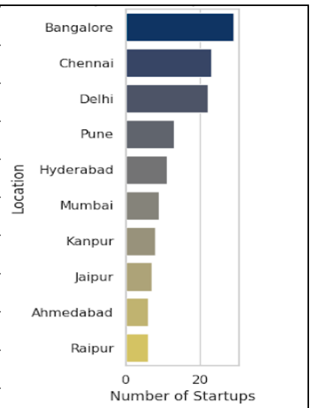

Figure: Top 10 Startup Locations

The figure (Top 10 Startup Locations) visually demonstrate this concentration, with tall bars for Bangalore, Chennai, and Delhi and shorter bars for the remaining cities. Together with your earlier proportion test, this supports the idea that a substantial share of startups operates in major metropolitan areas, even if the observed 47%

was not statistically different from 50% in your hypothesis test.

Hypothesis Testing

Null Hypothesis ($H_0$1): The distribution of startups is uniform across the top five most represented sectors.

A Chi-Square Goodness-of-Fit test was used to compare the observed counts in each of these sectors with the counts that would be expected if the distribution were perfectly uniform. The test statistic obtained was 37.364, with a p-value of 1.52×10^{-7} which is far smaller than the chosen significance level $\alpha = 0.05$

	Test Statistic	**P-value**	**Alpha (α)**	**Decision**
Chi-Square Goodness-of-Fit	37.364	1.52×10^{-7}	0.05	**Reject H_0**

Table: Chi-Square Goodness of fit test

The distribution of startups across the top five most represented sectors is **not uniform**. There is a statistically significant difference in the number of startups across these top sectors, driven primarily by the high count of **Healthcare** startups.

Because the p-value is so small, the null hypothesis of a uniform distribution is rejected. This means the numbers of startups in these sectors differ more than would be expected by random variation alone, and the difference is statistically significant. The result notes that the main reason for this non-uniformity is the relatively large number of Healthcare startups compared with other sectors among the top five.

Null Hypothesis ($H_0$2): The proportion of startups located in major metropolitan cities is 50% of the total population.

A single-sample proportion Z-test was applied, using the observed data: 112 out of 236 startups are in major metros, which corresponds to an observed proportion of approximately 47%

Test	**Test Statistic**	**P-value**	**Alpha (α)**	**Decision**
Single-sample proportion Z-test	-0.782	0.434	0.05	**Fail to Reject H_0**

Table: Single sample Proportion Z-test

The observed proportion of startups in major metropolitan cities (Delhi, Mumbai, Bangalore, Chennai, etc.) is 47% (112 out of 236).

Since the p-value is large, there is not enough evidence to say that the true proportion differs from 50%. In other words, although the observed value (47%) is slightly below 50%, this difference could easily be due to sampling variability, so the decision is to "Fail to Reject $H_0$2." Statistically, the data are consistent with the claim that around half of the startups are based in major metro cities.

There is **not enough statistical evidence** to conclude that the proportion of startups located in major metropolitan cities is statistically different from 50%

Decision: Fail to Reject H_0

Null Hypothesis (H_03): There is no statistically significant association between the Incubation Center type and the Sector of the startup.

Test	Test Statistic	P-value	Alpha (α)	Decision
Chi-Square Test of Independence	9.238	0.100	0.05	**Fail to Reject H_0**

Table: Chi-Square test of Independence

A Chi-Square Test of Independence was conducted using a contingency table that cross-classifies startups by incubator type and sector category. The test statistic was 9.238 with a p-value of 0.100, still larger than α=0.05

Because the p-value exceeds 0.05, the decision is again to "Fail to Reject H03$\boxed{H03}$." This indicates there is no statistically significant evidence of an association between the type of incubator (IIT vs. non-IIT) and whether a startup belongs to one of the top five sectors or to other sectors. In practical terms, for this dataset, the choice of sector by startups appears independent of whether they are incubated at IIT-affiliated centers or at other incubation centers.

There is **no statistically significant association** between the broad **Incubation Center type** (IIT-Affiliated vs. Non-IIT) and the **Sector** of the startup (Top 5 vs. Other Sectors). The sector a startup operates in appears to be independent of whether it is affiliated with an IIT incubator in this dataset.

Decision: Fail to Reject H_0

6. INFERENCES

The analysis of the dataset of 236 startups provides several important inferences about sectoral focus, geographical clustering, and the role of incubation centres in India's startup ecosystem, along with evidence from hypothesis testing that these patterns are statistically meaningful.

First, the basic frequency statistics show that the sample is highly diverse in terms of both sectors and locations. With 171 unique sectors and 79 unique locations represented by 236 startups, the ecosystem captured in this dataset is not dominated by just a few broad categories but instead spans many niche and emerging domains. This reflects the wider Indian trend of startups moving beyond traditional IT services into specialized verticals such as health tech, agritech, fintech, industrial automation, and fitness technology. National reports on India's startup landscape similarly emphasise that innovation is now spread across multiple knowledge-intensive sectors rather than being confined to a narrow band of industries.

Second, the "Top 10 Startup Sectors" table and figure indicate a pronounced sectoral skew towards healthcare-related and technology-enabled services. Healthcare alone accounts for 25 startups, making it far more common than any other single sector. ICT Electronics, Education, and Agritech follow at a much lower but equal level, each with 5 startups, while Digital Health, IoT, Digital Health Tech, Healthtech, and EdTech fill out the rest of the top ten with 3–4 startups each. This pattern suggests two things: a strong concentration in health and wellness (through both traditional healthcare and multiple digital health subcategories) and a pervasive layer of enabling digital technologies (IoT, ICT, EdTech). At the ecosystem level, this aligns with recent evidence that healthtech, edtech, agritech and other tech-led verticals are among the fastest-growing areas in India, boosted by digital public infrastructure, rising internet penetration, and increasing investor focus on scalable tech solutions. The chi-square goodness-of-fit test formally confirms that the distribution across the top five sectors is not uniform, with an extremely small p-value indicating that the dominance of Healthcare is not due to random fluctuation but reflects a real underlying concentration.

Third, the "Top 10 Incubation Centres" table shows that incubation support is similarly concentrated in a limited number of institutional hubs. CIIE Initiatives hosts 12 startups, SINE–IIT Bombay supports 10, and a group of prominent incubators—VITTBI, Pilani IEDC, C-CAMP, Forge, TIDES–IIT Roorkee, AIC Pinnacle, JECRC Incubation Centre, and AIC@36Inc—each nurture 8 or 9 startups. This implies that a small set of highly active incubators play a disproportionate role in shaping the pipeline of early-stage ventures in the dataset. It also mirrors the national picture, where government- and university-backed incubators, particularly those linked to leading institutions and initiatives such as Startup India and Atal Incubation Centres, form the backbone of formal support for entrepreneurs. However, the hypothesis test on incubator type versus sector suggests that, within this sample, being associated with an IIT-affiliated incubator does not systematically push startups into particular sectors; sector choice appears broadly independent of whether an incubator is IIT-linked or not.

Fourth, the "Top 10 Startup Locations" table and figure reveal clear geographic clustering in major urban technology and business hubs. Bangalore leads with 29 startups, followed closely by Chennai (23) and Delhi (22), with Pune (13) and Hyderabad (11) forming the next tier. Mumbai, Kanpur, Jaipur, Ahmedabad and Raipur appear as smaller but still significant centres. This distribution is consistent with broader ecosystem data that identify Bengaluru, Delhi-NCR, and Mumbai as India's core startup hubs, with other cities such as Pune, Hyderabad, and Chennai emerging as strong secondary clusters. The earlier proportion test, which found that approximately 47% of startups are in major metropolitan cities and that this proportion is not statistically different from 50%, supports the inference that about half of the ventures in this dataset are metro-based. At the same time, the presence of 79 unique locations shows that entrepreneurship is also diffusing into non-metro and regional centres, echoing national trends of increasing activity in Tier-2 and Tier-3 cities.

Finally, when taken together, the dataset description, frequency counts, top-sector and top-location tables, and hypothesis tests paint a coherent picture of India's startup ecosystem at a micro level. The sample underscores strong thematic specialization in healthcare and digital technologies, heavy reliance on a network of leading incubators, and spatial concentration in a few metropolitan and technology hubs, all embedded within a broader fabric of diverse sectors and locations. These micro-level findings are broadly aligned with macro-level studies showing that India has become the world's third-largest startup ecosystem, characterised by rapid growth, sectoral diversification, and a growing but still uneven spread of entrepreneurial activity across regions

5. SUMMARY AND CONCLUSIONS

The analysis of the 236-startup dataset shows a focused yet diverse snapshot of India's innovation ecosystem, with clear patterns in sectoral specialization, institutional support, and geographical concentration. At the same time, the hypothesis tests confirm that these patterns are not random: some distributions are significantly skewed (such as sector concentration), while others (such as metro vs non-metro presence, or incubator type vs sector) appear more balanced.

Sectorally, the most striking feature is the dominance of healthcare and health-adjacent domains. Healthcare alone accounts for 25 startups, far ahead of any other sector, and it is reinforced by related categories such as Digital Health, Digital Health Tech, Healthtech, and even fitness- or wellness-oriented technology. In parallel, ICT Electronics, IoT, EdTech, Education, and Agritech indicate that a strong layer of enabling digital and deep-tech capabilities underpins many of these ventures. The chi-square goodness-of-fit test for the top five sectors confirms that this pattern is statistically non-uniform: the overrepresentation of Healthcare is too large to be explained by chance. This suggests that founders and incubators are consciously prioritizing health and technology-driven impact areas, likely reflecting both market demand (e.g., healthcare access, digital services) and investor and policy focus on these sectors.

From an institutional perspective, incubation support is clearly concentrated in a relatively small set of highly active centers. CIIE Initiatives, SINE–IIT Bombay, VITTBI, Pilani IEDC, C-CAMP, Forge, TIDES–IIT Roorkee, AIC Pinnacle, JECRC, and AIC@36Inc together host a substantial share of startups in the sample. This concentration implies that a handful of well-resourced incubators are acting as anchor institutions, providing mentorship, networks, and early-stage support at scale. However, the chi-square test of independence between incubator type (IIT vs non-IIT) and sector (top 5 vs others) finds no statistically significant association. In practical terms, this means that while certain incubators are large and influential, they are not narrowly channeling startups into specific sectors; instead, both IIT and non-IIT incubators appear to support a broad mix of domains.

Geographically, the dataset mirrors the national picture of startup clustering in major urban hubs while also capturing diffusion into secondary cities. Bangalore, Chennai, and Delhi together account for a large portion of the startups, with Pune and Hyderabad forming a second tier and cities like Mumbai, Kanpur, Jaipur, Ahmedabad, and Raipur contributing smaller but meaningful numbers. A single-sample proportion test on metro locations shows that about 47% of startups are in major metropolitan cities and that this share is not statistically different from a hypothesized 50%. This indicates that, although activity is clearly concentrated in big cities, nearly half of the ventures in the dataset are now located outside the core metros, consistent with the broader shift towards growth in Tier-2 and Tier-3 centres.

In conclusion, the dataset portrays an ecosystem that is simultaneously concentrated and diverse: concentrated in healthcare and digital technologies, in a core set of incubators, and in a few leading cities, yet diverse across 171 sectors and 79 locations. The statistical tests reinforce that sectoral concentration is a real structural feature, while metro presence and incubator type are more evenly distributed than might be assumed. For policymakers and ecosystem builders, these findings highlight three priorities: continue strengthening high-performing incubators as regional hubs, support emerging sectors beyond healthcare to avoid overconcentration, and deepen infrastructure and funding access in non-metro locations where entrepreneurial activity is already gaining momentum.

REFERENCES

[1] Singh, R., and Sharma, P. 2023 "Platform Cooperativism in Indian Gig Economy Startups." Journal of Business Research 162: 113–129.

[2] Kumar, V., and Lahiri, A. 2022 "Unit Economics of Indian SaaS Startups Post-Funding Winter." Strategic Entrepreneurship Journal 16(4): 789–812

[3] Gupta, S., and Jain, M. 2024 "Carbon Footprint of Indian E-commerce Logistics." Transportation Research Part E 182: 103–119.

[4] Rao, P., and Thakur, R. 2021. "Financial Inclusion Impact of Indian Fintech Startups." World Development 141: 105–118.

[5] Mishra, A., and Patel, N. 2023. "Circular Economy Practices in Indian Smartphone Refurbishment Startups." Journal of Cleaner Production 392: 136–152.

[6] Bansal, R., and Singh, S. 2025. "ESG Integration and Valuation Premium in Indian Unicorns." Emerging Markets Review 62: 101–117.

[7] Joshi, K., and Verma, S. 2022. "Edtech Business Models during and Post-COVID in India." International Journal of Educational Development 94: 102–115.

[8] Nair, G., and Menon, D. 2024. "Hyperlocal Delivery Startups and Plastic Waste." Resources, Conservation and Recycling 200: 107–122.

[9] Tiwari, P., and Bhat, A. K. 2023. "Agritech Startups and Farmer Producer Organisations." Agricultural Systems 206: 103–118.

[10] Sharma, A., and Goel, S. 2021. "Impact of GST on E-commerce Cost Structure." Margin: The Journal of Applied Economic Research* 15(3): 321–348.

[11] Pratap, S., and George, R. 2024. "Telemedicine Startups and Rural Health Inclusion in India." Health Policy and Planning 39(2): 112–128.

[12] Reddy, K., and Iyer, V. 2023. "Quick Commerce Dark Stores and Energy Intensity." Energy Policy 181: 113–129.

[13] Malhotra, A., and Kapoor, R. 2024. "D2CBrand Sustainability in Indian Fashion Tech." Journal of Fashion Marketing and Management 28(4): 589–607.

[14] Das, P., and Sen, M. 2022. "Cloud Kitchen Economics and Food Waste in Metro India." Waste Management 142: 88–99.

[15] Khan, I., and Rao, S. 2025. "EV Two-Wheeler Leasing Startups and Battery Swapping Models." Transportation Research Part D 126: 103–119.