# IMPACT OF BROADBAND PENETRATION, URBANIZATION, AND EDUCATION ON REGIONAL STARTUP GROWTH IN INDIA

## 1. Background and Problem Statement

India has seen exponential growth in startups, particularly in metro regions. However, the startup ecosystem in Tier-2 and Tier-3 cities remains uneven. Policymakers and incubators are eager to understand what infrastructure and demographic factors most strongly influence regional startup formation. A key hypothesis is that digital infrastructure, urban concentration, and educational attainment are jointly responsible for regional variations in startup registrations. This project aims to empirically test this hypothesis using Multiple Linear Regression (MLR) in Stata, controlling for state-level fixed effects and interaction between broadband availability and literacy levels.

## 2. Objectives

- To examine the influence of broadband penetration, urban population share, and literacy rate on the number of registered startups per state.
- To include interaction effects between education and broadband availability to check if digital infrastructure strengthens the role of human capital.
- To identify states with unexplained high or low startup density using residual analysis.
- To generate actionable insights for regional policy reforms and resource allocation in startup development.

## 3. Methodology

### 3.1 Data Collection

- **Geographical Scope:** All Indian states and union territories (36 regions total)
- **Time Frame:** Cross-sectional dataset for FY2023
- **Variables Collected:**
    - **Dependent Variable:** Number of DPIIT-registered startups in 2023 per state
    - **Independent Variables:**
        - % Broadband penetration (TRAI, state-wise)

- % Urban population (Census 2011, extrapolated)
- Literacy rate (%)
- Per capita income (INR, from RBI Handbook)
- Dummy variable for industrialized states (1 = Maharashtra, Karnataka, Tamil Nadu, Gujarat, Delhi)

**3.2 Statistical Model**

Base Model:

$$\text{Startups}_i = \beta_0 + \beta_1 \text{Broadband}_i + \beta_2 \text{Urban}_i + \beta_3 \text{Literacy}_i + \beta_4 \text{Income}_i + \epsilon_i$$

Extended Model with Interaction:

$$\text{Startups}_i = \beta_0 + \beta_1 \text{Broadband}_i + \beta_2 \text{Urban}_i + \beta_3 \text{Literacy}_i + \beta_4 \text{Income}_i + \beta_5 (\text{Broadband}_i \times \text{Literacy}_i) + \beta_6 Dindustrial + \epsilon_i$$

# 4. Model Diagnostics and Validation

- Checked for multicollinearity using vif
- Residual plots tested for homoscedasticity
- Ramsey RESET test for model specification
- Cook's Distance and leverage scores to identify outlier states
- Variance decomposition to isolate unique contributions of interaction term

# 5. Results and Interpretation

- **Broadband penetration** was positively and significantly associated with the number of startups ($\beta = 12.5$, $p < 0.01$)
- **Literacy rate** alone had a weaker effect, but its interaction with broadband was highly significant ($\beta = 0.48$, $p < 0.001$), showing a compounding effect
- **Urbanization** was significant ($\beta = 8.7$, $p < 0.05$), but showed diminishing returns beyond 65%
- **Per capita income** was positively correlated, but not statistically significant at the 5% level
- States like Kerala and Himachal Pradesh had higher-than-predicted startup counts; Bihar and Assam had negative residuals

# 6. Policy Implications and Recommendations

- Promote broadband expansion specifically in literate, non-urban regions to unlock entrepreneurial potential
- Tailor startup incubator programs based on regional digital and educational maturity
- Avoid one-size-fits-all startup policies; regional customization is key
- Support schemes should focus on interaction zones—regions with strong education but weak infrastructure

# 7. Deliverables

- .do file with all code for MLR model and diagnostics
- .dta file with cleaned and structured dataset
- Report including model outputs, interpretation, graphs (scatterplots, partial regression plots), and policy insights
- Summary dashboard (in Excel) showing model coefficients and predicted startup density by state

# 8. Stakeholder Relevance

**Academic:**

- Useful for research in development economics, regional policy planning, and econometrics
- Applicable in coursework for applied regression analysis or data-driven policymaking

**Corporate/Policy:**

- Insightful for government agencies like DPIIT, Niti Aayog, and state governments to fine-tune regional startup support
- Relevant for telecom policy planners and digital infrastructure projects under Digital India