The Geography of Privatization: Analyzing Institutional Concentration in Indian Higher Education

Sector : Education

Team Details : Team G-1 (Section-D)

1. Gauri Mehrotra (2401010170)
2. Milind Bansal (2401010277)
3. Sambhav Kumar (2401010409)
4. Abhishek Verma (2401010023)
5. Abhishek Kumar Patel (2401010020)
6. Himank Kaushik (2401010187)

Institute : Newton School of Technology

Faculty : Mr. Archit Raj

Executive Summary

Problem Statement : Investigating whether the growth of private colleges in India is regionally concentrated in specific states or evenly distributed across the country.

Approach : A comparative analysis of higher education institutions in India, mapping institutional management types against state-level and regional demographics.

Key Insights :

- Private institutions account for 62.74% of total colleges, indicating private sector dominance.
- Growth of private universities is uneven and concentrated in specific states such as Telangana, Karnataka, and Maharashtra.
- Private institutions have driven the majority of expansion since 2000.
- Forecasting suggests continued growth, primarily led by private institutions.

Sector & Business Context

Sector Overview : India has one of the largest higher education systems globally, seeing a massive shift toward private ownership in the last decade.

Current Challenges : State-wise and region-wise disparities in the expansion of private universities remain a significant challenge, with growth concentrated in select states and urban centers rather than being evenly distributed across the country.

Why this problem was chosen : To determine if private market forces are naturally excluding specific geographic populations.

Problem Statement & Objectives

Problem Statement : Does the increase in private institutional participation correlate with higher geographic concentration in specific Indian states or locations?

Project Scope : Analysis of higher education institutions in India, with a focus on private vs. state and central universities.

Success Criteria : Identification of the correlation between private ownership and state-wise as well as rural–urban distribution of institutions.

Data Description

Dataset Source : Dataset Link

Data Structure : Relational tables including Institution names, State, District, and Management.

Columns : (14 columns)

- **State**
- **District**
- **University Type**
- **University Name**
- **College Name**
- **College Type**
- **Address**
- **Website**
- **Management**
- **Year of Establishment**

- **Specialised in**
- **Location**
- **Upload Year**
- **The Institute Added In Survey Year**

Data Size :

- **Original dataset** - 47590 rows
- **Working dataset** - 9000 rows

Limitations : Missing data for certain newly formed districts or private colleges not yet registered.

Data Cleaning & Preparation

- **Dataset Sampling:** Stratified the raw dataset and selected 9,000 rows as the working dataset using a Python script.(script reference in Appendix)
- **Handling Missing Values:** Replaced empty cells with "Not Specified" or "Not Defined" to ensure data completeness.
- **Data Type Formatting:** Converted columns to their appropriate data types for consistency and accurate analysis.
- **Data Splitting:** Separated University ID from University Name and College ID from College Name using a delimiter-based split.

KPI and Metric Framework

- ## KPI 1: Private College Share (%)

Formula:
Private College Share (%) = (Number of Private Colleges / Total Number of Colleges) × 100

Purpose:
This KPI indicates the level of privatisation and the dominance of private institutions.

Insight:
Private institutions form the majority of colleges, indicating strong private sector participation.

- ## KPI 2: Private-to-Government College Ratio

Formula:
Private-to-Government Ratio = Number of Private Colleges / Number of Government Colleges

Purpose:
Shows ownership balance and privatisation intensity.

Insight:
There are nearly 2 private colleges for every 1 government college.

- ## KPI 3: Private College Share by Location (Urban vs Rural)

  Formula:

  Urban Private Share (%)
  = (Private Colleges in Urban / Total Urban Colleges) × 100

  Rural Private Share (%)
  = (Private Colleges in Rural / Total Rural Colleges) × 100

  Purpose:
  Shows geographic distribution of privatisation.

  Insight:
  Private colleges dominate both rural and urban areas, with stronger presence in rural areas.

- ## KPI 4: Education Privatization Growth Trend

  Formula:
  Privatisation Trend = Private Colleges per Year / Total Colleges per Year

  Purpose:
  Identifies growth patterns and structural shifts.

  Insight:
  Private colleges have increased significantly over time.

- ## KPI 5: College Growth Forecast (2030)

  Method:
  Linear trend forecasting using establishment year data.

Purpose:
Shows expected future expansion.

Insight:
Higher education institutions are expected to continue growing.

- ## KPI 6: College Distribution Across States

Formula:
College Density = Count of Colleges per State

Purpose:
Identifies regional concentration.

Insight:
Some states have significantly higher college concentration.

- ## KPI 7: Ownership Shift Correlation (Private vs Government Growth)

Formula:
Correlation (Private, Government)

Purpose:
Shows structural expansion pattern.

Insight:
Both sectors are growing, but private growth is stronger.

- **KPI 8: Privatisation vs Urbanisation Correlation**

  Formula:
  Correlation (Private Colleges, Urban Colleges)

  Purpose:
  Shows geographic spread.

  Insight:
  Private colleges are not limited to urban areas.

Exploratory Data Analysis (EDA)

- **Trend Analysis:** A line chart showing the privatisation trend over the years.
- **Comparative Analysis:** A bar chart comparing the all states on the basis of their Private-to-Government ratio.
- **Distribution Analysis:** A bar chart illustrating the distribution of different types of colleges across rural and urban locations.
- **Correlation Analysis:**
  - The KPI cards indicate that privatisation has a **moderate association with geographic concentration** and a **strong positive growth trend over time**, while its relationship with **urban locality is very weak and slightly negative**, and the **variation across states and management types remains weak**, suggesting that privatisation growth is widespread rather than limited to specific regions or management structures.

Advanced Analysis

- ## Forecasting Analysis:
  - **Forecasting was performed using historical college establishment data to estimate future growth.**
    - The forecast indicates that the number of colleges is expected to increase steadily by 2030, reaching approximately 173 colleges per year.
    - This suggests continued expansion of higher education, driven primarily by private sector participation.
    - The decade-wise growth chart shows cyclical spikes, indicating that college establishment increases in phases rather than uniformly.
    - If the current trend continues, private institutions will remain the dominant contributors to expansion.

- ## Segmentation Analysis:

-> Segmentation was performed using -

- Management Type
- Location (Urban vs Rural)
- State-wise distribution
- College type

-> Key Insights -

- **Management Segmentation -**
  - Private colleges account for 62.74% of total colleges
  - Government colleges account for only 37.26%
  - Private-to-Government ratio is 1.9954, meaning almost 2 private colleges exist for every government college
- **Location Segmentation -**
  - Private colleges are more concentrated in urban areas

        - Rural areas have relatively fewer institutions, indicating access inequality
    - **State Segmentation -**
        - Some states show very high private-to-government ratios, indicating strong privatization
        - Other states rely more on government institutions
    - **College Type Segmentation -**
        - Affiliated colleges dominate the system
        - Autonomous and constituent colleges are fewer

- ## Root Cause Analysis:

Root cause analysis was conducted to understand why privatization is increasing.

-> Root Causes Identified:

- **Government capacity limitations -**
    - Government cannot expand fast enough to meet rising demand
- **Increasing demand for higher education -**
    - More students seeking degrees
    - Private sector filling the gap
- **Policy support -**
    - Liberal policies encouraging private investment in education

- ## Risk and Anomaly Analysis:

Risk analysis helps identify potential concerns.

-> Key Risks Identified:

- **Over-privatization -**
    - Private share already at 62.74%
    - May reduce affordability and accessibility
- **Regional imbalance -**
    - Some states have significantly higher private ratios
    - Leads to unequal access

-> Anomalies observed:

- Sudden spikes in certain decades in growth chart

- Certain states showing extremely high private ratios compared to others

- **Scenario Analysis**:
  - **Scenario 1: If current trend continues-**
    - Private colleges will dominate even more
    - Private share may exceed 70% by 2030
  - **Scenario 2: If government increases investment-**
    - Government college share will improve
    - Education accessibility will become more balanced
  - **Scenario 3: If private growth slows-**
    - Overall college growth rate will reduce

- **Final Conclusion:**

Advanced analysis reveals that India's higher education expansion is strongly privatization-driven, uneven across regions, and expected to grow further, highlighting the need for balanced policy intervention.

Dashboard Design

- **Dashboard Implementation**:

The dashboard was implemented using Google Sheets, leveraging pivot tables, calculated formulas, charts, and interactive slicers to analyse institutional concentration and privatisation trends in Indian higher education.

Key Google Sheets features used:

- Pivot Tables → for aggregation by state, management, location, and year
- Calculated Fields → for KPIs such as private share, ratios, and correlations

- Charts → line charts, bar charts, and pie charts
- Slicers / Filters → for interactive analysis

This enabled dynamic and user-driven exploration of the dataset.

• Dashboard Objective:

The primary objective of the dashboard is to:

- Analyse the extent of privatisation in Indian higher education
- Identify geographic concentration of institutions
- Compare private and government institutional presence
- Understand ownership trends over time
- Forecast future institutional growth

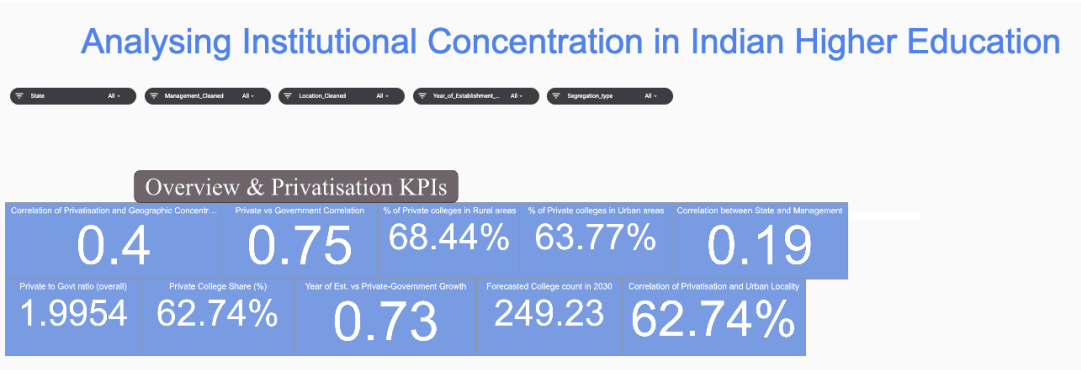
This dashboard helps decision-makers understand structural patterns and supports education planning.

• Dashboard View Structure:

The dashboard is organised into four major analytical sections:

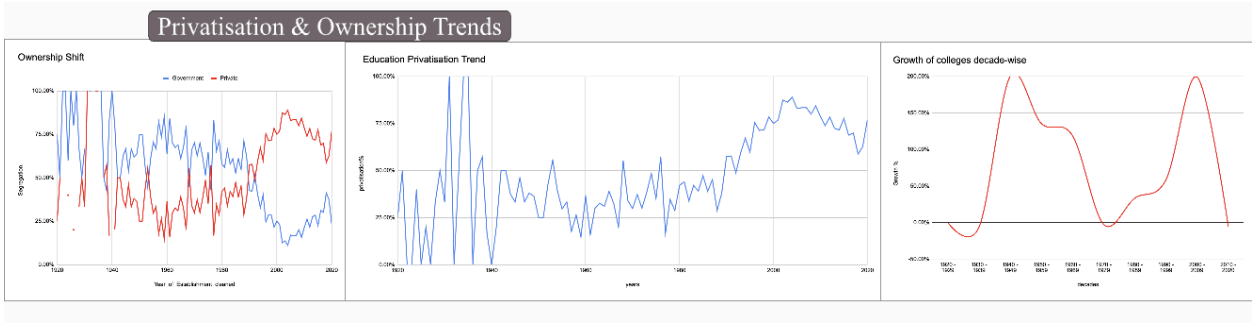
- **Overview & Privatisation KPIs :** This section provides high-level summary indicators such as:
    - Private College Share: 62.74%
    - Private-to-Government Ratio: 1.99
    - % Private Colleges in Rural Areas: 68.44%
    - % Private Colleges in Urban Areas: 63.77%
    - Forecasted College Count: 249.23

    These KPIs highlight the strong dominance of private institutions.



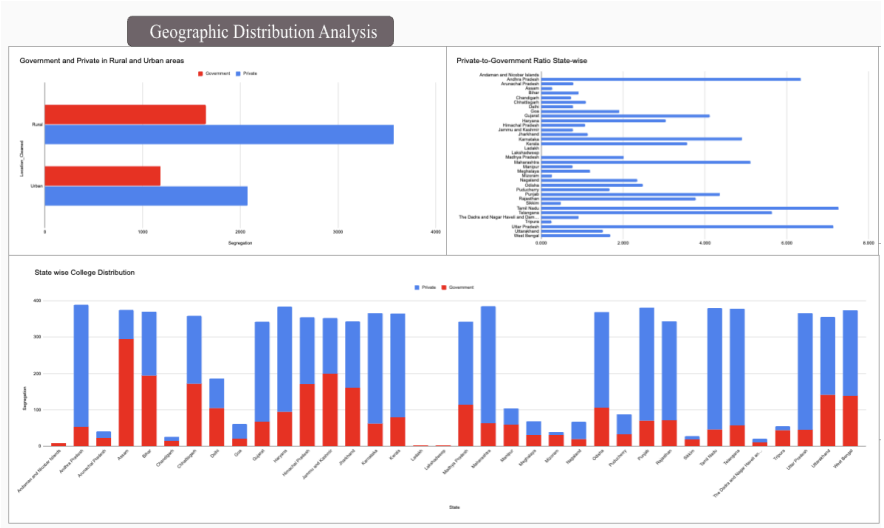
- **<u>Privatisation & Ownership Trends :</u>** This section uses line charts to show:
    - Growth of private vs government institutions over time
    - Ownership shift trends
    - Decade-wise institutional growth

<u>Insight:</u> Private institutions have increased significantly, especially after 2000.



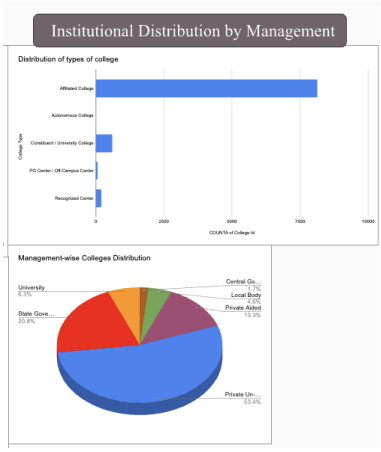
- **<u>Geographic Distribution Analysis:</u>** This section includes:
    - State-wise college distribution
    - Private-to-government ratio by state
    - Rural vs urban distribution

<u>Insight:</u> Certain states show high institutional concentration and private dominance.



- **Institutional Distribution by Management:** This section shows:
    - Management-wise distribution (Private, Government, Aided)
    - College type distribution

  Insight: Private unaided institutions form the majority.



- **Filters and Drilldowns:**

  Interactive filters allow users to dynamically analyse data by:

- State
- College Type
- Management Type
- Location (Rural / Urban)
- Year of Establishment

These filters enable detailed drill-down analysis.

- **Dashboard Explanation:**

The dashboard provides a comprehensive view of institutional concentration.

Key findings include:

- Private institutions dominate Indian higher education
- Institutional growth has accelerated in recent decades
- Rural areas also show strong private presence
- Institutional distribution varies significantly across states
- Private sector growth is expected to continue

The dashboard enables easy interpretation through visualisations and interactive exploration.

Insights Summary

- **Private institutions are the primary drivers of higher education expansion**
  With 62.74% of colleges under private management, policymakers must recognize that higher education growth in India is largely dependent on private sector participation rather than government expansion.

- **India's higher education system is structurally shifting toward privatization**

The Private-to-Government ratio of 1.99 indicates a significant structural shift, requiring regulatory frameworks to ensure quality and affordability in private institutions.

- **Private college growth is uneven and concentrated in specific states**
  State-wise analysis shows higher private concentration in select states, indicating the need for region-specific education planning rather than uniform national policies.

- **Certain time periods show rapid expansion, indicating policy influence**
  Growth spikes suggest that government policies directly influence higher education expansion, highlighting the importance of strategic planning.

Recommendations

- **Recommendation 1: Increase government colleges in states with high private dominance**
  - **Mapped Insight:** Private institutions dominate higher education, with a Private-to-Government ratio of 1.99.
  - **Recommendation:** The government should establish more public colleges in states where private institutions significantly outnumber government institutions.
  - **Business Impact:**
    - Improves affordability of higher education
    - Reduces over-dependence on private institutions
    - Ensures equal access to education
  - **Feasibility:**
    - High feasibility through phased expansion
    - Can be implemented via central and state government education programs

- **Recommendation 2: Focus on expanding higher education infrastructure in rural areas**
  - **Mapped Insight:** Private colleges are concentrated in urban areas, creating rural access gaps.
  - **Recommendation:** Government should prioritize establishing colleges in rural and semi-urban regions.
  - **Business Impact:**
    - Improves rural education access
    - Increases student enrollment
    - Supports regional economic development
  - **Feasibility:**
    - Moderate feasibility
    - Can be implemented through targeted rural education schemes

- **Recommendation 3: Implement stronger regulation and monitoring of private institutions**
  - **Mapped Insight:** Rapid privatization increases the need for regulatory oversight.
  - **Recommendation:** Introduce stricter quality standards, fee regulations, and accreditation policies for private institutions.
  - **Business Impact:**
    - Ensures education quality
    - Prevents excessive fee structures
    - Protects student interests
  - **Feasibility:**
    - High feasibility
    - Can be implemented through existing regulatory bodies such as UGC and AICTE

- **Recommendation 4: Promote Public-Private Partnerships (PPP) in education**
    - **Mapped Insight:**Private sector is driving most higher education expansion.
    - **Recommendation:**Encourage collaboration between government and private institutions.
    - **Business Impact:**
        - Faster infrastructure development
        - Reduced government investment burden
          Improved quality and innovation
    - **Feasibility:**
        - High feasibility
        - Already successfully used in multiple sectors

Conclusion : These recommendations aim to ensure balanced, accessible, and sustainable growth of higher education in India while addressing regional and privatization challenges.

Impact Estimation

- **Save Cost:** Optimizes government spending by identifying priority states for public college expansion.

- **Improve Efficiency:** Enables faster and more accurate education infrastructure planning using real-time insights.

- **Improve Service:** Improves access to higher education by highlighting underserved rural and state regions.

- **Reduce Risk:** Helps prevent excessive privatization and ensures balanced and affordable education growth.

Limitations

- **Data Issues:**
  - No data on students, faculty, or quality
  - Possible missing or incomplete records
  - No status (active/closed) of institutions
- **Assumption Risks:**
  - More colleges ≠ better access or quality
  - Rural/Urban classification assumed correct
  - All colleges assumed equal in size/capacity
- **Cannot Be Concluded:**
  - Education quality differences
  - Student outcomes or placements
  - Reasons behind private sector dominance

Future Scope

- **More Analysis:**
  - State-wise expansion analysis
  - Trend analysis over years
  - Specialisation growth analysis
- **New Data Needed:**
  - Student enrollment
  - Placement data
  - Faculty and infrastructure data

Conclusion

- Private colleges dominate overall
- Private sector drives higher education expansion
- Higher education has expanded geographically

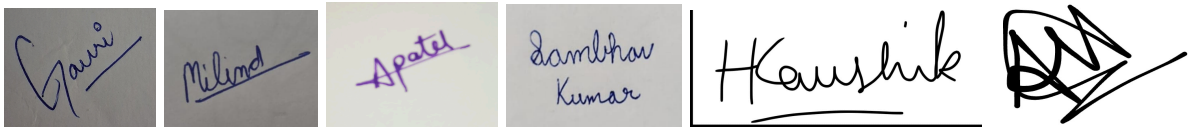
Contribution Matrix

Team Member	Dataset & Sourcing	Cleaning	KPI & Analysis	Dashboard	Report Writing	PPT	Overall Role
Gauri Mehrotra	✓	—	—	—	✓	✓	Team Lead
Milind Bansal	—	✓	✓	✓	—	—	Data Insights Lead
Sambhav Kumar	—	—	✓	✓	—	—	Dashboard Architect
Abhishek Verma	—	—	—	—	✓	✓	Presentation Designer

Abhishek Kumar Patel	✓	—	—	—	—	✓	Presentation Reviewer
Himank Kaushik	—	✓	—	—	—	—	Data Cleaning Help

Declaration: We confirm that the above contribution details are accurate and verifiable through version history and submitted artifacts.

Team Sign Block :



Appendix

Python Script for randomisation :

```
import pandas as pd
df = pd.read_csv("Cleaned College.csv")
TARGET = 9000
MIN_PER_STATE = 30
state_counts = df['State'].value_counts()
states = state_counts.index
# Step 1 — initial allocation
```

```
allocation = {state: min(MIN_PER_STATE, state_counts[state]) for state in states}
current_total = sum(allocation.values())
remaining = TARGET - current_total
# Step 2 — distribute remaining proportionally to states that still have data
capacity = {state: state_counts[state] - allocation[state] for state in states}
while remaining > 0:
    for state in states:
        if capacity[state] > 0:
            allocation[state] += 1
            capacity[state] -= 1
            remaining -= 1
            if remaining == 0:
                break
# Step 3 — sample
sampled_df = pd.concat([
    df[df['State'] == state].sample(n=allocation[state], random_state=42)
    for state in states
])
print("Final rows:", len(sampled_df))
print(sampled_df['State'].value_counts())
```

Data Dictionary :

Column Name	Data Type	Variable Type	Description	Example	Missing Values
State	**String**	**Categorical (Nominal)**	**Name of the State or Union Territory where the institute is located**	**Uttar Pradesh**	**Possible**
District	**String**	**Categorical (Nominal)**	**District where the institute is located**	**Jaipur**	**Possible**

University Type	String	Categorical (Nominal)	Type of university affiliating the college	State Public University	Few
University Name	String	Categorical (Nominal)	Name of the affiliating university	Pondicherry University	Few
College Name	String	Categorical (Nominal)	Official name of the college/institute	Tagore Government College of Education	No
College Type	String	Categorical (Nominal)	Institutional affiliation type	Affiliated College	Few
Address	String	Text	Full postal address of the institute	Port Blair, Andaman & Nicobar	Yes
Website	String (URL)	Text	Official website of the institute	www.jnrm.and.nic.in	Yes
Management	String	Categorical (Nominal)	Type of management controlling the institute	Private Un-Aided	No
Year of Establishment	Integer	Numerical (Discrete)	Year the institute was established	1983	Few
Specialised in	String	Categorical (Nominal)	Area of specialization of the institute	Engineering, Medical	Yes
Location	String	Categorical (Nominal)	Type of geographical location	Urban / Rural	Few

Upload Year	**Integer**	**Numerical (Discrete)**	**Year when institute record was uploaded**	**2020**	**Few**
The Institute Added In Survey Year	**Integer**	**Numerical (Discrete)**	**Year when institute was added to the survey database**	**2019**	**Few**