

Datathon for Data Driven Policy Innovation

Driving Evidence based interventions in Public Policy

Participant Brief

About the Dataset:

The Unified District Information System for Education Plus is an educational management information system under the Department of School Education & Literacy, Ministry of Education under the Government of India. UDISE+ functions as a central platform that enables respective schools to efficiently record and submit data related to its profile (infrastructure & facilities), individual students and teachers' details. This is done in real-time mode by all recognised schools that provide formal, and special education in any of the school categories (that is, from pre-primary to higher secondary levels) throughout the country each academic year. The recorded data is validated by the MIS and other designated officials at the block, district and state levels; and monitored at four levels, including the national.

This makes UDISE+ one of the largest management information systems for credible information collection mechanisms and serves as the official statistics of the Ministry of Education.

To know more:

<https://udiseplus.gov.in/#/en/home>

Refer to the UDISE booklet to understand the data collection system, limitations of UDISE, features of UDISE+ including data capturing process, data mapping etc.

Refer to the Data Capture 2024-25 document to understand what details have been captured at each level including school, states, and union territories.

Dataset:

You can find the UDISE 2023-24 and 2024-25 datasets in the Education Dataset folder. Refer to the Data Sharing Portal Schema documents to understand the schema for all the profiles.

Please download other publicly available data sets such as NFHS that may seem relevant for your analysis.

Each team is required to choose and respond to **at least two prompts** from the below provided list to structure their analysis and policy recommendations:

Data Analysis Prompts:

1. Which states show consistent improvement over the last three years? Which factors strongly influence high dropout rates?
2. Do dropout and retention patterns differ between urban and rural districts and when disaggregated by gender and caste?
3. How does availability of teachers and/or functional toilets influence retention?
4. How do pupil-teacher ratio and school facilities vary with enrolment levels?
5. Is there a correlation between poor school infrastructure and low student enrolment or high dropout rates?
6. How do household income, parental education, and employment status relate to student enrolment and dropout rates?
7. Is there a relationship between child nutrition or health indicators (e.g., anemia) and school attendance?
8. What is the impact of school type or management? Do schools with improved facilities (libraries, labs, digital access) show higher retention and transition rates?
9. How do enrolment and retention differ between boys and girls especially in aspirational districts?
10. What are the bottlenecks observed in transitioning from Upper Primary -> Secondary & Secondary -> Higher Secondary?

Key Deliverables:

Each team must upload the following materials to the submission link.

1. Document or Slide Presentation with:
 - Brief explanation of your modelling approach.
 - Rationale behind indicator selection and assumptions.
 - Minimum of 3 key insights or policy recommendations.
2. Link to Dashboard or Visualization
 - Developed using Tableau, PowerBI, or any preferred tool.
 - Must effectively communicate key findings to non-technical decision-makers.

Your primary audience is a District Magistrate or Policy Officer. So focus on clarity, context, and feasibility.

All the best :)