**CONCEPT NOTE  
Analysing School Dropout Rate data to address the problem of Educational Exclusion** (SDG 4: Quality Education)

**CONCEPT OF THE PROJECT**

Educational exclusion refers to being denied access to education, or being limited in ability to fully participate in it. This project aims to analyse School Dropout Rate data to understand the trends and possible factors leading to the same. By leveraging data analysis tools and methodologies, the project seeks to propose actionable solutions that align with Sustainable Development Goal 4 (SDG 4): Quality Education. This SDG aims to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.

.

**PROBLEM STATEMENT**

Some states in India are witnessing a higher dropout rate especially at the secondary level due to variety of reasons such as poverty, lack of engagement, financial obligation towards family (low-income groups). Children living in remote areas are more likely to drop out of their academics due to poor infrastructure and lack of basic amenities. Dropout rates have far reached impacts for both the individual and the society. Limited work potential, increased crime rates, strained social safety nets are all directly related to dropout population. Despite various measures, controlling dropout rate still remains a challenge due to the lack of precise data and effective policy implementation. This project seeks to address this problem by analysing school dropout data to identify sources and trends, and by proposing targeted interventions to improve the prevailing conditions.

**OBJECTIVE OF THE PROJECT**

The primary objective of this project is to analyse school dropout data to identify the major sources and trends of educational exclusion and to propose data-driven solutions that can help reduce the higher drop-out rates. The specific objectives are:

* To collect and analyse school dropout data from reliable sources.
* To identify the primary reasons of academic negligence in women and specific in each states.
* To understand the temporal and spatial trends of refined data.
* To propose actionable solutions and policy recommendations to mitigate Educational Exclusion.
* To assess the potential impact of these solutions on achieving SDG 4.

**DATA SOURCES USED**The project will use school dropout datasets from the following sources:

1. Education For All in India:Dropout Rates in Schools in India: An Analysis of UDISE+ 2021-22 Data.( https://educationforallinindia.com/dropout-rates-in-schools-in-india)
2. Open Government Data (OGD) Platform India (https://data.gov.in/catalog/dropout-rate)
3. Statistica : Average Annual drop-outs in schools in India . (https://www.statista.com/statistics/661218/average-annual-drop-outs-in-schools-india/)

**FEATURES**

The key features of the dataset will include:

* Level of Education to which the student belongs: Primary, Upper Primary and Secondary
* Gender of the student: Male or Female
* Year of Dropout (Data from recent years i.e 2019-2022)
* States: State-wise records of school dropout data.

**TOOL FOR ANALYSIS**

The following tools and technologies will be used for data analysis:

1. Ms Excel: For data cleaning, analysis, and visualization using in-built functions and charts.
2. Collab: Google Collab for executing python codes for data interpretation.
3. Matplotlib: For creative visualization and development of Interactive charts.
4. QGIS: For spatial analysis of dropout trends.
5. Tableau: For creating interactive dashboards and visualizations to present the findings.

**HYPOTHESIS**

The hypothesis of the project is that by analysing school dropout rate data, we can identify key factors that contribute to educational exclusion. This will allow us to develop targeted interventions that address these factors and ultimately reduce dropout rates and promote educational inclusion. This hypothesis suggests that there's a connection between specific factors found in the data (e.g., socioeconomic background, dropouts in different level of education which can be primary secondary etc, location) and the likelihood of a student dropping out. By understanding these connections, we can bring changes like tutoring programs, financial aid, strict regulatory commissions (to check for child labour), targeted Government schemes etc., that addresses these root causes and keep students engaged in school. Also, by studying the Data available we can predict the Dropouts in specific backgrounds, remote locations etc. and provide predictive analysis that will be vital to predict and reduce dropouts in the targeted regions with higher dropout rates. Additionally, specific temporal and spatial trends in School dropout rates can be identified and addressed effectively.

**METHODOLOGY**

The project will be conducted in the following phases:

Data Collection:

* Gather school dropout data from the above-mentioned sources.
* Compile supplementary data to support the analysis.
* Data Cleaning and Preprocessing.
* Handle missing values, outliers, and inconsistencies in the data.
* Standardize data formats and integrate datasets from different sources.

Exploratory Data Analysis (EDA)

* Perform descriptive statistical analysis to understand the school dropout rates among various communities, state.
* Visualize temporal trends (annual) and spatial distributions using charts and maps.

Source Identification

* Use correlation analysis and regression models to identify communities at a potential risk of school dropout
* Analyse impact of various factors like household income and communities most affected by dropouts.

Solution Development:

* Based on the analysis, propose solutions such as implementing policies and pouring government resource at places which are more affected.
* Assess the feasibility and potential impact of these solutions.

Reporting and Presentation:

* Compile the findings into a comprehensive report.
* Create visualizations and interactive dashboards to present the results.
* Develop policy briefs and recommendations for stakeholders (including Government).

**PROBABLE OUTCOME**

The expected outcomes of the project are:

* Comprehensive Analysis: Analysis of data over past few years in various states to observe trends and draw inference.
* Determining Hotspots: States in which dropout rates are higher serve as the hotspot.
* Target specific solutions: Data-driven solutions and policy recommendations to reduce school-dropout
* Impact Assessment: Evaluation of the potential impact of proposed solutions on achieving SDG 4.
* Awareness and Engagement: Increased awareness among policymakers and the stakeholders about the issue and drawing public attention towards this problem to pressurize concerned authorities

By addressing educational exclusion through data analysis and evidence-based solutions, this project will contribute to creating inclusive Education system, aligning with the objectives of SDG 4: inclusive and equitable quality education and promote lifelong learning opportunities for all.