**TN Marginal Workers Assessment**

**Project Definition:**

The project involves analyzing the demographic characteristics of marginal workers in Tamil Nadu based on their age, industrial category, and sex. The objective is to perform a socioeconomic analysis and create visualizations to represent the distribution of marginal workers across different categories. This project includes defining objectives, designing the analysis approach, selecting appropriate visualization types, and performing the analysis using Python and data visualization libraries.

**Design Thinking:**

Project Objectives: such as analyzing marginal worker demographics, understanding age and gender distribution, and exploring industrial categories.

Demographic Analysis:

Specific Tasks:

Collect data related to age, gender, education, and residence of marginal workers.

Explore the distribution of marginal workers across different age groups and genders.

Examine the educational backgrounds of marginal workers.

Assess the geographic distribution of marginal workers within Tamil Nadu.

Age and Gender Distribution:

Specific Tasks:

Create visualizations (e.g., histograms, bar charts) to illustrate the age distribution of marginal workers.

Analyze the gender composition of marginal workers and any gender-related trends.

Identify age and gender disparities within specific industrial categories.

Industrial Categories Exploration:

Specific Tasks:

Collect data on the types of industries or sectors where marginal workers are engaged.

Categorize and classify these industries into meaningful groups.

Analyze the distribution of marginal workers across different industrial categories.

Assess the level of informality within these industries.

Analysis Approach: Extract the data into a suitable format for analysis, such as CSV, Excel, or a database, and Remove duplicates, Verify data integrity and Standardize data those are steps for cleaning and Apply statistical and analytical techniques such as descriptive statistics, regression analysis, clustering, or classification, depending on our objectives.

Data Extraction and Import: Import the data into IBM Cognos in a suitable format for analysis, such as CSV, Excel, or by connecting to a database.

Remove Duplicates: Utilize IBM Cognos data cleansing tools to identify and eliminate duplicate records from our dataset.

Verify Data Integrity: Use data validation features in IBM Cognos to check for data integrity issues, ensuring that data conforms to expected standards.

Standardize Data: Implement data standardization techniques using IBM Cognos to ensure consistency in categorical data.

Analytical Techniques:Depending on your objectives, leverage IBM Cognos' analytics capabilities, which include:

Descriptive statistics: Generate summary statistics and metrics.

Regression analysis: Build regression models to explore relationships.

Clustering: Segment data into clusters.

Classification: Develop classification models for predictive analysis.

Visualization Selection: Based on the given dataset we will plan to visualize the data ( e.g, bar charts, pie charts, heatmaps) and provide required and meaningful report and dashboards by using IBM Cognos Analytics.