**Assessment of Marginal Workers in TamilNadu-A**

**Socioeconomic Analysis Phase-2**

**Name:** Hariprasath k

**Email id**:harikappini@gmail.com

**2.1 Short Explanation about the question (Assessment of marginal workers in Tamil Nadu):**

The question pertains to the assessment of marginal workers in the state of Tamil Nadu, India, using a specific dataset. Marginal workers are individuals who are engaged in irregular and low-paying employment. The goal is to analyze this dataset to gain insights into the characteristics and distribution of marginal workers in Tamil Nadu.

**2.2 Dataset and its details:**

The dataset was obtained from the official government data portal of Tamil Nadu: [https://tn.data.gov.in/resource/marginal-workers-classified-age-industrial-category-and-sex-scheduled-caste-2011-tamil](https://tn.data.gov.in/resource/marginal-workers-classified-age-industrial-category-and-sex-scheduled-caste-2011-tamil).

**2.3 Details about columns (columns which you're going to use):**

The dataset contains a wide range of columns. To assess marginal workers, you may want to focus on specific columns relevant to this analysis. Possible columns of interest could include "Total/ Rural/ Urban," "Age group," and various "Industrial Category" columns that describe the type of work. The specific columns to use will depend on your research questions and objectives.

**2.4 Details of libraries to be used and how to download:**

You can use popular data analysis libraries like Pandas, Matplotlib, and NumPy in Python to work with this dataset. To download the dataset, you can visit the provided URL and follow the download instructions on the data portal.

**2.5 How to train and test:**

Training and testing are typically associated with machine learning models. If you intend to build predictive models using this data, you would split the dataset into a training set and a testing set. For descriptive analysis, you can directly explore the dataset without the need for training and testing.

**2.6 Rest of the explanation:**

The rest of the analysis would involve data preprocessing, exploratory data analysis (EDA), and visualization to understand the characteristics of marginal workers in Tamil Nadu. You can derive insights, identify trends, and patterns in the data. Additionally, you might perform statistical analysis or apply machine learning techniques depending on the specific objectives of your assessment.

**2.7 What metrics used for the accuracy check:**

The choice of metrics for accuracy checks would depend on the type of analysis you are performing. For descriptive analysis, metrics such as mean, median, and standard deviation might be relevant. If you are building predictive models, common metrics include accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC), depending on the nature of the problem (classification or regression). The choice of metrics should align with the objectives of your analysis.