|  |
| --- |
| **TN (TAMIL NADU)**  **Marginal Workers Assessment** |

|  |
| --- |
| SUBMITTED BY ,  POLI TEJESWAR REDDY  au723921244040  politeja3@gmail.com |

Abstract :

**Project Definition:**The project involves analysing 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 According to our dataset we will design it.

**Design thinking:**

1. **Project Objectives:**

**a**. **Objective 1**: Analysing Marginal Worker Demographics

- Define the first specific objective related to demographics. For example: "To analyse the demographic profile of marginal workers, including their age, gender, educational background, and rural/urban distribution."

**b. Objective 2**: Understanding Age and Gender Distribution

- Specify the second objective focused on age and gender distribution: "To examine the age and gender distribution among marginal workers, identifying trends and variations in different regions of Tamil Nadu."

**c. Objective 3**: Exploring Industrial Categories

- State the third objective regarding industrial categories: "To explore the types of industries and sectors where marginal workers are employed, categorizing them based on formal and informal sectors."

**2. Analysis Approach:**

**1.Data Requirements**:

Marginal workers (workers who work less than 6 months) from the industries of all categories.

**2. Data Extraction:**

**a. Data Source Identification:**

Data is extracted from :https://tn.data.gov.in/resource/marginal-workers-classified-age-industrial-category-and-sex-scheduled-caste-2011-tamil

**b. Data Collection and Download:**

Collect the necessary data from the identified sources. This may involve downloading datasets in various formats, such as CSV, Excel, or database exports.

**3. Data Cleaning:**

**a. Data Cleaning Plan:**

Develop a data cleaning plan that outlines the specific actions you'll take to address issues like missing data, duplicates, outliers, and inconsistencies.

**b. Handling Missing Data:** Decide on a strategy for dealing with missing data. You can choose to impute missing values, remove rows with missing values, or use domain-specific methods.

**c. Handling Duplicates:**

Identify and remove duplicate records if they exist in your dataset.

**d. Outlier Treatment:**

Determine how to handle outliers in your dataset. You may choose to remove them or transform them based on the nature of your analysis.

**4. Data Analysis:**

**a. Statistical Analysis:**

Conduct statistical tests and analyses that are relevant to your research questions and hypotheses.

**4.Visualization Selection:**

Choosing the right visualization types which exactly fit to our data set.

1. **Pie chart**: Pie charts are suitable for displaying the parts of a whole. You can use them to show the percentage distribution of gender within a population, for instance. Be cautious with pie charts as they can be less effective at conveying precise information compared to bar charts for some audiences.
2. **Bar chart:** Use bar charts when comparing discrete categories such as age groups, education levels, or gender.