**PROJECT TITLE**

## **TN Marginal Workers Assessment**

**TEAM MEMBERS**

**PRAVEEN V 2021115076**

**RATISH MS 2021115085**

**SANJAY 2021115001**

**SARAVANAN 2021115325**

**HEMANTH P 2021115305**

**PHASE 1**

**1. Project Definition:**

The project involves analyzing the demographic characteristics of marginal workers in TamilNadu 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.

1. **Design Thinking for Analyzing Demographic Characteristics of Marginal Workers in TamilNadu:**

**2.1 Analysis Objectives:**

Before delving into the data analysis, it's essential to establish clear objectives for our project. In this phase, we define the following analysis objectives to provide a well defined direction for our analysis, focusing on key demographic factors related to marginal workers:

**\***Compare the age distribution of marginal workers within different industrial categories in Tamil Nadu.

**\***Examine the gender distribution among marginal workers across various age groups and industries in Tamil Nadu.

**2.2 Data Collection:**

To proceed with the analysis, we need to gather the necessary data. We will obtain the dataset containing information on demographic characteristics of marginal workers in Tamil Nadu. Data sources may include government reports, surveys, or other relevant sources. Data cleaning and preprocessing will be essential to handle any missing values, format inconsistencies, and outliers.

**Dataset Link:**  
**[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" \t "https://courses.myclass.skillup.online/courses/course-v1:IBM+DAC101+2023_B1/courseware/d8660830b7ec4f2e8158584fd8319a7d/f07fd2afa2f94df0803864c408bf4545/[object Object])**

**2.3 Visualization Strategy:**

Effective data visualization is crucial for conveying insights. In this phase, we plan how to visualize the demographic characteristics of marginal workers using Python and data visualization libraries. Our visualization strategy includes:

\* Selecting appropriate chart types (e.g., bar charts, pie charts, scatter plots) to represent age and gender distribution.Creating visualizations that facilitate comparisons across different industrial categories and age groups.

\*Ensuring visualizations are interactive and can be filtered by specific demographics or industrial categories.

\* Adding labels, legends, and tooltips for clarity and context.

Python and relevant data visualization libraries, such as Matplotlib or Seaborn, will be employed to create dynamic visualizations that aid in understanding the demographic data**.**

**2.4 Insights Generation:**

The primary objective of this project is to derive valuable insights from the analysis of demographic data pertaining to marginal workers in Tamil Nadu. Insights may include:

\* Identifying age groups with a higher concentration of marginal workers in specific industrial categories.

\* Revealing any gender disparities in the distribution of marginal workers within different age brackets and industries.

\* Recognizing any notable variations or patterns in the demographic composition of marginal workers.

\* Drawing connections between demographic trends and potential policy recommendations or interventions to improve the conditions of marginal workers in Tamil Nadu.

These insights will be derived through a combination of statistical analysis and visual examination of the demographic data. They will serve as valuable information for policymakers, government agencies, and organizations aiming to address the needs of marginal workers in Tamil Nadu.

**3. Next Steps:**

The subsequent phases of this project will encompass data preprocessing, exploratory data analysis (EDA), creating visualizations, conducting statistical analysis, and generating insights to achieve the defined objectives. Close collaboration among team members is crucial for the project's success.

**4. Timeline:**

**A tentative timeline for the project is outlined below:**

**Data Collection and Preprocessing: 2 weeks**

\* During this phase, we will collect and clean the demographic data related to marginal workers in Tamil Nadu. This includes handling missing values, data formatting, and addressing any outliers.

**Exploratory Data Analysis (EDA): 3 weeks**

\* EDA involves exploring the dataset, identifying key trends, and conducting preliminary statistical analysis. This phase will provide a deeper understanding of the data before visualization and in depth analysis.

**Visualization Design and Implementation: 4 weeks**

\* In this phase, we will design and create visualizations using Python and data visualization libraries. These visualizations will represent the distribution of marginal workers across different demographic categories, including age, industrial categories, and gender**.**

**Statistical Analysis: 4 weeks**

\* The statistical analysis phase will involve conducting relevant statistical tests and analyses to fulfill the project's objectives. This includes comparing demographic distributions and identifying any significant patterns or disparities.

**Insights Generation: 3 weeks**

\* Insights will be generated by interpreting the results of the statistical analysis and drawing meaningful conclusions regarding the demographic characteristics of marginal workers in Tamil Nadu.

**Documentation and Reporting: 2 weeks**

\* Comprehensive documentation of the analysis process, findings, and insights will be prepared. A final report summarizing the project's outcomes will be created**.**

**Review and Finalization:1 week**

\* The project team will conduct a final review of all project components, ensuring that the analysis aligns with the defined objectives and that the documentation is complete and accurate.

This timeline is subject to adjustments based on the complexity of the data, unforeseen challenges, and collaboration among team members. Regular communication and updates will be essential to keep the project on track and ensure its successful completion.