

COLLEGE CODE : 8107

COURSE :DATA ANALYTICS WITH COGNOS

PHASE I: PROJECT SUBMISSION PART I

PROJECT TITLE: Assessment of marginal workers in
Tamil Nadu – A socioeconomic Analysis

TEAM MEMBERS DETAILS:

JAYASHREE. M
QUINY CHRISTABEL. R
SHINORA THOMAS. T
SHREENITHE. R

Assessment of marginal workers in Tamil Nadu – A socioeconomic Analysis

Problem 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.

Understanding the problem:

The analysis aims to understand factors such as age, gender, occupation, and migration patterns among marginalized workers. It also involves evaluating their income levels, education, healthcare accessibility, and utilization of social welfare programs. The goal is to identify disparities, challenges, and opportunities faced by these workers, providing insights for informed policies and interventions to improve their overall quality of life and socioeconomic status. The analysis involves data extraction, cleaning, and analysis, followed by visualization and interpretation to derive meaningful insights for stakeholders and policymakers.

Design Thinking:

- Objectives:
 - Analyse Marginal Worker Demographics:
 - Subgroups Identification: Identifying specific subgroups within the marginalized worker population, such as agricultural labourers, construction workers, domestic helpers, etc.

- **Demographic Profiling:** Creating a detailed demographic profile including age, gender, ethnicity, and geographical distribution of these subgroups.
- **Migration Patterns:** Exploring migration patterns within Tamil Nadu and from other states.
- **Understand Age and Gender Distribution:**
 - **Age Distribution:** Examining the age distribution within different categories of marginalized workers to understand workforce composition and aging trends.
 - **Gender Analysis:** Analysing the gender distribution, exploring the challenges and opportunities faced by male and female marginalized workers.
 - **Impact of Age and Gender:** Understanding how age and gender influence employment opportunities, wages, and access to social welfare programs.
- **Explore Industrial Categories:**
 - **Occupational Analysis:** To Categorise marginalized workers based on their occupations, including agricultural labour, construction, domestic work, etc.
 - **Industrial Distribution:** Investigate the distribution of these occupational categories across various industries in Tamil Nadu.
 - **Income Disparities:** Examining income disparities within and between different industrial categories to identify wage gaps and disparities in economic opportunities.
- **Study Social Welfare Program Utilization:**
 - **Effectiveness Assessment:** Evaluate the effectiveness of existing social welfare programs in improving the socioeconomic status of these workers.
 - **Challenges in Access:** Identify challenges faced by marginalized workers in accessing and benefiting from social welfare initiatives.

- **Analysis Approach:**

- Step 1: Data Extraction

- Determine relevant sources such as government databases, surveys, academic research, or NGO reports focusing on labor and

marginalized communities in Tamil Nadu. Obtain permission and access to the selected datasets.

- Collect quantitative data related to demographics, employment, education, healthcare, and social welfare programs.

Step 2: Data Cleaning

- Merge data from various sources into a unified dataset, ensuring compatibility and consistency in variable formats.
- Identify and handle missing or incomplete data points using appropriate methods such as imputation or data removal, ensuring minimal impact on overall analysis.
- Identify outliers in the dataset that could skew the analysis. Decide whether to remove outliers or transform them based on the context of the analysis.
- Validate the cleaned dataset to ensure accuracy, consistency, and completeness.

Step 3: Data Analysis

- Calculate basic statistics (mean, median, standard deviation) for key variables to understand the dataset's characteristics.
- Use visualizations (bar charts, pie charts) and statistical methods to analyse demographics, age, gender, and migration patterns.
- Evaluate the relationship between education levels and employment opportunities.
- Compare wages across different occupations and industries using statistical tests.

Step 4: Visualising and Reporting

- Create visualizations (charts, graphs, heat maps) to present the insights clearly and effectively.
- Prepare a presentation summarizing the analysis for stakeholders.
- Communicate the findings effectively, highlighting important trends and policy implications.

- Visualisation Selection:

- Use histograms or bar charts to display the frequency distribution of different age groups among marginalized workers.
- Represent the proportion of male and female workers using

- Display the number or percentage of workers from different ethnic or caste groups using a bar chart for easy comparison.
- If you want to represent additional data (like population size), a bubble map can show both the geographical distribution and the magnitude of the worker population.
- If analysing migration patterns within and outside Tamil Nadu, a flow map can demonstrate the movement of workers between different regions.
- Violin plots combine aspects of box plots and kernel density plots, providing a more detailed view of income distribution, especially when comparing multiple groups.
- Use a heat map to visualize the correlation between educational attainment and income levels among marginalized workers.