**Team members:**

**K. Swetha**

**N. Sandhya**

**K. Pooja**

**M. Priyadarshini**

**Assessment of marginal workers in Tamil Nadu**

**INTRODUCTION**

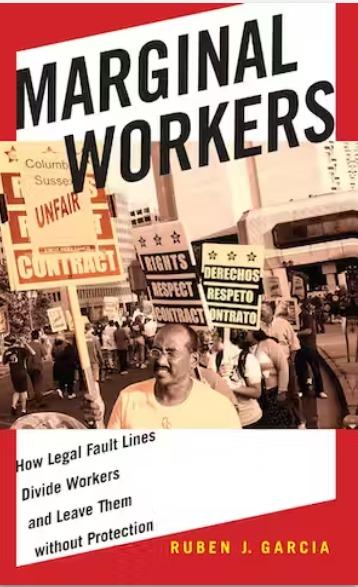
Marginal workers in Tamil Nadu, as in many other parts of India, are a significant

segment of the labour force. They represent individuals who are employed in various economic

activities for only part of the year, often due to seasonal, casual, temporary, or

underemployment reasons. Marginal workers in Tamil Nadu play a crucial role in various

sectors, including agriculture, construction, and informal labour markets.



**DATA PROCESSING PROCEDURE**

To create a data visualization and perform machine learning analysis on a marginal workers in Tamil Nadu dataset using Python libraries like pandas, NumPy, matplotlib, seaborn, and scikit-learn for

decision tree classification, follow these steps:

**1. Data Collection and Pre-processing:** Obtain the marginal workers in Tamil Nadu with relevant

parameters Import necessary Python libraries: pandas, NumPy, matplotlib, seaborn, and

scikit-learn. Load and pre-process the dataset using pandas to handle missing values, outliers,

and data cleaning.

**2. Data Visualization**: Utilize matplotlib and seaborn to create visualizations of the data.

Some common plots include scatter plots, histograms, and box plots to understand data

distributions.

**3. Correlation Analysis:** Use pandas to calculate the correlation between different parameters

in the dataset. Create correlation matrices and visualize them using heatmap plots from

seaborn to identify relationships between variables.

**4. Machine Learning Preparation:** Select the target variable (e.g., marginal workers )and

features (marginal workers parameters) for the machine learning model. Split the dataset into

training and testing sets.

**5.Decision Tree Classifier:** Train a decision tree classifier using the Decision Tree Classifier from scikit-learn. Fit the model to the training data and evaluate its performance on the

testing data.

**6. Model Evaluation:** Calculate performance metrics such as accuracy, precision, recall, and F1-score to assess the model's classification performance.

7**. Visualization of Decision Tree:** Visualize the decision tree structure using tools provided by

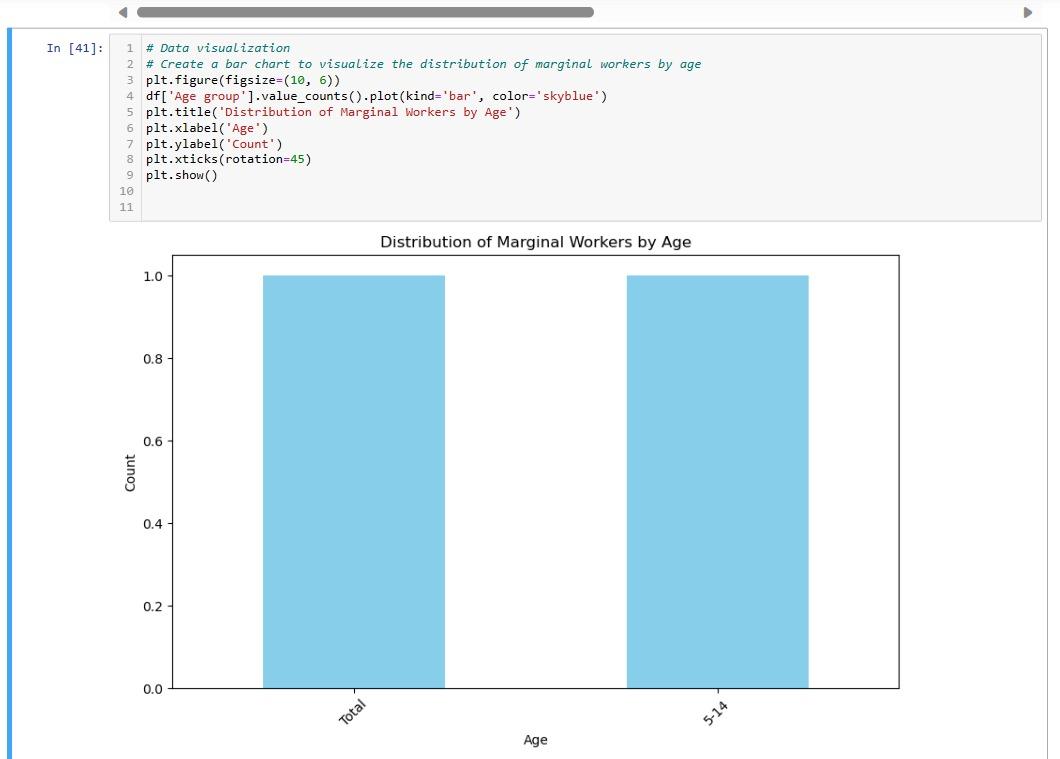
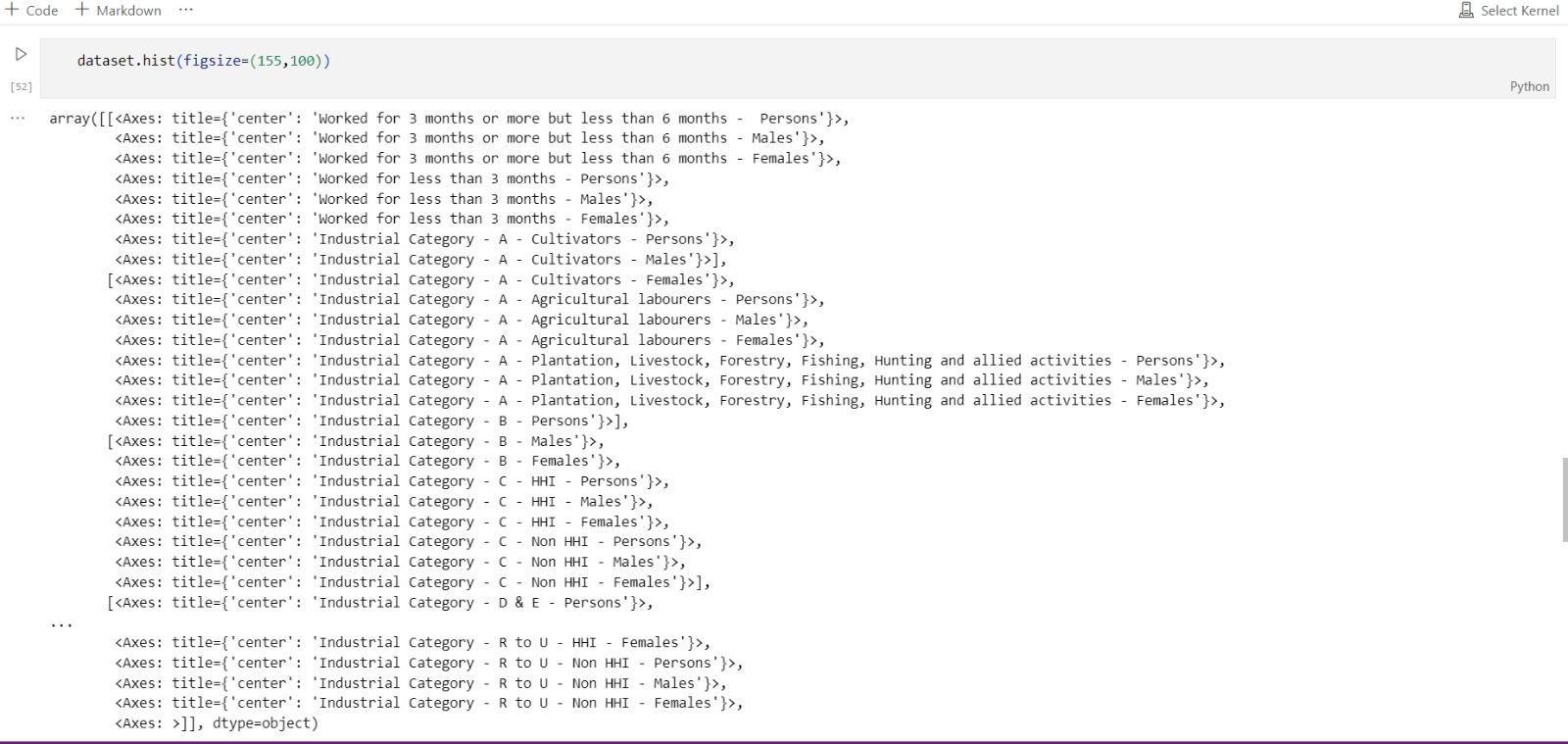
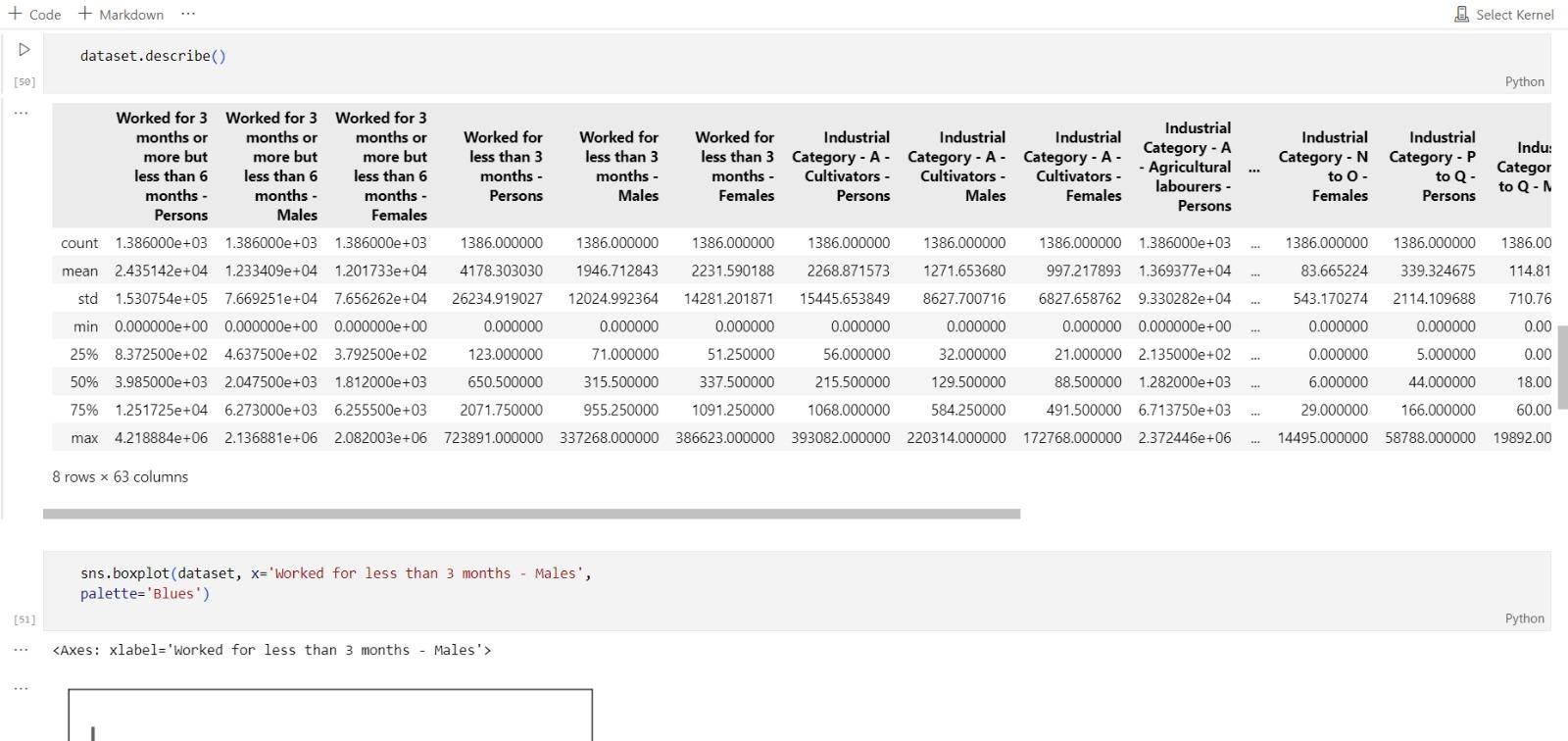
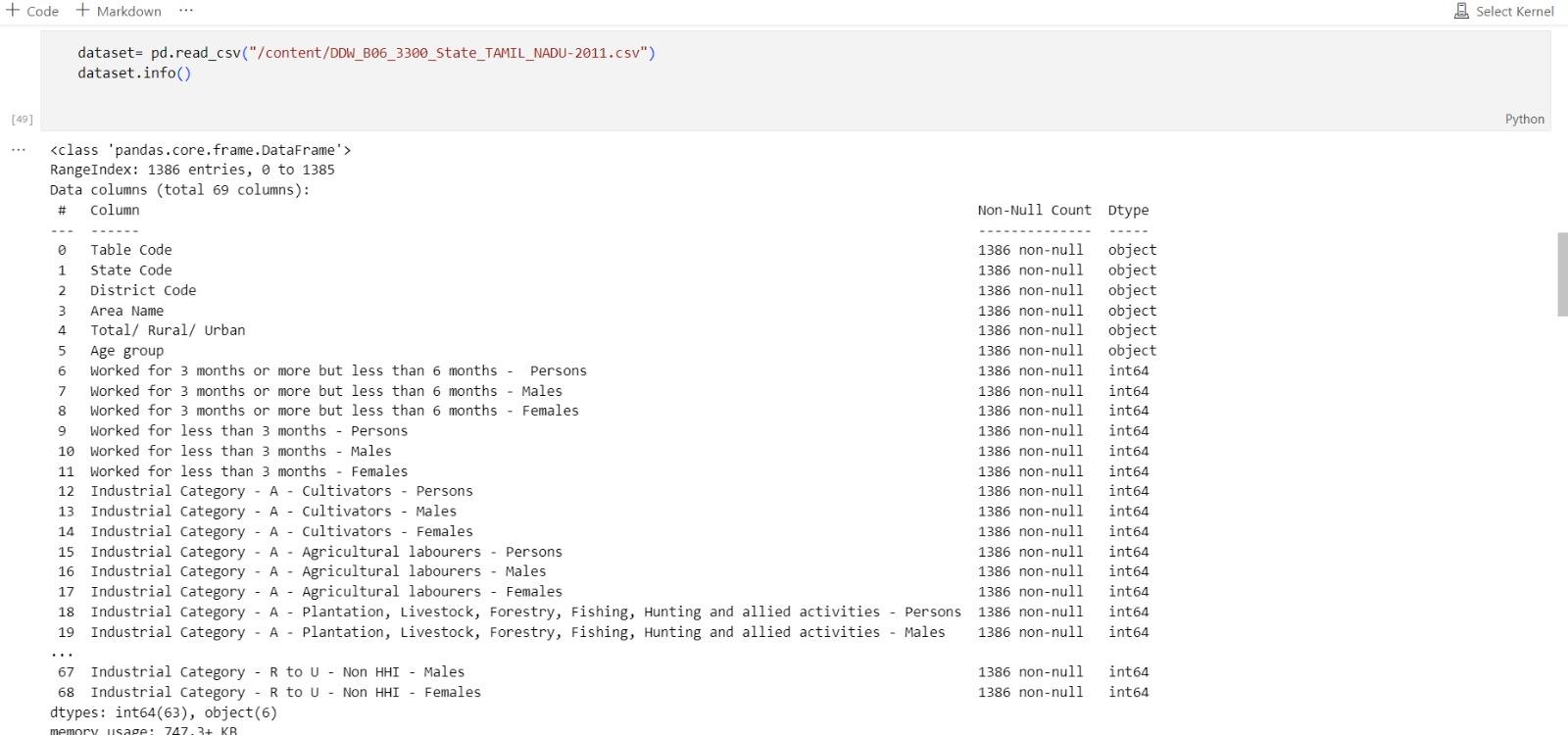
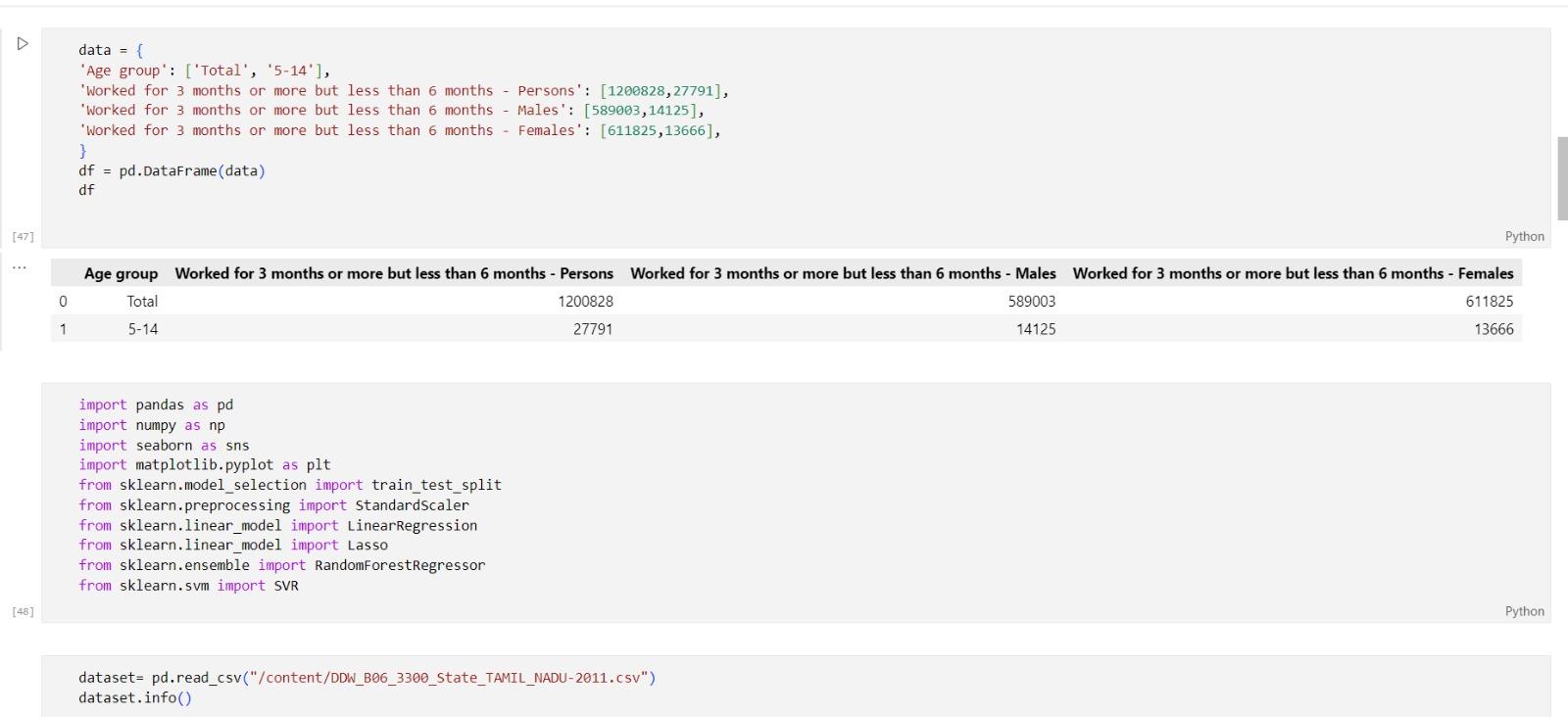
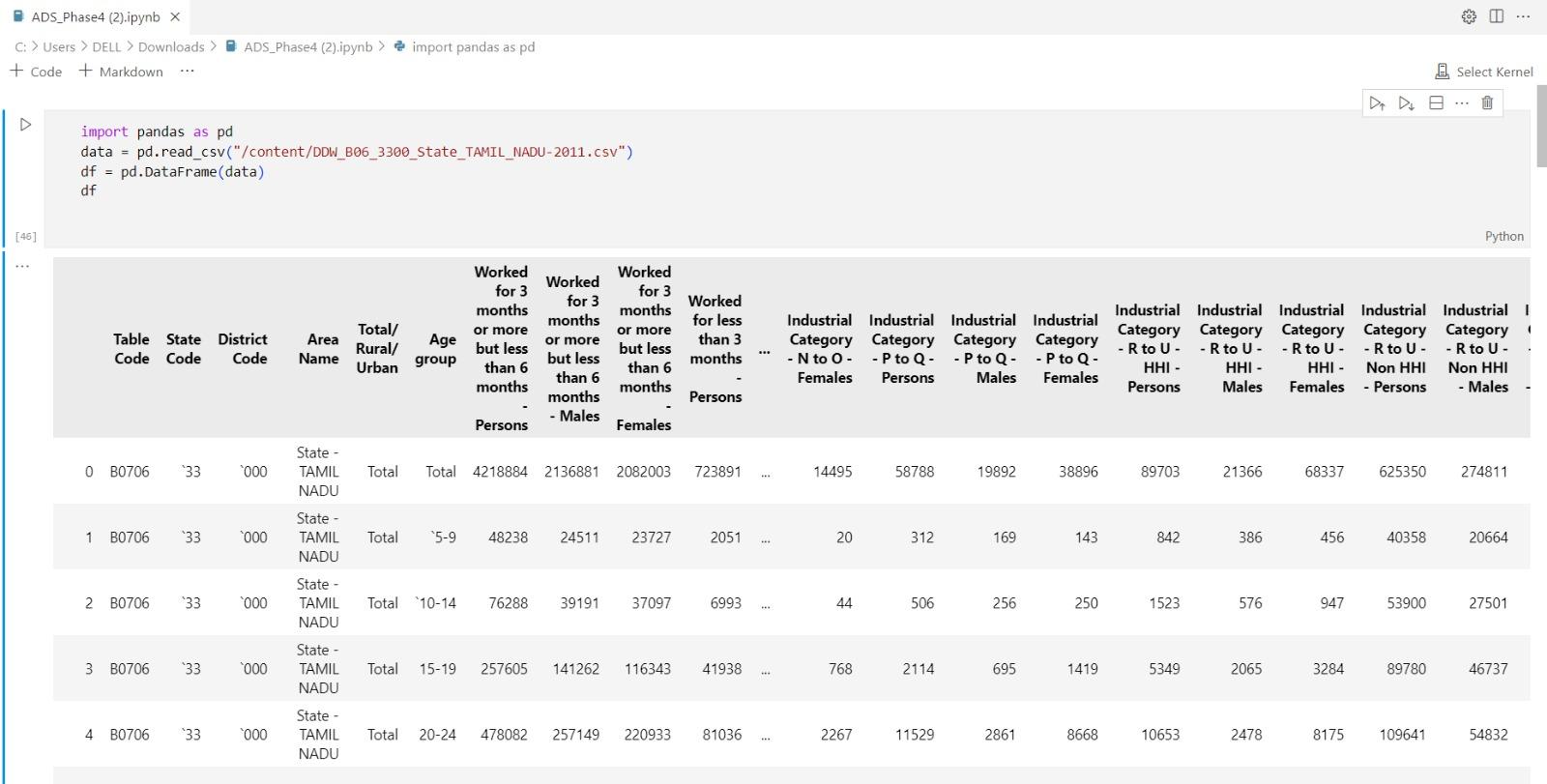
scikit-learn, such as the plot tree function.

**Data Visualization:**

When representing the final output of an assessment of marginal workers in Tamil Nadu, you'll want to choose data visualization types that effectively convey the socioeconomic problems these workers face.

**Bar charts :** Use bar charts to show a comparison of different socioeconomic indicators

among marginal workers, such as income levels, education levels, or access to healthcare. You can have clustered or stacked bar charts to make comparisons across different subgroups or regions



**Conclusion** :

In conclusion, addressing the challenges faced by marginal workers is a crucial

endeavour that requires a multifaceted approach. These workers, often engaged in irregular and insecure employment, deserve attention and support to enhance their well-being and livelihoods.

Implementing effective solutions involves understanding their unique needs, advocating for their rights, and providing access to education and skill development programs.