DATA ANALYTICS WITH COGNOS

PHASE 5**:** PROJECT DOCUMENTATION AND SUBMISSION

***PROJECT TITLE: MARGINAL WORKERS ASSESSMENT IN TN SOCIO ECONOMIC ANALYSIS***

**Phase 5: Project Documentation & Submission** **:**

In this part you will document your project and prepare it for submission.

Document the Assessment of Marginal Workers project and prepare it for submission

**Documentation:**

Describe the project's objectives, analysis approach, visualization types, and code implementation.

Include example outputs of data analysis and visualizations.

Explain how the analysis provides insights into the demographic characteristics of marginal workers in Tamil Nadu

**MARGINAL WORKERS:**

Marginal workers are those whose employment is insecure, poorly paid, and unprotected. They often work in informal sectors such as construction, agriculture, and domestic work. Marginal workers are vulnerable to exploitation and abuse, and they often lack access to social security benefits.

***Objective of marginal worker assessment in TN socio economic analysis:***

*The objective of assessing marginal workers in the context of Tamil Nadu's socio-economic analysis is to gain insights into the labour force dynamics, employment patterns, and economic conditions of this specific segment of the population.*

*Marginal workers typically refer to individuals who are employed for a major part of the year but work for shorter durations, and they often face economic vulnerabilities.*

***The assessment of marginal workers can help policymakers and researchers:***

*1. Understand the extent of underemployment and seasonal employment in the state.*

*2. Identify the economic challenges and vulnerabilities faced by marginal workers.*

*3. Develop targeted policies and interventions to improve the socio-economic conditions of this group.*

*4. Ensure inclusive and equitable economic growth, with a focus on improving the well-being of marginalized laborers.*

*5. Monitor changes in the status and conditions of marginal workers over time.*

***ANALYZE APPROACH:***

*The assessment of marginal workers in Tamil Nadu's socio-economic analysis aims to provide a more comprehensive understanding of labour dynamics and inform policies that promote inclusive development and economic security for this segment of the workforce.*

*Analyzing the socio-economic status of marginal workers in Tamil Nadu (TN) requires a structured approach. Here's a general analysis framework:*

***1. Data Collection:***

*- Gather relevant data from government sources, surveys, and research reports.*

*- Collect data on income, education, employment, health, and living conditions.*

***2. Define Marginal Workers:***

*- Clearly define who qualifies as "marginal workers" in the context of your analysis. This could include daily wage laborers, seasonal workers, or other specific categories.*

***3. Demographic Analysis:***

*- Analyze the demographic characteristics of marginal workers, such as age, gender, and location.*

***4. Economic Status Assessment:***

*- Examine the income levels, employment patterns, and job security of marginal workers.*

*- Assess their access to financial services and social safety nets.*

***5. Education and Skill Levels:***

*- Evaluate the educational background and skill sets of marginal workers.*

*- Determine if there are opportunities for skill development.*

***6. Health and Well-being:***

*- Study the health status and access to healthcare of marginal workers.*

*- Assess the impact of their working conditions on health.*

***7. Living Conditions:***

*- Analyze the housing, sanitation, and overall living conditions of marginal workers.*

*- Look for issues related to housing security and basic amenities.*

***8. Government Policies and Support:***

*- Review existing government policies and programs aimed at helping marginal workers.*

*- Assess the effectiveness of these policies.*

***9. Income Disparities:***

*- Investigate income disparities among marginal workers based on factors like gender, location, and type of work.*

***10. Qualitative Research:***

*- Conduct interviews or surveys to gather qualitative insights from marginal workers themselves.*

*- Understand their perspectives and challenges.*

***11. Comparative Analysis:***

*- Compare the socio-economic status of marginal workers in TN to other states or regions in India for context.*

***12. Recommendations:***

*- Based on the analysis, develop recommendations for improving the socio-economic status of marginal workers.*

*- Advocate for policy changes, skill development programs, and social support measures.*

***13. Continuous Monitoring:***

*- Establish a framework for ongoing monitoring and evaluation of the socio-economic conditions of marginal workers in TN.*

***VISUALIZATION TYPES :***

*1. Bar Charts*

*2. Pie Charts*

*3. Heatmaps*

*4. Scatter Plots:*

*5. Line Graphs:*

*6. Choropleth Maps:*

*7. Stacked Area Charts:*

*8. Box Plots*

*9. Sankey Diagrams:*

*10. Word Clouds:*

*11. Radar Charts:*

*12. Infographics:*

*13. Dashboards:*

***CODE IMPLEMENTATION:***

**1. Import necessary libraries:**

Import pandas as pd

Import numpy as np

from sklearn .model \_selection import train \_test \_split

from sklearn. preprocessing import StandardScaler

**2. Load the dataset:**

df=pd.read\_csv(r'C:\Users\ranji\Desktop\marginalworkers.csv')

Pd.read()

#Checkformissingvalues

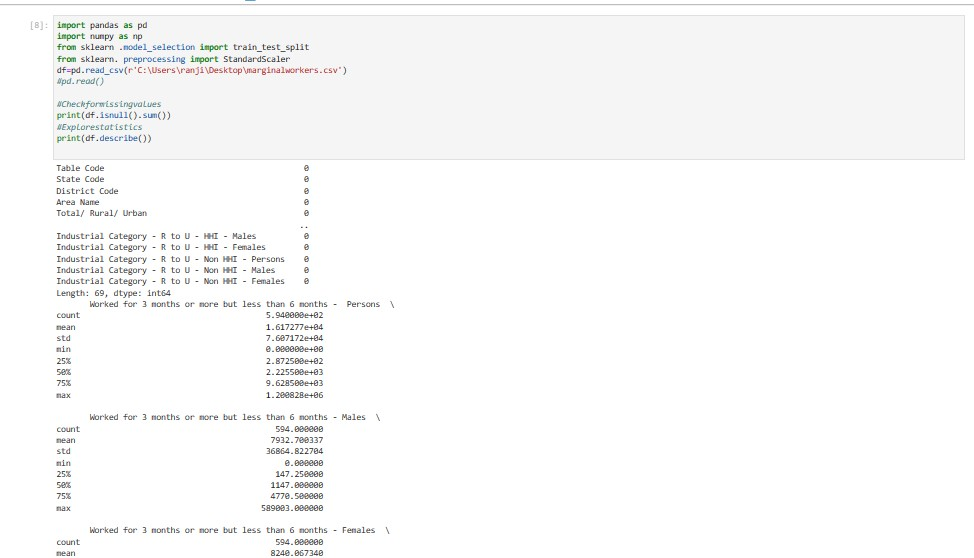
print(df.isnull().sum())

#Explorestatistics

print(df.describe())

#Visualizethedata(e.g.,histograms,scatterplots,etc.)

**OUTPUT:**



**# Step 1: Load the dataset**

Data=pd.read\_csv('C:\Users\ranji\Desktop\marginalworkers.csv"csv')

**# Step 2: Exploratory Data Analysis (EDA)**

print("--- Exploratory Data Analysis ---")

print("1. Checking for Missing Values:")

missing\_values = data.isnull().sum()

print(missing\_values)

print("\n2. Descriptive Statistics:")

description = data.describe()

print(description)

**# Step 3: Feature Engineering**

print("\n--- Feature Engineering ---")

# Separate features and target variable

X = data .drop('price', axis=1)

y = data['price']

# Define which columns should be one-hot encoded (categorical)

categorical\_ cols = [' Avg. Area House Age']

**# Step 4: Data Splitting**

print("\n--- Data Splitting ---")

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2,random\_state=42)

print(f"X\_train shape: {X\_train.shape}")

print(f"X\_test shape: {X\_test.shape}")

print(f"y\_train shape: {y\_train.shape}")

print(f"y\_test shape: {y\_test.shape}")

**# Step 5: Preprocessing and Feature Scaling using Pipeline**

print("\n--- Feature Scaling ---")

model = Pipeline([('preprocessor', preprocessor),])

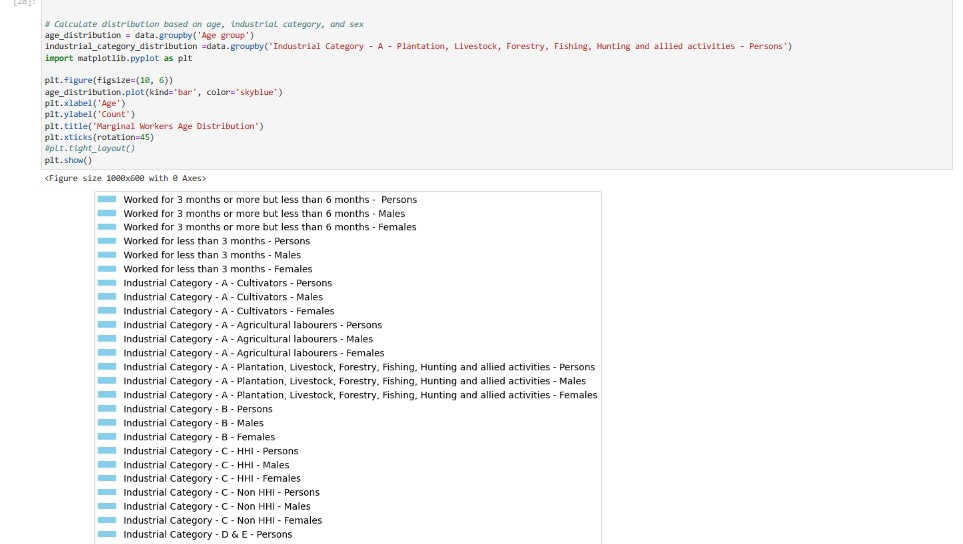
# Fit the preprocessing pipeline on the training data

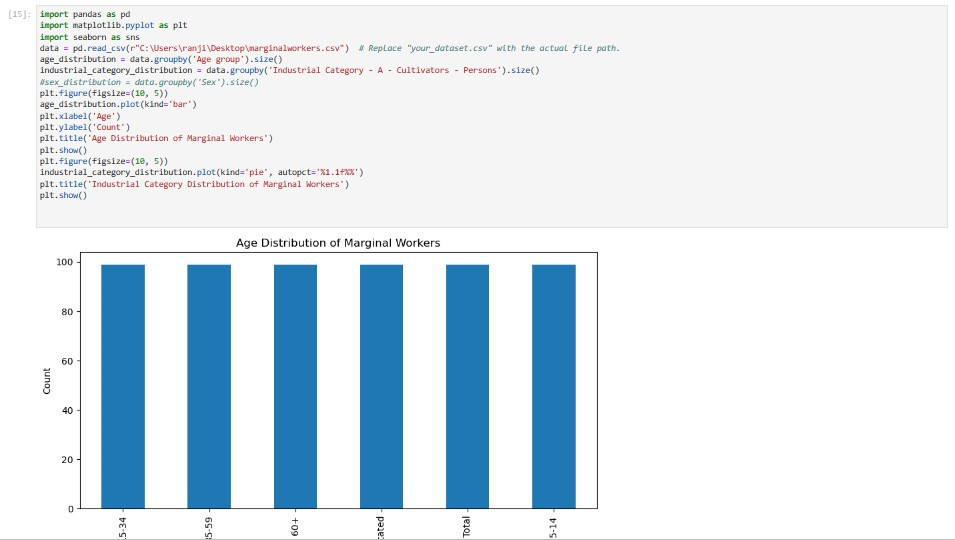
X\_train = model.fit\_transform(X\_train)

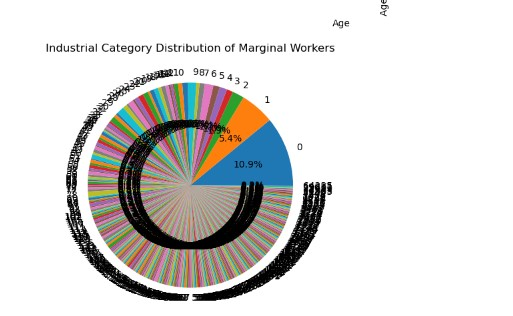
# Transform the testing data using the fitted pipeline

X\_test = model.transform(X\_test)

print("--- Preprocessing Complete! ---")







**TEAMMATES:**

1. M KEERTHIGA
2. S RANJITHA
3. K VASUKI
4. K KAYALVIZHI