**Team Members:**

**K. Swetha**

**N. Sandhya**

**M. Priyadharshini**

**K. Pooja**

**Assessment of Marginal Workers in Tamilnadu**

Marginal workers, often referring to those on the fringes of the job market, can benefit from data science in various ways. Data analysis can help identify employment trends, enabling policymakers to create targeted programs. Additionally, data-driven platforms can match marginal workers with suitable job opportunities, fostering better employment prospects. Training programs can be optimized through data analysis, aligning skill development with market demands, enhancing the employability of marginalized individuals. Overall, data science plays a vital role in creating more inclusive job markets for marginal workers. let's delve into more detail about how data science can empower marginal workers:

1. Skill Development Programs: Data science can analyze local job markets to identify the most in-demand skills. Training programs can then be tailored to teach marginalized individuals these skills, enhancing their employability.

2. Job Matching Platforms: Data-driven platforms can match marginal workers with suitable job opportunities. By analyzing resumes, job requirements, and employer preferences, algorithms can recommend jobs that align with the worker's skills and experience, increasing the chances of finding suitable employment.

3. Predictive Analytics for Employment Trends: Data science can analyze historical employment data to predict future job trends. This information helps policymakers and organizations create targeted initiatives to support marginalized communities, ensuring that training programs focus on skills relevant to upcoming job opportunities.

4. Personalized Career Counseling: Data analysis can be used to provide personalized career counseling. By understanding an individual's skills, qualifications, and aspirations, counselors can offer tailored advice on career paths, training programs, and job opportunities, increasing the chances of successful employment.

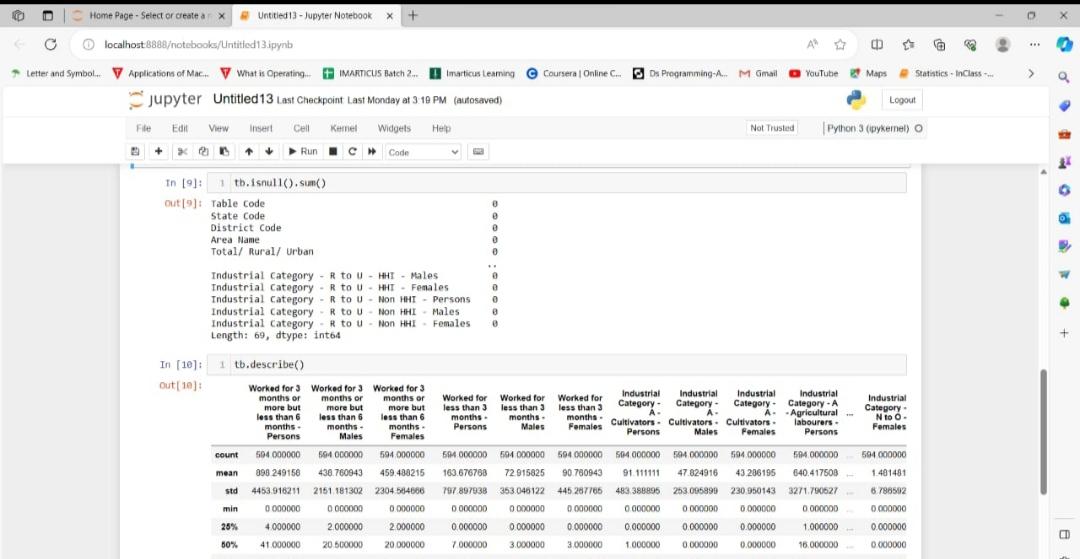
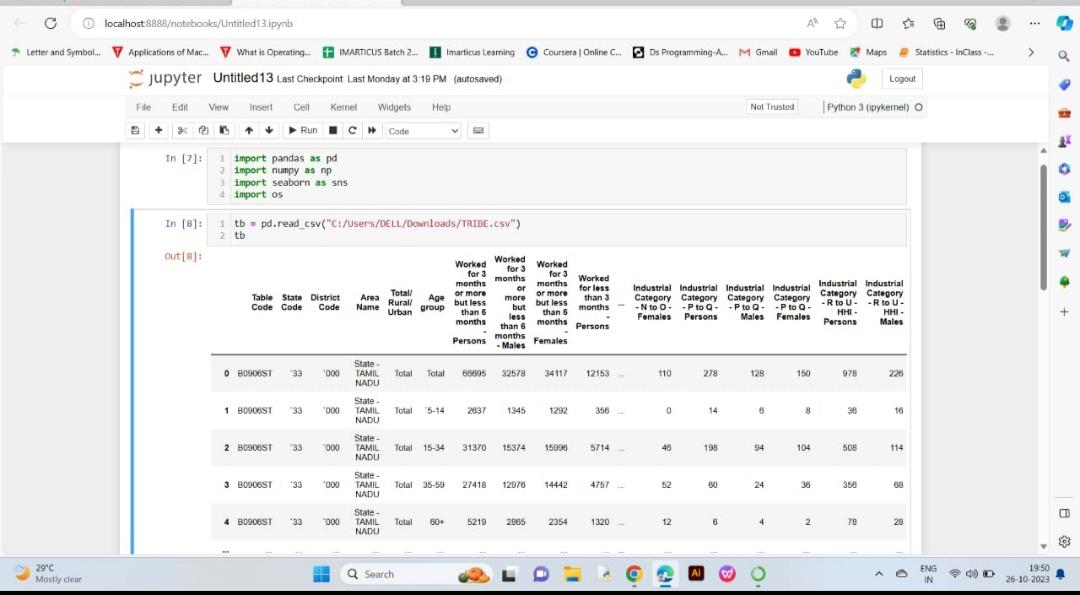
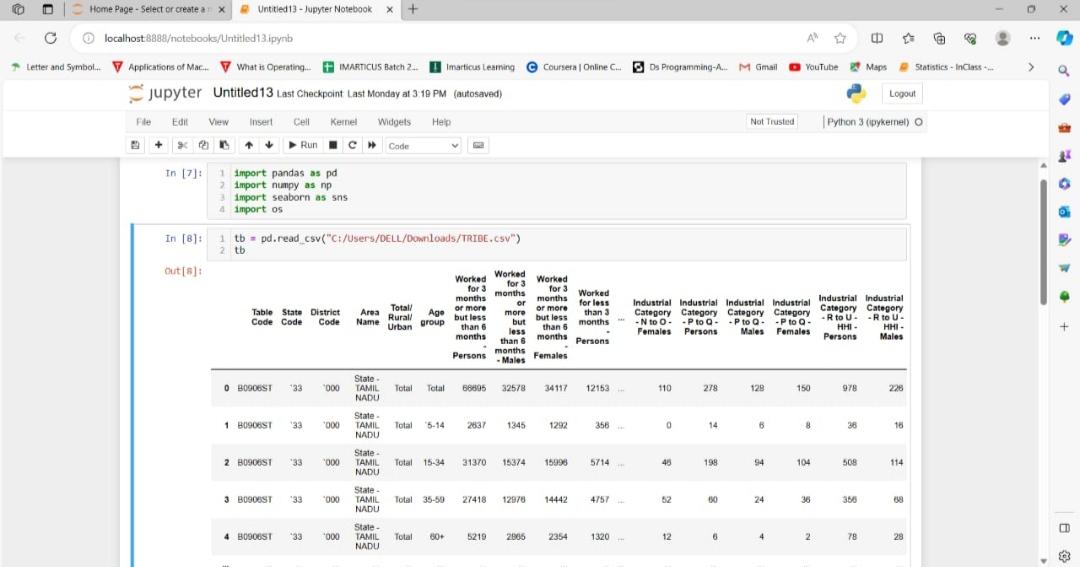
5. Entrepreneurship Support: For marginal workers interested in entrepreneurship, data science can provide insights into market demands. Analyzing consumer behavior and market trends can help aspiring entrepreneurs make informed decisions about the products or services they want to offer, increasing their chances of establishing successful businesses.

6. Workforce Diversity Analysis: Employers can use data science to analyze their workforce diversity. By understanding the demographics of their employees, companies can implement inclusive hiring practices, ensuring that marginalized communities are not discriminated against in the hiring process.

7. Performance Monitoring and Feedback: In workplaces, data science can be used to monitor employee performance and provide constructive feedback. For marginal workers, this feedback loop is crucial for skill improvement and career growth, enhancing their long-term employability prospects.

8. Policy Formulation: Governments and NGOs can use data science to assess the effectiveness of existing policies aimed at supporting marginalized workers. By analyzing the outcomes of various initiatives, policymakers can refine existing programs or create new ones that are more impactful and responsive to the needs of the marginalized workforce.

In summary, data science equips policymakers, organizations, and individuals with valuable insights, enabling tailored support for marginal workers. This targeted approach increases their chances of finding meaningful employment, thereby improving their quality of life and contributing to a more inclusive society.



**Program:**

Import necessary libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

# Load your dataset

data = pd.read\_csv("marginal\_workers\_data.csv")

# Basic Data Exploration

# Display the first few rows of the dataset

print(data.head())

# Summary statistics

print(data.describe())

# Check for missing values

print(data.isnull().sum())

# Data Cleaning

# Handle missing values (if any)

data = data.dropna()

# Data Visualization

# Distribution of marginal workers by age

plt.figure(figsize=(8, 6))

sns.histplot(data['age'], bins=20, kde=True)

plt.title("Age Distribution of Marginal Workers")

plt.xlabel("Age")

plt.ylabel("Count")

plt.show()

# Gender distribution of marginal workers

plt.figure(figsize=(6, 6))

sns.countplot(data['gender'])

plt.title("Gender Distribution of Marginal Workers")

plt.xlabel("Gender")

plt.ylabel("Count")

plt.show()

# Education level of marginal workers

plt.figure(figsize=(8, 6))

sns.countplot(data['education\_level'])

plt.title("Education Level of Marginal Workers")

plt.xlabel("Education Level")

plt.ylabel("Count")

plt.xticks(rotation=45)

plt.show()

# Employment status of marginal workers

plt.figure(figsize=(8, 6))

sns.countplot(data['employment\_status'])

plt.title("Employment Status of Marginal Workers")

plt.xlabel("Employment Status")

plt.ylabel("Count")

plt.xticks(rotation=45)

plt.show()

# Correlation between variables

correlation\_matrix = data.corr()

plt.figure(figsize=(10, 8))

sns.heatmap(correlation\_matrix, annot=True, cmap="coolwarm")

plt.title("Correlation Matrix")

plt.show()

# Data Analysis

# Grouping and summarizing data

age\_groups = data.groupby('age\_group')['income'].mean()

print(age\_groups)

# Hypothesis testing (e.g., t-tests, ANOVA) to analyze differences between groups

# Additional exploratory analyses as needed

# Conclusion and Insights

# The insights and conclusions will depend on the specific research questions and findings from the EDA process.

# Save or export the cleaned data if needed

data.to\_csv("cleaned\_marginal\_workers\_data.csv", index=False)