**TN Marginal Workers Assessment**

**Phase 4: Development Part 2**

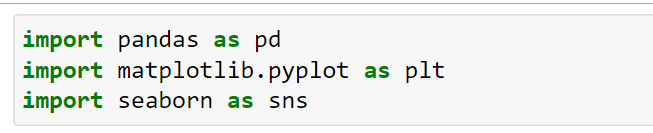
In this part you will continue building your project.

Perform the demographic analysis and create visualizations.

Calculate the distribution of marginal workers based on age, industrial category, and sex using data aggregation and manipulation.

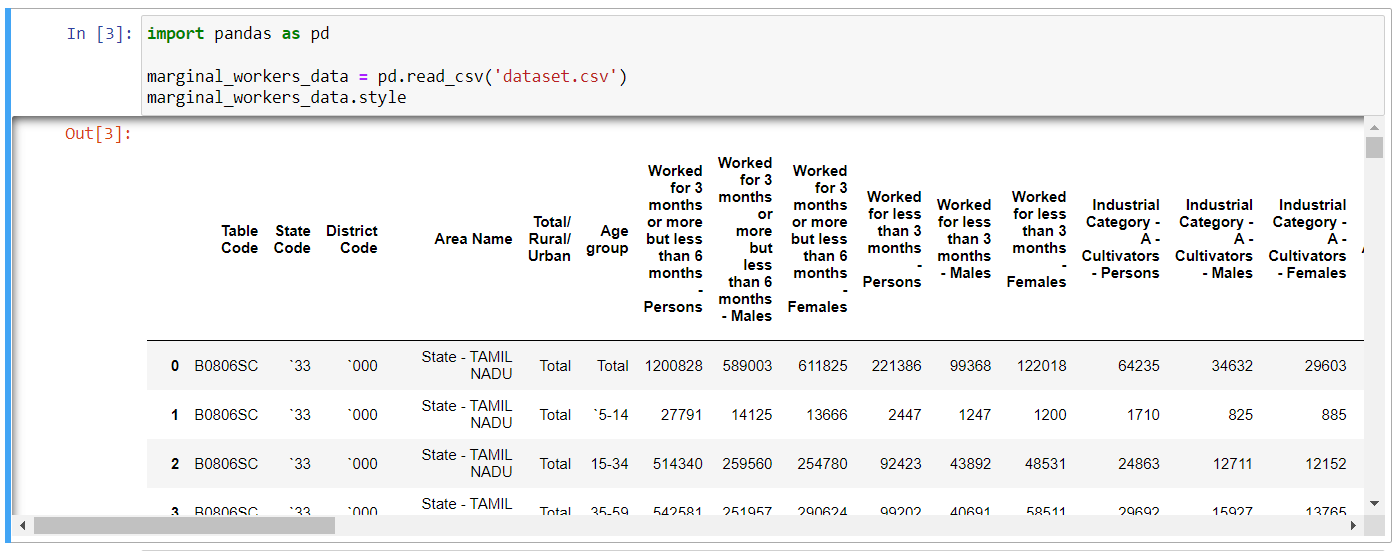
Create visualizations using data visualization libraries (e.g., Matplotlib, Seaborn).

We have to import the module to perform the demographic analysis and to calculate the distribution of marginal workers based on age, industrial category and sex



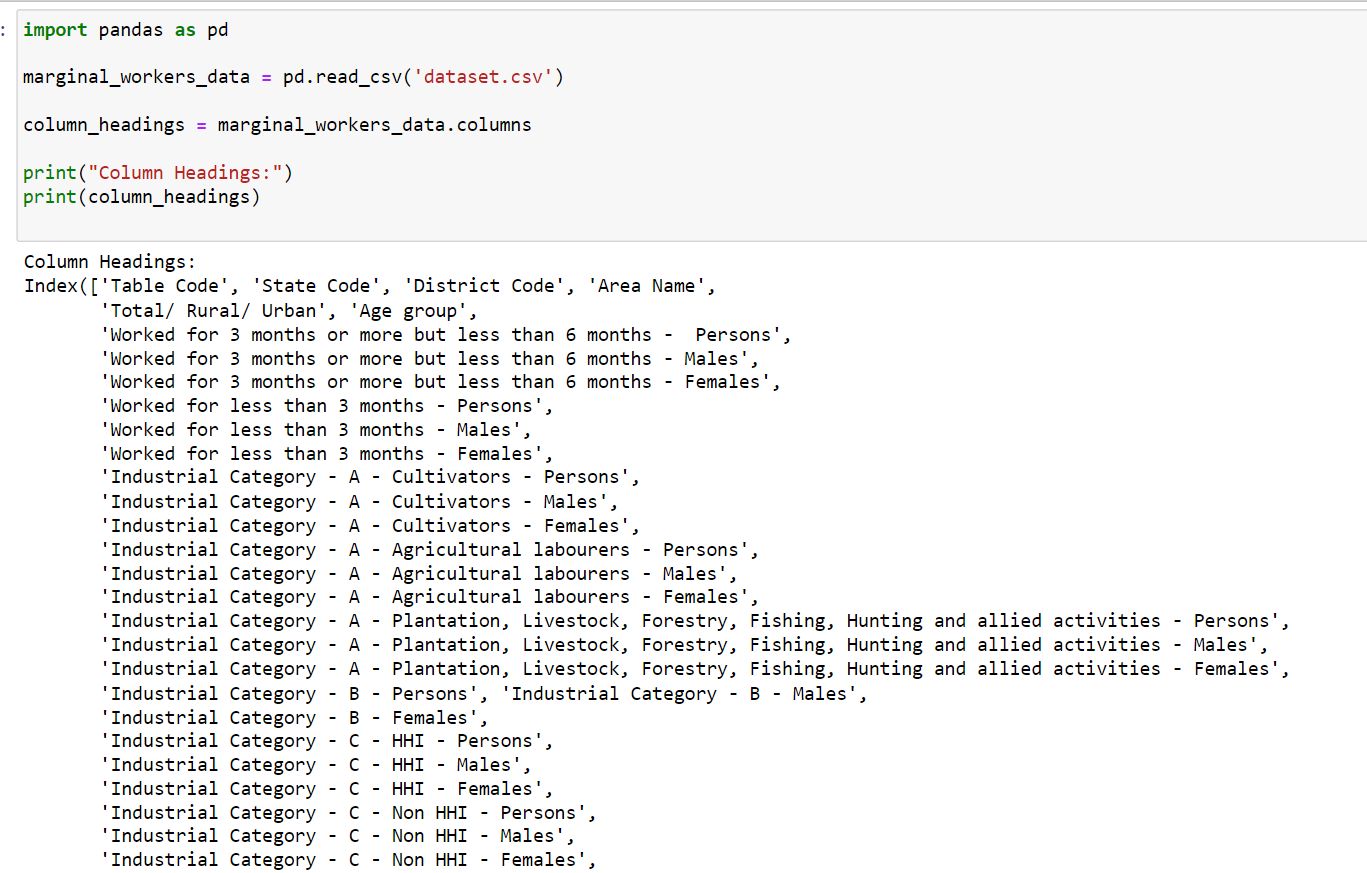
We have to do the data pre-processing

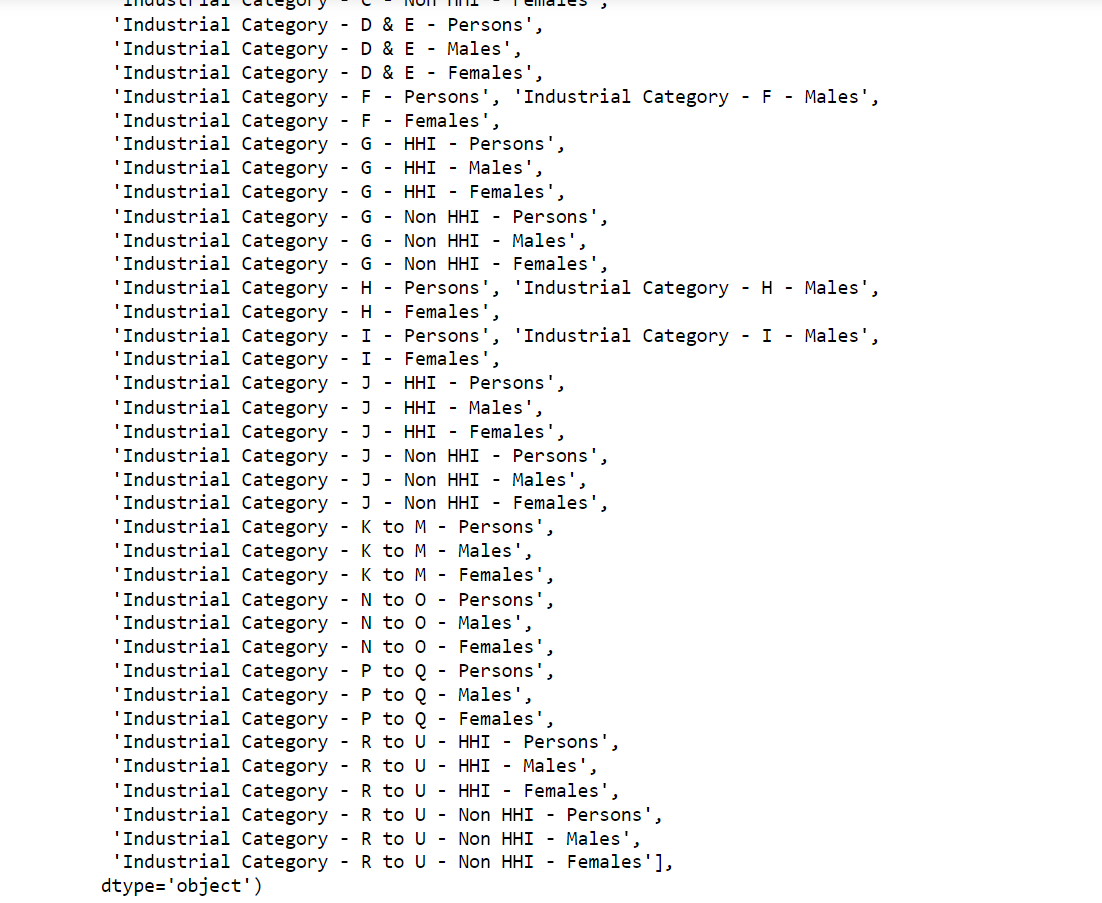
So that we use the pandas for loading the dataset



For finding the distribution process we have to know the column in the dataset

By that we can calculate the analysis





To calculate the distribution (Age, Industrial category and Sex)

**Age Distribution in marginal workers**

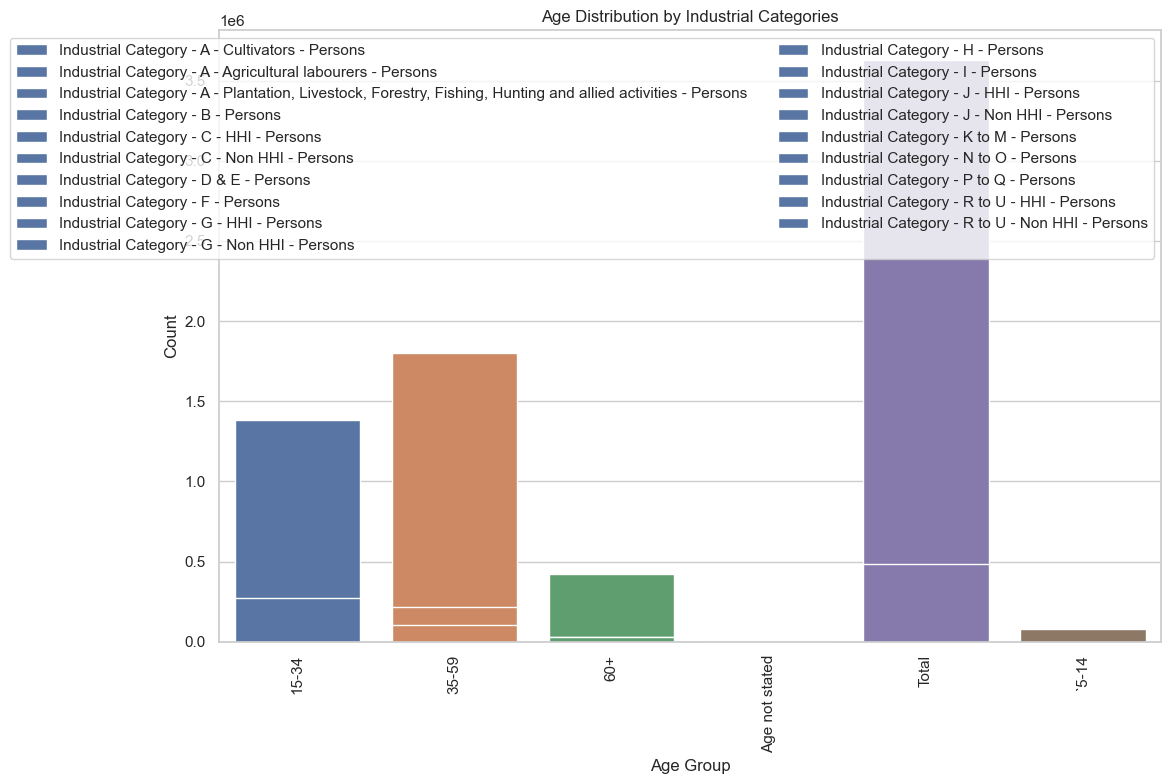
In the age distribution we use all the industrial category so that we can analysis which age group is has a maximum contribution in the marginal workers

**INPUT:**



**OUTPUT:**

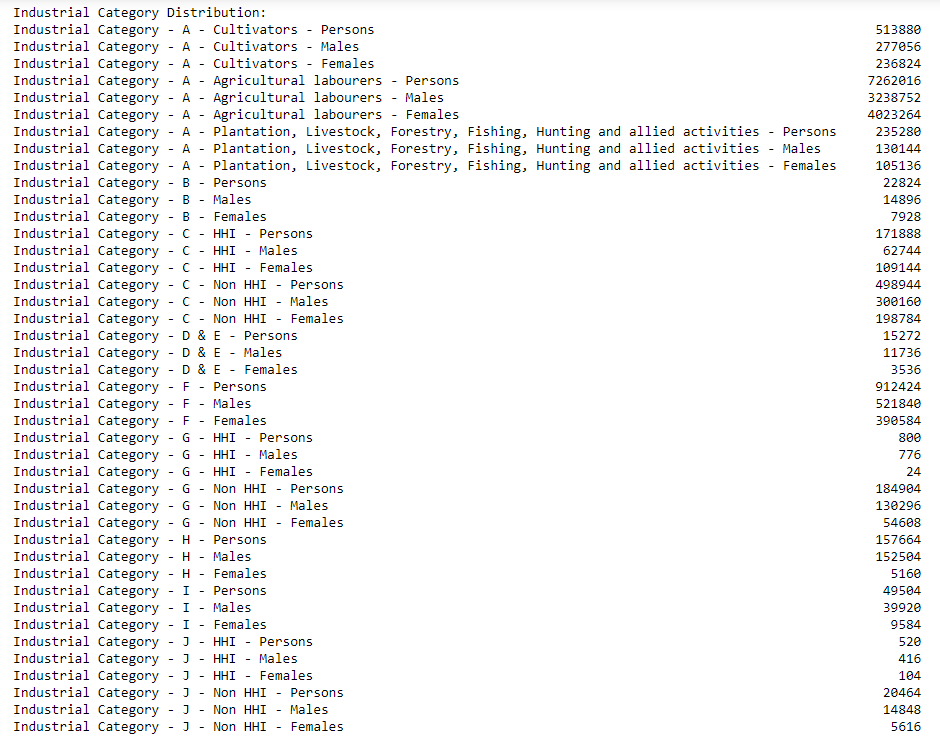
The visualization it represent the age distribution with age group, count and the industrial category.

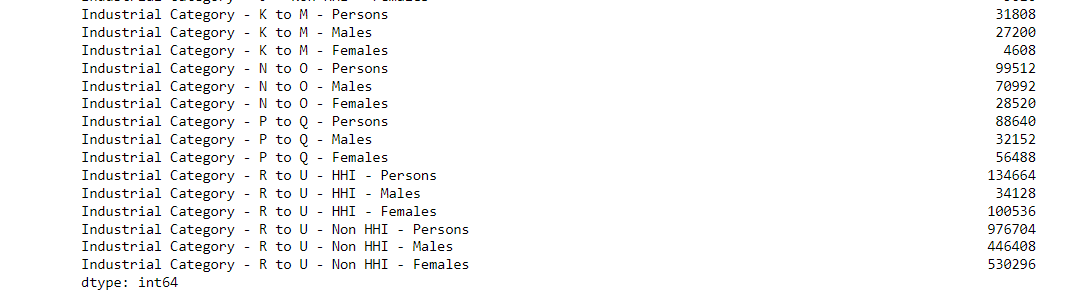


**Industrial category distribution**

**INPUT:**





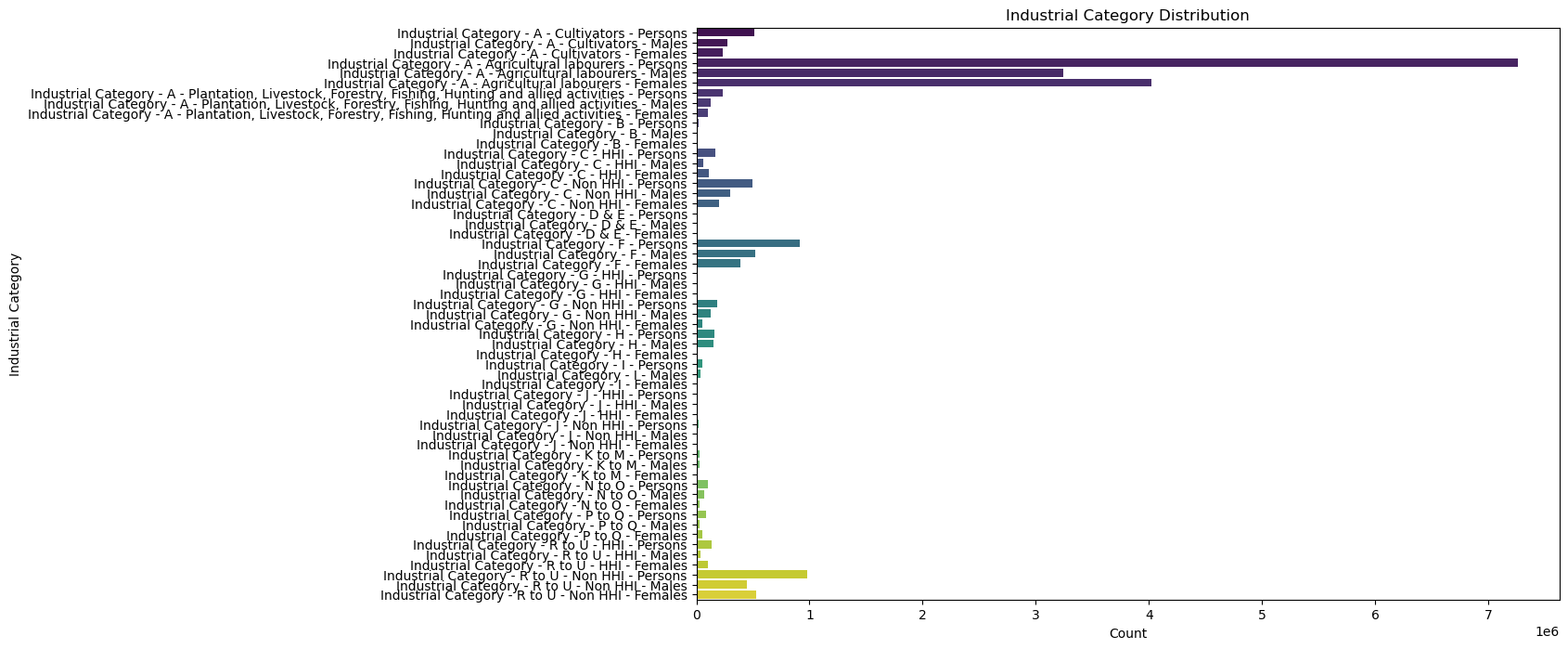


To visualize the industrial category we use the seaborn, matplotlib module.



For representation of the industrial category distribution we use the bar chart for visualization

**OUTPUT:**



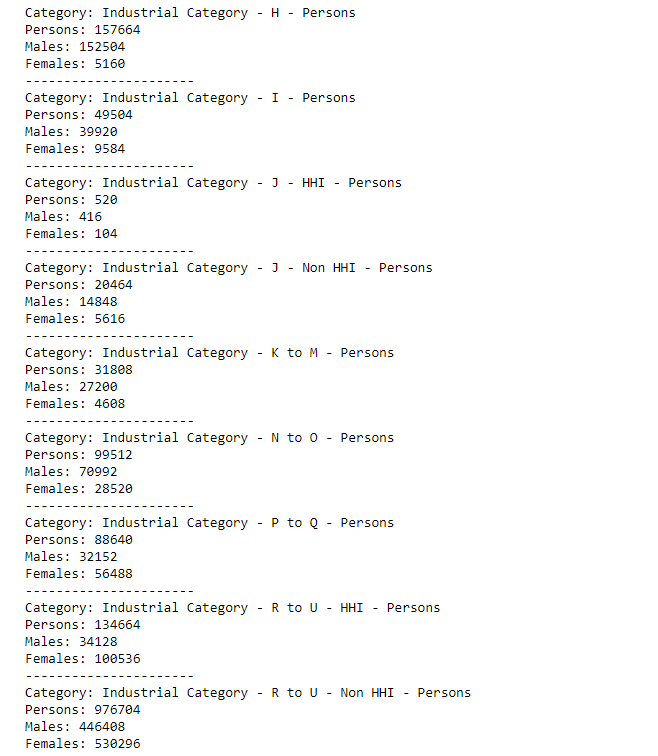
**Sex distribution in marginal workers.**

For this distribution we have calculate how many male and female are working in each category.

**INPUT**

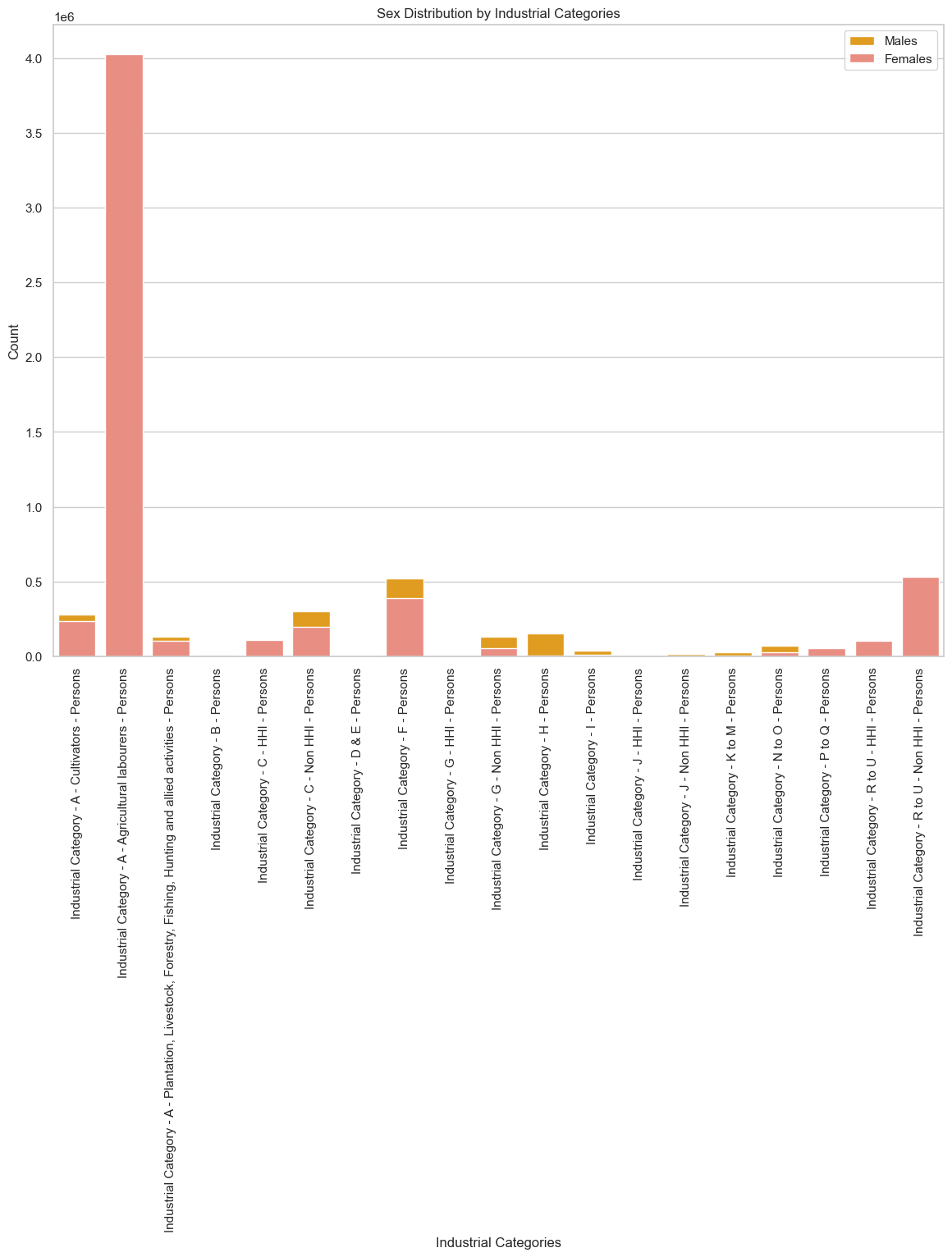






For visualize the distribution of industrial workers.





**Source code:**

**1. Import the module**

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

**2. Loading the dataset**

import pandas as pd

marginal\_workers\_data = pd.read\_csv('dataset.csv')

marginal\_workers\_data.style

**3. Display the column heading in the dataset**

marginal\_workers\_data = pd.read\_csv('dataset.csv')

column\_headings = marginal\_workers\_data.columns

print("Column Headings:")

print(column\_headings)

**4. Calculating the distribution for age and visualize it**

age\_industrial\_columns = [

'Age group',

'Industrial Category - A - Cultivators - Persons',

'Industrial Category - A - Agricultural labourers - Persons',

'Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Persons',

'Industrial Category - B - Persons',

'Industrial Category - C - HHI - Persons',

'Industrial Category - C - Non HHI - Persons',

'Industrial Category - D & E - Persons',

'Industrial Category - F - Persons',

'Industrial Category - G - HHI - Persons',

'Industrial Category - G - Non HHI - Persons',

'Industrial Category - H - Persons',

'Industrial Category - I - Persons',

'Industrial Category - J - HHI - Persons',

'Industrial Category - J - Non HHI - Persons',

'Industrial Category - K to M - Persons',

'Industrial Category - N to O - Persons',

'Industrial Category - P to Q - Persons',

'Industrial Category - R to U - HHI - Persons',

'Industrial Category - R to U - Non HHI - Persons'

]

age\_industrial\_data = marginal\_workers\_data[age\_industrial\_columns]

age\_distribution = age\_industrial\_data.groupby('Age group').sum()

plt.figure(figsize=(12, 8))

sns.set(style="whitegrid")

for column in age\_distribution.columns:

sns.barplot(x=age\_distribution.index, y=column, data=age\_distribution, label=column)

plt.xlabel('Age Group')

plt.ylabel('Count')

plt.title('Age Distribution by Industrial Categories')

plt.xticks(rotation=90)

plt.legend(ncol=2, loc='upper right')

plt.tight\_layout()

plt.show()

**5. Calculating the industrial category**

import pandas as pd

industrial\_categories = [

'Industrial Category - A - Cultivators - Persons', 'Industrial Category - A - Cultivators - Males', 'Industrial Category - A - Cultivators - Females',

'Industrial Category - A - Agricultural labourers - Persons', 'Industrial Category - A - Agricultural labourers - Males', 'Industrial Category - A - Agricultural labourers - Females',

'Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Persons', 'Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males', 'Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Females',

'Industrial Category - B - Persons', 'Industrial Category - B - Males', 'Industrial Category - B - Females',

'Industrial Category - C - HHI - Persons', 'Industrial Category - C - HHI - Males', 'Industrial Category - C - HHI - Females',

'Industrial Category - C - Non HHI - Persons', 'Industrial Category - C - Non HHI - Males', 'Industrial Category - C - Non HHI - Females',

'Industrial Category - D & E - Persons', 'Industrial Category - D & E - Males', 'Industrial Category - D & E - Females',

'Industrial Category - F - Persons', 'Industrial Category - F - Males', 'Industrial Category - F - Females',

'Industrial Category - G - HHI - Persons', 'Industrial Category - G - HHI - Males', 'Industrial Category - G - HHI - Females',

'Industrial Category - G - Non HHI - Persons', 'Industrial Category - G - Non HHI - Males', 'Industrial Category - G - Non HHI - Females',

'Industrial Category - H - Persons', 'Industrial Category - H - Males', 'Industrial Category - H - Females',

'Industrial Category - I - Persons', 'Industrial Category - I - Males', 'Industrial Category - I - Females',

'Industrial Category - J - HHI - Persons', 'Industrial Category - J - HHI - Males', 'Industrial Category - J - HHI - Females',

'Industrial Category - J - Non HHI - Persons', 'Industrial Category - J - Non HHI - Males', 'Industrial Category - J - Non HHI - Females',

'Industrial Category - K to M - Persons', 'Industrial Category - K to M - Males', 'Industrial Category - K to M - Females',

'Industrial Category - N to O - Persons', 'Industrial Category - N to O - Males', 'Industrial Category - N to O - Females',

'Industrial Category - P to Q - Persons', 'Industrial Category - P to Q - Males', 'Industrial Category - P to Q - Females',

'Industrial Category - R to U - HHI - Persons', 'Industrial Category - R to U - HHI - Males', 'Industrial Category - R to U - HHI - Females',

'Industrial Category - R to U - Non HHI - Persons', 'Industrial Category - R to U - Non HHI - Males', 'Industrial Category - R to U - Non HHI - Females'

]

industrial\_category\_distribution = marginal\_workers\_data[industrial\_categories].sum()

print("Industrial Category Distribution:")

print(industrial\_category\_distribution)

**5.1: for visualization the distribution**

category\_distribution\_df = pd.DataFrame(industrial\_category\_distribution, columns=['Count'])

category\_distribution\_df.reset\_index(inplace=True)

category\_distribution\_df.rename(columns={'index': 'Industrial Category'}, inplace=True)

plt.figure(figsize=(12, 8))

sns.barplot(x='Count', y='Industrial Category', data=category\_distribution\_df, orient='h', palette='viridis')

plt.xlabel('Count')

plt.ylabel('Industrial Category')

plt.title('Industrial Category Distribution')

plt.show()

**6. Calculating the sex distribution**

industrial\_categories\_pairs = [

('Industrial Category - A - Cultivators - Persons', 'Industrial Category - A - Cultivators - Males', 'Industrial Category - A - Cultivators - Females'),

('Industrial Category - A - Agricultural labourers - Persons', 'Industrial Category - A - Agricultural labourers - Males', 'Industrial Category - A - Agricultural labourers - Females'),

('Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Persons', 'Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males', 'Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Females'),

('Industrial Category - B - Persons', 'Industrial Category - B - Males', 'Industrial Category - B - Females'),

('Industrial Category - C - HHI - Persons', 'Industrial Category - C - HHI - Males', 'Industrial Category - C - HHI - Females'),

('Industrial Category - C - Non HHI - Persons', 'Industrial Category - C - Non HHI - Males', 'Industrial Category - C - Non HHI - Females'),

('Industrial Category - D & E - Persons', 'Industrial Category - D & E - Males', 'Industrial Category - D & E - Females'),

('Industrial Category - F - Persons', 'Industrial Category - F - Males', 'Industrial Category - F - Females'),

('Industrial Category - G - HHI - Persons', 'Industrial Category - G - HHI - Males', 'Industrial Category - G - HHI - Females'),

('Industrial Category - G - Non HHI - Persons', 'Industrial Category - G - Non HHI - Males', 'Industrial Category - G - Non HHI - Females'),

('Industrial Category - H - Persons', 'Industrial Category - H - Males', 'Industrial Category - H - Females'),

('Industrial Category - I - Persons', 'Industrial Category - I - Males', 'Industrial Category - I - Females'),

('Industrial Category - J - HHI - Persons', 'Industrial Category - J - HHI - Males', 'Industrial Category - J - HHI - Females'),

('Industrial Category - J - Non HHI - Persons', 'Industrial Category - J - Non HHI - Males', 'Industrial Category - J - Non HHI - Females'),

('Industrial Category - K to M - Persons', 'Industrial Category - K to M - Males', 'Industrial Category - K to M - Females'),

('Industrial Category - N to O - Persons', 'Industrial Category - N to O - Males', 'Industrial Category - N to O - Females'),

('Industrial Category - P to Q - Persons', 'Industrial Category - P to Q - Males', 'Industrial Category - P to Q - Females'),

('Industrial Category - R to U - HHI - Persons', 'Industrial Category - R to U - HHI - Males', 'Industrial Category - R to U - HHI - Females'),

('Industrial Category - R to U - Non HHI - Persons', 'Industrial Category - R to U - Non HHI - Males', 'Industrial Category - R to U - Non HHI - Females')

]

sex\_distribution = {}

for category\_pair in industrial\_categories\_pairs:

total\_persons = marginal\_workers\_data[category\_pair[0]].sum()

total\_males = marginal\_workers\_data[category\_pair[1]].sum()

total\_females = marginal\_workers\_data[category\_pair[2]].sum()

sex\_distribution[category\_pair[0]] = {

'Persons': total\_persons,

'Males': total\_males,

'Females': total\_females

}

for category, counts in sex\_distribution.items():

print(f"Category: {category}")

print(f"Persons: {counts['Persons']}")

print(f"Males: {counts['Males']}")

print(f"Females: {counts['Females']}")

print("----------------------")

**6.1 For visualize the sex distribution**

sex\_distribution\_df = pd.DataFrame.from\_dict(sex\_distribution, orient='index')

plt.figure(figsize=(12, 8))

sns.set(style="whitegrid")

sns.barplot(x=sex\_distribution\_df.index, y='Males', data=sex\_distribution\_df, label='Males', color='orange')

sns.barplot(x=sex\_distribution\_df.index, y='Females', data=sex\_distribution\_df, label='Females', color='salmon')

plt.xlabel('Industrial Categories')

plt.ylabel('Count')

plt.title('Sex Distribution by Industrial Categories')

plt.xticks(rotation=90)

plt.legend()

plt.show()