

Assessment of Marginal Workers in Tamil Nadu

Domain Name	Applied Data Science
Project Name	Assessment of Marginal Workers in Tamil Nadu-A Socioeconomic Analysis(ADS)
Date	31-10-2023
Team ID	3896
Team Members Name	VENKATESH K SUBASH S SUBESHAN M VENKATESAN M

Table of Content

1	Introduction
2	Problem Statement
3	Introduction
4	Objectives of the Project
5	Analysis Approach
6	Visualization Analysis
7	Code implementation
8	Overall Analysis
9	Conclusion

Problem statement :

"Assessment of Marginal workers in Tamil Nadu", A Socioeconomic Analysis: Analyse the demographic characteristic of Marginal workers based on age, industry category, and sex, Create visualization such as bar chart, pie chart, or heatmaps to represent the distribution across different categories.

Introduction:

This project, "Assessment of Marginal Workers in Tamil Nadu," serves as a comprehensive socioeconomic analysis aimed at examining the demographic characteristics of marginal workers in the state of Tamil Nadu, India. Marginal workers, defined as those employed for less than six months in a year, often face unique socioeconomic challenges. Understanding their demographic attributes, including age, industry category, and gender, is critical for policymakers, researchers, and organizations seeking to develop targeted interventions and policies to improve the well-being of this vulnerable population.

In pursuit of these objectives, this documentation outlines the methodology, data sources, data analysis techniques, and visualizations employed to gain insights into the demographic characteristics of marginal workers in Tamil Nadu.

Through the presentation of bar charts, pie charts, heatmaps, and other visual representations, we offer a comprehensive overview of the distribution of marginal workers across various categories. These insights aim to inform evidence-based decision-making, policy formulation, and interventions that can enhance the socioeconomic conditions and opportunities for marginal workers in the region.

Objectives of the project:

The objectives of the project, "Assessment of Marginal Workers in Tamil Nadu,"

- 1. Demographic Analysis:** To examine and understand the demographic characteristics of marginal workers in Tamil Nadu, with a focus on age, industry category, and gender.
- 2. Age Distribution:** To analyse the age distribution of marginal workers, providing insights into the age groups most affected by seasonal or temporary employment.
- 3. Industry Category:** To categorize marginal workers based on the industry or economic sector in which they are employed, offering a breakdown of their occupational diversity.
- 4. Gender Disparities:** To investigate and reveal any gender disparities within the population of marginal workers, contributing to a more nuanced understanding of their socioeconomic challenges.
- 5. Visualization:** To create visualizations such as bar charts, pie charts, and heatmaps to represent the distribution of marginal workers across different demographic and industry categories, making the data more accessible and actionable for stakeholders and policymakers.

6. Informed Decision-Making: To provide policymakers, researchers, and organizations with data-driven insights that can inform evidence-based decision-making, targeted interventions, and the formulation of policies aimed at improving the well-being and opportunities for marginal workers in Tamil Nadu.

By achieving these objectives, the project aims to contribute valuable information for addressing the unique socioeconomic needs of marginal workers and enhancing their quality of life in the region.

Analysis Approach:

1. Age Distribution Analysis:

- Create a histogram or bar chart to visualize the age distribution of marginal workers.
- Calculate relevant statistics, such as mean, median, and standard deviation, to better understand the central tendency and variability of age among marginal workers.

2. Industry Category Categorization:

- Group marginal workers based on industry categories, such as agriculture, manufacturing, construction, services, etc.
- Calculate the percentage distribution of workers in each category.

3. Gender Disparities Analysis:

- Analyse the gender distribution among marginal workers and calculate the proportion of male and female workers.
- Visualize the gender distribution using charts like pie charts or bar charts.

Visualization Analysis:

To analyse the demographic characteristics of marginal workers in Tamil Nadu, several visualization types can be employed to effectively represent the distribution of these workers across different categories. Here are some visualization types that can be used in the analysis:

1. Bar Charts:

- Bar charts can represent the distribution of marginal workers in different categories such as age groups, industry categories, and gender. They provide a clear visual comparison of the number or proportion of workers in each category.

2. Pie Charts:

- Pie charts are useful for showing the proportion or percentage of marginal workers in various industry categories, age groups, or gender. They provide a quick overview of the distribution.

3. Heatmaps:

- Heatmaps can be used to display the concentration or density of marginal workers across different regions or districts in Tamil Nadu. They provide a geographic view of where these workers are most prevalent.

4. Scatterplots:

- Scatterplots are valuable for examining correlations between variables, such as age and educational qualifications, or age and income. They can help identify patterns or trends within the data.

Code Implementation:

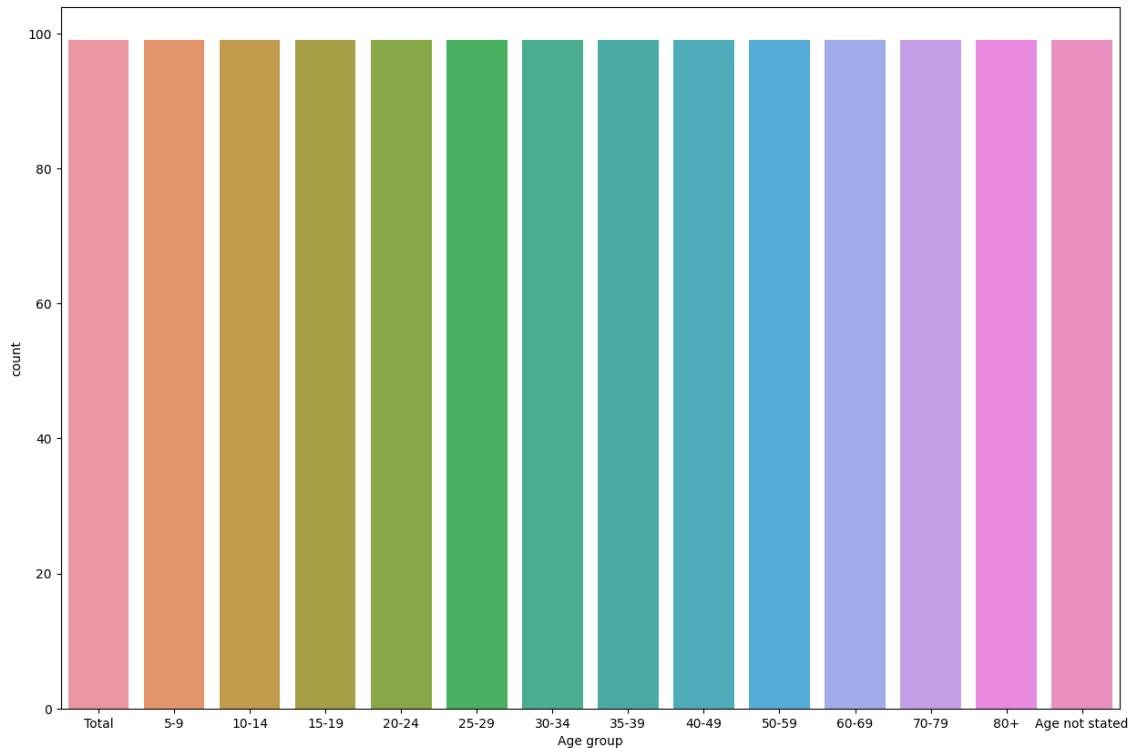
```
In [17]: 1 dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1386 entries, 0 to 1385
Data columns (total 69 columns):
 #   Column          Dtype  
 0   Table Code      object  
 1   State Code      int64  
 2   District Code   object  
 3   Area Name       object  
 4   Total/ Rural/ Urban object 
 5   Age group       object  
 6   ...
```

Data Visualization

```
In [18]: 1 plt.figure(figsize=(15,10))
2 sns.countplot(x='Age group',data=dataset)
```

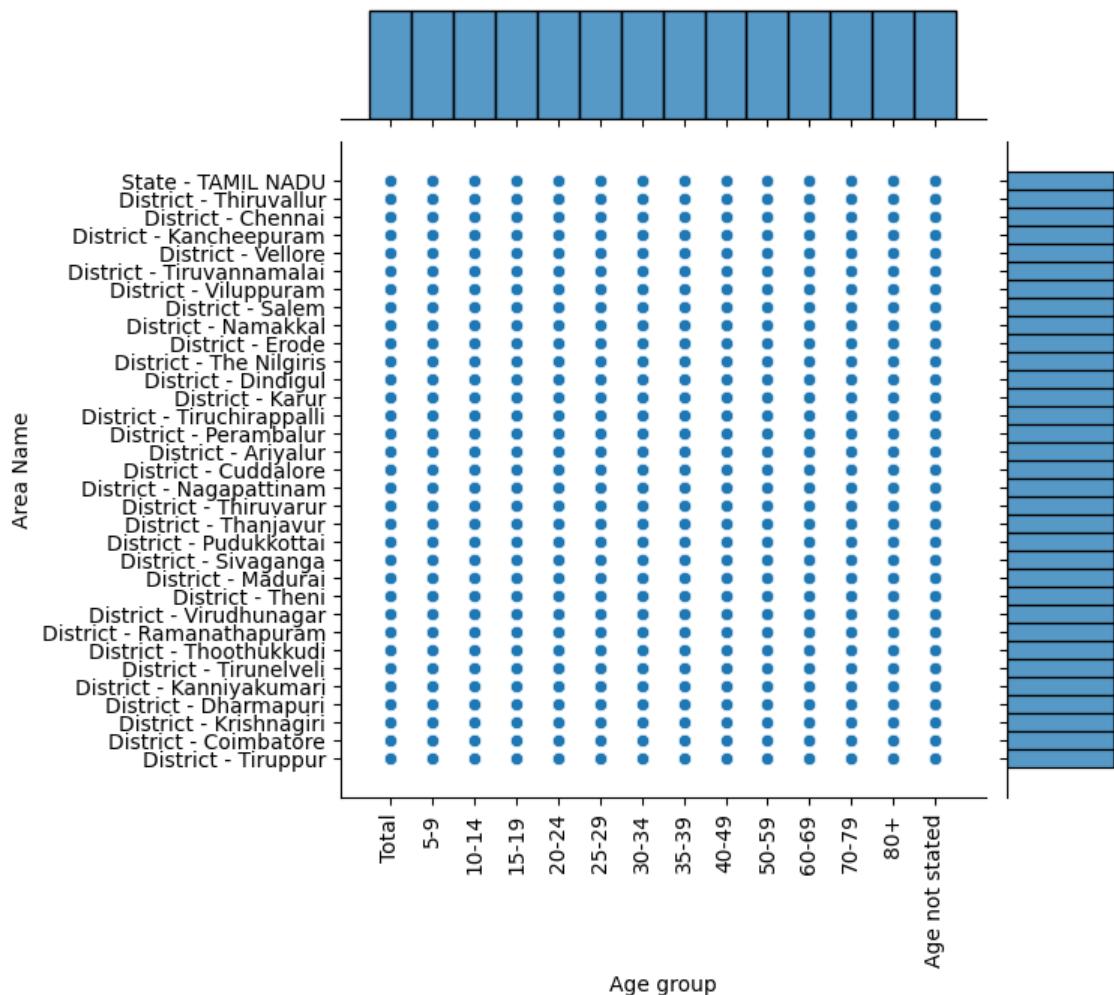
```
Out[18]: <AxesSubplot: xlabel='Age group', ylabel='count'>
```



```
In [19]: 1 plt.figure(figsize=(10,10))
2 sns.jointplot(dataset,x='Age group',y='Area Name')
3 plt.xticks(rotation=90)
```

```
Out[19]: ([0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13],
[Text(0, 0, 'Total'),
Text(1, 0, '5-9'),
Text(2, 0, '10-14'),
Text(3, 0, '15-19'),
Text(4, 0, '20-24'),
Text(5, 0, '25-29'),
Text(6, 0, '30-34'),
Text(7, 0, '35-39'),
Text(8, 0, '40-49'),
Text(9, 0, '50-59'),
Text(10, 0, '60-69'),
Text(11, 0, '70-79'),
Text(12, 0, '80+'),
Text(13, 0, 'Age not stated')])
```

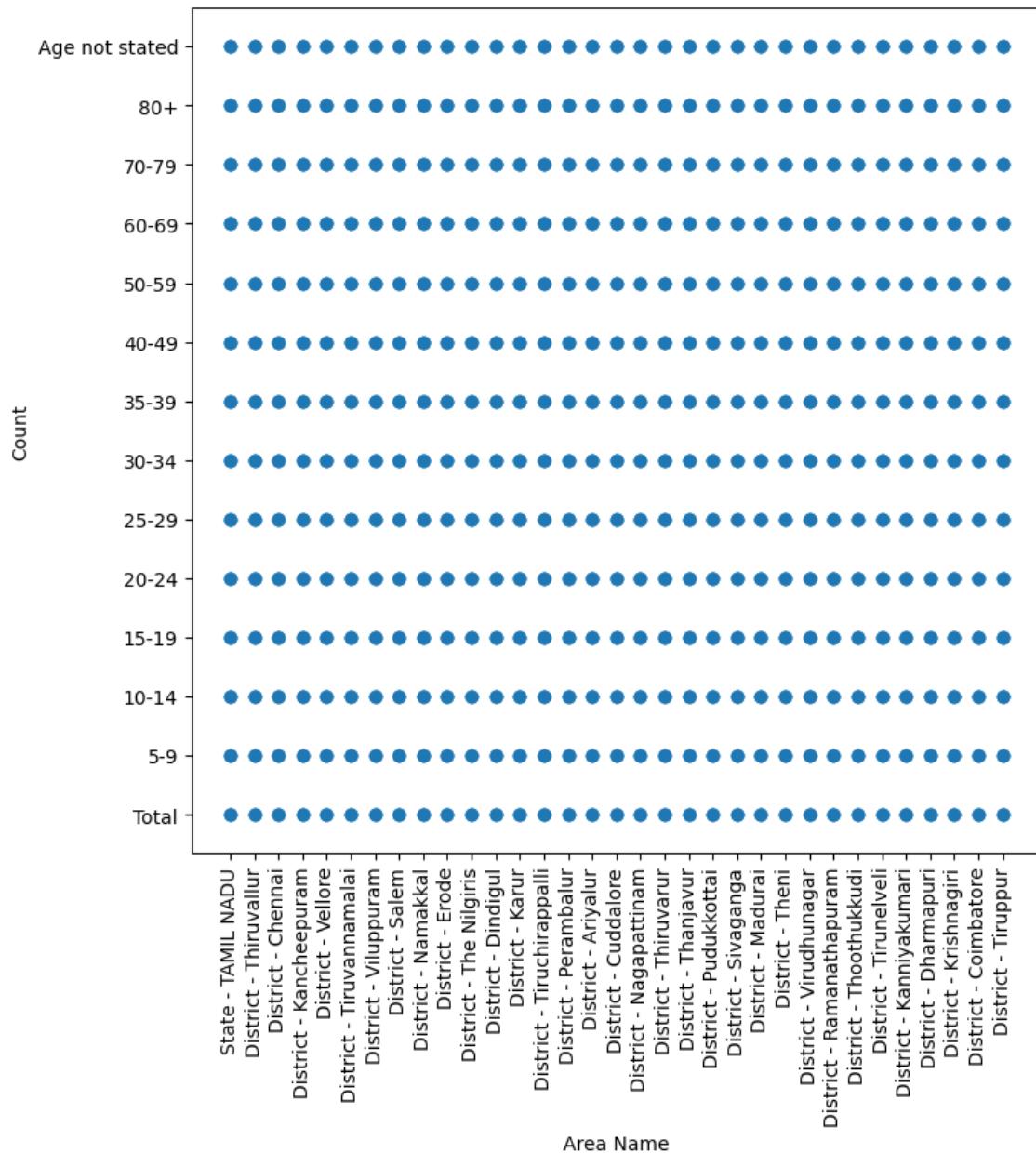
<Figure size 1000x1000 with 0 Axes>



In [20]:

```
1 plt.figure(figsize=(8,8))
2 plt.scatter(dataset[ 'Area Name' ],dataset[ 'Age group' ])
3 plt.xticks(rotation=90)
4 plt.xlabel("Area Name")
5 plt.ylabel("Count")
6
```

Out[20]: Text(0, 0.5, 'Count')



Analysis by Age group and Industrial category by persons

```
In [21]: 1 IndustryCategorybyPerson = dataset[['Industrial Category - A - Cultiva  
2 'Industrial Category - A - Agricultu  
3 'Industrial Category - A - Plantatio  
4 'Industrial Category - B - Persons',  
5 'Industrial Category - C - HHI - Per  
6 'Industrial Category - C - Non HHI -  
7 'Industrial Category - D & E - Perso  
8 'Industrial Category - F - Persons',  
9 'Industrial Category - G - HHI - Per  
10 'Industrial Category - G - Non HHI -  
11 'Industrial Category - H - Persons',  
12 'Industrial Category - I - Persons',  
13 'Industrial Category - J - HHI - Per  
14 'Industrial Category - J - Non HHI -  
15 'Industrial Category - K to M - Pers  
16 'Industrial Category - N to O - Pers  
17 'Industrial Category - P to Q - Pers  
18 'Industrial Category - R to U - HHI  
19 'Industrial Category - R to U - Non
```

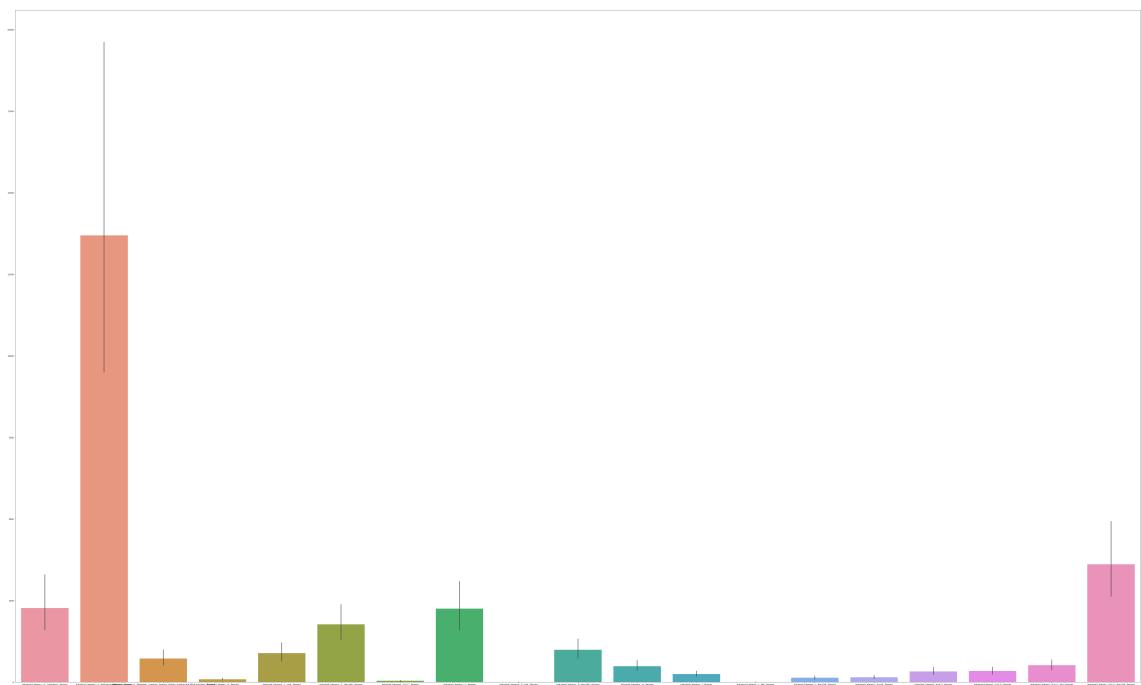
```
In [22]: 1 IndustryCategorybyPerson.head()
```

Out[22]:

	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
0	393082	2372446	125099	14979	154133	306528	7137	390275
1	3363	4169	198	6	228	246	6	264
2	5072	13939	834	70	996	2808	24	1522
3	17864	102106	4613	689	6642	28826	319	18661
4	33647	216966	9171	1490	14366	50791	1032	41762

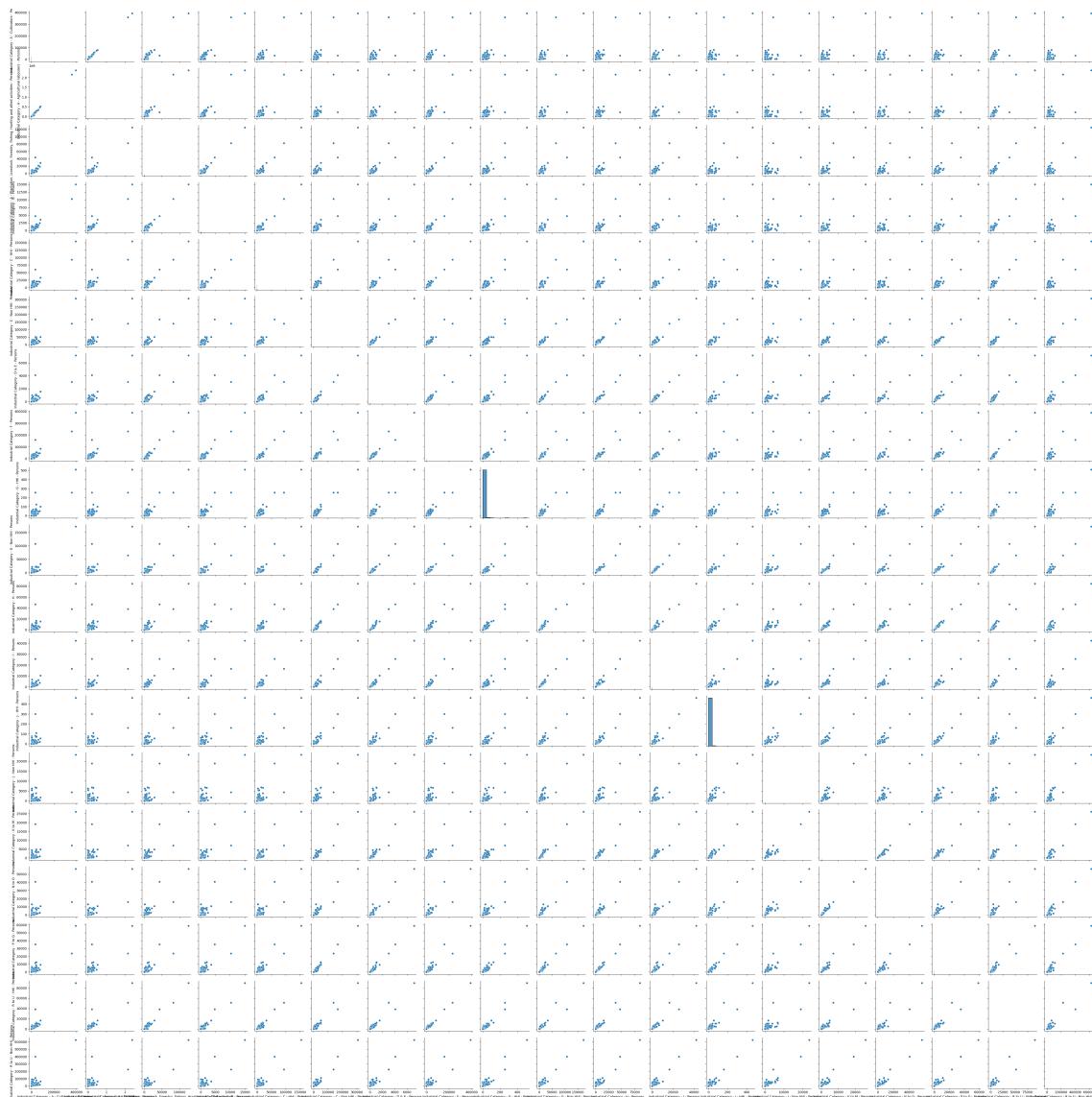
```
In [23]: 1 plt.figure(figsize=(100,60))
2 sns.barplot(IndustryCategorybyPerson)
```

Out[23]: <AxesSubplot: >



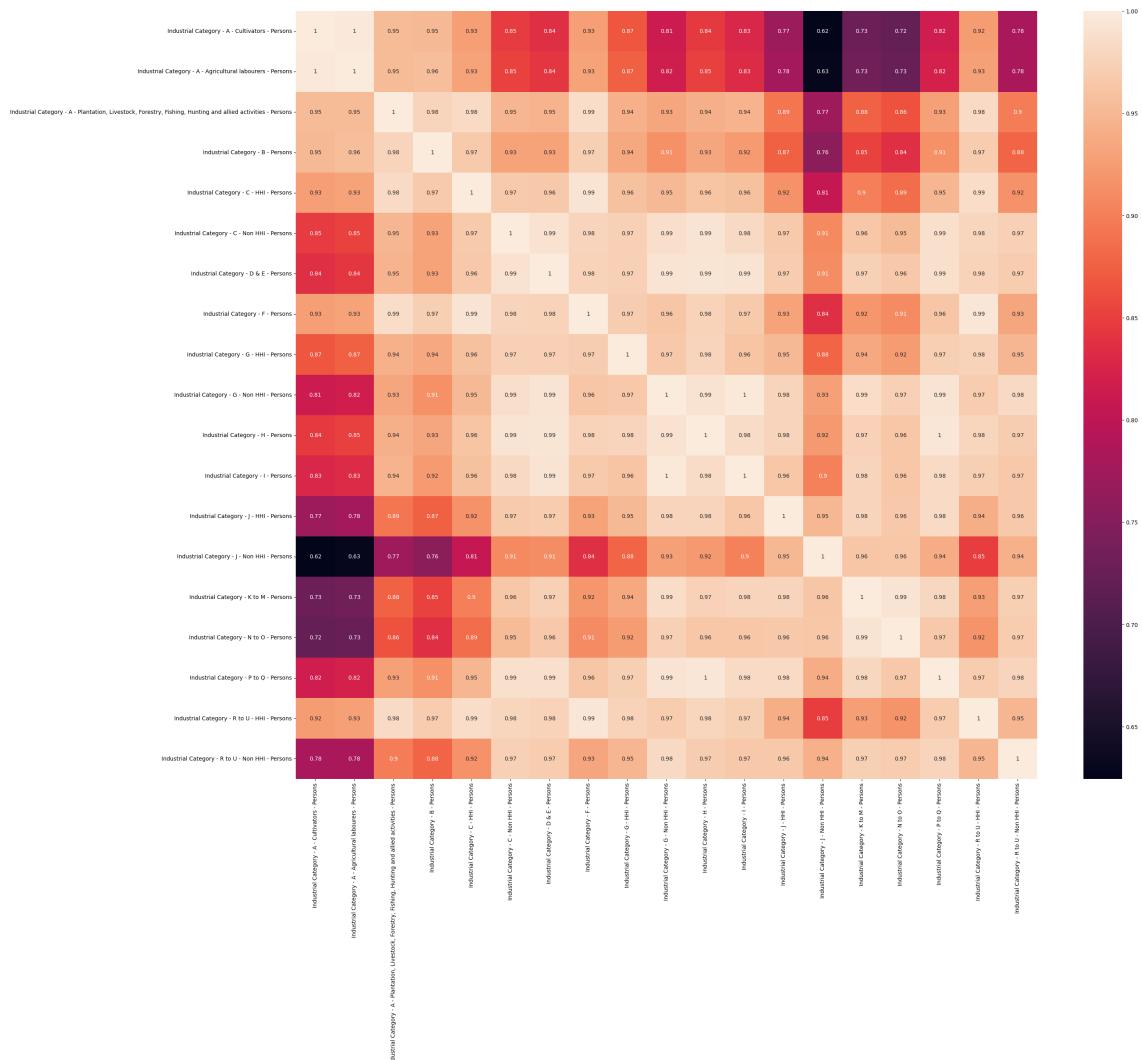
```
In [24]: 1 sns.pairplot(IndustryCategorybyPerson)
```

```
Out[24]: <seaborn.axisgrid.PairGrid at 0x1a99131d250>
```



```
In [25]: 1 plt.figure(figsize=(30,25))
2 sns.heatmap(IndustryCategorybyPerson.corr(), annot=True)
```

Out[25]: <AxesSubplot: >



Analysis by Age group and Industrial category by Males

```
In [26]: 1 IndustryCategorybyMale = dataset[['Industrial Category - A - Cultivato  
2 'Industrial Category - A - Agricultu  
3 'Industrial Category - A - Plantatio  
4 'Industrial Category - B - Males',  
5 'Industrial Category - C - HHI - Mal  
6 'Industrial Category - C - Non HHI -  
7 'Industrial Category - D & E - Males  
8 'Industrial Category - F - Males',  
9 'Industrial Category - G - HHI - Mal  
10 'Industrial Category - G - Non HHI -  
11 'Industrial Category - H - Males',  
12 'Industrial Category - I - Males',  
13 'Industrial Category - J - HHI - Mal  
14 'Industrial Category - J - Non HHI -  
15 'Industrial Category - K to M - Male  
16 'Industrial Category - N to O - Male  
17 'Industrial Category - P to Q - Male  
18 'Industrial Category - R to U - HHI  
19 'Industrial Category - R to U - Non
```

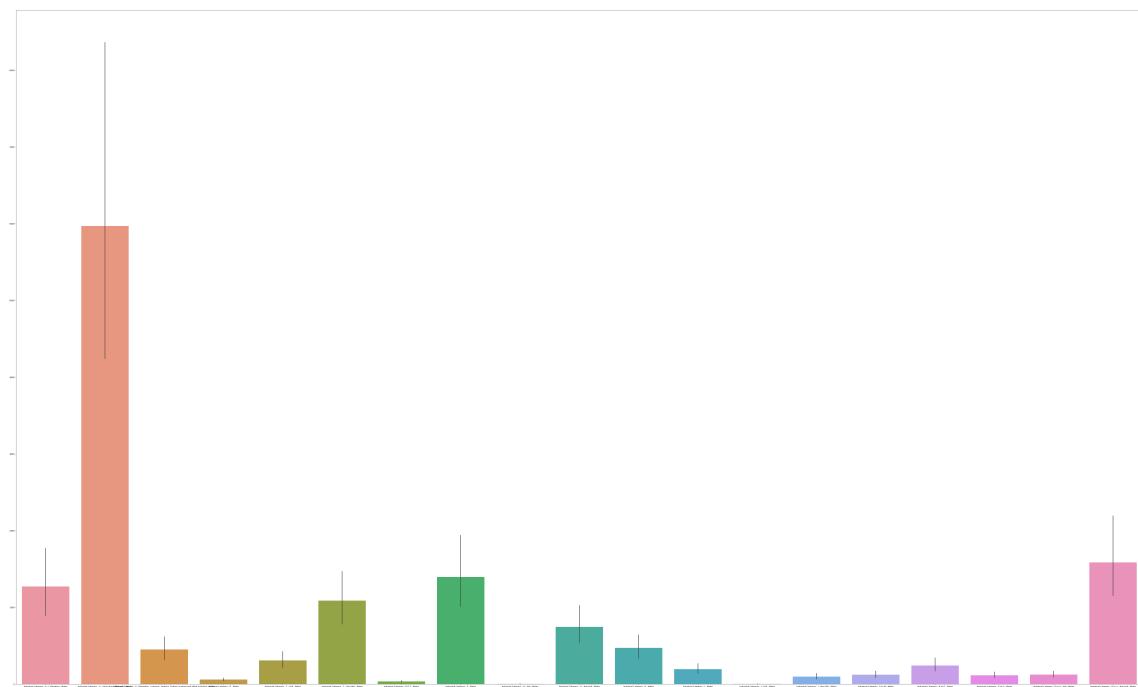
```
In [27]: 1 IndustryCategorybyMale.head()
```

Out[27]:

	Industrial Category - A - Cultivators - Males	Industrial Category - A - Agricultural labourers - Males	Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males	Industrial Category - B - Males	Industrial Category - C - HHI - Males	Industrial Category - C - Non HHI - Males	Industrial Category - D & E - Males	Industrial Category - F - Males
0	220314	1034184	78052	10290	53418	188464	6003	241619
1	1592	1992	106	0	113	127	4	176
2	2486	6974	464	39	394	1416	14	1081
3	9336	51763	3157	544	2027	14958	285	15695
4	16930	98959	6302	1092	4374	32445	931	32398

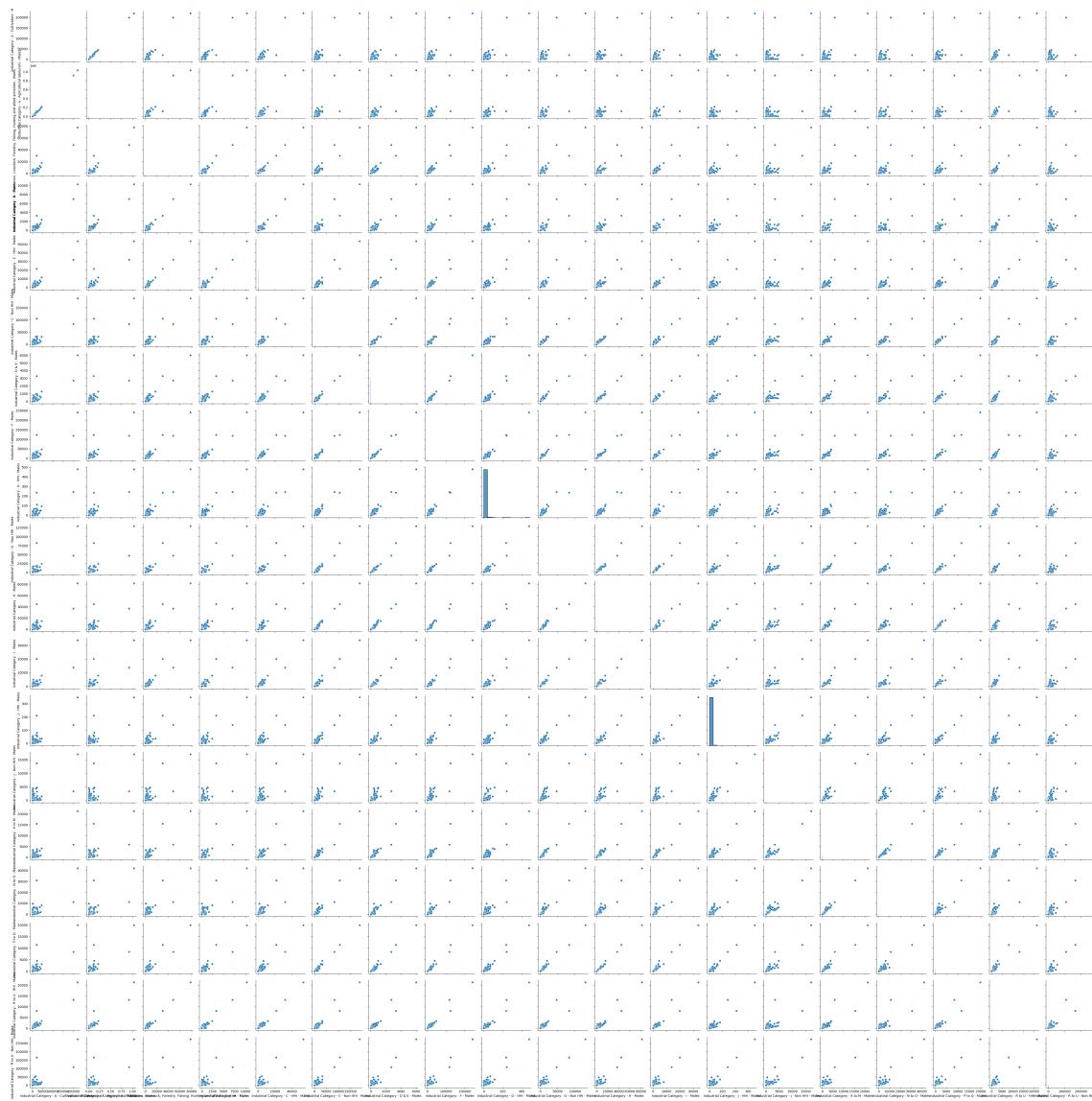
```
In [28]: 1 plt.figure(figsize=(100,60))
2 sns.barplot(IndustryCategorybyMale)
```

Out[28]: <AxesSubplot: >



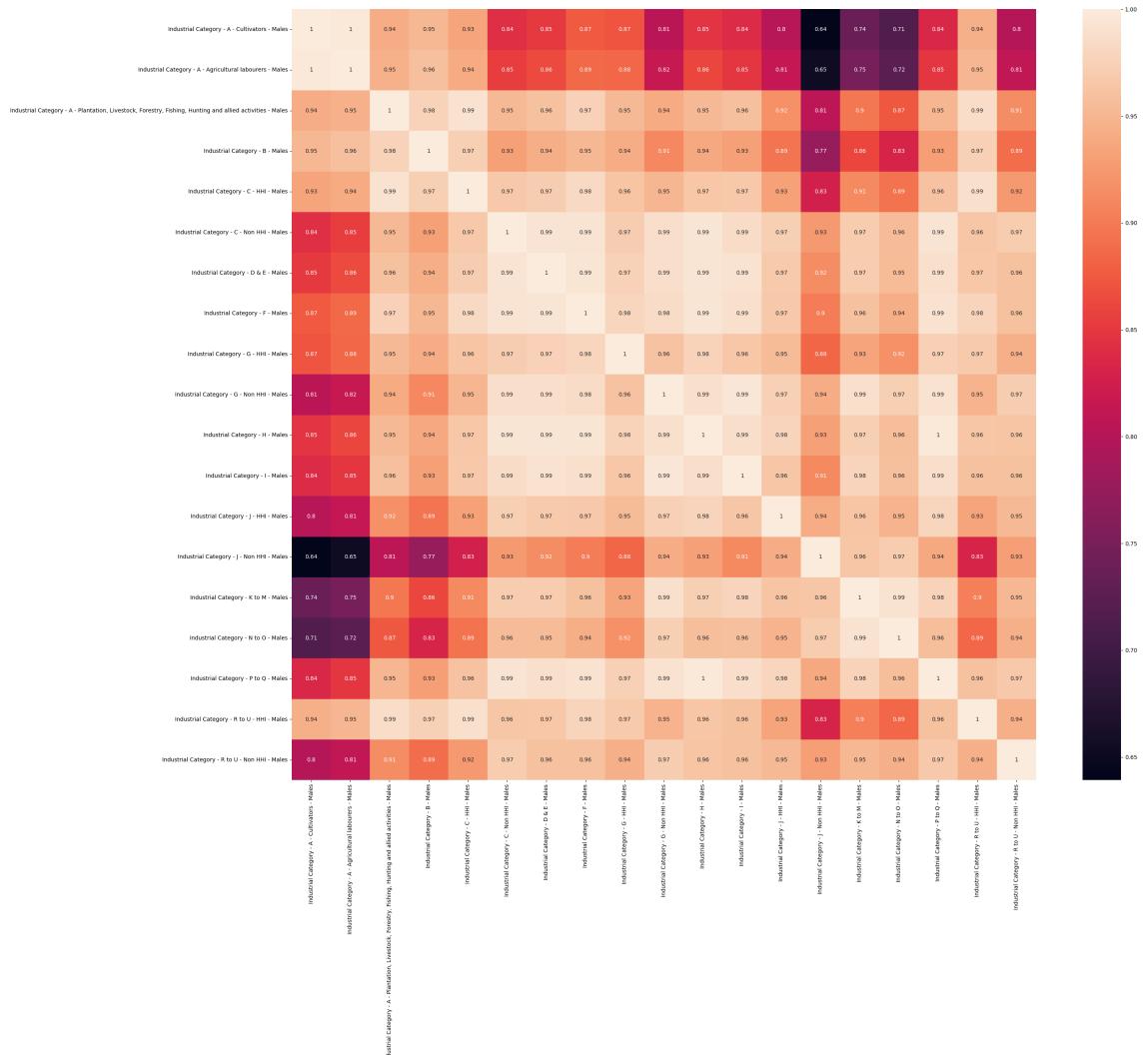
```
In [29]: 1 sns.pairplot(IndustryCategorybyMale)
```

```
Out[29]: <seaborn.axisgrid.PairGrid at 0x1a9c5bba950>
```



```
In [30]: 1 plt.figure(figsize=(30,25))
2 sns.heatmap(IndustryCategorybyMale.corr(), annot=True)
```

Out[30]: <AxesSubplot: >



Analysis by Age group and Industrial category by Females

```
In [31]: 1 IndustryCategorybyFemale = dataset[['Industrial Category - A - Cultiva  
2 'Industrial Category - A - Agricultu  
3 'Industrial Category - A - Plantatio  
4 'Industrial Category - B - Females',  
5 'Industrial Category - C - HHI - Fem  
6 'Industrial Category - C - Non HHI -  
7 'Industrial Category - D & E - Femal  
8 'Industrial Category - F - Females',  
9 'Industrial Category - G - HHI - Fem  
10 'Industrial Category - G - Non HHI -  
11 'Industrial Category - H - Females',  
12 'Industrial Category - I - Females',  
13 'Industrial Category - J - HHI - Fem  
14 'Industrial Category - J - Non HHI -  
15 'Industrial Category - K to M - Fema  
16 'Industrial Category - N to O - Fema  
17 'Industrial Category - P to Q - Fema  
18 'Industrial Category - R to U - HHI  
19 'Industrial Category - R to U - Non
```

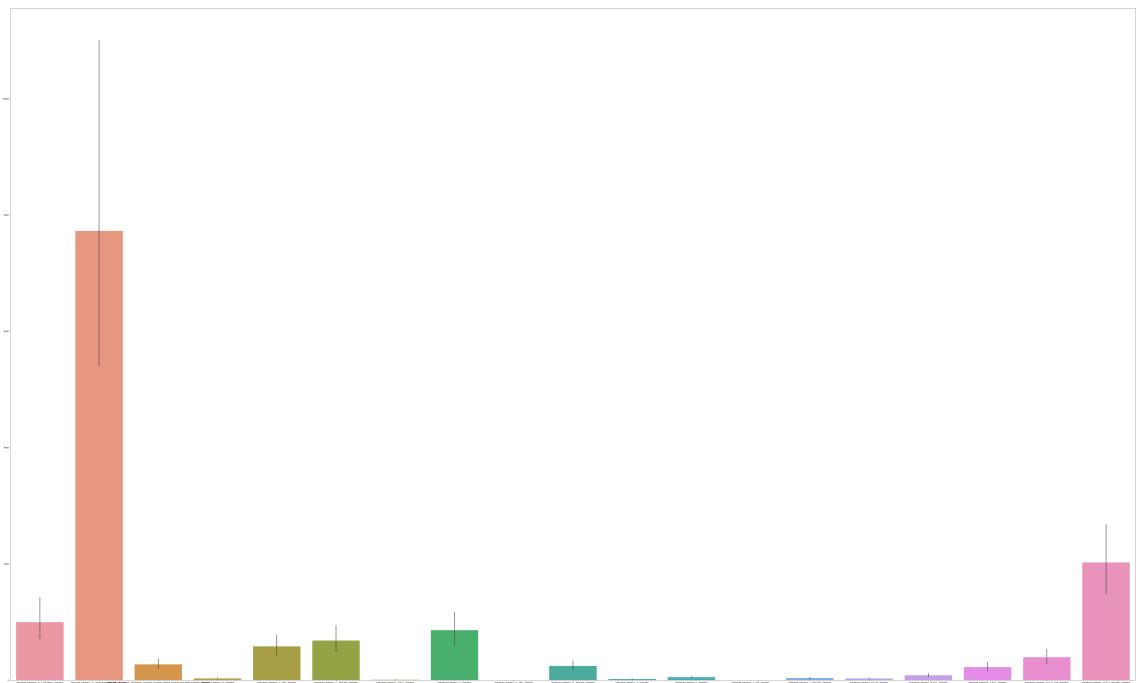
```
In [32]: 1 IndustryCategorybyFemale.head()
```

Out[32]:

	Industrial Category - A - Cultivators - Females	Industrial Category - A - Agricultural labourers - Females	Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Females	Industrial Category - B - Females	Industrial Category - C - HHI - Females	Industrial Category - C - Non HHI - Females	Industrial Category - D & E - Females	Industrial Category - F - Females
0	172768	1338262	47047	4689	100715	118064	1134	148656
1	1771	2177	92	6	115	119	2	88
2	2586	6965	370	31	602	1392	10	441
3	8528	50343	1456	145	4615	13868	34	2966
4	16717	118007	2869	398	9992	18346	101	9364

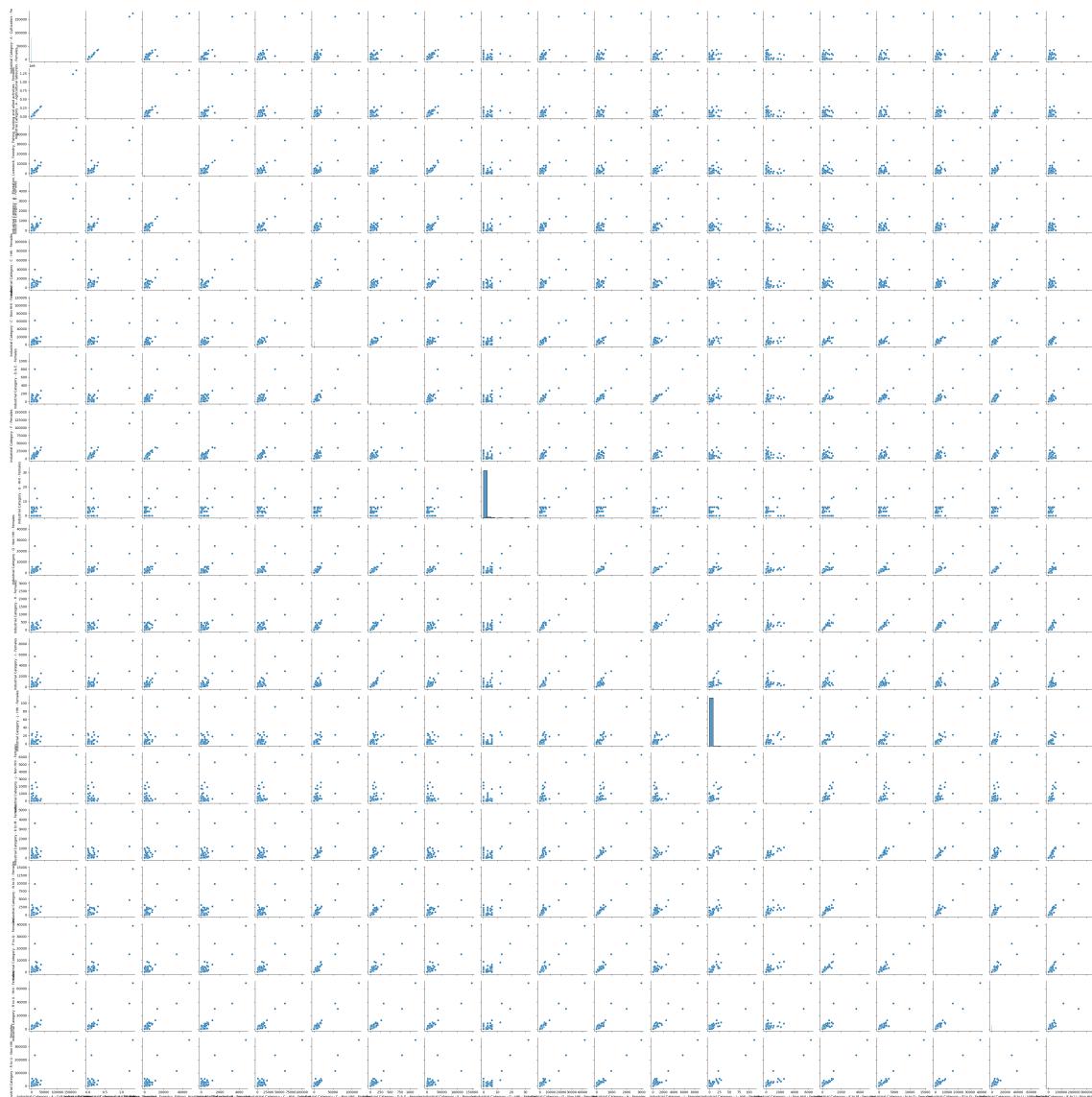
```
In [33]: 1 plt.figure(figsize=(100,60))
2 sns.barplot(IndustryCategorybyFemale)
```

Out[33]: <AxesSubplot: >



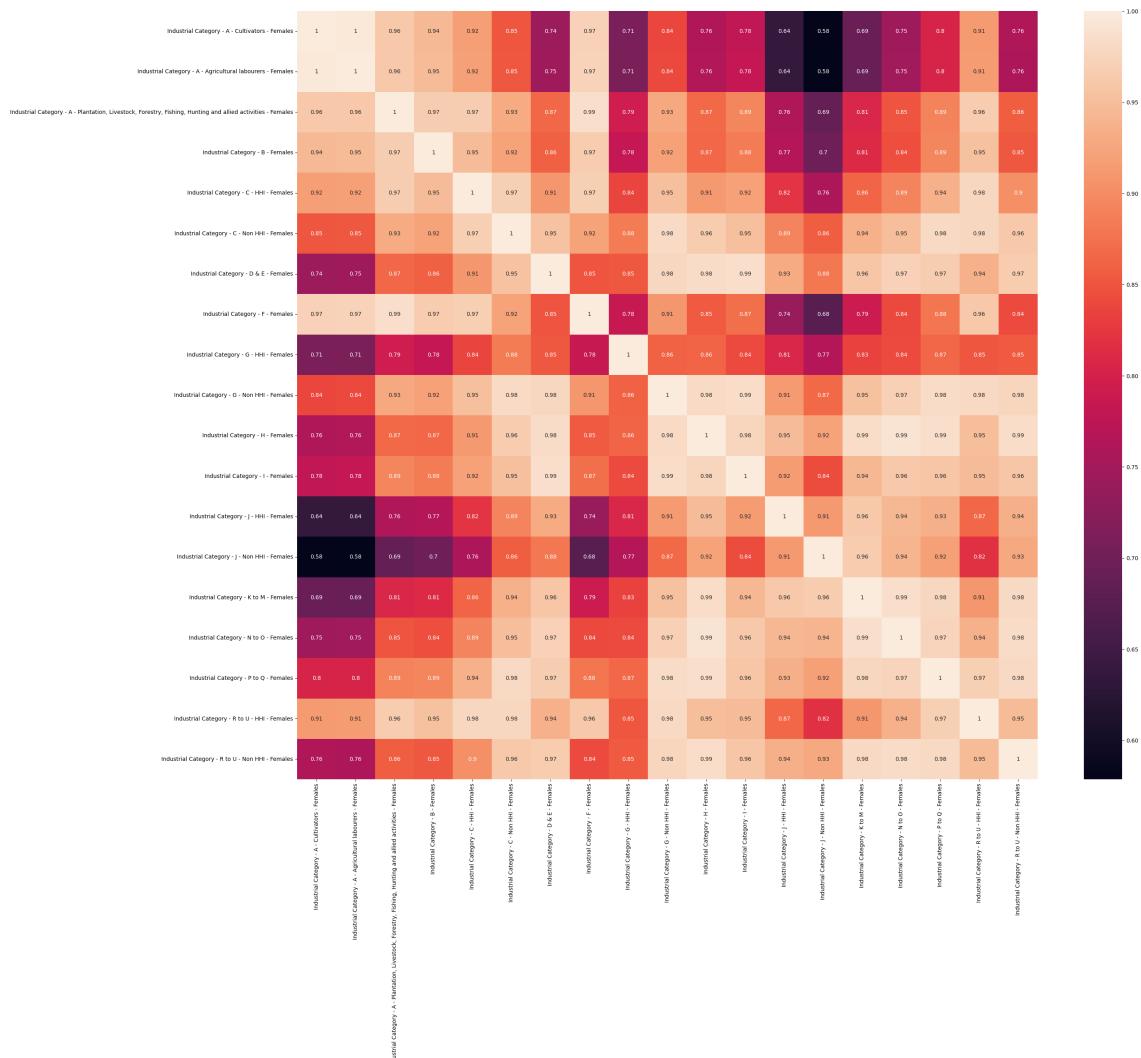
```
In [34]: 1 sns.pairplot(IndustryCategorybyFemale)
```

```
Out[34]: <seaborn.axisgrid.PairGrid at 0x1a9d3058d90>
```



```
In [35]: 1 plt.figure(figsize=(30,25))
2 sns.heatmap(IndustryCategorybyFemale.corr(), annot=True)
```

Out[35]: <AxesSubplot: >



Visualization of Industrial Category by Age group and Area Name

Industrial Category A

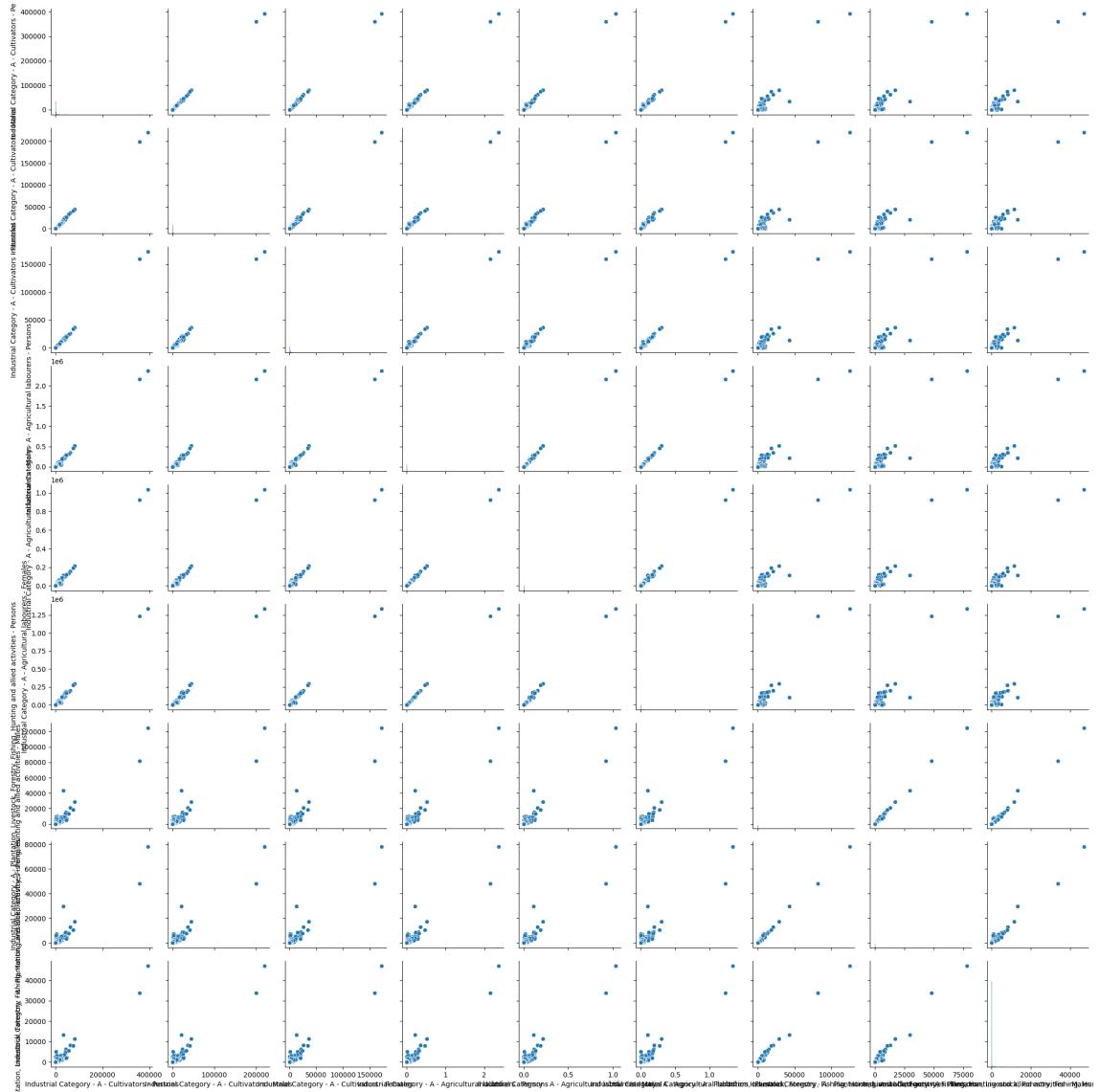
```
In [36]: 1 IndustrialCategoryA = dataset[['Industrial Category - A - Cultivators',
2 'Industrial Category - A - Cultivators - Males',
3 'Industrial Category - A - Cultivators - Females',
4 'Industrial Category - A - Agricultural labourers - Persons',
5 'Industrial Category - A - Agricultural labourers - Males',
6 'Industrial Category - A - Agricultural labourers - Females',
7 'Industrial Category - A - Plantation, Livestock, Forestry, Fis
8 'Industrial Category - A - Plantation, Livestock, Forestry, Fis
9 'Industrial Category - A - Plantation, Livestock, Forestry, Fis
```

Pair Plot

```
In [37]: 1 plt.figure(figsize=(10,10))
2 sns.pairplot(IndustrialCategoryA)
```

Out[37]: <seaborn.axisgrid.PairGrid at 0x1aa6ecc2d90>

<Figure size 1000x1000 with 0 Axes>



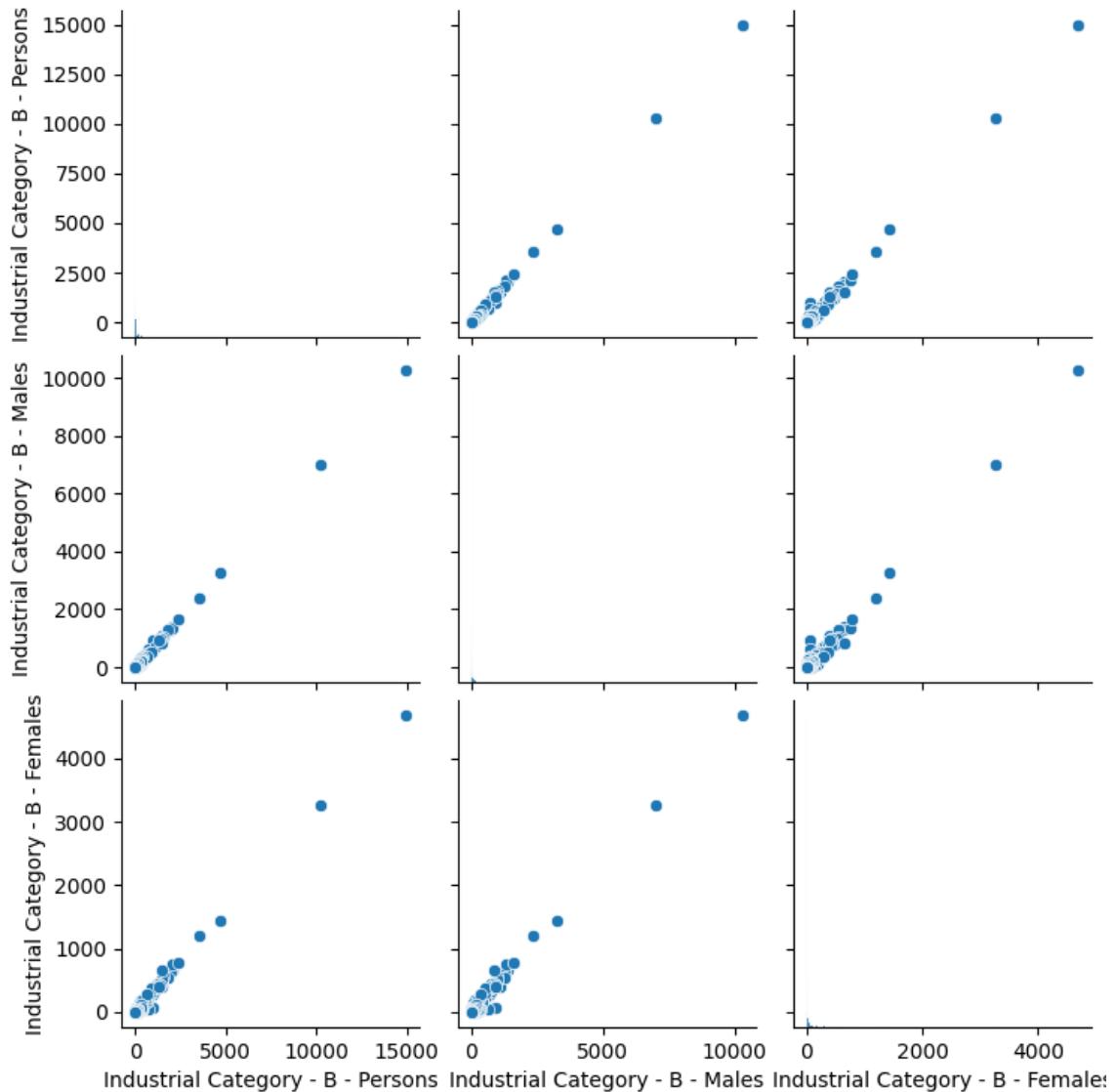
Industrial Category B

```
In [38]: 1 IndustrialCategoryB = dataset[['Industrial Category - B - Persons', 'I
2 'Industrial Category - B - Females']]
```

Pair Plot

```
In [39]: 1 sns.pairplot(IndustrialCategoryB)
          2 plt.xticks(rotation=30)
```

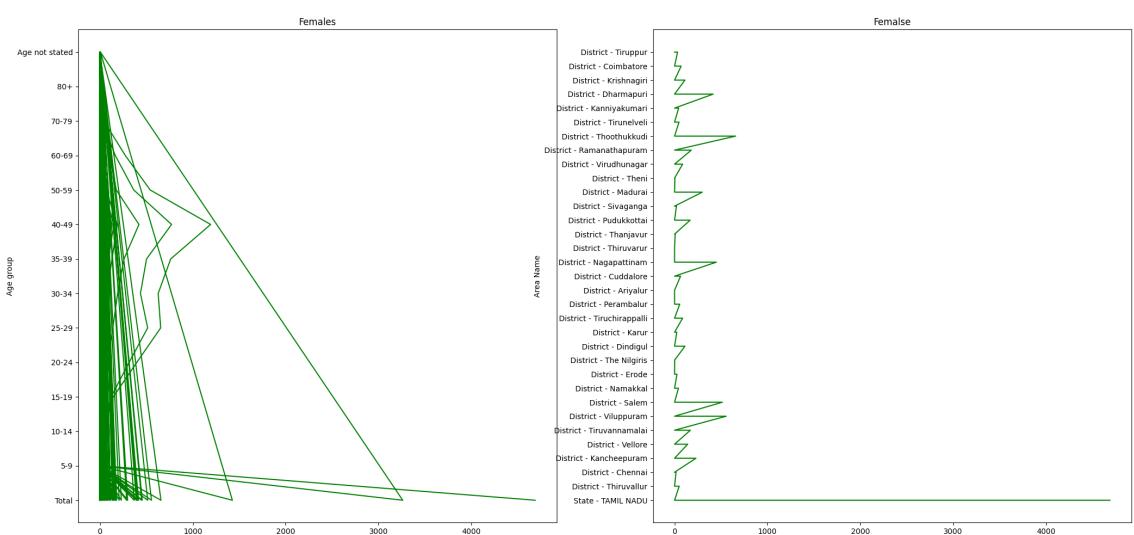
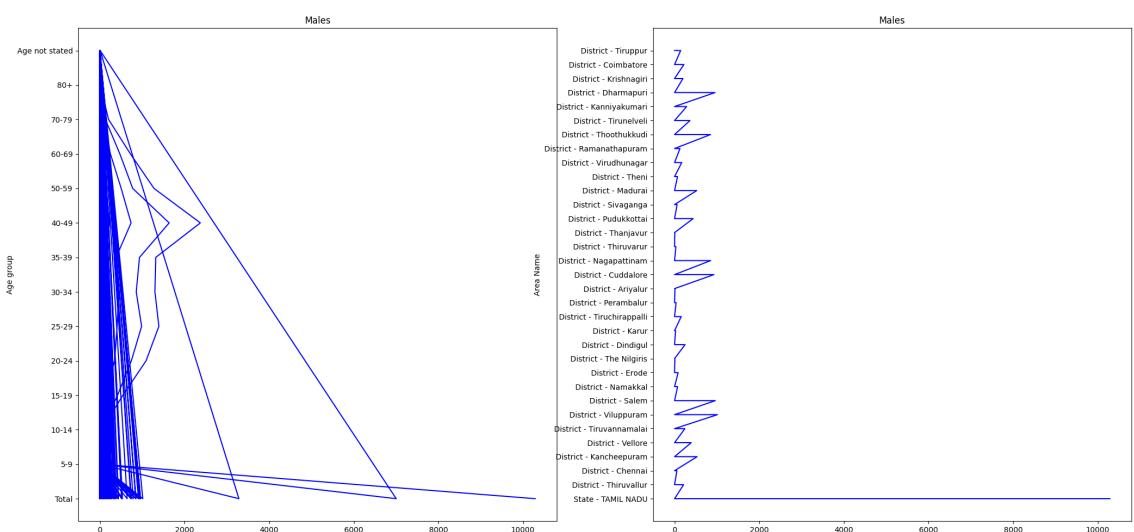
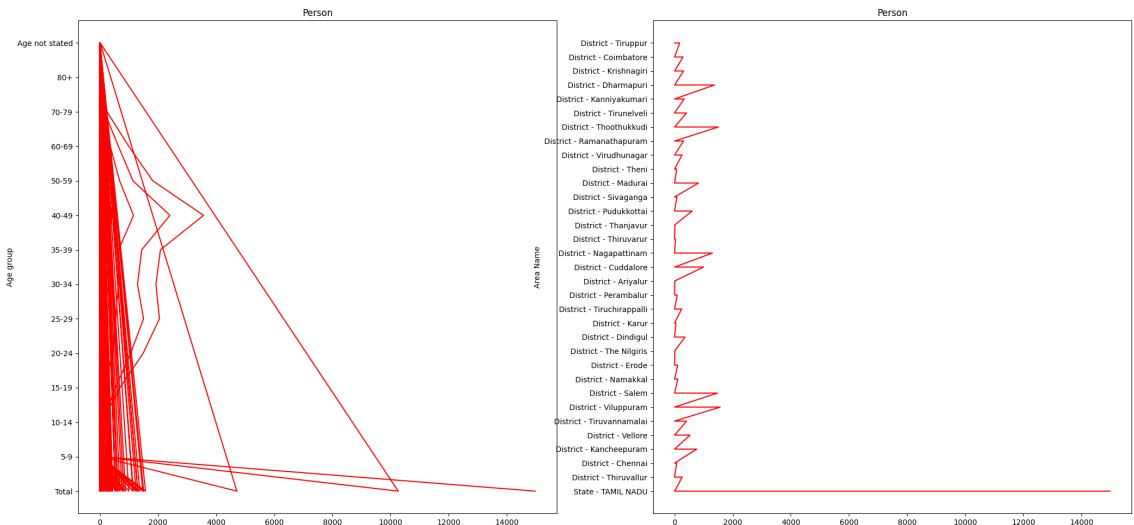
```
Out[39]: (array([-2000., 0., 2000., 4000., 6000.]),
 [Text(-2000.0, 0, '-2000'),
 Text(0.0, 0, '0'),
 Text(2000.0, 0, '2000'),
 Text(4000.0, 0, '4000'),
 Text(6000.0, 0, '6000')])
```



SUB plot Comparing Age group and Area Name

```
In [40]: 1 fig, ax = plt.subplots(figsize=(25,40),nrows=3, ncols=2)
2 ax[0, 0].plot(IndustrialCategoryB['Industrial Category - B - Persons'])
3 ax[0, 1].plot(IndustrialCategoryB['Industrial Category - B - Persons'])
4 ax[1, 0].plot(IndustrialCategoryB['Industrial Category - B - Males'],
5 ax[1, 1].plot(IndustrialCategoryB['Industrial Category - B - Males']),
6 ax[2, 0].plot(IndustrialCategoryB['Industrial Category - B - Females'])
7 ax[2, 1].plot(IndustrialCategoryB['Industrial Category - B - Females'])
8
9 ax[0,0].set_ylabel("Age group")
10 ax[0,1].set_ylabel("Area Name")
11 ax[0,0].set_title("Person")
12 ax[0,1].set_title("Person")
13
14
15 ax[1,0].set_ylabel("Age group")
16 ax[1,1].set_ylabel("Area Name")
17 ax[1,0].set_title("Males")
18 ax[1,1].set_title("Males")
19
20
21 ax[2,0].set_ylabel("Age group")
22 ax[2,1].set_ylabel("Area Name")
23 ax[2,0].set_title("Females")
24 ax[2,1].set_title("Femalse")
25
26
27
```

Out[40]: Text(0.5, 1.0, 'Femalse')



Industrial Category c

In [41]:

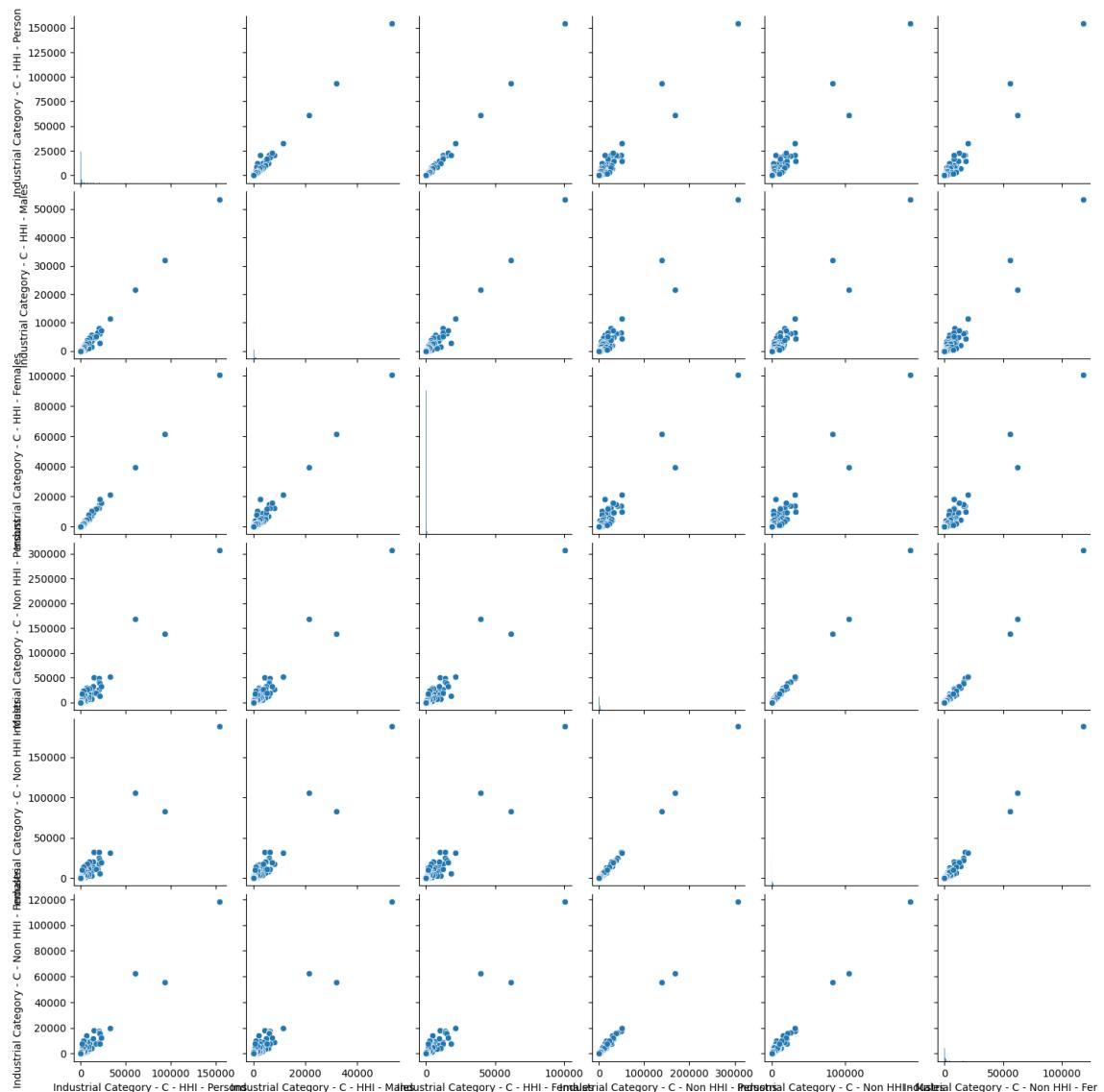
```
1 IndustrialCategoryC = dataset[['Industrial Category - C - HHI - Person  
2             'Industrial Category - C - HHI - Males',  
3             'Industrial Category - C - HHI - Females',  
4             'Industrial Category - C - Non HHI - Persons',  
5             'Industrial Category - C - Non HHI - Males',  
6             'Industrial Category - C - Non HHI - Females']]
```

Pair Plot

In [42]:

```
1 sns.pairplot(IndustrialCategoryC)  
2 plt.xticks(rotation=90)
```

Out[42]: (array([-50000., 0., 50000., 100000., 150000.]),
[Text(-50000.0, 0, '-50000'),
Text(0.0, 0, '0'),
Text(50000.0, 0, '50000'),
Text(100000.0, 0, '100000'),
Text(150000.0, 0, '150000')])

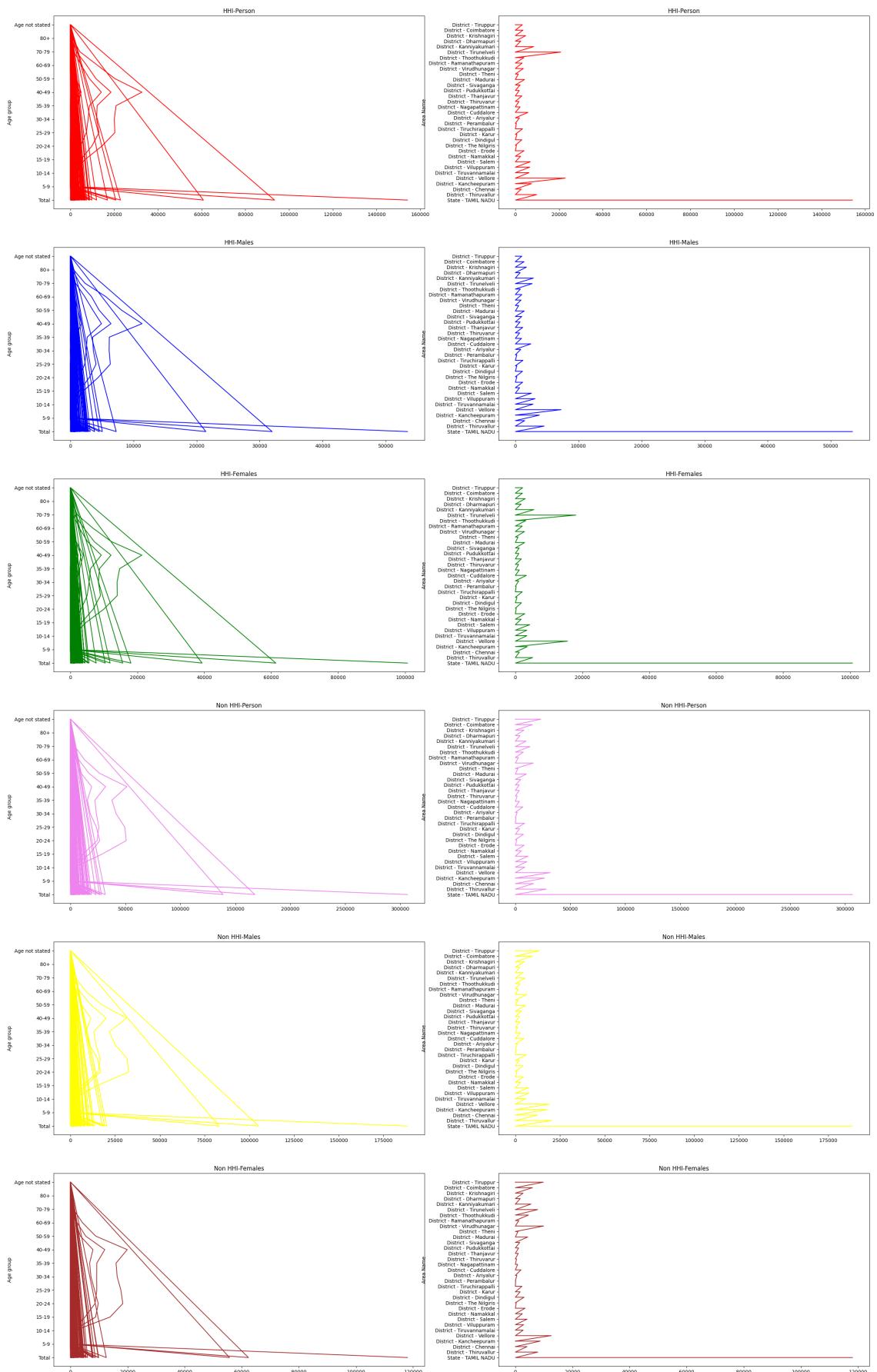


Sub plot

In [43]:

```
1 fig , ax =plt.subplots(figsize=(30,50),nrows=6,ncols=2)
2 ax[0, 0].plot(dataset['Industrial Category - C - HHI - Persons'], data
3 ax[0, 1].plot(dataset['Industrial Category - C - HHI - Persons'], data
4
5 ax[0,0].set_ylabel("Age group")
6 ax[0,1].set_ylabel("Area Name")
7 ax[0,0].set_title("HHI-Person")
8 ax[0,1].set_title("HHI-Person")
9
10 ax[1, 0].plot(dataset['Industrial Category - C - HHI - Males'], data
11 ax[1, 1].plot(dataset['Industrial Category - C - HHI - Males'], data
12 ax[1,0].set_ylabel("Age group")
13 ax[1,1].set_ylabel("Area Name")
14 ax[1,0].set_title("HHI-Males")
15 ax[1,1].set_title("HHI-Males")
16
17
18 ax[2, 0].plot(dataset['Industrial Category - C - HHI - Females'], data
19 ax[2, 1].plot(dataset['Industrial Category - C - HHI - Females'], data
20 ax[2,0].set_ylabel("Age group")
21 ax[2,1].set_ylabel("Area Name")
22 ax[2,0].set_title("HHI-Females")
23 ax[2,1].set_title("HHI-Females")
24
25 ax[3, 0].plot(dataset['Industrial Category - C - Non HHI - Persons'], data
26 ax[3, 1].plot(dataset['Industrial Category - C - Non HHI - Persons'], data
27 ax[3,0].set_ylabel("Age group")
28 ax[3,1].set_ylabel("Area Name")
29 ax[3,0].set_title("Non HHI-Person")
30 ax[3,1].set_title("Non HHI-Person")
31
32 ax[4, 0].plot(dataset['Industrial Category - C - Non HHI - Males'], da
33 ax[4, 1].plot(dataset['Industrial Category - C - Non HHI - Males'], da
34 ax[4,0].set_ylabel("Age group")
35 ax[4,1].set_ylabel("Area Name")
36 ax[4,0].set_title("Non HHI-Males")
37 ax[4,1].set_title("Non HHI-Males")
38
39 ax[5, 0].plot(dataset['Industrial Category - C - Non HHI - Females'], da
40 ax[5, 1].plot(dataset['Industrial Category - C - Non HHI - Females'], da
41 ax[5,0].set_ylabel("Age group")
42 ax[5,1].set_ylabel("Area Name")
43 ax[5,0].set_title("Non HHI-Females")
44 ax[5,1].set_title("Non HHI-Females")
```

Out[43]: Text(0.5, 1.0, 'Non HHI-Females')



Calculate the distribution of marginal workers based on age, industrial category, and sex

In [44]: 1 dataset.columns

```
Out[44]: Index(['Table Code', 'State Code', 'District Code', 'Area Name',
       'Total/ Rural/ Urban', 'Age group',
       'Worked for 3 months or more but less than 6 months - Persons',
       'Worked for 3 months or more but less than 6 months - Males',
       'Worked for 3 months or more but less than 6 months - Females',
       'Worked for less than 3 months - Persons',
       'Worked for less than 3 months - Males',
       'Worked for less than 3 months - Females',
       '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'],
dtype='object')
```

By Persons

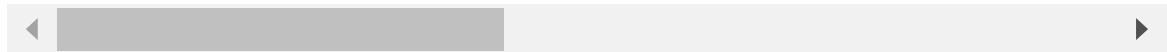
In [45]:

```
1 marginal_worker_criteria_person = (dataset['Worked for 3 months or more'] & dataset['Worked for less than 3 months'])
2
3
4 # Filter the dataset to only include marginal workers
5 marginal_workers_person = dataset[marginal_worker_criteria_person]
6
7 # Group the data by age group, industrial category, and sex
8 grouped = marginal_workers_person.groupby(['Age group', 'Total/ Rural/ Urban', 'Industrial Category - A - B - C - D - E - F - G - H - I - J - K - L - M - N to O - P to R to S - T to Z'])
9
10
11
12
13
14
15
16
17
18
19
20 ])
21
22 # Calculate the count of marginal workers in each group
23 marginal_worker_counts_person = grouped.size().reset_index(name='Count')
24
25 # Calculate the total count of marginal workers for all groups
26 total_marginal_workers_person = marginal_worker_counts_person['Count']
27
28 # Calculate the distribution of marginal workers
29 marginal_worker_counts_person['Distribution'] = (marginal_worker_counts_person['Count'] / total_marginal_workers_person)
30
31
32 # Print or export the results
33 marginal_worker_counts_person
```

Out[45]:

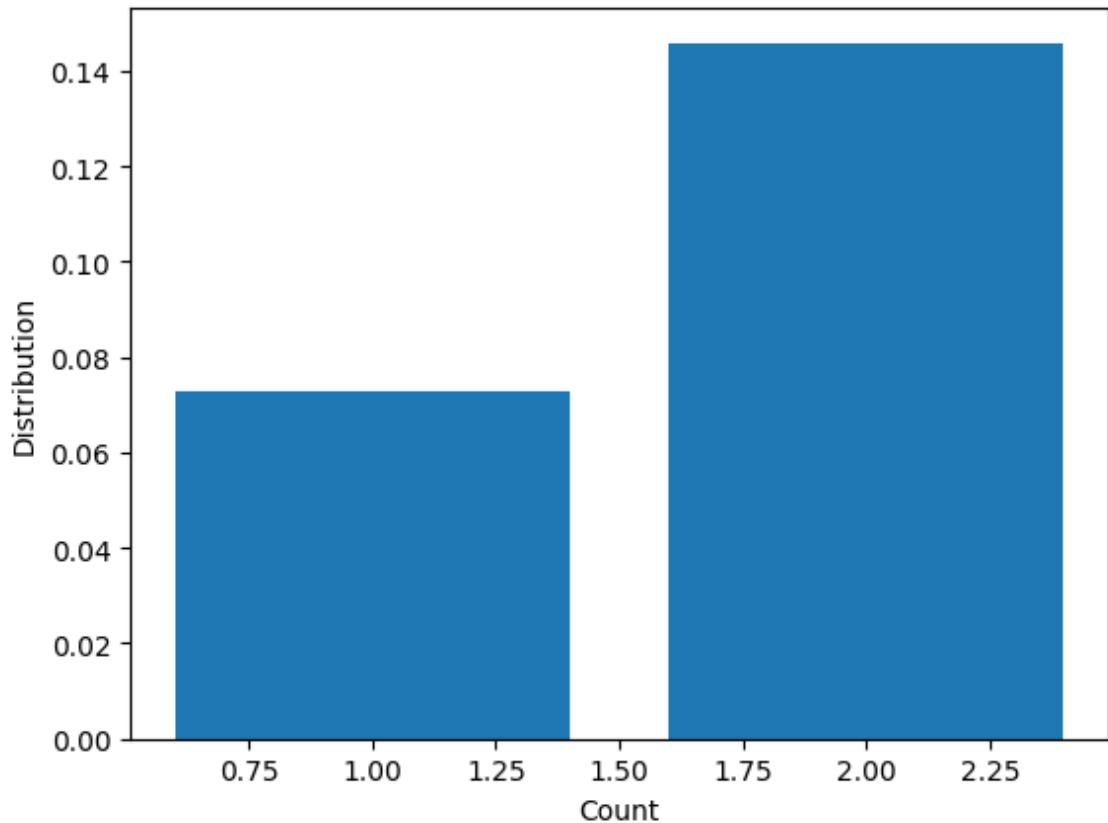
						Industrial Category - A				
	Age group	Total/ Rural/ Urban	Industrial Category - A - Cultivators - Persons	Industrial Category - A - Agricultural labourers - Persons	Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Persons	Industrial Category - B - Persons	Industrial Category - C - HHI - Persons	Industrial Category - C - Non HHI - Persons	Industrial Category - C - Non HHI - Persons	In C - F
0	10-14	Rural	11	147	8	0	0	0	14	
1	10-14	Rural	19	13	2	0	0	6	8	
2	10-14	Rural	31	31	8	0	0	8	6	
3	10-14	Rural	45	110	6	0	0	6	18	
4	10-14	Rural	51	147	10	0	0	12	42	
...
1365	Total	Urban	2155	9687	1923	322	3143	14292		
1366	Total	Urban	2175	16398	6655	290	6732	8314		
1367	Total	Urban	2253	6465	2193	94	4837	16719		
1368	Total	Urban	4244	3423	1129	71	2675	16593		
1369	Total	Urban	33799	217288	43198	4713	60789	167693		

1370 rows × 23 columns



```
In [46]: 1 plt.bar(marginal_worker_counts_person['Count'],marginal_worker_counts_
2 plt.xlabel('Count')
3 plt.ylabel("Distribution")
```

```
Out[46]: Text(0, 0.5, 'Distribution')
```



By Males

In [47]:

```
1 marginal_worker_criteria_Males = (dataset['Worked for 3 months or more'
2 dataset['Worked for less than 3 mont'
3
4 # Filter the dataset to only include marginal workers
5 marginal_workers_Males = dataset[marginal_worker_criteria_Males]
6
7 # Group the data by age group, industrial category, and sex
8 grouped = marginal_workers_Males.groupby(['Age group', 'Total/ Rural/
9
10 # Calculate the count of marginal workers in each group
11 marginal_worker_counts_Males = grouped.size().reset_index(name='Count')
12
13 # Calculate the total count of marginal workers for all groups
14 total_marginal_workers_Males = marginal_worker_counts_Males['Count'].sum()
15
16 # Calculate the distribution of marginal workers
17 marginal_worker_counts_Males['Distribution'] = (marginal_worker_counts_
18
19 # Print or export the results
20 marginal_worker_counts_Males
```

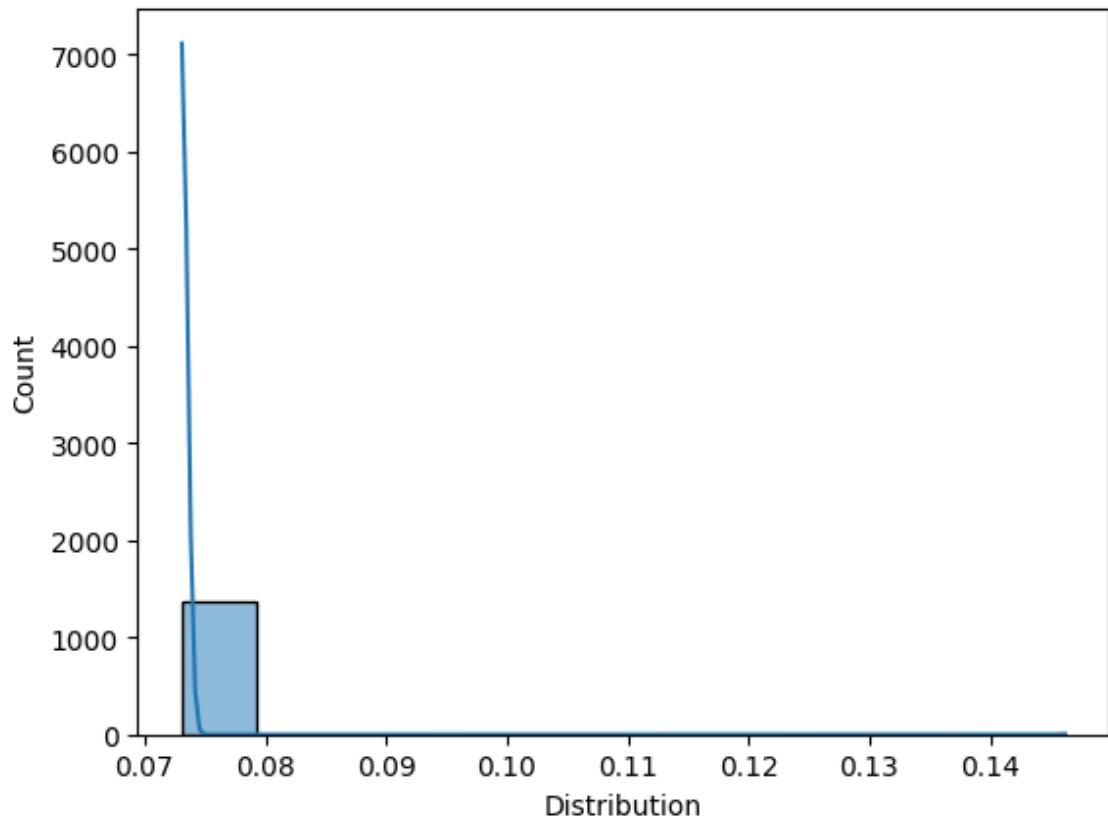
Out[47]:

Age group	Total/ Rural/ Urban	Industrial Category - A - Cultivators - Males	Industrial Category - A - Agricultural labourers - Males	Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males	Industrial Category - A -				Industrial Category - C - Non HHI - Males	In C -
					Industrial Category - B - Males	Industrial Category - C - HHI - Males	Industrial Category - C - HHI - Males	Industrial Category - C - HHI - Males		
0	10-14	Rural	7	4	2	0	0	0	6	
1	10-14	Rural	8	76	2	0	0	0	2	
2	10-14	Rural	9	18	2	0	0	6	6	
3	10-14	Rural	22	49	0	0	0	0	14	
4	10-14	Rural	25	79	0	0	0	6	0	
...
1362	Total	Urban	1348	5288	1492	207	1803	9763		
1363	Total	Urban	1394	3529	1605	84	2403	12527		
1364	Total	Urban	1723	13310	5953	258	2398	3551		
1365	Total	Urban	2188	1830	1010	55	1477	12498		
1366	Total	Urban	20874	112532	29883	3286	21456	105343		

1367 rows × 23 columns

```
In [48]: 1 sns.histplot(marginal_worker_counts_Males['Distribution'],kde= True)
```

```
Out[48]: <AxesSubplot: xlabel='Distribution', ylabel='Count'>
```



By Females

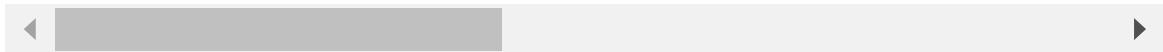
In [49]:

```
1 marginal_worker_criteria_Females = (dataset['Worked for 3 months or mo
2 dataset['Worked for less than 3 mont
3
4 # Filter the dataset to only include marginal workers
5 marginal_workers_Females = dataset[marginal_worker_criteria_Females]
6
7 # Group the data by age group, industrial category, and sex
8 grouped = marginal_workers_Females.groupby(['Age group', 'Total/ Rural
9             'Industrial Category - A -
10            'Industrial Category - A -
11            'Industrial Category - B -
12            'Industrial Category - C -
13            'Industrial Category - G -
14            'Industrial Category - G -
15            'Industrial Category - I -
16            'Industrial Category - J -
17            'Industrial Category - N t
18            'Industrial Category - R t
19
20 # Calculate the count of marginal workers in each group
21 marginal_worker_counts_Females = grouped.size().reset_index(name='Coun
22
23 # Calculate the total count of marginal workers for all groups
24 total_marginal_workers_Females = marginal_worker_counts_Females['Count
25
26 # Calculate the distribution of marginal workers
27 marginal_worker_counts_Females['Distribution'] = (marginal_worker_coun
28
29 # Print or export the results
30 marginal_worker_counts_Females
```

Out[49]:

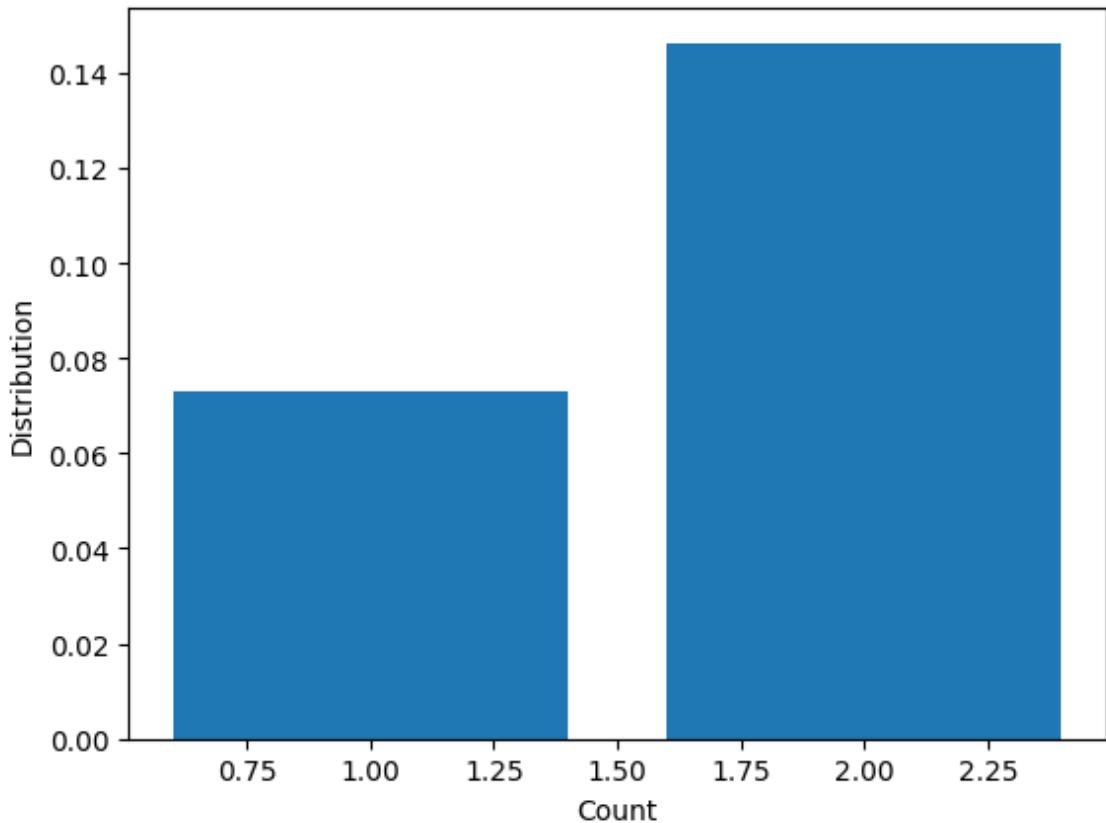
						Industrial Category - A				
	Age group	Total/ Rural/ Urban	Industrial Category - A - Cultivators - Females	Industrial Category - A - Agricultural labourers - Females	Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Females	Industrial Category - B - Females	Industrial Category - C - HHI - Females	Industrial Category - C - Non HHI - Females	Industrial Category - C - Non HHI - Females	In C - F
0	10-14	Rural	3	71	6	0	0	0	12	
1	10-14	Rural	12	9	0	0	0	6	2	
2	10-14	Rural	22	13	6	0	0	2	0	
3	10-14	Rural	23	61	6	0	0	6	4	
4	10-14	Rural	26	68	10	0	0	6	42	
...	
1361	Total	Urban	807	4399	431	115	1340	4529		
1362	Total	Urban	812	4888	1478	29	1213	4097		
1363	Total	Urban	859	2936	588	10	2434	4192		
1364	Total	Urban	2056	1593	119	16	1198	4095		
1365	Total	Urban	12925	104756	13315	1427	39333	62350		

1366 rows × 23 columns



```
In [50]: 1 plt.bar(marginal_worker_counts_Females['Count'],marginal_worker_counts_Females['Distribution'])
2 plt.xlabel("Count")
3 plt.ylabel("Distribution")
```

```
Out[50]: Text(0, 0.5, 'Distribution')
```



Age distribution

```
In [51]: 1 dataset['Age group'].value_counts()
```

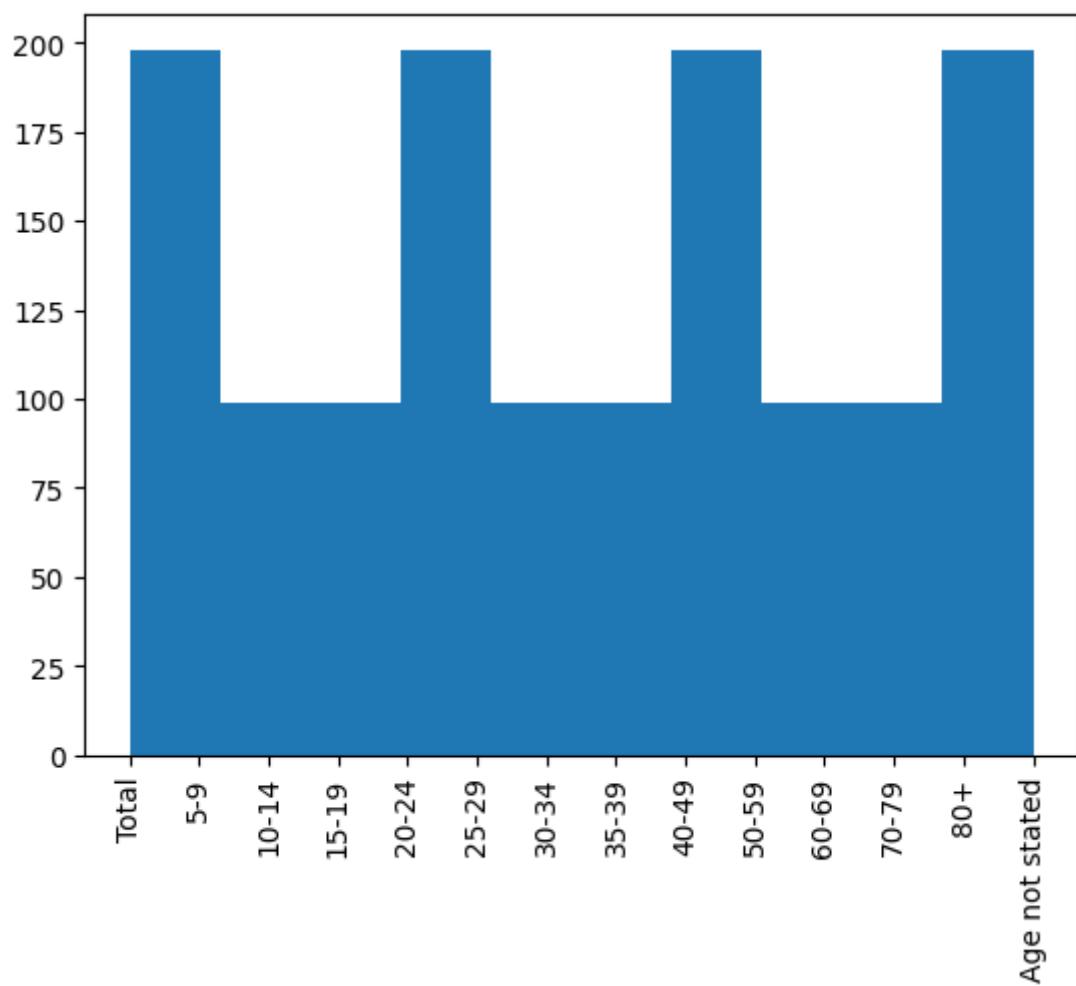
```
Out[51]: Total          99
5-9            99
10-14          99
15-19          99
20-24          99
25-29          99
30-34          99
35-39          99
40-49          99
50-59          99
60-69          99
70-79          99
80+            99
Age not stated  99
Name: Age group, dtype: int64
```

```
In [52]: 1 dataset['Age group'].mode()
```

```
Out[52]: 0          10-14
1          15-19
2          20-24
3          25-29
4          30-34
5          35-39
6          40-49
7          5-9
8          50-59
9          60-69
10         70-79
11         80+
12    Age not stated
13         Total
Name: Age group, dtype: object
```

```
In [53]: 1 plt.hist(dataset[ 'Age group' ])
2 plt.xticks(rotation=90)
```

```
Out[53]: ([0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13],
[Text(0, 0, 'Total'),
Text(1, 0, '5-9'),
Text(2, 0, '10-14'),
Text(3, 0, '15-19'),
Text(4, 0, '20-24'),
Text(5, 0, '25-29'),
Text(6, 0, '30-34'),
Text(7, 0, '35-39'),
Text(8, 0, '40-49'),
Text(9, 0, '50-59'),
Text(10, 0, '60-69'),
Text(11, 0, '70-79'),
Text(12, 0, '80+'),
Text(13, 0, 'Age not stated')])
```



From the analysis:

The analysis of demographic characteristics of marginal workers in Tamil Nadu provides valuable insights into their socioeconomic status and living conditions. Here's how the analysis provides these insights:

1. Age Distribution:

- By analyzing the age distribution of marginal workers through techniques like histograms or bar charts, the analysis can reveal the dominant age groups among this population. Insights gained from this include whether younger or older individuals are more likely to be engaged in temporary or seasonal employment.

- For example, if the analysis shows that a significant portion of marginal workers falls within the age group of 20-40 years, it suggests that this demographic group is more susceptible to seasonal employment. This insight can guide policies aimed at supporting these age groups during non-employment periods.

2. Industry Category:

- Categorizing and visualizing marginal workers by industry category, such as agriculture, manufacturing, construction, and services, offers insights into the diversity of their occupational choices.

- It can help identify which industries employ a substantial number of marginal workers, providing a basis for targeted interventions and policies to improve working conditions in these sectors.

3. Gender Disparities:

- Analyzing the gender distribution among marginal workers helps uncover potential gender disparities within this workforce. It highlights whether men and women are equally represented or if there is a significant gender imbalance.

- Gender disparities could indicate unequal access to employment opportunities or variations in working conditions, wages, and benefits. The analysis can be used to address gender-specific challenges faced by marginal workers.

4. Geographic Distribution:

- If geospatial data is available and utilized in the analysis, it can reveal the geographic distribution of marginal workers across different regions or districts within Tamil Nadu.

- This insight helps policymakers target specific areas that have a higher concentration of marginal workers, potentially indicating areas with greater economic challenges.

5. Correlations and Trends:

- Through the analysis, correlations and trends can be identified. For instance, the analysis may reveal that younger marginal workers are more likely to be employed in the agricultural sector while older workers are more common in construction.
- Such findings can inform strategies to provide skill development or training programs tailored to the needs of specific age groups and industries.

6. Data-Driven Decision-Making:

- By providing visual representations and statistical insights, the analysis empowers policymakers, organizations, and researchers to make informed decisions and design interventions to improve the socioeconomic conditions of marginal workers.
- It ensures that resources are allocated effectively and efficiently, addressing the specific needs of different demographic groups within the marginal worker population.

In summary, the analysis of demographic characteristics of marginal workers in Tamil Nadu offers a comprehensive understanding of their age distribution, industry preferences, gender disparities, geographical distribution, and potential correlations. These insights are essential for crafting targeted policies, programs, and initiatives aimed at improving the socioeconomic conditions and livelihoods of this vulnerable workforce.

Conclusion:

In conclusion, the analysis of the demographic characteristics of marginal workers in Tamil Nadu has provided valuable insights into this vulnerable population, shedding light on their socioeconomic status and living conditions. By examining their age distribution, industry category preferences, and gender disparities, this analysis has generated actionable findings and implications for policymakers, organizations, and researchers.

The key takeaways from the analysis include:

- 1. Age Insights:** The dominant age groups among marginal workers have been identified, offering guidance for designing age-specific interventions to support individuals in their pursuit of seasonal or temporary employment.
- 2. Industry Diversity:** The categorization and visualization of marginal workers by industry category have revealed the diversity of their occupational choices. This information can be used to target sectors with a high prevalence of marginal workers for improving working conditions and opportunities.
- 3. Gender Disparities:** By analyzing gender distribution, potential gender disparities within the marginal worker population have been exposed. Addressing these disparities can help ensure that both men and women have equal access to employment opportunities and equitable working conditions.

4. Geographic Concentration: The geographical distribution of marginal workers across different regions has been highlighted, offering insights into areas that may require specific attention and support to uplift the economic conditions of marginal workers.

5. Correlations and Trends: The analysis has unveiled correlations and trends within the data, enabling the development of tailored programs and initiatives to meet the unique needs of different demographic groups and industries.

This comprehensive analysis, conducted through data visualization and statistical techniques, equips stakeholders with the knowledge needed to make informed, evidence-based decisions. It serves as a foundation for the formulation of policies and interventions aimed at enhancing the well-being and livelihoods of marginal workers in Tamil Nadu. As we move forward, the insights gained from this analysis can be used to guide and drive positive change, ensuring that marginal workers are better supported and integrated into the economic and social fabric of the region.