TN MARGINAL WORKERS ASSESSMENT

PHASE 4: DEVELOPMENT PART-2

INTRODUCTION:

- The Demographic analysis and to create visualizations for the given dataset of marginal workers is the process of this phase.
- In this we need to determine the workers age group,industrial category, and gender distribution.
- To represent the required dataset information in visualization format by using the data visualization libraries such as matplotlib,seaborn.

DEMOGRAPHIC ANALYSIS:

To perform the demographic ananlysis using the given dataset are as the following steps,they are:

- Load the data
- Explore the dataset
- Filter required or relevant variables
- Create visualizations
- Condust comparative analysis

Load the data:

- Load the dataset of the given dataset by using the python libraries such as pandas.

Explore the data:

- Explore the content and structure of given dataset and check the missing values in the given dataset.

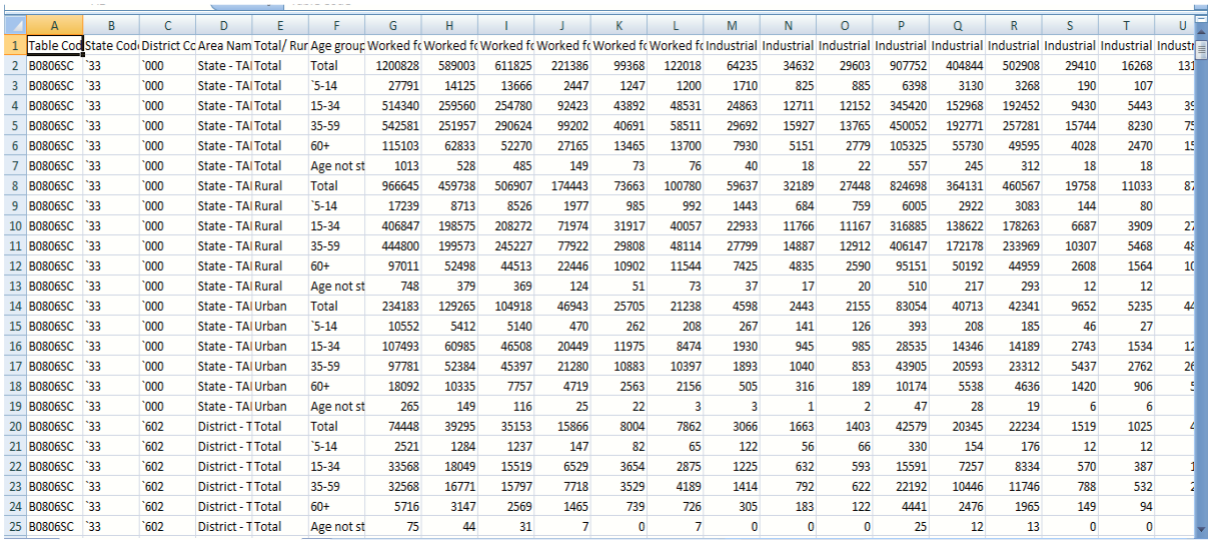
Filter relevant variables:

- Identify the required demographic variables such as age,gender and industrial categories based on the dataset .

Create visualizations:

- Represent the given dataset in the format of histograms,bar charts,pie charts and so on.
- By the demographic analysis we identify the age distribution,industrial categories and gender distribution of marginal workers .

DATASET LINK: ***Dataset Link:*** ***https://tn.data.gov.in/resource/marginal-workers-classified-age-industrial-category-and-sex-scheduled-caste-2011-tamil***



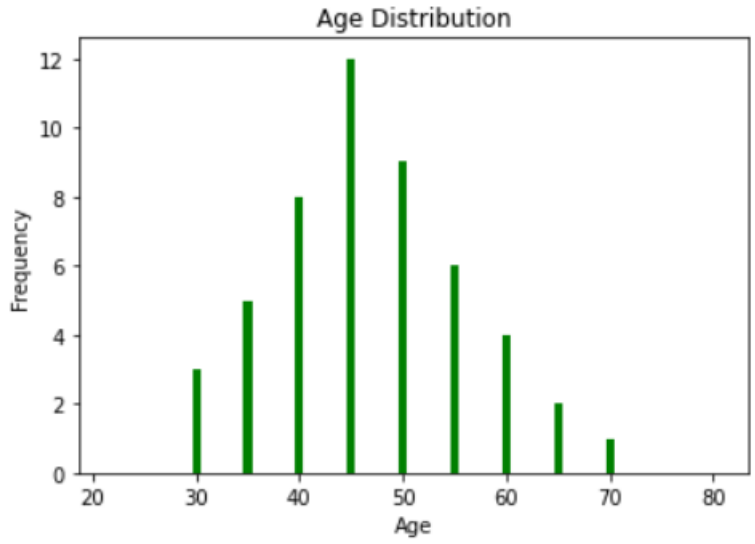
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Table Cod	State Cod	District Cc	Area Nam	Total/ Rur	Age group	Worked fc	Worked fc	Worked fc	Worked fc	Worked fc	Worked fc	Industrial	Industrial	Industrial	Industrial	Industrial	Industrial	Industrial	Industrial	Industr
2	B0806SC	`33	`000	State - TA	Total	Total	1200828	589003	611825	221386	99368	122018	64235	34632	29603	907752	404844	502908	29410	16268	131
3	B0806SC	`33	`000	State - TA	Total	`5-14	27791	14125	13666	2447	1247	1200	1710	825	885	6398	3130	3268	190	107	
4	B0806SC	`33	`000	State - TA	Total	15-34	514340	259560	254780	92423	43892	48531	24863	12711	12152	345420	152968	192452	9430	5443	39
5	B0806SC	`33	`000	State - TA	Total	35-59	542581	251957	290624	99202	40691	58511	29692	15927	13765	450052	192771	257281	15744	8230	79
6	B0806SC	`33	`000	State - TA	Total	60+	115103	62833	52270	27165	13465	13700	7930	5151	2779	105325	55730	49595	4028	2470	15
7	B0806SC	`33	`000	State - TA	Total	Age not st	1013	528	485	149	73	76	40	18	22	557	245	312	18	18	
8	B0806SC	`33	`000	State - TA	Rural	Total	966645	459738	506907	174443	73663	100780	59637	32189	27448	824698	364131	460567	19758	11033	87
9	B0806SC	`33	`000	State - TA	Rural	`5-14	17239	8713	8526	1977	985	992	1443	684	759	6005	2922	3083	144	80	
10	B0806SC	`33	`000	State - TA	Rural	15-34	406847	198575	208272	71974	31917	40057	22933	11766	11167	316885	138622	178263	6687	3909	27
11	B0806SC	`33	`000	State - TA	Rural	35-59	444800	199573	245227	77922	29808	48114	27799	14887	12912	406147	172178	233969	10307	5468	48
12	B0806SC	`33	`000	State - TA	Rural	60+	97011	52498	44513	22446	10902	11544	7425	4835	2590	95151	50192	44959	2608	1564	10
13	B0806SC	`33	`000	State - TA	Rural	Age not st	748	379	369	124	51	73	37	17	20	510	217	293	12	12	
14	B0806SC	`33	`000	State - TA	Urban	Total	234183	129265	104918	46943	25705	21238	4598	2443	2155	83054	40713	42341	9652	5235	44
15	B0806SC	`33	`000	State - TA	Urban	`5-14	10552	5412	5140	470	262	208	267	141	126	393	208	185	46	27	
16	B0806SC	`33	`000	State - TA	Urban	15-34	107493	60985	46508	20449	11975	8474	1930	945	985	28535	14346	14189	2743	1534	12
17	B0806SC	`33	`000	State - TA	Urban	35-59	97781	52384	45397	21280	10883	10397	1893	1040	853	43905	20593	23312	5437	2762	26
18	B0806SC	`33	`000	State - TA	Urban	60+	18092	10335	7757	4719	2563	2156	505	316	189	10174	5538	4636	1420	906	5
19	B0806SC	`33	`000	State - TA	Urban	Age not st	265	149	116	25	22	3	3	1	2	47	28	19	6	6	
20	B0806SC	`33	`602	District - T	Total	Total	74448	39295	35153	15866	8004	7862	3066	1663	1403	42579	20345	22234	1519	1025	4
21	B0806SC	`33	`602	District - T	Total	`5-14	2521	1284	1237	147	82	65	122	56	66	330	154	176	12	12	
22	B0806SC	`33	`602	District - T	Total	15-34	33568	18049	15519	6529	3654	2875	1225	632	593	15591	7257	8334	570	387	1
23	B0806SC	`33	`602	District - T	Total	35-59	32568	16771	15797	7718	3529	4189	1414	792	622	22192	10446	11746	788	532	2
24	B0806SC	`33	`602	District - T	Total	60+	5716	3147	2569	1465	739	726	305	183	122	4441	2476	1965	149	94	
25	B0806SC	`33	`602	District - T	Total	Age not st	75	44	31	7	0	7	0	0	0	25	12	13	0	0	

DISTRIBUTION OF MARGINAL WORKERS:

AGE DISTRIBUTION ANALYSIS:

Representation of age distribution in data visualization are given below:

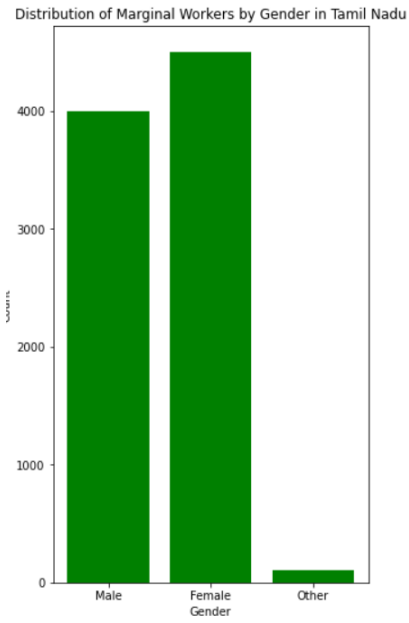
```
import matplotlib.pyplot as plt
# ages
ages = [22, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80]
# Frequency of each age group
age_groups = [0, 0, 3, 5, 8, 12, 9, 6, 4, 2, 1, 0, 0]
# Plotting the age distribution
plt.bar(ages, age_groups, color='green')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.title('Age Distribution')
plt.show()
```



GENDER DISTRIBUTION ANALYSIS:

Representation of data visualization for marginal workers code is below:

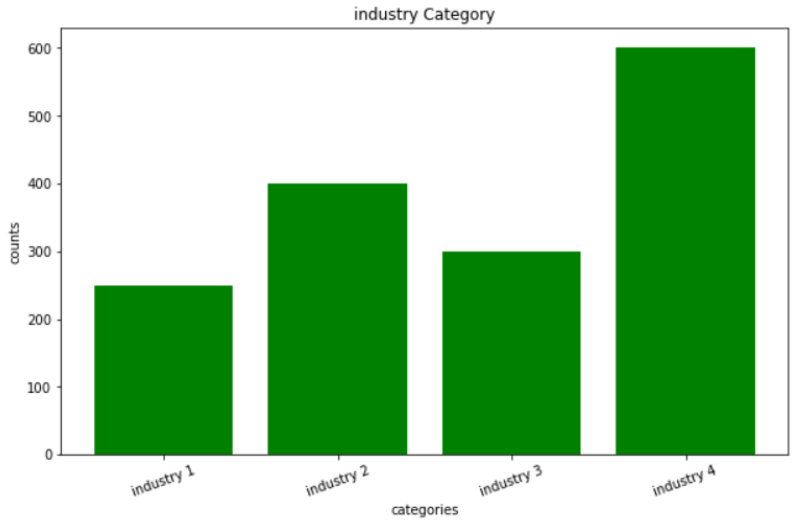
```
import matplotlib.pyplot as plt
#the given data
gender = ['Male', 'Female', 'Other']
count = [4000, 4500, 100]
plt.figure(figsize=(5, 9))
#create bar chart
plt.bar(gender, count, color='green')
plt.xlabel('Gender')
plt.ylabel('Count')
plt.title('Distribution of Marginal Workers by Gender in Tamil Nadu')
```



INDUSTRIAL CATEGORY ANALYSIS:

Representation of industrial categories of the given dataset are below as :

```
import matplotlib.pyplot as plt
#the given data
industry_categories = ['industry 1', 'industry 2', 'industry 3', 'industry 4']
counts = [250, 400, 300, 600]
# Create a bar chart
plt.figure(figsize=(10, 6))
plt.bar(categories, counts, color='green')
plt.title('industry Category')
plt.xlabel('categories')
plt.ylabel('counts')
plt.xticks(rotation=20)
plt.show()
plt.show()
```



VISUALIZATION OF DATASET:

In order to create the visualization from the given dataset ,below are the following steps :

- Use the necessary libraries
- Import the necessary libraries
- Prepare the dataset
- Choose the pyplot
- Create the visualization
- Customize the visualization
- Display the visualization

Use the necessary libraries:

- For the given dataset use the popular visualization libraries such as matplotlib,seaborn and plotly.

Import the necessary library:

- Starting by the library as importing libraries

```
#import matplotlib.pyplot as plt
```

Prepare your dataset:

- Make sure that the data is ready for the visualization.organize it as a format that can be easily processed the chosen visualization library.

Choose a plot type:

- Depending on your dataset choose the plot type like line plot,scatter plot,bar plot or histogram.

Create the visualization:

- Use the appropriate method provided by the library to create the plot and customize it as needed.

```
#plt.plot(x,y)#to create a line plot
```

Customize the visualization:

- Add labels,titles,legends,and other customizations to make your visualization more informative and visually appealing.

Display the visualizations:

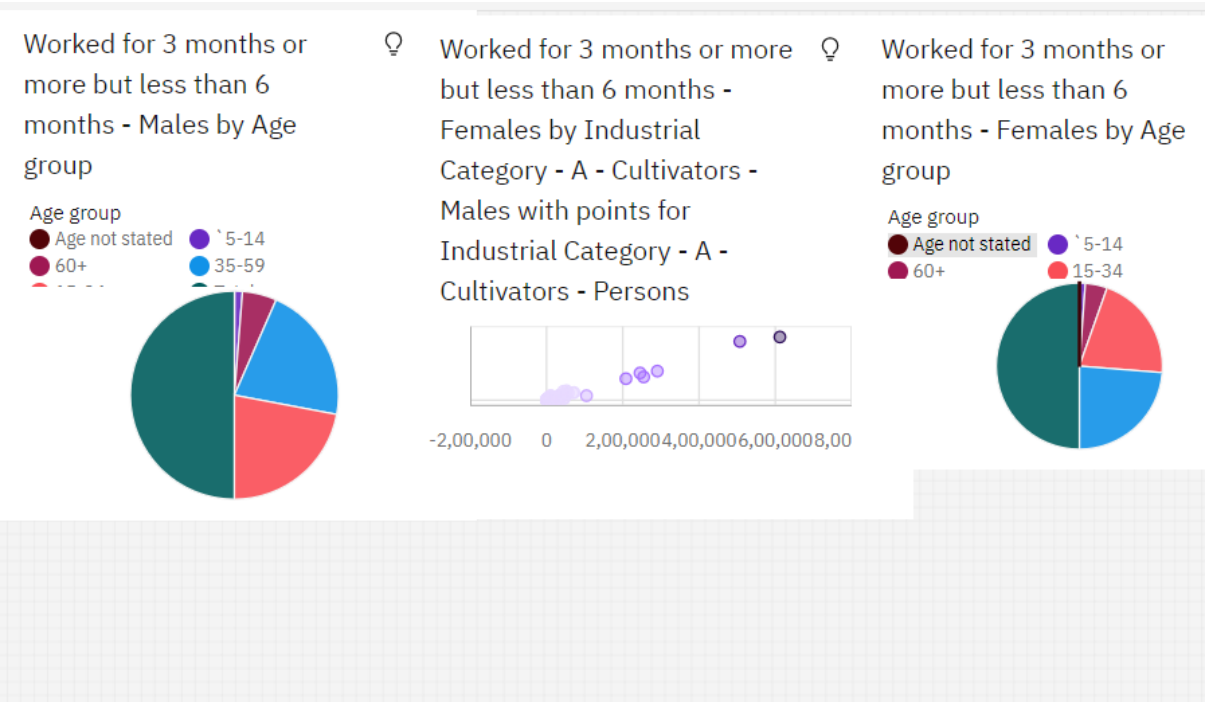
- Use the "plt.show" to display the created visualization.

PROGRAM FOR DATA VISUALIZATION:

The data visualization for the given dataset of age distribution,gender distribution and industrial categories code is given below as follow as:

```
import pandas as pd

import matplotlib.pyplot as plt

# Load the dataset

data = pd.read_csv(' https://tn.data.gov.in/resource/marginal-workers-classified-age-industrial-category-and-sex-scheduled-caste-2011-tamil'')

# Filter data for marginal workers

marginal_workers = data[data['age'] == 'age group]

# Calculate the distribution based on age, industrial category, and sex

age_distribution = marginal_workers['age'].value_counts()

category_distribution = marginal_workers['industrial_category'].value_counts()

sex_distribution = marginal_workers['sex'].value_counts()

# Create visualizations

plt.figure(figsize=(10, 6))

plt.subplot(1, 3, 1)

age_distribution.plot(kind='pie', title='Age Distribution of Marginal Workers')
```

```
plt.xlabel('Age')
plt.ylabel('Count')
plt.subplot(1, 3, 2)
category_distribution.plot(kind='plot', title='Industrial Category Distribution of Marginal Workers')
plt.xlabel('Industrial Category')
plt.ylabel('Count')
plt.subplot(1, 3, 3)
sex_distribution.plot(kind='pie', title='Sex Distribution of Marginal Workers')
plt.xlabel('Sex')
plt.ylabel('Count')
plt.tight_layout()
plt.show()
```



Conclusion:

In this document the process of demographic analysis and creating the data visualizations using various libraries such as matplotlib,seaborn are done by using the given dataset of TN Marginal TamilNadu dataset.we observe the marginal workers multiple age groups,examination of industrial categories and gender distribution,these are all express the vital process of marginal workers.These are the concepts explained and visualized in this phase of TN Marginal Workers TamilNadu in India.