TN MARGINAL WORKERS ASSESSMENT

(A Socioeconomic Analysis)

INTRODUCTION:

In this phase we are going to calculate the distribution of marginal workers based on age, industrial category, and sex using data aggregation and manipulation and create visualization using python libraries in this dataset.

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

Finding the frequency of people based on age and their district for particular industrial Category using aggregate function:

```
import pandas as pd

impor matplotlib.pyplot as plt

data=pd.read_csv("marginal-workers.csv")

d=data.groupby(["Area Name" ,"Age group "])["Industrial Category -A-Cultivators-Persons"].mean().unstack()

print(d)
```

```
In [2]: runfile('C:/Users/thana/Desktop/k/p2.py', wdir='C:/Users/thana/Desktop/k')
Age group                          15-34  ...        `5-14
Area Name                                 ...
District - Ariyalur           400.666667  ...    21.333333
District - Chennai            252.666667  ...    60.666667
District - Coimbatore         176.000000  ...    31.333333
District - Cuddalore         1893.333333  ...   167.333333
District - Dharmapuri         150.666667  ...    10.666667
District - Dindigul           266.000000  ...    17.333333
District - Erode               92.666667  ...    14.666667
District - Kancheepuram      1099.333333  ...    74.000000
District - Kanniyakumari       19.333333  ...     1.333333
District - Karur               36.000000  ...     3.333333
District - Krishnagiri        398.666667  ...    29.333333
District - Madurai            394.000000  ...    21.333333
District - Nagapattinam       524.000000  ...    25.333333
District - Namakkal           165.333333  ...    20.666667
District - Perambalur         833.333333  ...    65.333333
District - Pudukkottai        622.666667  ...    18.666667
District - Ramanathapuram    1236.666667  ...    74.666667
District - Salem              196.000000  ...    11.333333
District - Sivaganga          741.333333  ...    28.000000
District - Thanjavur          307.333333  ...    14.666667
District - The Nilgiris        70.666667  ...     8.000000
District - Theni              112.000000  ...     4.000000
District - Thiruvallur        816.666667  ...    81.333333
District - Thiruvarur         527.333333  ...    35.333333
District - Thoothukkudi       196.000000  ...    14.666667
District - Tiruchirappalli    308.666667  ...    16.666667
District - Tirunelveli        402.000000  ...    24.000000
District - Tiruppur           152.000000  ...    19.333333
District - Tiruvannamalai    1224.000000  ...    58.000000
District - Vellore            534.666667  ...    37.333333
District - Viluppuram        2148.000000  ...   109.333333
District - Virudhunagar       277.333333  ...    20.666667
State - TAMIL NADU          16575.333333  ...  1140.000000

[33 rows x 6 columns]
```

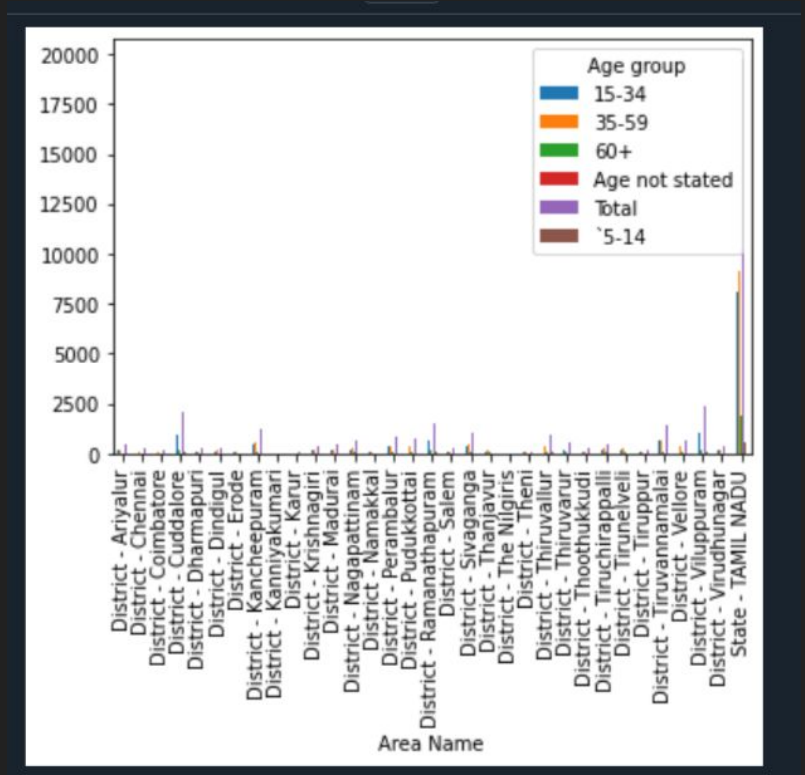
FINDING THE TOTAL NUMBER OF MARGINAL WORKERS BASED ON AGE AND DISTRICT:

```
import pandas as pd

impor matplotlib.pyplot as plt

data=pd.read_csv(“marginal-workers
.csv”)

d=data.groupby([“Age group”,”Area
Name”])

print(d.sum())
```

```
In [25]: runfile('C:/Users/thana/Desktop/k/p2.py', wdir='C:/Users/thana/Desktop/k')
                                                    Table Code  ... Industrial Category -
R to U - Non HHI - Females
Age group Area Name                                             ...
15-34     District - Ariyalur        B0806SCB0806SCB0806SC  ...
404
          District - Chennai         B0806SCB0806SCB0806SC  ...
11366
          District - Coimbatore      B0806SCB0806SCB0806SC  ...
3562
          District - Cuddalore       B0806SCB0806SCB0806SC  ...
4736
          District - Dharmapuri      B0806SCB0806SCB0806SC  ...
656
...                                                    ...  ...
...
`5-14     District - Tiruvannamalai  B0806SCB0806SCB0806SC  ...
636
          District - Vellore         B0806SCB0806SCB0806SC  ...
808
          District - Viluppuram      B0806SCB0806SCB0806SC  ...
1428
          District - Virudhunagar    B0806SCB0806SCB0806SC  ...
428
          State - TAMIL NADU         B0806SCB0806SCB0806SC  ...
19062

[198 rows x 67 columns]
```

Finding the frequency of people based on age and their district for particular industrial Category :

```
import pandas as pd

impor matplotlib.pyplot as plt

data=pd.read_csv(“marginal-workers.csv”)

d=data.groupby([“Area Name” ,”Age group “])[“Industrial Category -A-Cultivators-Persons”].mean().unstack()

d.plot(kind=’bar’)

Plt show()
```

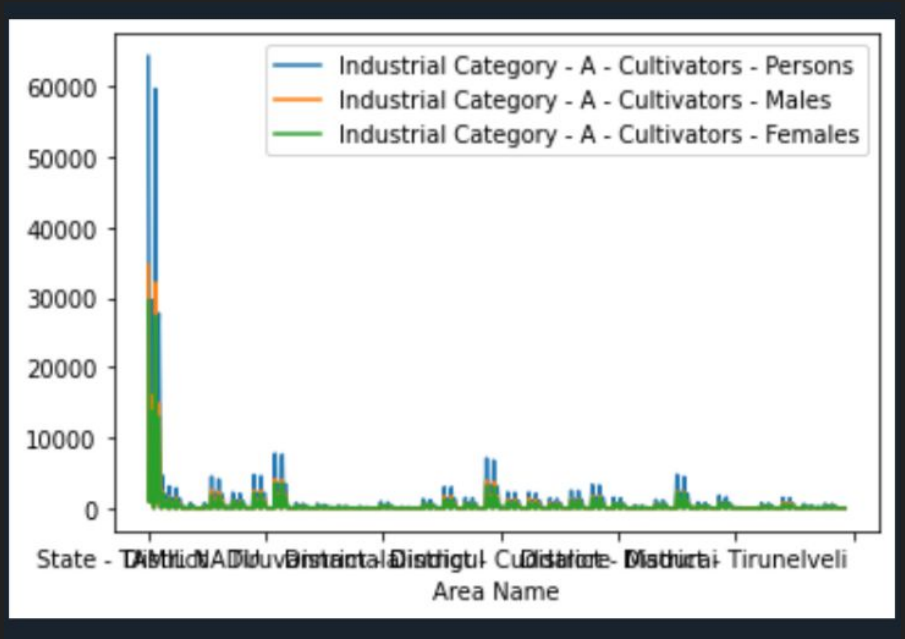


FINDING THE TOTAL NUMBER OF MARGINAL WORKERS BASED ON AGE AND IN A PARTICULAR DISTRICT:

```
import pandas as pd

impor matplotlib.pyplot as plt

data=pd.read_csv(“marginal-workers.cs
v”)

d=data.groupby([data[“Area
Name”]==”District-Ariyalur”,”Age
group”])

d.plot(kind=’bar’

Plt.show()
```

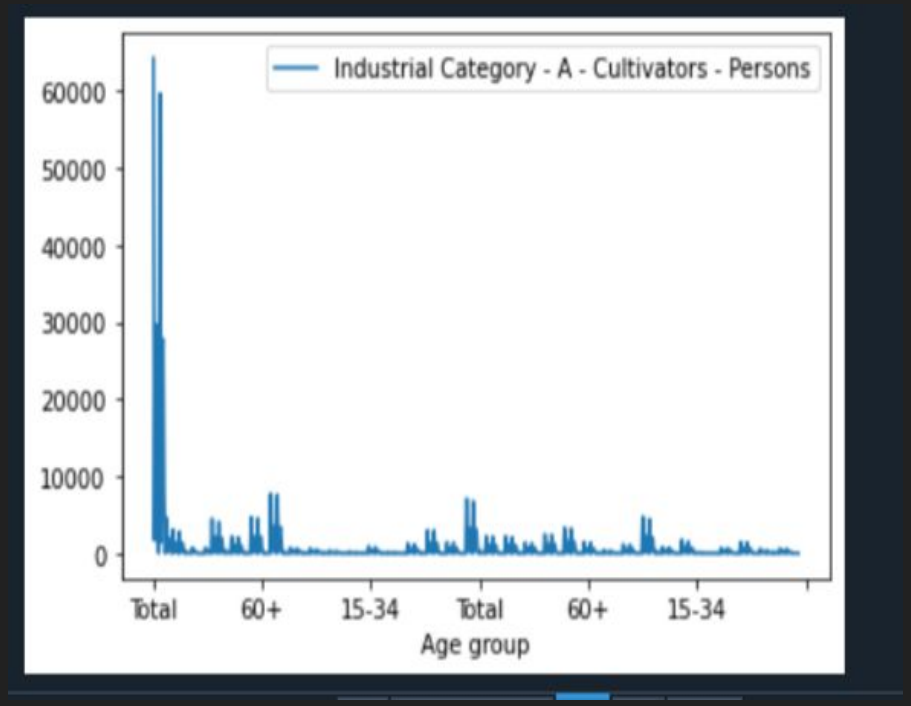
DISTRIBUTION OF MARGINAL WORKERS BASED ON SEX BY THEIR DISTRICT IN PARTICULAR INDUSTRIAL CATEGORY:

```
import pandas as pd

impor matplotlib.pyplot as plt

data=pd.read_csv(“marginal-workers.csv”)

data.plot(x=”Area Name”,y=[“Industrial Category -A-Cultivators-Persons”, Industrial Category -A-Cultivators-Males”,”Industrial Category -A-Cultivators-Females”])

Plt.show()
```



NUMBER OF MARGINAL WORKERS IN PARTICULAR CATEGORY BASED ON AGE:

```
import pandas as pd

impor matplotlib.pyplot as plt

data=pd.read_csv(“marginal-workers.csv”)

data.plot(x=”Age group”,y=”Industrial Category -A-Cultivators-Females”)

plt.show()
```



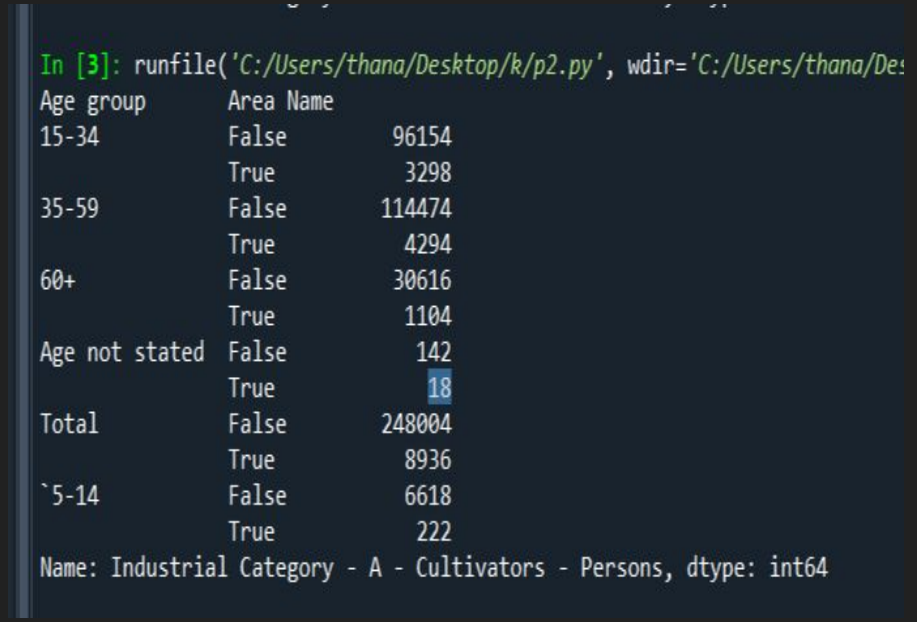
USING AGGREGATE FUNCTION IN INDUSTRIAL CATEGORY:

```
import pandas as pd

import matplotlib.pyplot as plt

data=pd.read_csv("marginal
-workers")

Print(data["Industrial
Category
-A-Cultivators-Persons"].
describe ())
```

```
In [2]: runfile('C:/Users/thana/Desktop/k/p2.py', wdir='C:/Users/thana/Deskt
count      594.000000
mean       865.117845
std       4274.458077
min          0.000000
25%          9.000000
50%         69.500000
75%        466.000000
max      64235.000000
Name: Industrial Category - A - Cultivators - Persons, dtype: float64
```

USING DESCRIBE FUNCTION FOR PARTICULAR INDUSTRIAL CATEGORY IN PARTICULAR DISTRICT:

```
import pandas as pd

import matplotlib.pyplot as plt

data=pd.read_csv("marginal -workers")

d=data.groupby(["Age
group"],data["Area Name"]=="District
-Kancheepuram"])["Industrial Category
-A-Cultivators-Persons"].sum()

print (d)
```



```
In [3]: runfile('C:/Users/thana/Desktop/k/p2.py', wdir='C:/Users/thana/Des
Age group       Area Name
15-34           False        96154
                True          3298
35-59           False       114474
                True          4294
60+             False        30616
                True          1104
Age not stated  False          142
                True            18
Total           False       248004
                True          8936
`5-14           False         6618
                True           222
Name: Industrial Category - A - Cultivators - Persons, dtype: int64
```

Thank you….