

## **Assessment of Marginal Workers in Tamil Nadu**

### **Objective:**

The objective of this assessment is to understand the employment and socio-economic conditions of marginal workers in the state of Tamil Nadu, India.

### **Introduction:**

The assessment of marginal workers in Tamil Nadu aims to provide insights into the employment patterns and socio-economic status of individuals who fall under the category of "marginal workers". This category typically includes individuals engaged in irregular and low-paying work.

### **Demographic Analysis:**

Demographic analysis involves studying and interpreting the characteristics of a population. In this context, you're focusing on the distribution of marginal workers based on age, industrial category, and sex.

### **Age Distribution:**

Age distribution refers to the spread of individuals across different age groups in a given population. Analyzing the age distribution of marginal workers can provide insights into which age groups are more likely to be considered as marginal workers.

### **Industrial Category:**

Industrial category analysis categorizes workers based on the sectors they are employed in. This information is crucial for understanding the specific industries where marginal workers are prevalent. Industries can range from agriculture to manufacturing, services, and more.

### **Sex Distribution:**

Analyzing the distribution of marginal workers based on sex provides insights into any gender disparities that may exist among marginal workers. This can be crucial for identifying potential areas of inequality and designing targeted interventions.

### **Data Aggregation and Manipulation:**

Data aggregation involves combining and summarizing data to create useful information. This is typically done using aggregation functions like sum, mean, count, etc.

### **Grouping Data:**

Grouping data involves splitting the dataset into groups based on one or more categorical variables (e.g., age, industry category, sex). This is done using functions like `'groupby'` in pandas.

### **Aggregating Data:**

Aggregating data involves applying a function to each group to obtain a single value that represents the group. For example, you might use the `'sum'` function to calculate the total number of marginal workers in each age group.

### **Data Visualization Libraries:**

Data visualization libraries are tools that enable the creation of visual representations of data. Two commonly used libraries are:

#### **Matplotlib:**

Matplotlib is a widely used data visualization library in Python. It provides a wide range of plots, including line plots, bar plots, scatter plots, histograms, etc. These visualizations help in presenting data in a clear and understandable manner.

#### **Seaborn:**

Seaborn is built on top of Matplotlib and provides a high-level interface for drawing attractive and informative statistical graphics. It simplifies the process of creating complex visualizations and provides a variety of aesthetically pleasing styles.

### **General process:**

1. **Data Preparation:** Ensure your dataset is structured with relevant columns (age, industry category, sex) and includes information about marginal workers.
2. **Data Loading:** Use a library like pandas to load the data into your Python environment.
3. **Data Aggregation and Manipulation:** Use functions like `'groupby'` to group and aggregate the data based on age, industrial category, and sex.
4. **Visualization:** Utilize Matplotlib and Seaborn to create visualizations based on the aggregated data. For example, bar charts, pie charts, etc., are effective for displaying demographic distributions.
5. **Interpretation:** Analyze the visualizations to draw meaningful insights about the distribution of marginal workers based on age, industrial category, and sex.

## Analysis:

```
print('The total number of person in Tamil Nadu who are working more than 3 Months and less than 6 months')
data = pd.read_csv('dataset1.csv')
persons=data['Persons'].sum()
print(persons)
print('The total number of Male Workers in Tamil Nadu who are working more than 3 Months and less than 6 months')
Male=data['Males'].sum()
print(Male)
print('The total number of Female Workers in Tamil Nadu who are working more than 3 Months and less than 6 months')
Female=data['Females'].sum()
print(Female)
```

```
➡ The total number of person in Tamil Nadu who are working more than 3 Months and less than 6 months
9606631.0
The total number of Male Workers in Tamil Nadu who are working more than 3 Months and less than 6 months
4712032.0
The total number of Female Workers in Tamil Nadu who are working more than 3 Months and less than 6 months
4894609.0
```

## Workers working more than 3 months and less than 6 months:

➡ Total number of male workers in each district of Tamil Nadu who are working more than 3 months and less than 6 months

state	male
District - Ariyalur	9926.0
District - Chennai	19313.0
District - Coimbatore	13530.0
District - Cuddalore	52389.0
District - Dharmapuri	7357.0
District - Dindigul	14174.0
District - Erode	7765.0
District - Kancheepuram	47269.0
District - Kanniyakumari	2886.0
District - Karur	3308.0
District - Krishnagiri	10275.0
District - Madurai	15115.0
District - Nagapattinam	36360.0
District - Namakkal	6808.0
District - Perambalur	7835.0
District - Pudukkottai	9903.0
District - Ramanathapuram	9824.0
District - Salem	11407.0
District - Sivaganga	13618.0
District - Thanjavur	16607.0
District - The Nilgiris	5400.0
District - Theni	7357.0
District - Thiruvallur	39295.0
District - Thiruvarur	28159.0
District - Thoothukkudi	9746.0
District - Tiruchirappalli	12793.0
District - Tirunelveli	17240.0
District - Tiruppur	11040.0
District - Tiruvannamalai	31028.0
District - Vellore	36863.0
District - Viluppuram	65063.0
District - Virudhunagar	9350.0
State - TAMIL NADU	589003.0

➡ Total number of female workers in each district of Tamil Nadu who are working more than 3 months and less than 6 months

state	female
District - Ariyalur	11787.0
District - Chennai	14435.0
District - Coimbatore	14167.0
District - Cuddalore	56844.0
District - Dharmapuri	7735.0
District - Dindigul	16353.0
District - Erode	9116.0
District - Kancheepuram	44746.0
District - Kanniyakumari	2036.0
District - Karur	4264.0
District - Krishnagiri	9642.0
District - Madurai	16237.0
District - Nagapattinam	35205.0
District - Namakkal	8859.0
District - Perambalur	8962.0
District - Pudukkottai	11631.0
District - Ramanathapuram	11151.0
District - Salem	14009.0
District - Sivaganga	13962.0
District - Thanjavur	16433.0
District - The Nilgiris	4571.0
District - Theni	7108.0
District - Thiruvallur	35153.0
District - Thiruvarur	28709.0
District - Thoothukkudi	11339.0
District - Tiruchirappalli	14017.0
District - Tirunelveli	18204.0
District - Tiruppur	12224.0
District - Tiruvannamalai	34448.0
District - Vellore	34832.0
District - Viluppuram	71964.0
District - Virudhunagar	11682.0
State - TAMIL NADU	611825.0

✓  
0s

▶ Total number of workers in each district of Tamil Nadu who are working more than 3 months and less than 6 months

state	person
District - Ariyalur	21713.0
District - Chennai	33748.0
District - Coimbatore	27697.0
District - Cuddalore	109233.0
District - Dharmapuri	15092.0
District - Dindigul	30527.0
District - Erode	16881.0
District - Kancheepuram	92015.0
District - Kanniyakumari	4922.0
District - Karur	7572.0
District - Krishnagiri	19917.0
District - Madurai	31352.0
District - Nagapattinam	71565.0
District - Namakkal	15667.0
District - Perambalur	16797.0
District - Pudukkottai	21534.0
District - Ramanathapuram	20975.0
District - Salem	25416.0
District - Sivaganga	27580.0
District - Thanjavur	33040.0
District - The Nilgiris	9971.0
District - Theni	14465.0
District - Thiruvallur	74448.0
District - Thiruvarur	56868.0
District - Thoothukkudi	21085.0
District - Tiruchirappalli	26810.0
District - Tirunelveli	35444.0
District - Tiruppur	23264.0
District - Tiruvannamalai	65476.0
District - Vellore	71695.0
District - Viluppuram	137027.0
District - Virudhunagar	21032.0
State - TAMIL NADU	1200828.0

### Workers working less than 3 months:

➡ Total number of male workers in each district of Tamil Nadu who are working less than 3 months

state	male1
District - Ariyalur	1770.0
District - Chennai	3954.0
District - Coimbatore	2058.0
District - Cuddalore	9994.0
District - Dharmapuri	1104.0
District - Dindigul	2186.0
District - Erode	1419.0
District - Kancheepuram	8655.0
District - Kanniyakumari	636.0
District - Karur	642.0
District - Krishnagiri	2053.0
District - Madurai	2891.0
District - Nagapattinam	4923.0
District - Namakkal	547.0
District - Perambalur	613.0
District - Pudukkottai	1412.0
District - Ramanathapuram	1671.0
District - Salem	1937.0
District - Sivaganga	1893.0
District - Thanjavur	2782.0
District - The Nilgiris	849.0
District - Theni	1365.0
District - Thiruvallur	8004.0
District - Thiruvarur	3405.0
District - Thoothukkudi	979.0
District - Tiruchirappalli	1915.0
District - Tirunelveli	3131.0
District - Tiruppur	1474.0
District - Tiruvannamalai	5832.0
District - Vellore	7753.0
District - Viluppuram	9800.0
District - Virudhunagar	1721.0
State - TAMIL NADU	99368.0

Total number of female workers in each district of Tamil Nadu  
who are working less than 3 months

state	female1
District - Ariyalur	2709.0
District - Chennai	2397.0
District - Coimbatore	2585.0
District - Cuddalore	12445.0
District - Dharmapuri	1676.0
District - Dindigul	3309.0
District - Erode	1838.0
District - Kancheepuram	9037.0
District - Kanniyakumari	612.0
District - Karur	930.0
District - Krishnagiri	2046.0
District - Madurai	3330.0
District - Nagapattinam	6897.0
District - Namakkal	1154.0
District - Perambalur	1128.0
District - Pudukkottai	2155.0
District - Ramanathapuram	1898.0
District - Salem	2587.0
District - Sivaganga	2658.0
District - Thanjavur	3174.0
District - The Nilgiris	754.0
District - Theni	1553.0
District - Thiruvallur	7862.0
District - Thiruvavur	5110.0
District - Thoothukkudi	1410.0
District - Tiruchirappalli	2527.0
District - Tirunelveli	3647.0
District - Tiruppur	1836.0
District - Tiruvannamalai	7871.0
District - Vellore	9093.0
District - Viluppuram	13387.0
District - Virudhunagar	2403.0
State - TAMIL NADU	122018.0

## Male and Female Cultivators:

➞ Total number of male cultivators in each district of Tamil Nadu

state	cultivator_male
District - Ariyalur	760.0
District - Chennai	327.0
District - Coimbatore	242.0
District - Cuddalore	3886.0
District - Dharmapuri	356.0
District - Dindigul	443.0
District - Erode	169.0
District - Kancheepuram	2578.0
District - Kanniyakumari	60.0
District - Karur	96.0
District - Krishnagiri	871.0
District - Madurai	739.0
District - Nagapattinam	1313.0
District - Namakkal	276.0
District - Perambalur	1707.0
District - Pudukkottai	1305.0
District - Ramanathapuram	2463.0
District - Salem	338.0
District - Sivaganga	1779.0
District - Thanjavur	821.0
District - The Nilgiris	117.0
District - Theni	205.0
District - Thiruvallur	1663.0
District - Thiruvarur	1366.0
District - Thoothukkudi	393.0
District - Tiruchirappalli	607.0
District - Tirunelveli	916.0
District - Tiruppur	314.0
District - Tiruvannamalai	2582.0
District - Vellore	1163.0
District - Viluppuram	4182.0
District - Virudhunagar	595.0
State - TAMIL NADU	34632.0



Total number of female cultivators in each district of Tamil Nadu

state	cultivator_female
District - Ariyalur	688.0
District - Chennai	413.0
District - Coimbatore	334.0
District - Cuddalore	3192.0
District - Dharmapuri	361.0
District - Dindigul	473.0
District - Erode	180.0
District - Kancheepuram	1890.0
District - Kanniyakumari	47.0
District - Karur	81.0
District - Krishnagiri	584.0
District - Madurai	741.0
District - Nagapattinam	944.0
District - Namakkal	363.0
District - Perambalur	1275.0
District - Pudukkottai	1172.0
District - Ramanathapuram	2281.0
District - Salem	410.0
District - Sivaganga	1522.0
District - Thanjavur	583.0
District - The Nilgiris	108.0
District - Theni	195.0
District - Thiruvallur	1403.0
District - Thiruvarur	865.0
District - Thoothukkudi	410.0
District - Tiruchirappalli	687.0
District - Tirunelveli	849.0
District - Tiruppur	307.0
District - Tiruvannamalai	2089.0
District - Vellore	1026.0
District - Viluppuram	3560.0
District - Virudhunagar	570.0
State - TAMIL NADU	29603.0



## Visualization Based on Categories:

### Cultivation:

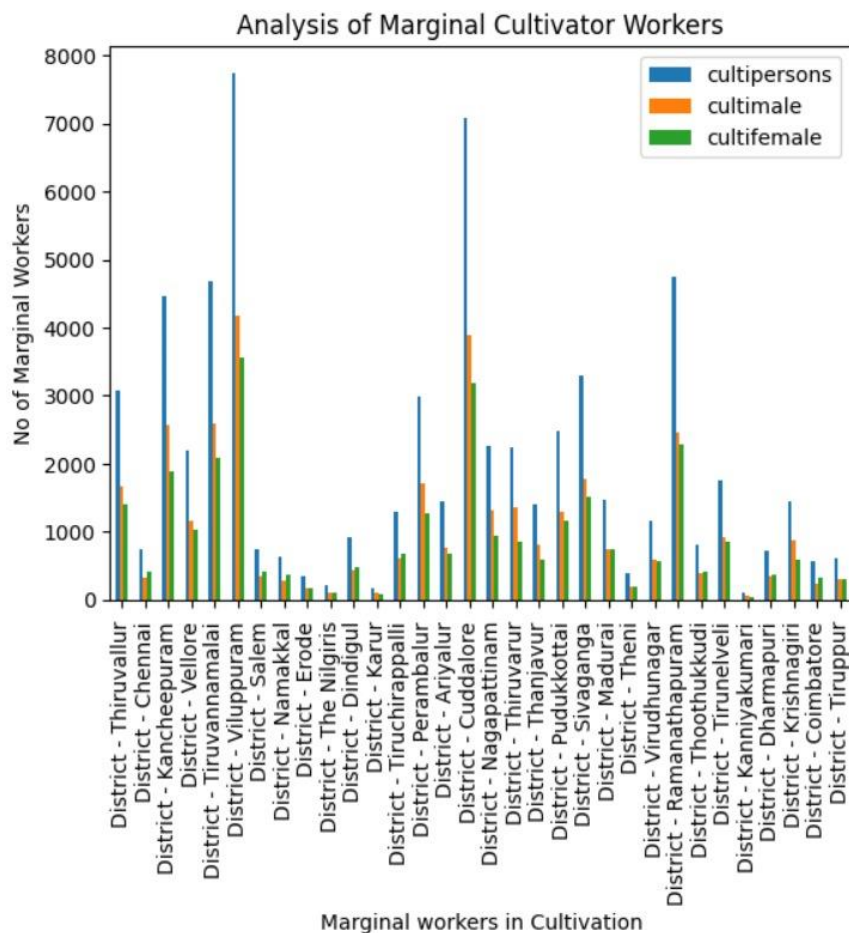
```
import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("modifiedDatasetMarginal2.csv")

df = df[df['TRU'] == 'Total']
df = df[df['Age'] == 'Total']
selected_column=df[['cultipersons','cultimale','cultifemale','StateName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis')

plt.title('Analysis of Marginal Cultivator Workers')

plt.xlabel('Marginal workers in Cultivation')

plt.ylabel('No of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['StateName'])
#Display the figure
plt.show()
```



## Agriculture:

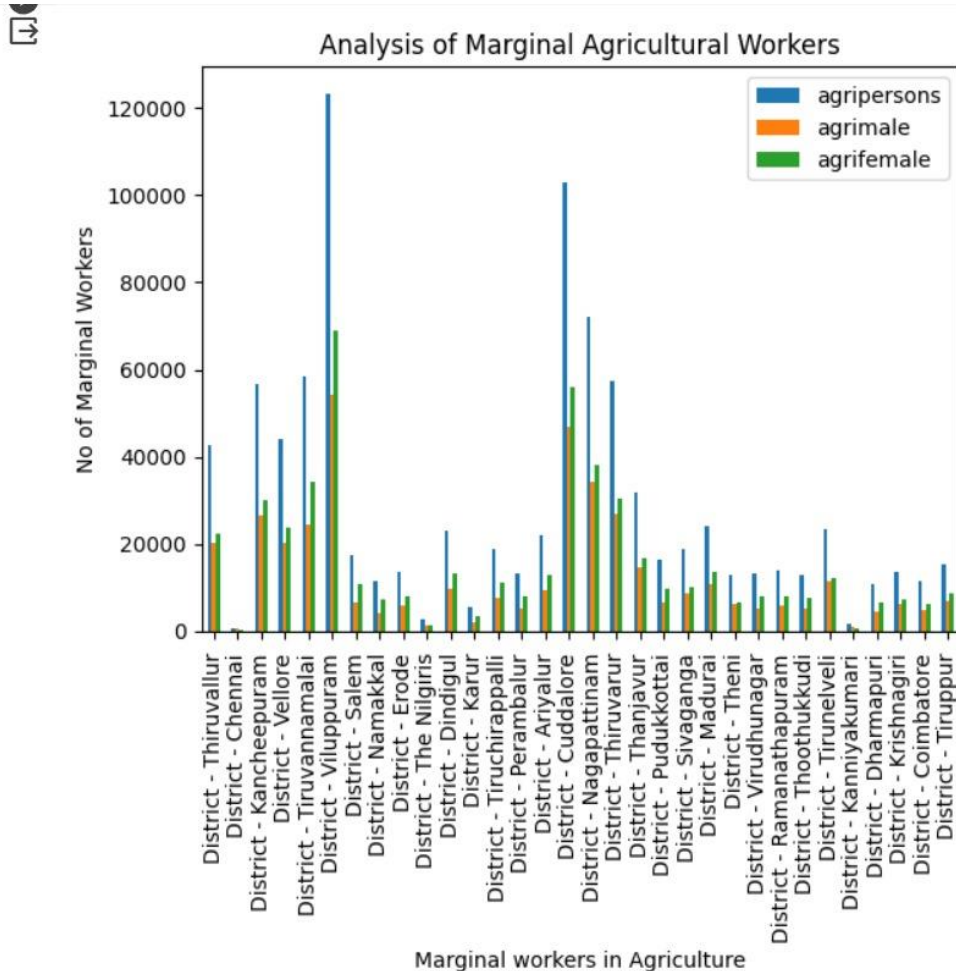
```
import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("modifiedDatasetMarginal2.csv")

df = df[df['TRU'] == 'Total']
df = df[df['Age'] == 'Total']
selected_column=df[['agripersons','agrimale','agrifemale','StateName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis')

plt.title('Analysis of Marginal Agricultural Workers')

plt.xlabel('Marginal workers in Agriculture')

plt.ylabel('No of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['StateName'])
#Display the figure
plt.show()
```



## Plantation, Livestock, Forestry, Fishing, Hunting and allied activities:

```
import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("modifiedDatasetMarginal2.csv")

df = df[df['TRU'] == 'Total']
df = df[df['Age'] == 'Total']
selected_column=df[['plantpersons','plantmale','plantfemale','StateName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis')

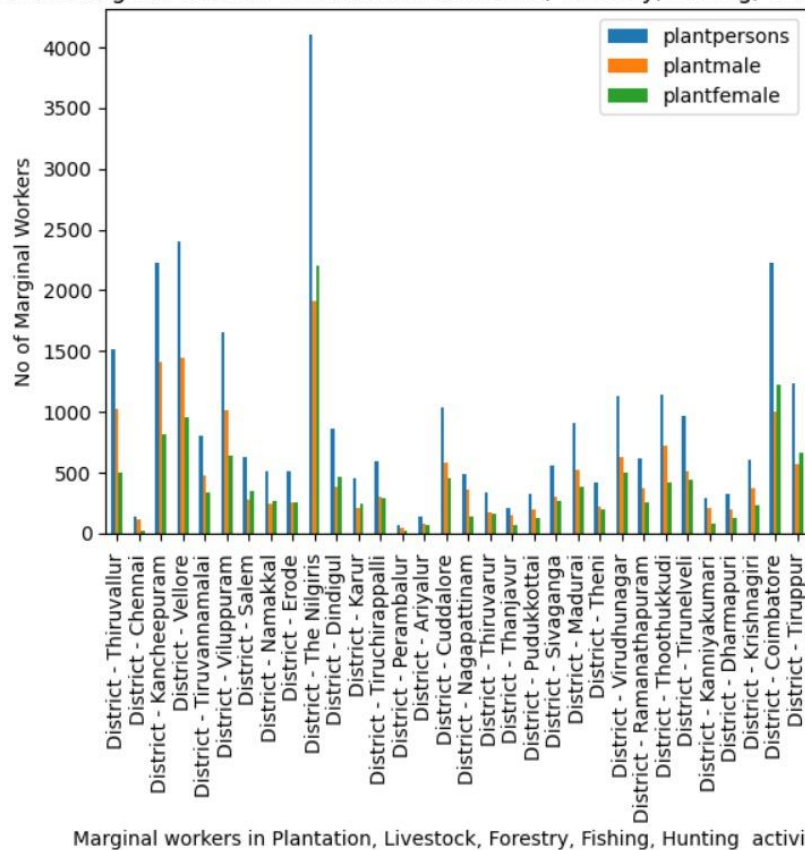
plt.title('Analysis of Marginal Workers in Plantation Livestock, Forestry, Fishing, Hunting activities')

plt.xlabel('Marginal workers in Plantation, Livestock, Forestry, Fishing, Hunting activities')

plt.ylabel('No of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['StateName'])
#Display the figure
plt.show()
```



Analysis of Marginal Workers in Plantation Livestock, Forestry, Fishing, Hunting activities



## Visualization Based on Ages:

```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/Marginal_workers.csv")

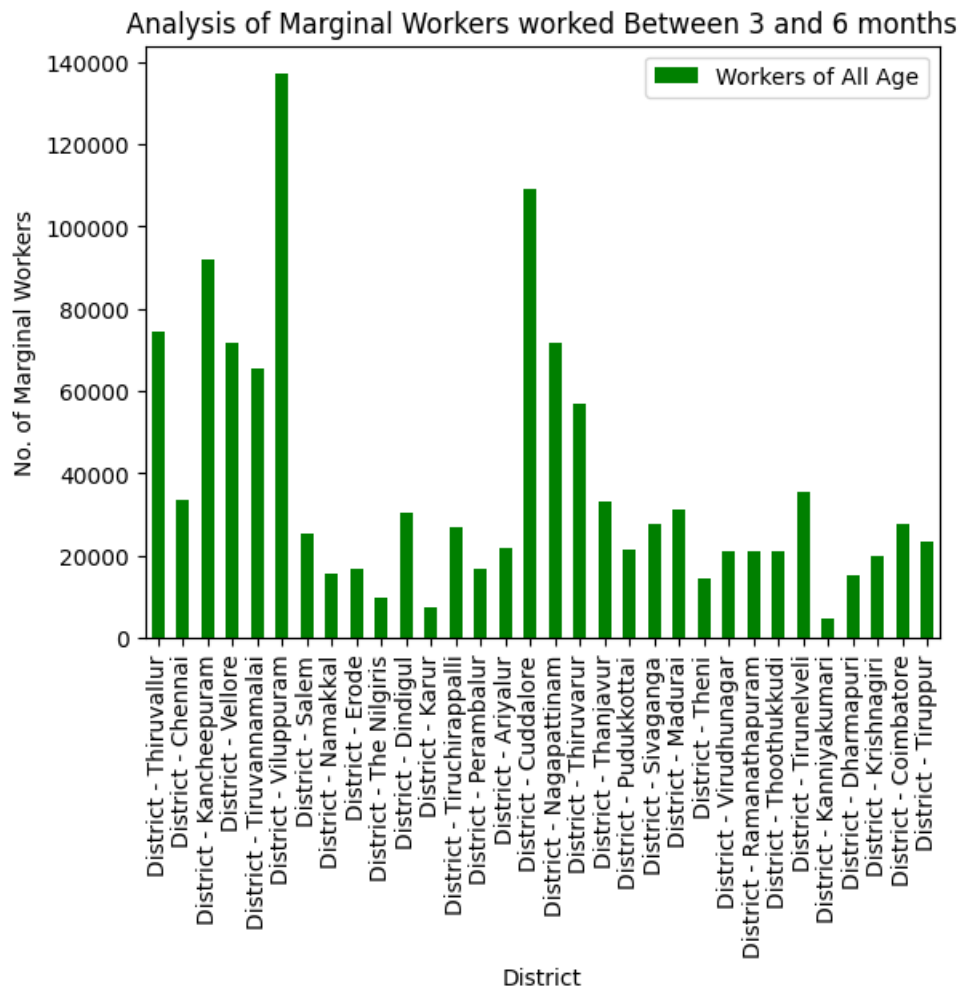
df = df[df['AgeGroup'] == 'Total']
selected_column=df[['Worked_Between_3and6_months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='green')

plt.title('Analysis of Marginal Workers worked Between 3 and 6 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of All Age"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/Marginal_workers.csv")

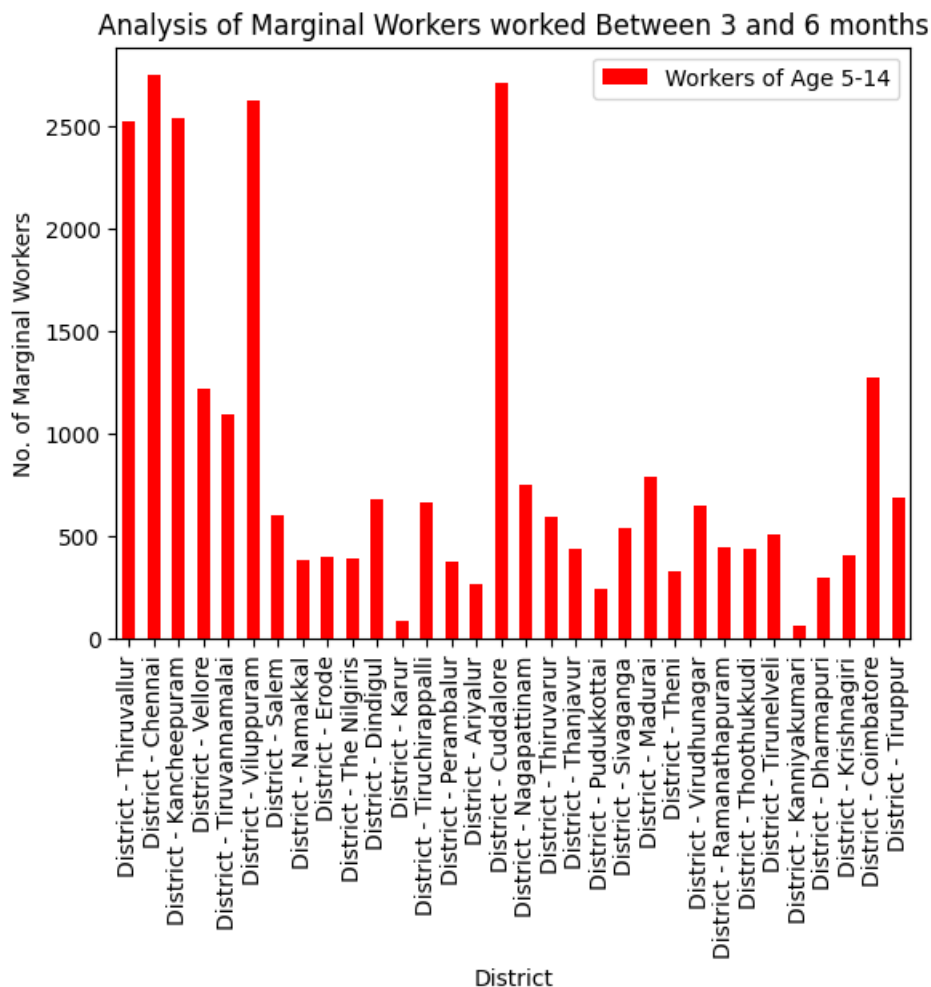
df = df[df['AgeGroup'] == '5-14']
selected_column=df[['Worked_Between_3and6_months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='red')

plt.title('Analysis of Marginal Workers worked Between 3 and 6 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 5-14"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/Marginal_workers.csv")

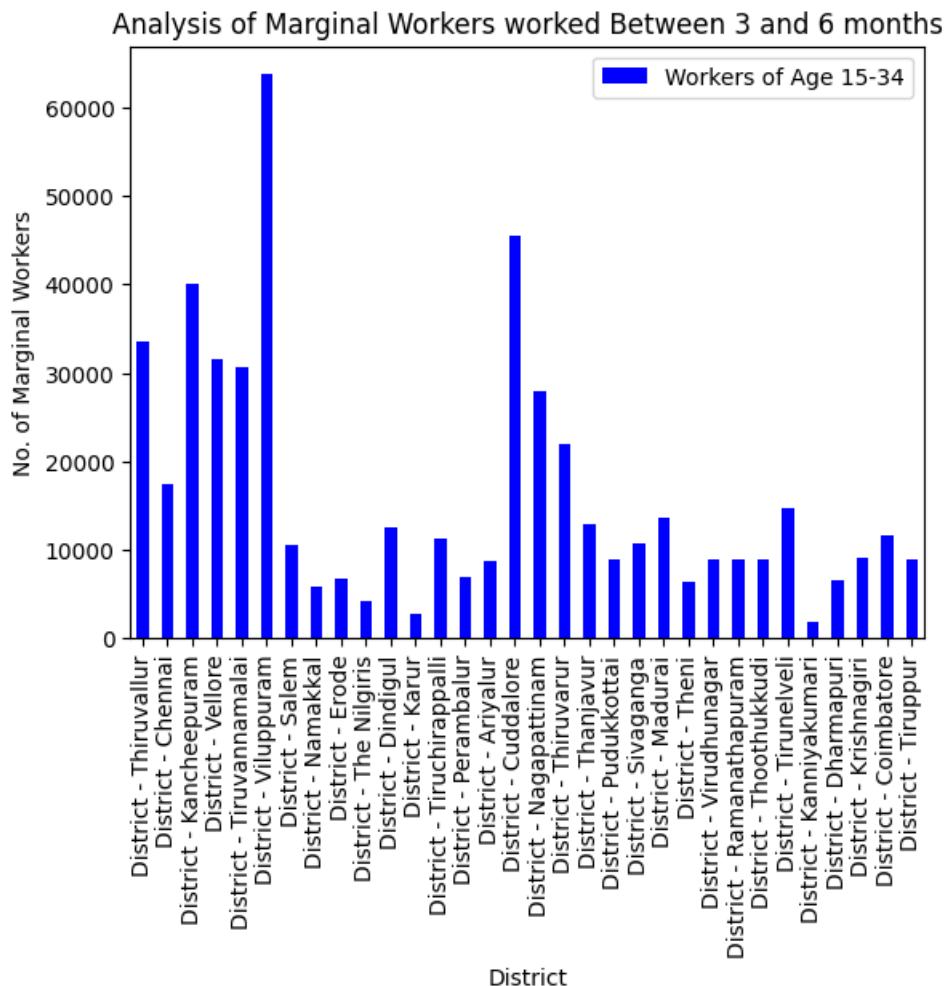
df = df[df['AgeGroup'] == '15-34']
selected_column=df[['Worked_Between_3and6_months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='blue')

plt.title('Analysis of Marginal Workers worked Between 3 and 6 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 15-34"])
#Display the figure
plt.show()

```





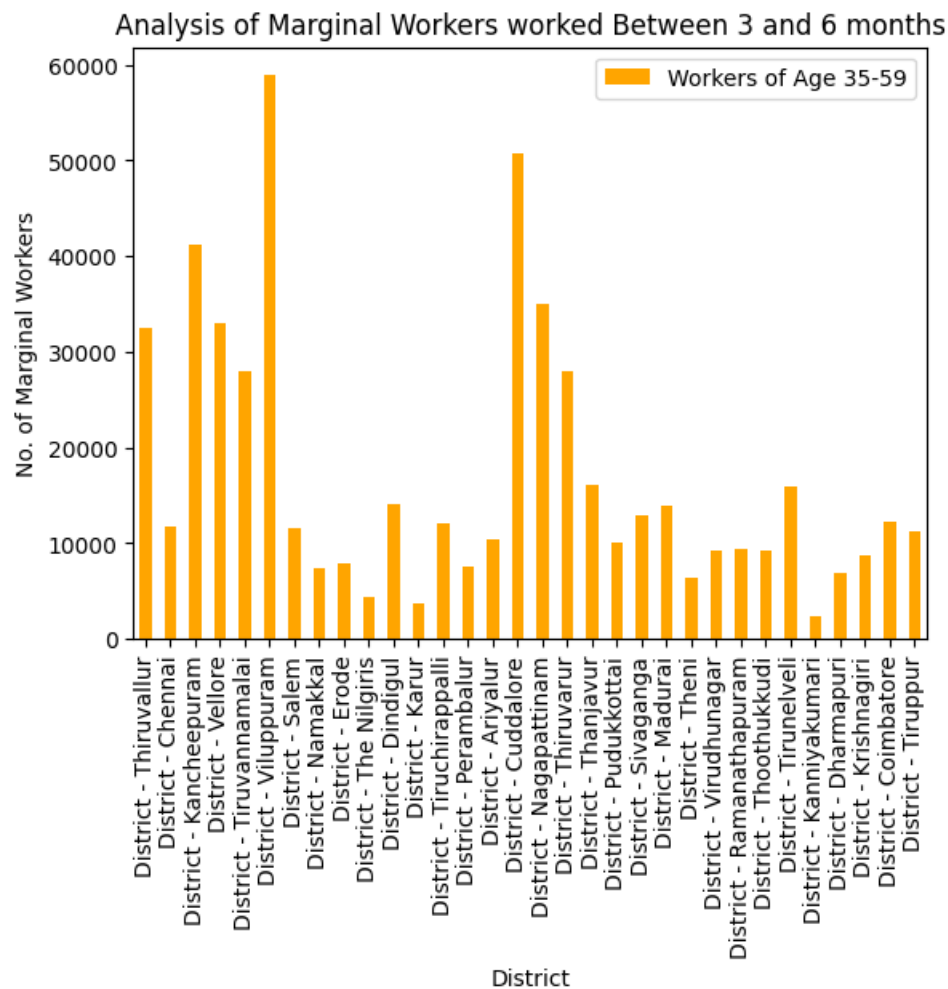
```
[ ] import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/Marginal_workers.csv")

df = df[df['AgeGroup'] == '35-59']
selected_column=df[['Worked_Between_3and6_months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='orange')

plt.title('Analysis of Marginal Workers worked Between 3 and 6 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 35-59"])
#Display the figure
plt.show()
```





```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/Marginal_workers.csv")

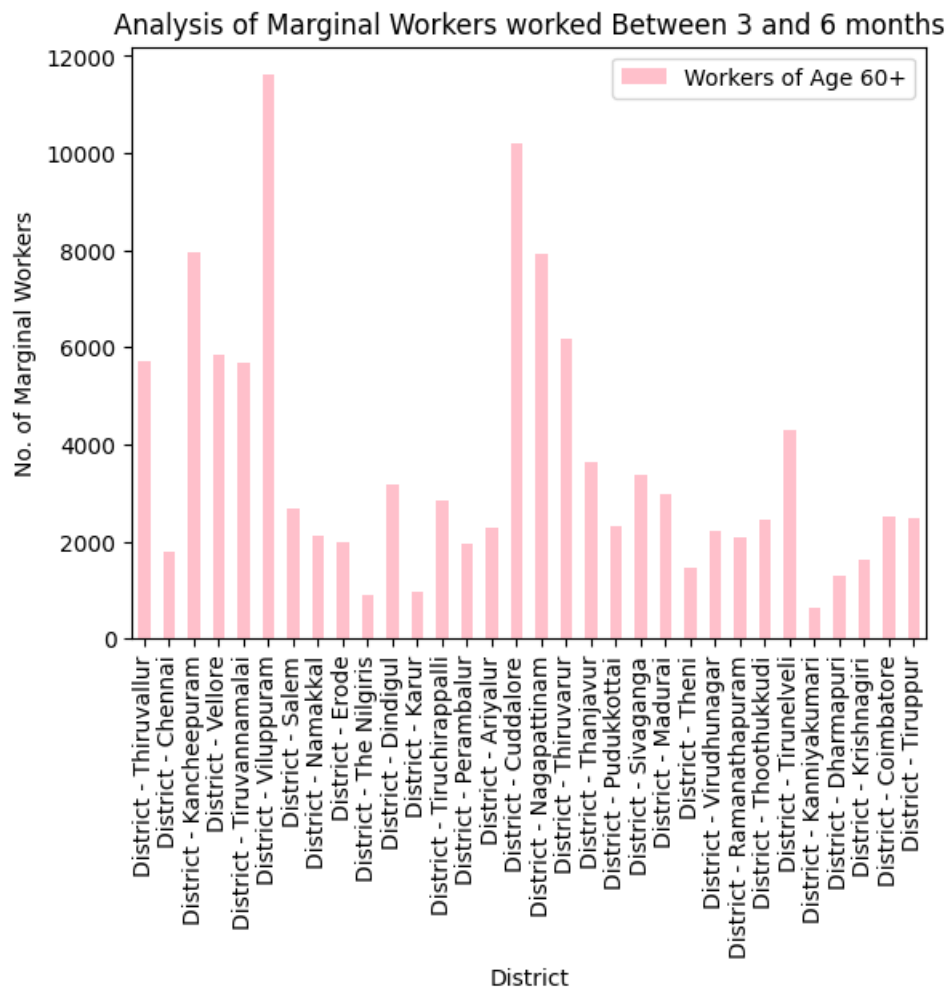
df = df[df['AgeGroup'] == '60+']
selected_column=df[['Worked_Between_3and6_months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='pink')

plt.title('Analysis of Marginal Workers worked Between 3 and 6 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 60+"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/Marginal_workers.csv")

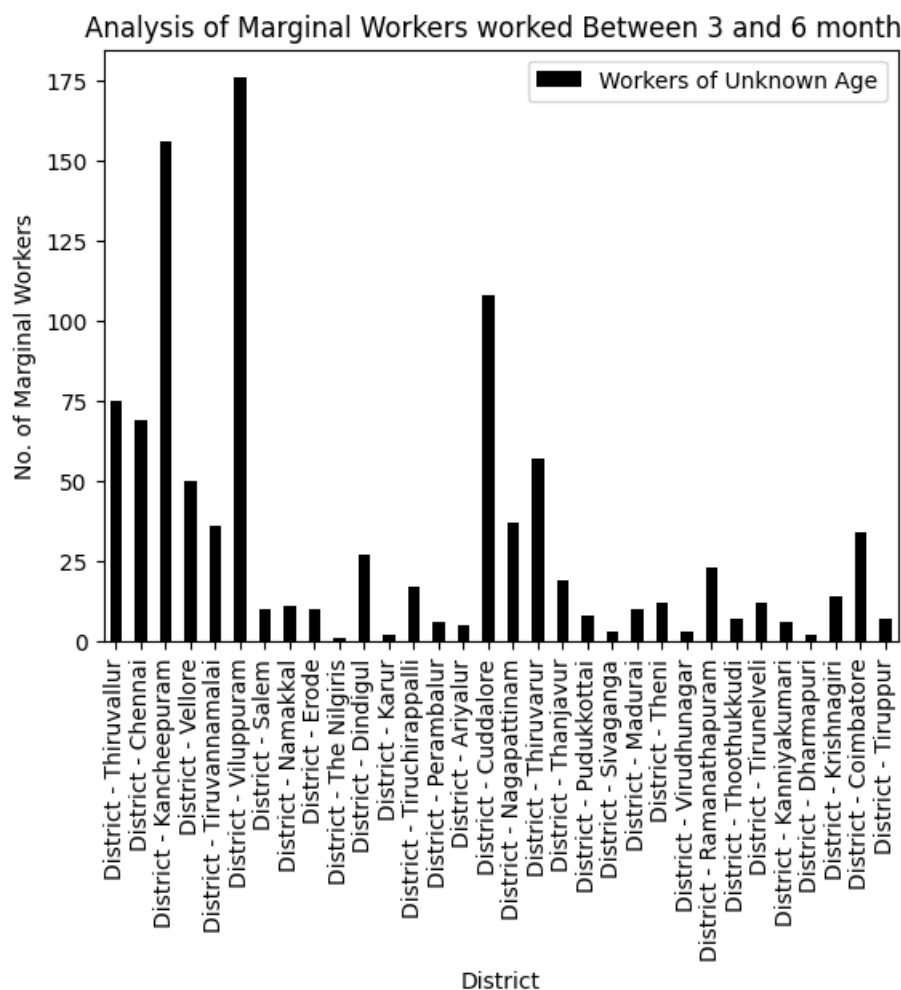
df = df[df['AgeGroup'] == 'Age not stated']
selected_column=df[['Worked_Between_3and6_months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='black')

plt.title('Analysis of Marginal Workers worked Between 3 and 6 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Unknown Age"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/marginal_less3.csv")

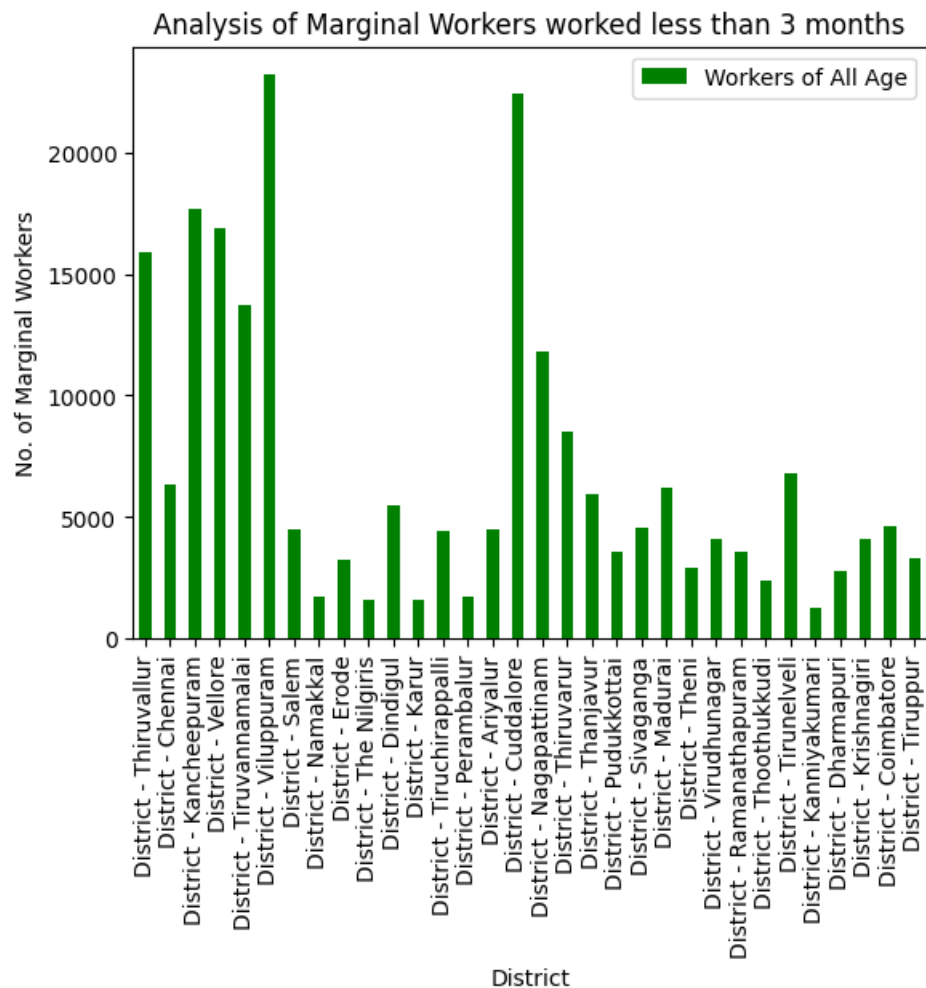
df = df[df['AgeGroup'] == 'Total']
selected_column=df[['Worked_less_than_3Months','AreaName']]
selected_column.plot(kind='bar',title='Marginal Workers Analysis',color='green')

plt.title('Analysis of Marginal Workers worked less than 3 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of All Age"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/marginal_less3.csv")

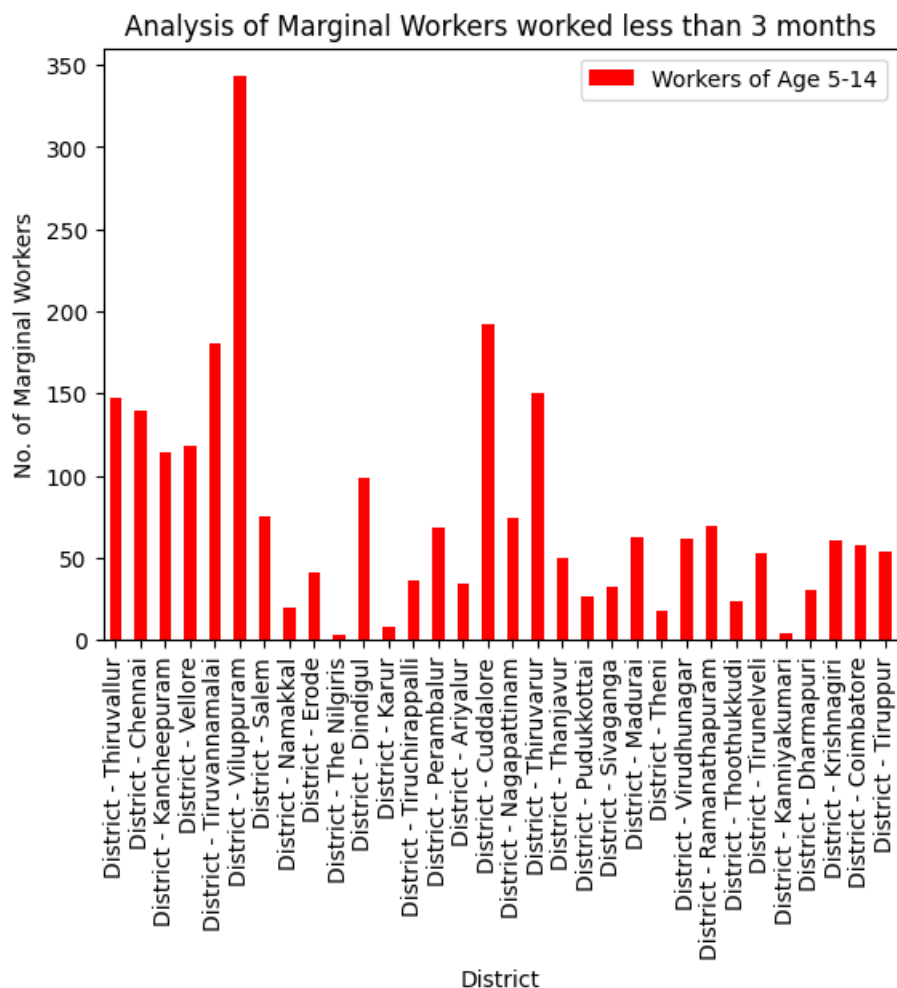
df = df[df['AgeGroup'] == '5-14']
selected_column=df[['Worked_less_than_3Months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='red')

plt.title('Analysis of Marginal Workers worked less than 3 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 5-14"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/marginal_less3.csv")

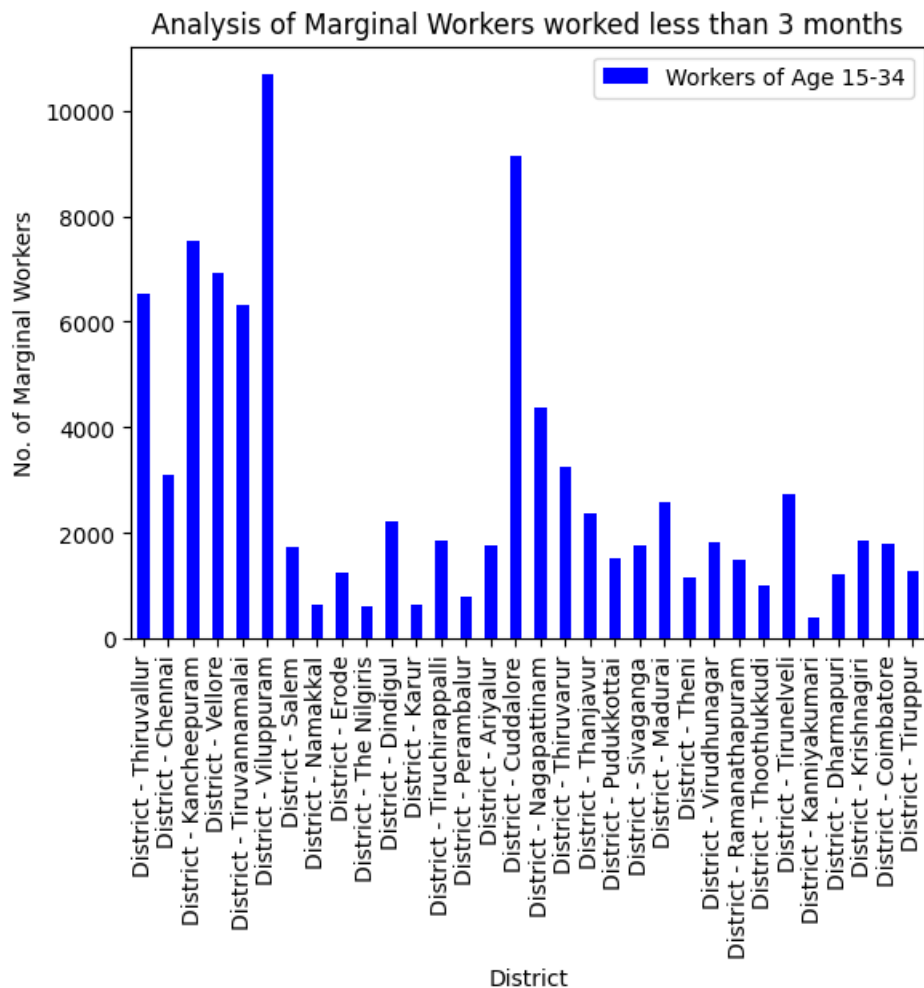
df = df[df['AgeGroup'] == '15-34']
selected_column=df[['Worked_less_than_3Months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='blue')

plt.title('Analysis of Marginal Workers worked less than 3 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 15-34"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/marginal_less3.csv")

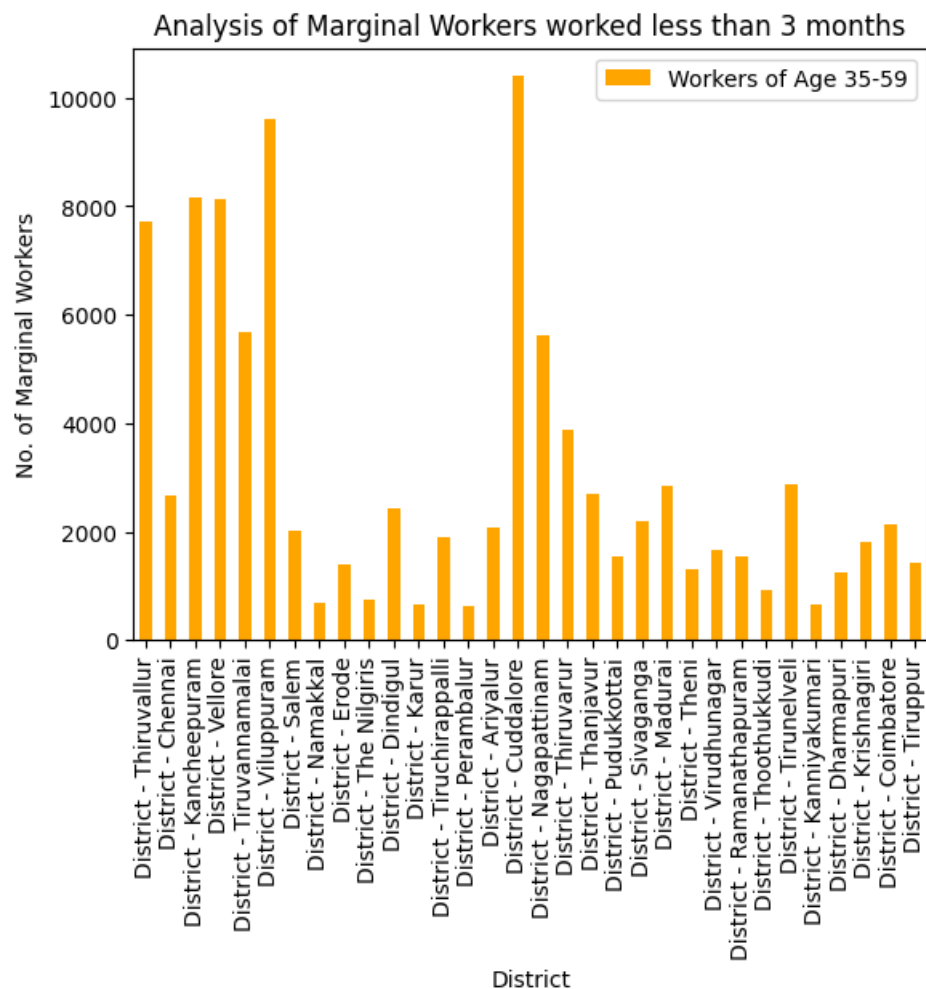
df = df[df['AgeGroup'] == '35-59']
selected_column=df[['Worked_less_than_3Months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='orange')

plt.title('Analysis of Marginal Workers worked less than 3 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 35-59"])
#Display the figure
plt.show()

```



```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/marginal_less3.csv")

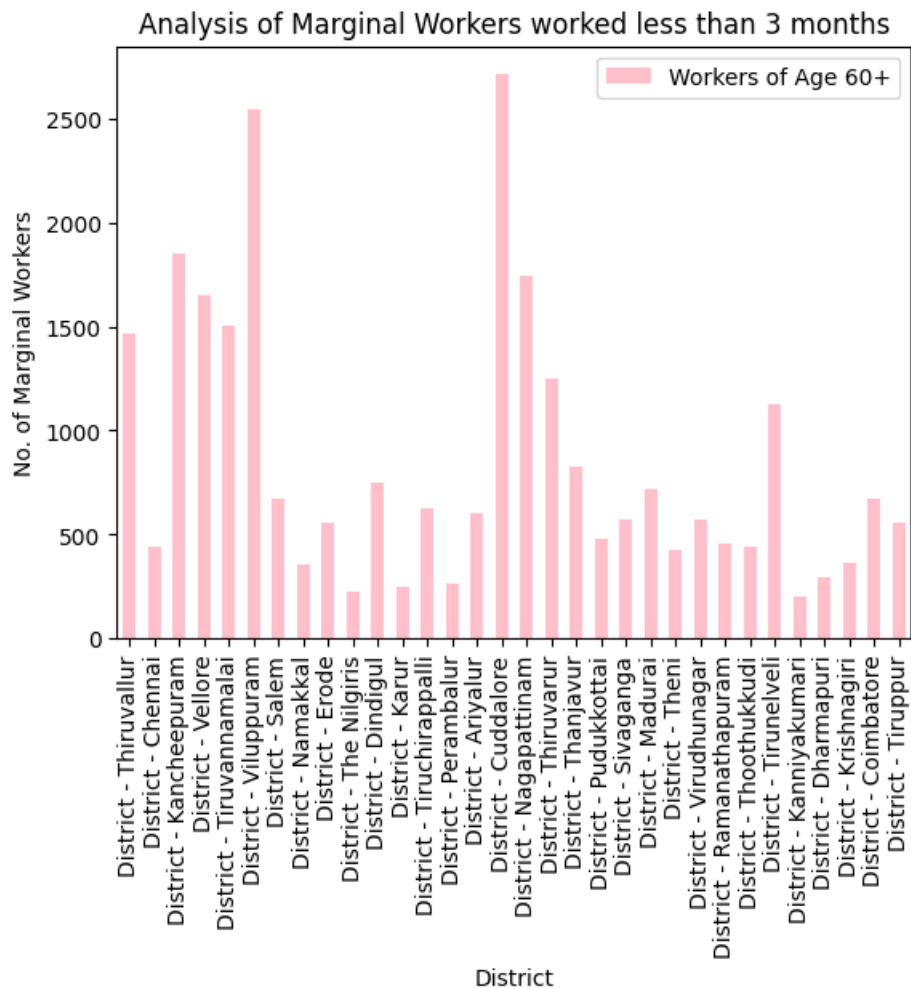
df = df[df['AgeGroup'] == '60+']
selected_column=df[['Worked_less_than_3Months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='pink')

plt.title('Analysis of Marginal Workers worked less than 3 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Age 60+"])
#Display the figure
plt.show()

```





```

import pandas as pd
import matplotlib.pyplot as plt
# reads "modifiedDatasetMarginal2.csv" to df by giving path to the file
df=pd.read_csv("/content/marginal_less3.csv")

df = df[df['AgeGroup'] == 'Age not stated']
selected_column=df[['Worked_less_than_3Months','AreaName']]
selected_column.plot(kind='bar',title='Marginal_Workers Analysis',color='black')

plt.title('Analysis of Marginal Workers worked less than 3 months')

plt.xlabel('District')

plt.ylabel('No. of Marginal Workers')
plt.xticks(range(len(selected_column)), selected_column['AreaName'])
plt.legend(["Workers of Unknown Age"])
#Display the figure
plt.show()

```

