

PHASE 4: DEVELOPMENT PART 2

PROJECT TITLE: ASSESSMENT OF MARGINAL WORKERS IN TAMILNADU-A SOCIO ECONOMIC ANALYSIS

TEAM MEMBERS:

SIVA GURU.K

NITHISH.S

MANIBARATHI

KARTHIKEYAN

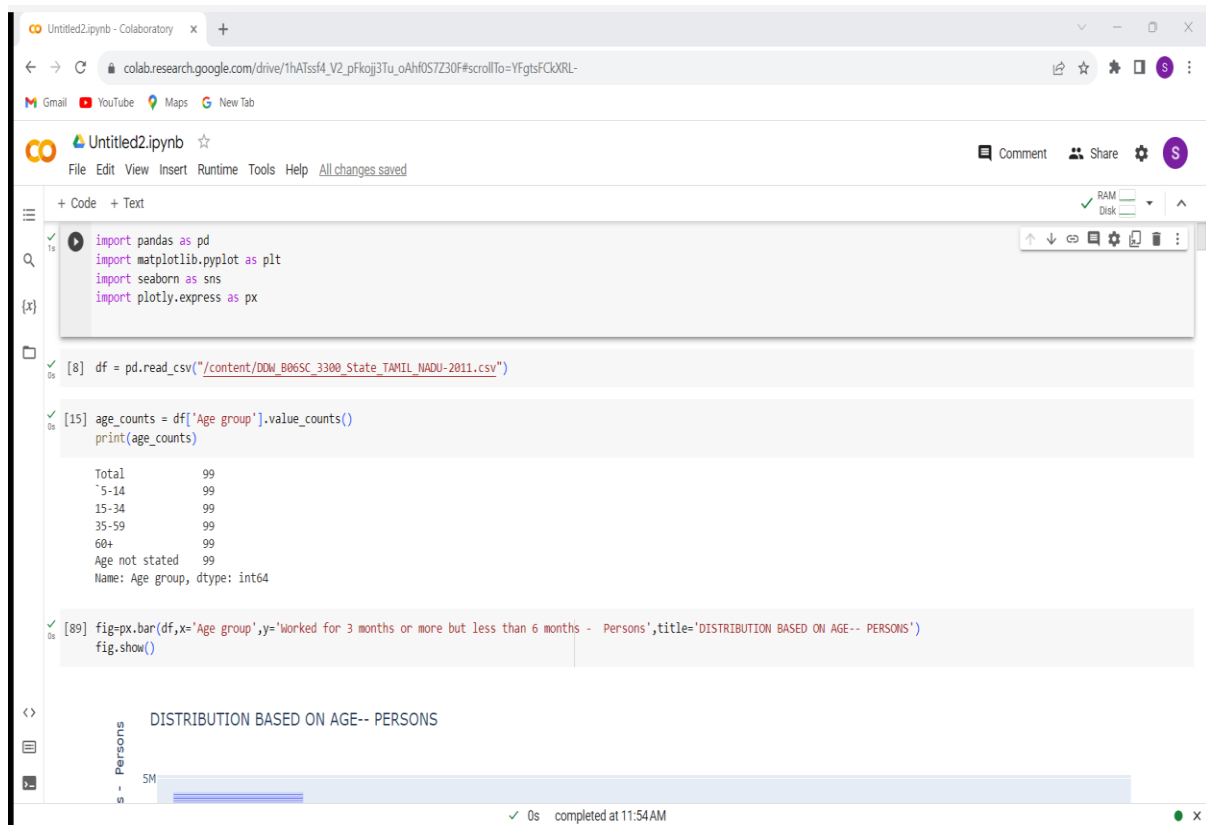
DEVELOPMENT PART 2: DEMOGRAPHIC ANALYSIS AND CREATING VISUALIZATIONS USING IBM COGNOS

OBJECTIVES:

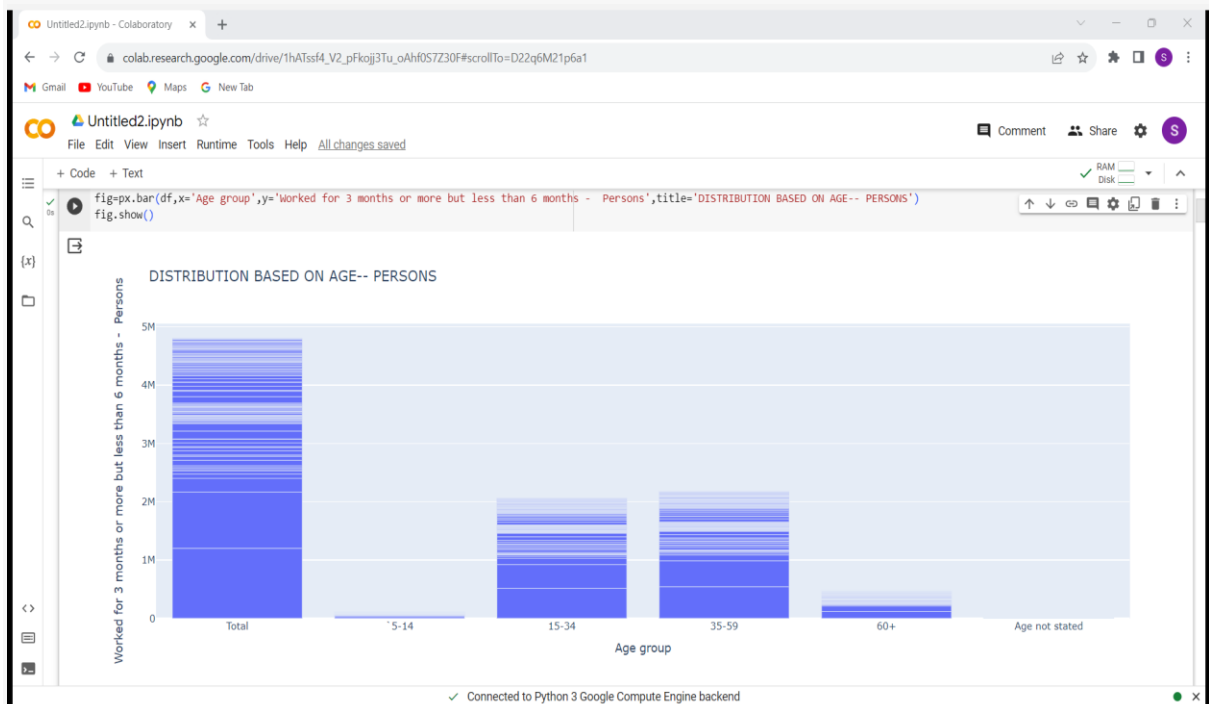
- Perform demographic analysis on the marginal workers dataset.
- Create visualizations using data visualization libraries (e.g., Matplotlib, Seaborn) to show the distribution of marginal workers based on age, industrial category, and sex.

DISTRIBUTION MARGINAL WORKERS BASED ON AGE:

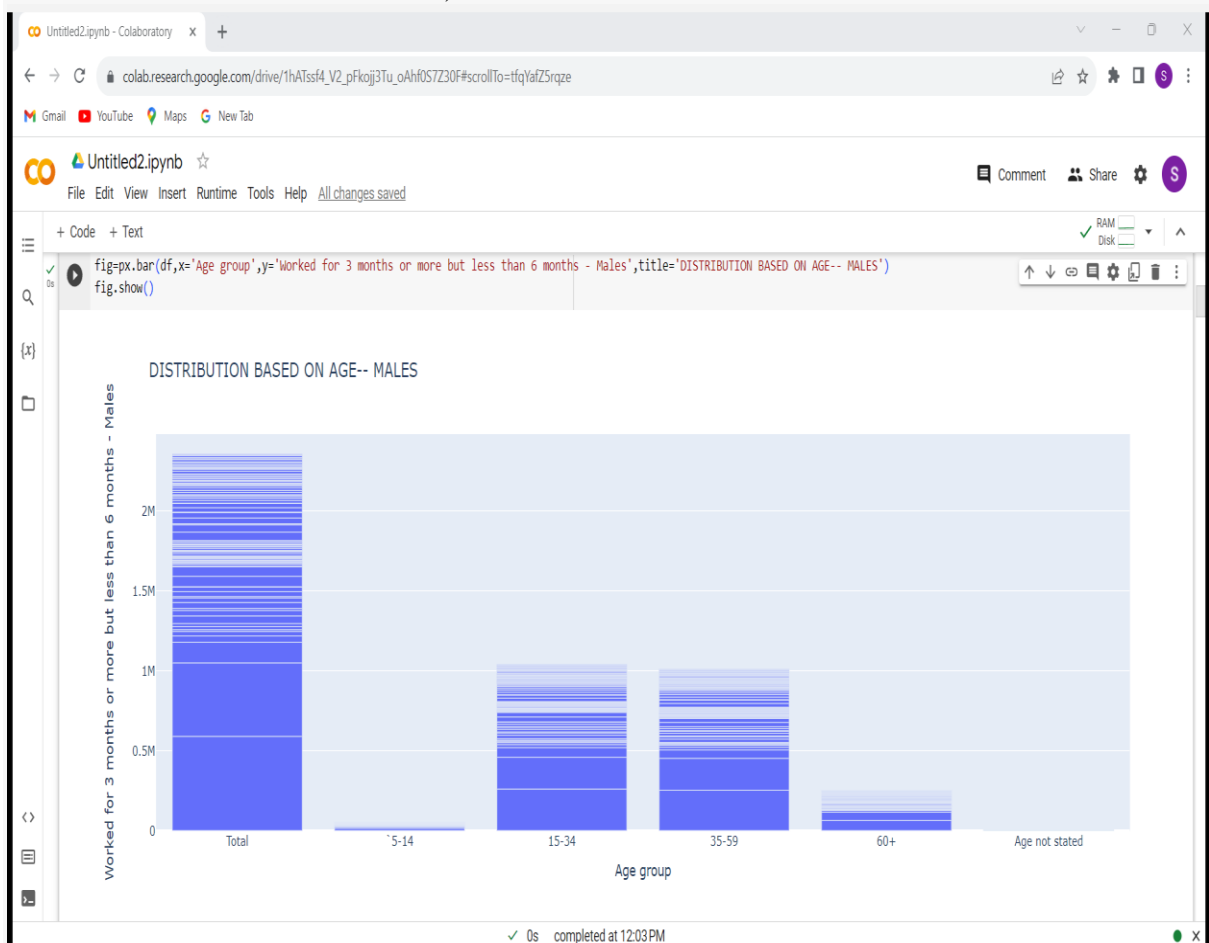
DEMOGRAPHIC ANALYSIS:



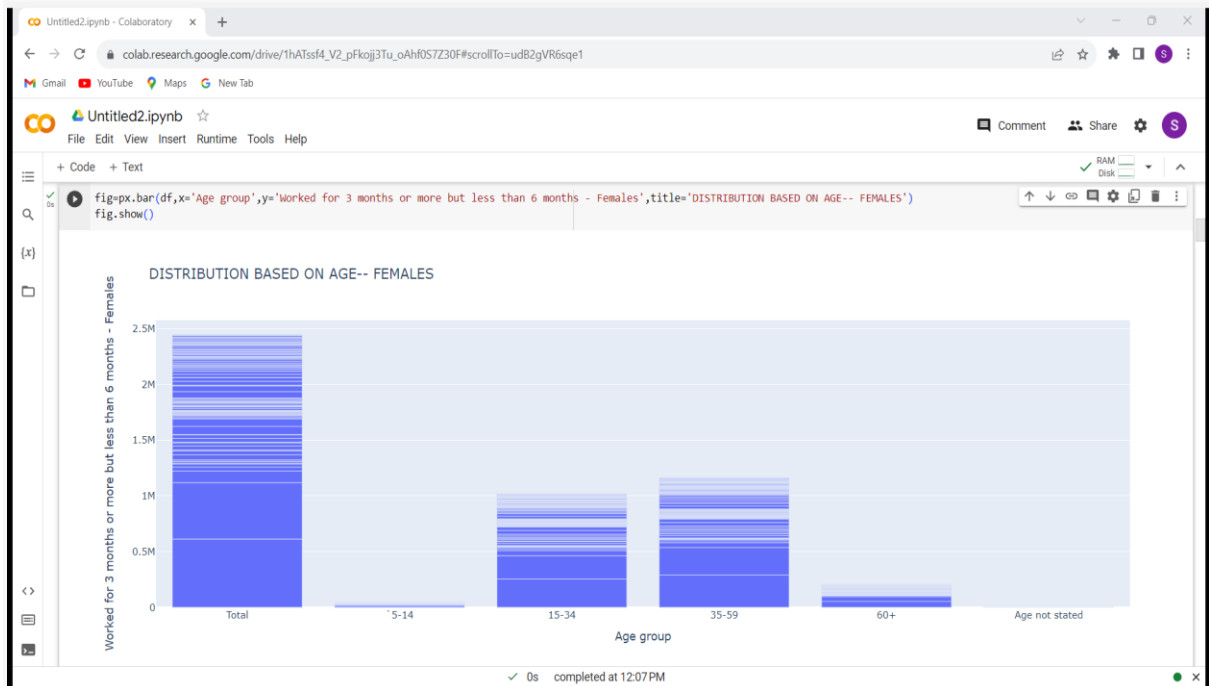
VISUALIZATION(Worked for 3 months or more but less than 6 months Persons WITH RESPECT TO AGE GROUP):



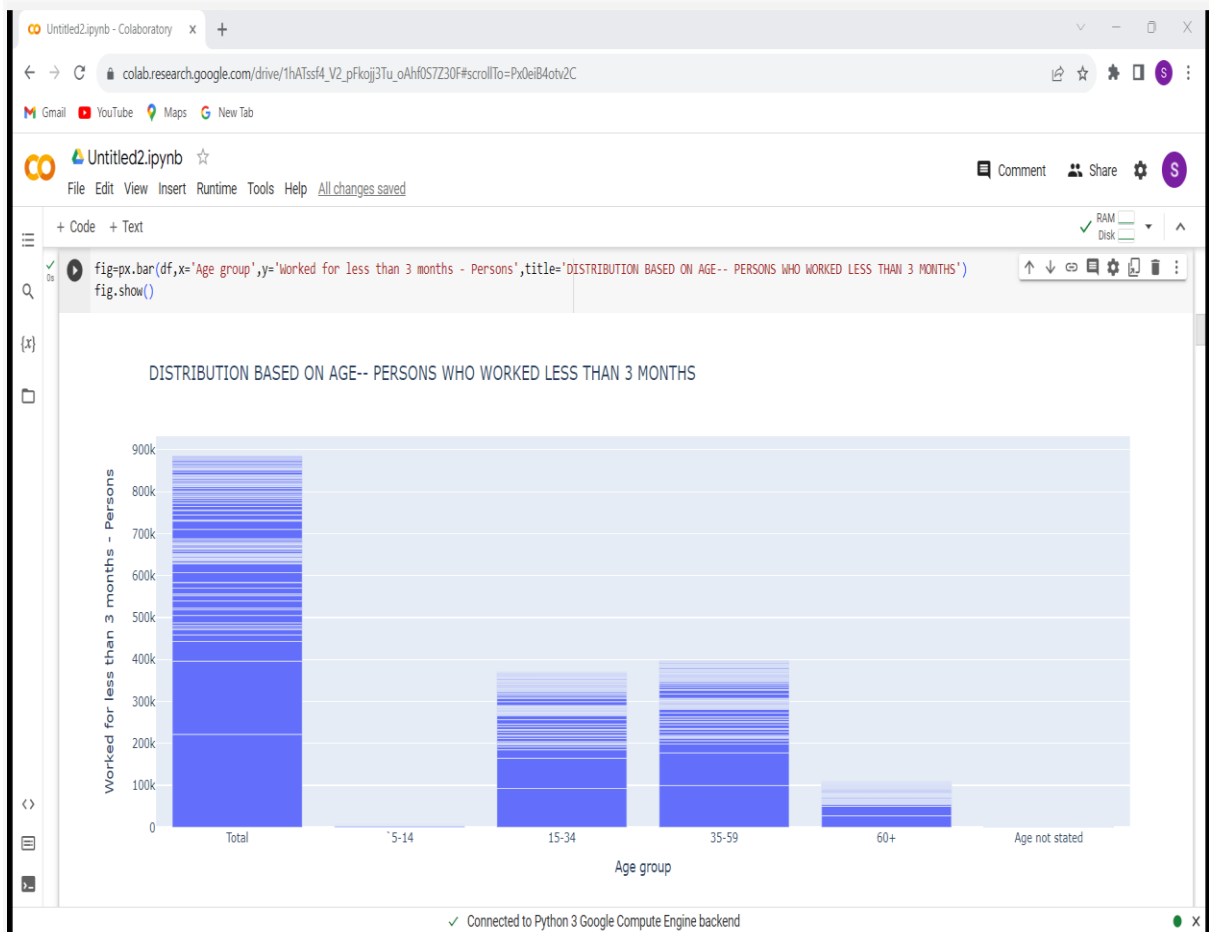
VISUALIZATION(Worked for 3 months or more but less than 6 months Males WITH RESPECT TO AGE GROUP):



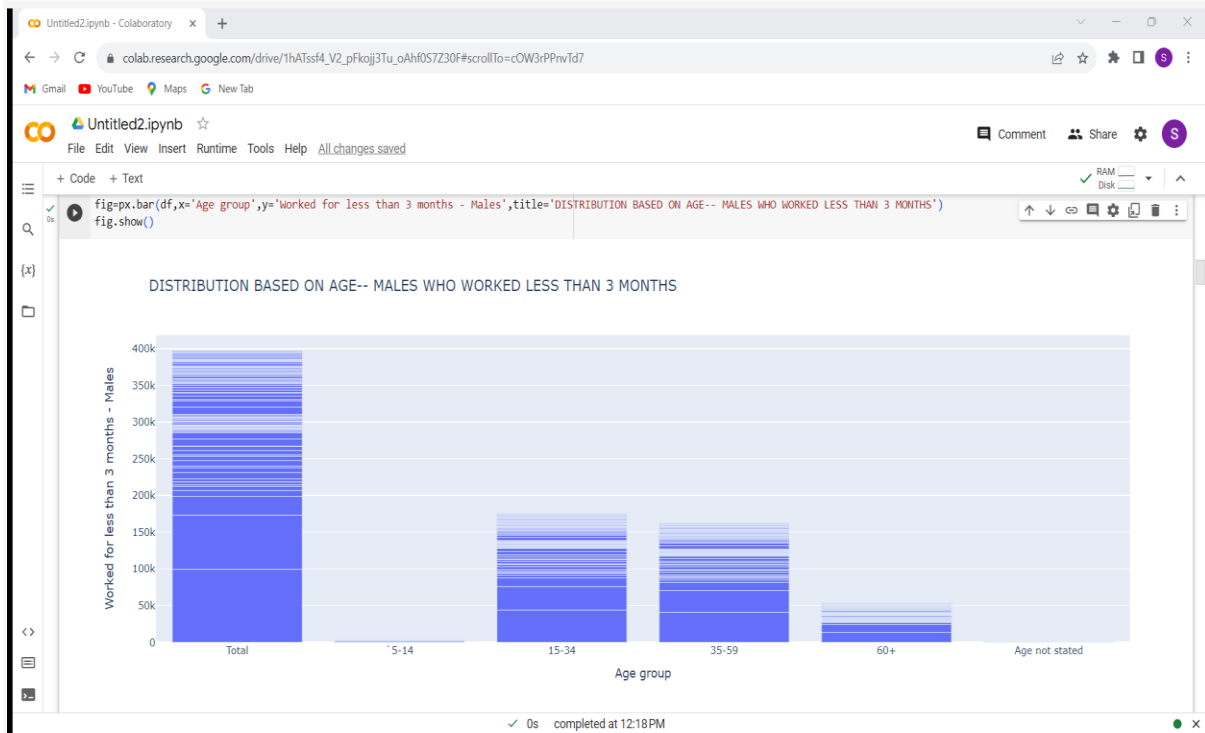
VISUALIZATION(Worked for 3 months or more but less than 6 months Females WITH RESPECT TO AGE GROUP):



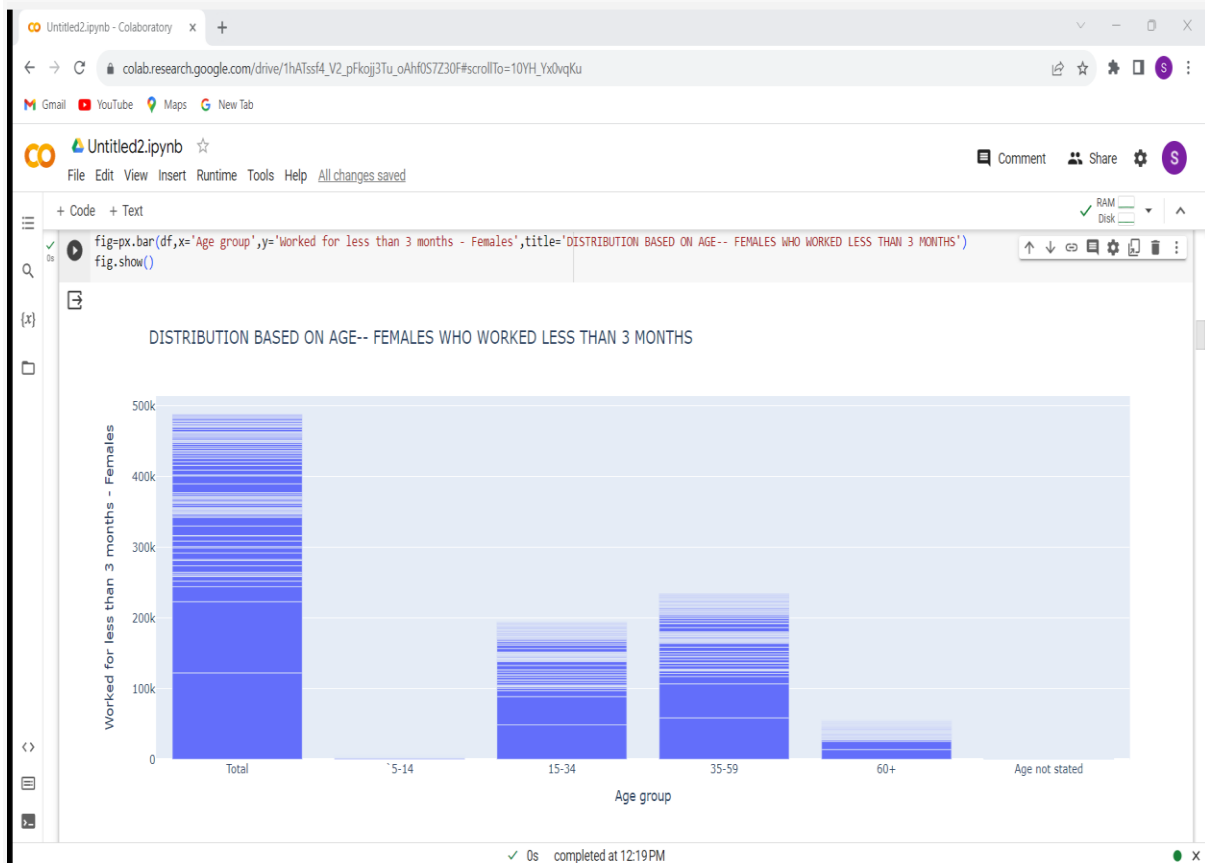
VISUALIZATION(Worked less than 3 months-Persons WITH RESPECT TO AGE GROUP):



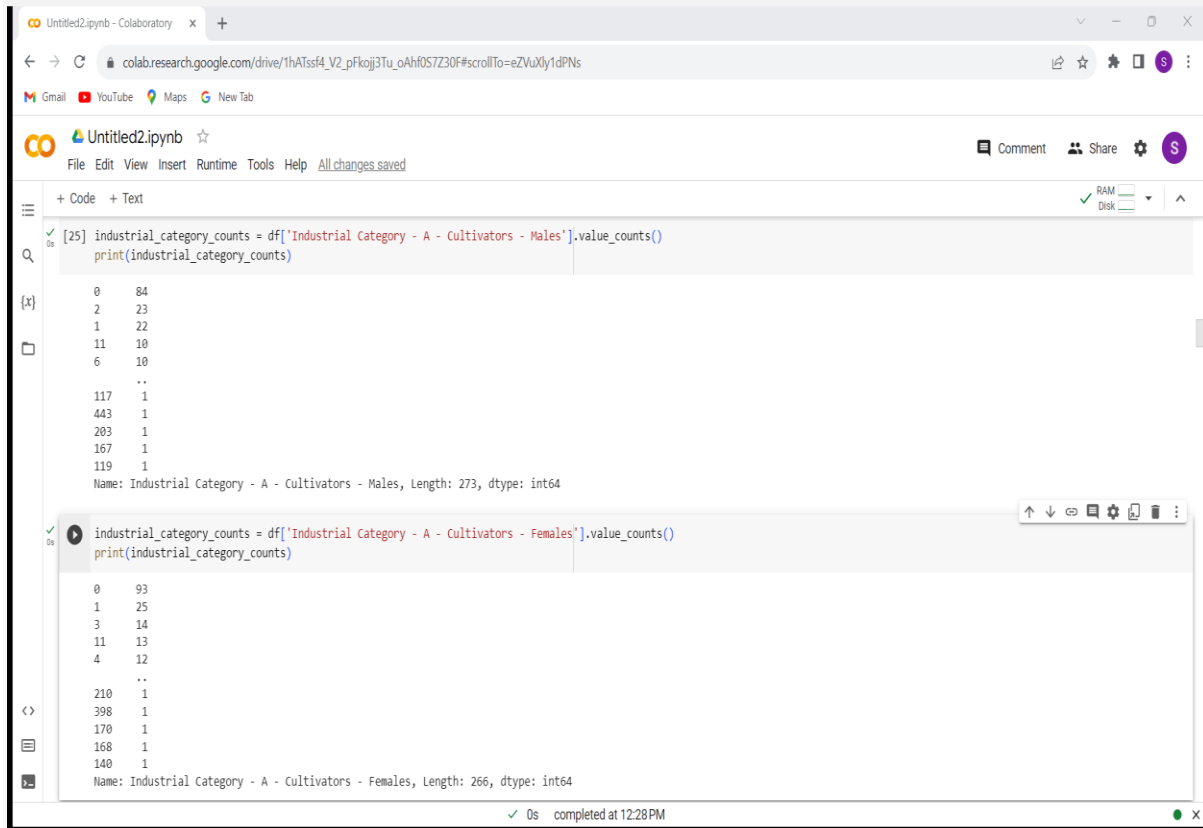
VISUALIZATION(Worked less than 3 months-Males WITH RESPECT TO AGE GROUP):



VISUALIZATION(Worked less than 3 months-Females WITH RESPECT TO AGE GROUP):



DISTRIBUTION MARGINAL WORKERS BASED ON INDUSTRIAL CATEGORY: DEMOGRAPHIC ANALYSIS(CULTIVATORS):



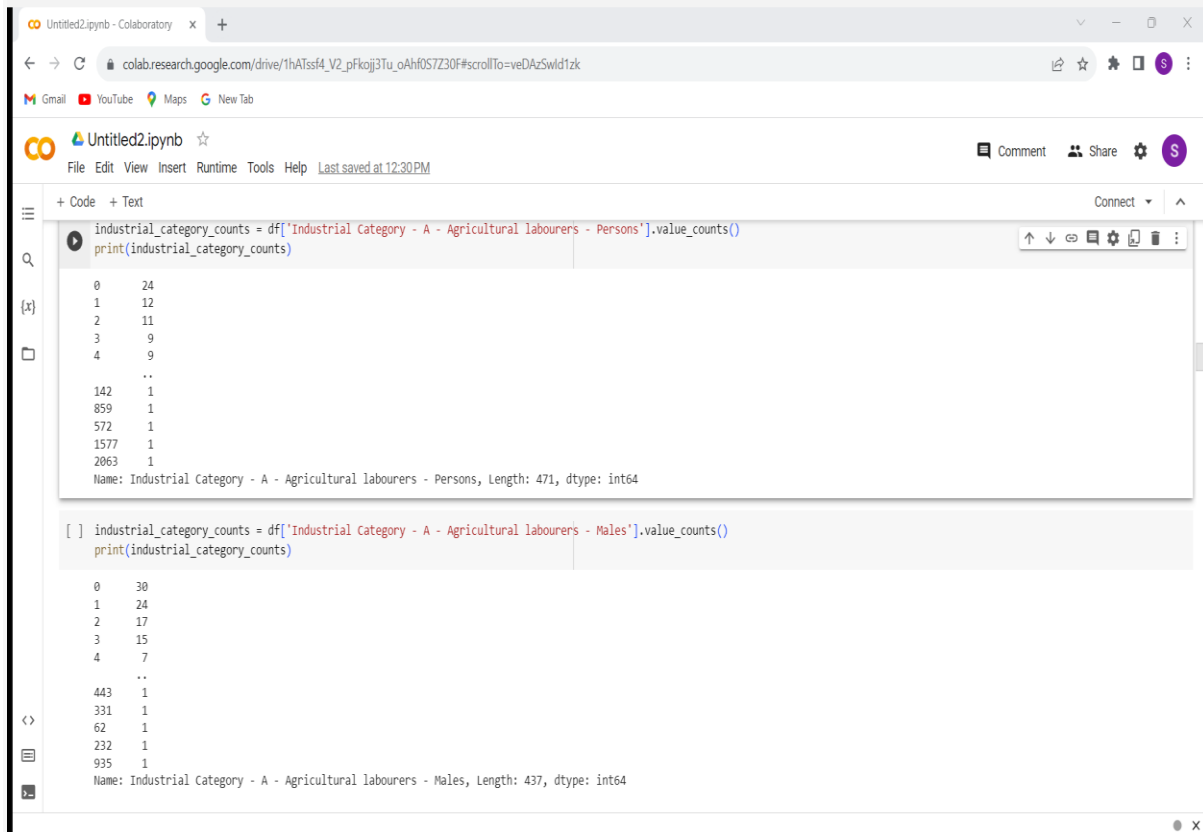
```
[25] industrial_category_counts = df['Industrial Category - A - Cultivators - Males'].value_counts()
      print(industrial_category_counts)

0      84
2      23
1      22
11     10
6      10
..
117     1
443     1
203     1
167     1
119     1
Name: Industrial Category - A - Cultivators - Males, Length: 273, dtype: int64
```

```
industrial_category_counts = df['Industrial Category - A - Cultivators - Females'].value_counts()
print(industrial_category_counts)

0      93
1      25
3      14
11     13
4      12
..
210     1
398     1
170     1
168     1
140     1
Name: Industrial Category - A - Cultivators - Females, Length: 266, dtype: int64
```

DEMOGRAPHIC ANALYSIS(AGRICULTUAL LABORS):



```
industrial_category_counts = df['Industrial Category - A - Agricultural labourers - Persons'].value_counts()
print(industrial_category_counts)

0      24
1      12
2      11
3       9
4       9
..
142     1
859     1
572     1
1577    1
2063    1
Name: Industrial Category - A - Agricultural labourers - Persons, Length: 471, dtype: int64
```

```
[ ] industrial_category_counts = df['Industrial Category - A - Agricultural labourers - Males'].value_counts()
    print(industrial_category_counts)

0      30
1      24
2      17
3      15
4       7
..
443     1
331     1
62      1
232     1
935     1
Name: Industrial Category - A - Agricultural labourers - Males, Length: 437, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY -A):

```
Untitled2.ipynb - Collaboratory
colab.research.google.com/drive/1hATss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=Kr7_uhkoZtyqd
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help Saving...
+ Code + Text
[ ] industrial_category_counts = df['Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Persons'].value_counts()
print(industrial_category_counts)
0      154
6       17
12       9
8         9
14         9
...
1224      1
2332      1
1586      1
870       1
35        1
Name: Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Persons, Length: 256, dtype: int64

industrial_category_counts = df['Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males'].value_counts()
print(industrial_category_counts)
0      168
6       22
12       14
18       11
8        10
...
49        1
102       1
254       1
132       1
84        1
Name: Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males, Length: 224, dtype: int64
```

```
Untitled2.ipynb - Collaboratory
colab.research.google.com/drive/1hATss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=bSDXWlQogBX2
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help
+ Code + Text
[ ] industrial_category_counts = df['Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males'].value_counts()
print(industrial_category_counts)
0      168
6       22
12       14
18       11
8        10
...
49        1
102       1
254       1
132       1
84        1
Name: Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Males, Length: 224, dtype: int64

industrial_category_counts = df['Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Females'].value_counts()
print(industrial_category_counts)
0      106
12       17
2         1
34       12
8         11
...
349       1
566       1
364       1
640       1
29        1
Name: Industrial Category - A - Plantation, Livestock, Forestry, Fishing, Hunting and allied activities - Females, Length: 196, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY B):

```
Untitled2.ipynb - Collaboratory
colab.research.google.com/drive/1hATss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=FLMkSbARgwtid
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help Saving...
+ Code + Text
[ ] industrial_category_counts = df['Industrial Category - B - Persons'].value_counts()
print(industrial_category_counts)
0      322
6       46
3       18
12      14
2       12
...
57        1
28        1
294       1
161       1
46        1
Name: Industrial Category - B - Persons, Length: 98, dtype: int64

industrial_category_counts = df['Industrial Category - B - Males'].value_counts()
print(industrial_category_counts)
0      344
6       54
12      23
3       18
9        9
...
16        1
7         1
36        1
130       1
31        1
Name: Industrial Category - B - Males, Length: 77, dtype: int64
```

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hAtssf4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=flMkSbARgwtD
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help
+ Code + Text
Industrial_category_counts = df['Industrial Category - B - Males'].value_counts()
print(industrial_category_counts)

0      344
6       54
12      23
3       18
9        9
...
16        1
7         1
35        1
130       1
31        1
Name: Industrial Category - B - Males, Length: 77, dtype: int64

[ ] Industrial_category_counts = df['Industrial Category - B - Females'].value_counts()
print(industrial_category_counts)

0      401
3       24
2       24
6       16
8       14
...
15        1
144       1
78        1
114       1
11        1
Name: Industrial Category - B - Females, Length: 65, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY C):

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hAtssf4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=8cdEN8mohueC
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help Saving...
+ Code + Text
Industrial_category_counts = df['Industrial Category - C - HHI - Persons'].value_counts()
print(industrial_category_counts)

0      157
6       15
2       11
3       10
12      10
...
169       1
232       1
79        1
47         1
171        1
Name: Industrial Category - C - HHI - Persons, Length: 236, dtype: int64

Industrial_category_counts = df['Industrial Category - C - HHI - Males'].value_counts()
print(industrial_category_counts)

0      196
6       33
12      20
18       12
8       10
...
263       1
7         1
732       1
588       1
166       1
Name: Industrial Category - C - HHI - Males, Length: 155, dtype: int64
```

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hAtssf4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=KMA-0oxlh49Z
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
Industrial_category_counts = df['Industrial Category - C - HHI - Males'].value_counts()
print(industrial_category_counts)

0      196
6       33
12      20
18       12
8       10
...
263       1
7         1
732       1
588       1
166       1
Name: Industrial Category - C - HHI - Males, Length: 155, dtype: int64

Industrial_category_counts = df['Industrial Category - C - HHI - Females'].value_counts()
print(industrial_category_counts)

0      177
6       21
14      14
2       13
8       12
...
230       1
426       1
176       1
470       1
83        1
Name: Industrial Category - C - HHI - Females, Length: 187, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY D&E):

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hATss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=DNCC4ykgtIN6
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
+ Code + Text
Connect
[ ] industrial_category_counts = df['Industrial Category - D & E - Persons'].value_counts()
print(industrial_category_counts)

0      300
6       40
2       23
12      21
14       17
...
105      1
67       1
58       1
91       1
19       1
Name: Industrial Category - D & E - Persons, Length: 89, dtype: int64

industrial_category_counts = df['Industrial Category - D & E - Males'].value_counts()
print(industrial_category_counts)

0      330
6       45
3       25
12      19
8       15
...
91       1
61       1
74       1
58       1
64       1
Name: Industrial Category - D & E - Males, Length: 73, dtype: int64
```

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hATss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=FQ1bl_D5i0WS
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
+ Code + Text
Connect
[ ] industrial_category_counts = df['Industrial Category - D & E - Females'].value_counts()
print(industrial_category_counts)

0      293
2       58
6       27
12      15
5       10
3       10
7        7
4        6
10       6
14       6
8        5
9        4
22       4
16       4
27       3
11       3
19       3
18       3
24       2
34       2
28       2
26       2
15       1
21       1
17       1
13       1
442      1
58       1
62       1
33       1
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY F):

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hATss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=238nIRxj/Wf
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
+ Code + Text
Connect
[ ] industrial_category_counts = df['Industrial Category - F - Persons'].value_counts()
print(industrial_category_counts)

0       99
6       18
2       16
8        8
12       8
..
1263     1
116      1
172      1
494      1
95       1
Name: Industrial Category - F - Persons, Length: 353, dtype: int64

industrial_category_counts = df['Industrial Category - F - Males'].value_counts()
print(industrial_category_counts)

0       112
6        27
2        14
12       11
20        9
...
358       1
1579      1
15        1
736       1
47        1
Name: Industrial Category - F - Males, Length: 318, dtype: int64
```



```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hATssf4_V2_pFkoj3Tu_oAh0S7Z30F#scrollTo=HlJgdU9ajOtm

Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
Industrial_category_counts = df['Industrial Category - F - Males'].value_counts()
print(Industrial_category_counts)

0      112
6       27
2       14
12      11
28       0
...
858      1
1579     1
15       1
736      1
47       1
Name: Industrial Category - F - Males, Length: 318, dtype: int64

Industrial_category_counts = df['Industrial Category - F - Females'].value_counts()
print(Industrial_category_counts)

0      146
2       19
6       14
4        9
7         0
...
264      1
186      1
210      1
1931     1
277      1
Name: Industrial Category - F - Females, Length: 288, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY H):

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hATssf4_V2_pFkoj3Tu_oAh0S7Z30F#scrollTo=1KIZkGsfjig-

Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text

industrial_category_counts = df['Industrial Category - H - Persons'].value_counts()
print(Industrial_category_counts)

0      199
6       41
3       17
12      10
2         9
...
221      1
188      1
126      1
113      1
116      1
Name: Industrial Category - H - Persons, Length: 217, dtype: int64

industrial_category_counts = df['Industrial Category - H - Males'].value_counts()
print(Industrial_category_counts)

0      204
6       47
3       17
12      13
18        7
...
300      1
124      1
136      1
131      1
29       1
Name: Industrial Category - H - Males, Length: 207, dtype: int64
```

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hATssf4_V2_pFkoj3Tu_oAh0S7Z30F#scrollTo=cqUNMDj8Vo

Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text

industrial_category_counts = df['Industrial Category - H - Females'].value_counts()
print(Industrial_category_counts)

0      375
6       41
2       41
8       18
16      13
12       9
18       9
3        8
4        7
5         5
10        5
14         5
11         4
9          4
24         4
17         3
110        2
46         2
13         2
7          2
28         2
22         2
41         2
58         2
27         1
40         1
19         1
20         1
32         1
25         1
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY I):

```
Untitled2.ipynb - Colaboratory
colabresearch.google.com/drive/1hAtssf4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=vjPiqjorkIoF

Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
[ ] industrial_category_counts = df['Industrial Category - I - Persons'].value_counts()
print(industrial_category_counts)

0      238
6       24
8       16
2       13
3       13
...
233      1
276      1
533      1
147      1
53       1
Name: Industrial Category - I - Persons, Length: 144, dtype: int64

industrial_category_counts = df['Industrial Category - I - Males'].value_counts()
print(industrial_category_counts)

0      256
6       42
12      19
3       16
18      12
...
71       1
157      1
91       1
191      1
7        1
Name: Industrial Category - I - Males, Length: 133, dtype: int64
```

```
Untitled2.ipynb - Colaboratory
colabresearch.google.com/drive/1hAtssf4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=vjPiqjorkIoF

Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
industrial_category_counts = df['Industrial Category - I - Males'].value_counts()
print(industrial_category_counts)

0      256
6       42
12      19
3       16
18      12
...
71       1
157      1
91       1
191      1
7        1
Name: Industrial Category - I - Males, Length: 133, dtype: int64

[ ] industrial_category_counts = df['Industrial Category - I - Females'].value_counts()
print(industrial_category_counts)

0      301
6       53
2       47
4       29
8       19
...
104      1
106      1
69       1
88       1
41       1
Name: Industrial Category - I - Females, Length: 64, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORK K&M):

```
Untitled2.ipynb - Colaboratory
colabresearch.google.com/drive/1hAtssf4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=ax1SoYf5koxp

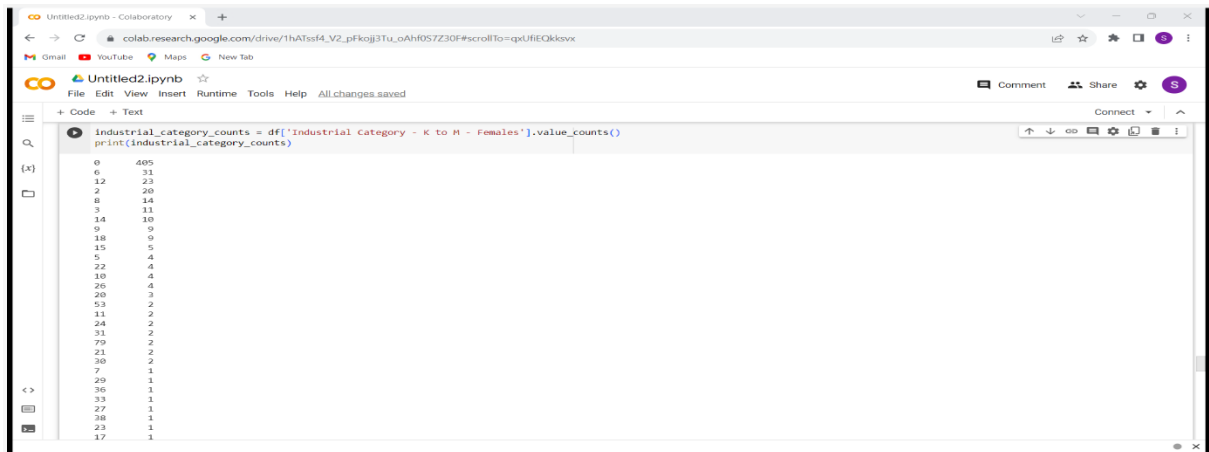
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
industrial_category_counts = df['Industrial Category - K to M - Persons'].value_counts()
print(industrial_category_counts)

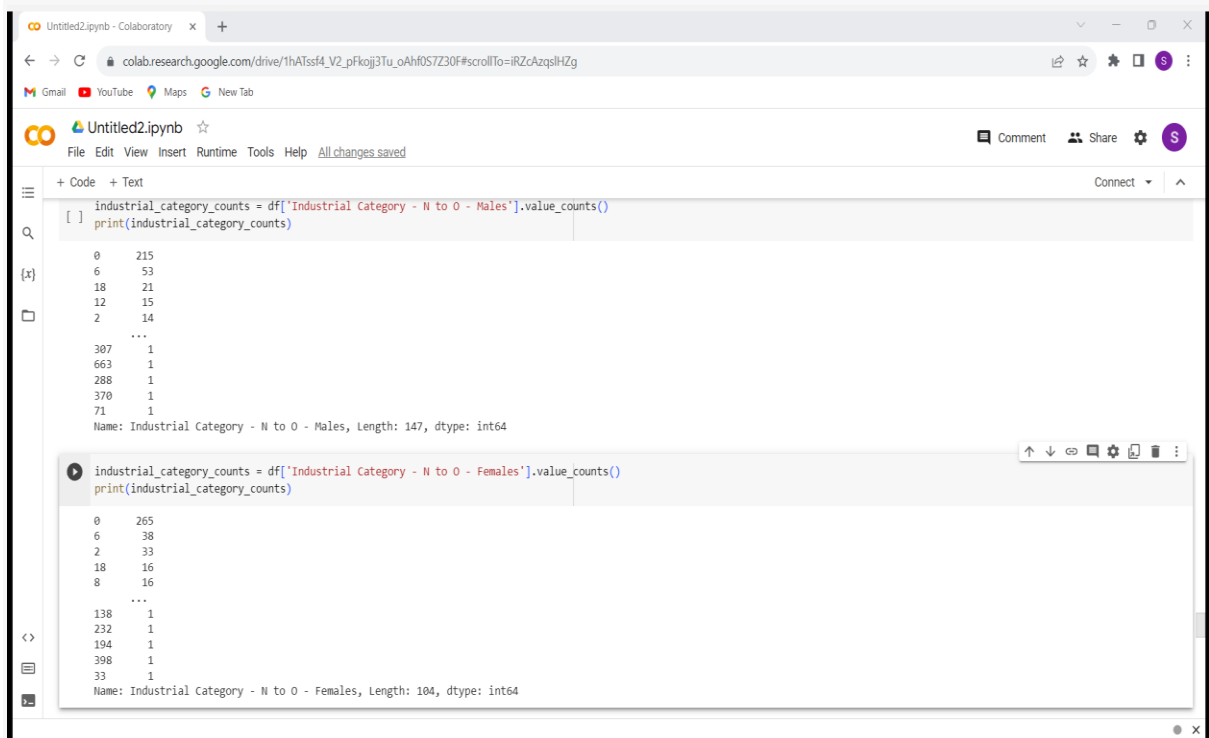
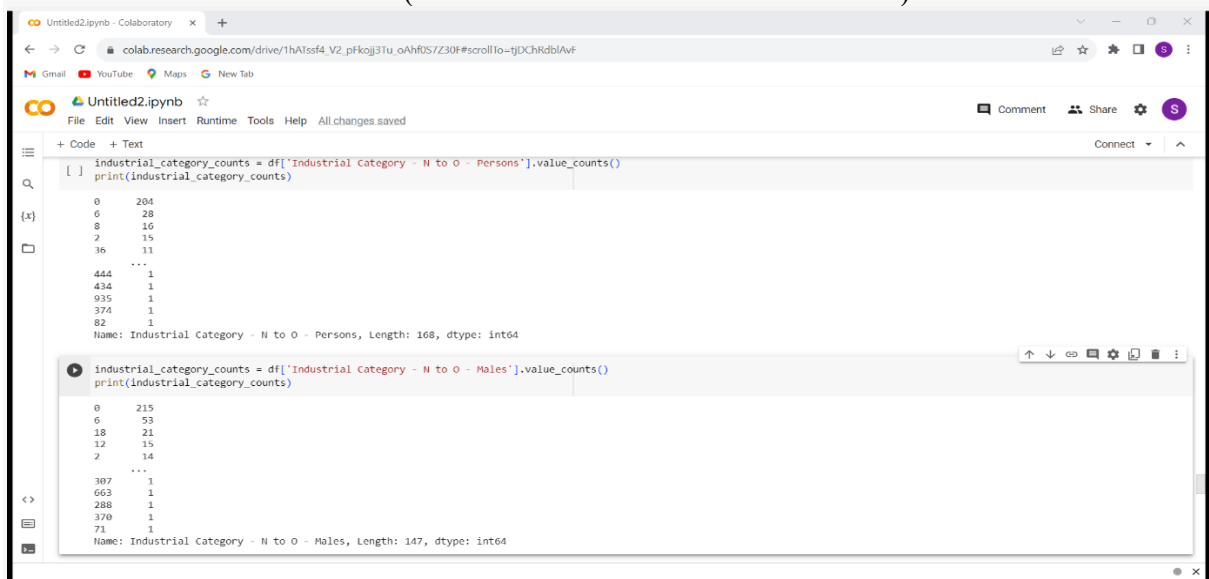
0      265
6       32
12      24
18      21
24      14
...
139      1
187      1
337      1
13       1
41       1
Name: Industrial Category - K to M - Persons, Length: 114, dtype: int64

industrial_category_counts = df['Industrial Category - K to M - Males'].value_counts()
print(industrial_category_counts)

0      268
6       39
12      26
18      24
3       16
...
80       1
187      1
88       1
118      1
23       1
Name: Industrial Category - K to M - Males, Length: 107, dtype: int64
```



DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY N TO O):



DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY P TO Q):

```
Untitled2.ipynb - Colaboratory
colabresearch.google.com/drive/1hAtss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=02ARUVZwVhE
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Connect
+ Code + Text
[ ] industrial_category_counts = df['Industrial Category - P to Q - Persons'].value_counts()
print(industrial_category_counts)

0      285
2       22
6       22
8       11
3       10
...
782      1
72       1
100      1
276      1
35       1
Name: Industrial Category - P to Q - Persons, length: 181, dtype: int64

industrial_category_counts = df['Industrial Category - P to Q - Males'].value_counts()
print(industrial_category_counts)

0      262
6       34
12      16
18      15
20      15
...
187      1
33       1
113      1
149      1
39       1
Name: Industrial Category - P to Q - Males, length: 111, dtype: int64
```

```
Untitled2.ipynb - Colaboratory
colabresearch.google.com/drive/1hAtss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=04WB9F1Z7Yoy
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Connect
+ Code + Text
[ ] industrial_category_counts = df['Industrial Category - P to Q - Males'].value_counts()
print(industrial_category_counts)

0      262
6       34
12      16
18      15
20      15
...
187      1
33       1
113      1
149      1
39       1
Name: Industrial Category - P to Q - Males, length: 111, dtype: int64

industrial_category_counts = df['Industrial Category - P to Q - Females'].value_counts()
print(industrial_category_counts)

0      233
2       25
6       21
18      12
3       10
...
332      1
458      1
250      1
246      1
23       1
Name: Industrial Category - P to Q - Females, length: 158, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY R TO U):

```
Untitled2.ipynb - Colaboratory
colabresearch.google.com/drive/1hAtss4_V2_pFkoj3Tu_oAhf0S7Z30F#scrollTo=_58qnhHulotD
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Connect
+ Code + Text
[ ] industrial_category_counts = df['Industrial Category - R to U - HHI - Persons'].value_counts()
print(industrial_category_counts)

0      150
6       21
8       15
14      11
26      11
...
340      1
746      1
185      1
480      1
154      1
Name: Industrial Category - R to U - HHI - Persons, length: 204, dtype: int64

industrial_category_counts = df['Industrial Category - R to U - HHI - Males'].value_counts()
print(industrial_category_counts)

0      214
6       55
12      27
14      15
18      15
...
131      1
293      1
98       1
154      1
73       1
Name: Industrial Category - R to U - HHI - Males, length: 120, dtype: int64
```

DEMOGRAPHIC ANALYSIS(INDUSTRIAL CATEGORY (R TO U NON HHI):

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hAtss4_V2_pFkoj3Tu_oAhR0S7Z30F#scrollTo=5_jidi1hsP2
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
+ Code + Text
Industrial_category_counts = df['Industrial Category - R to U - Non HHI - Persons'].value_counts()
print(Industrial_category_counts)

0      62
6      14
2       0
68       5
12       5
..
88       1
500      1
258      1
1154     1
81       1
Name: Industrial Category - R to U - Non HHI - Persons, Length: 397, dtype: int64

[ ] Industrial_category_counts = df['Industrial Category - R to U - Non HHI - Males'].value_counts()
print(Industrial_category_counts)

0      81
12     10
6      10
3       7
60      7
..
102     1
317     1
69      1
886     1
247     1
Name: Industrial Category - R to U - Non HHI - Males, Length: 314, dtype: int64
```

```
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1hAtss4_V2_pFkoj3Tu_oAhR0S7Z30F#scrollTo=dayUBrF6mVpU
Gmail YouTube Maps New Tab
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
+ Code + Text
Industrial_category_counts = df['Industrial Category - R to U - Non HHI - Males'].value_counts()
print(Industrial_category_counts)

0      81
12     10
6      10
3       7
60      7
..
102     1
317     1
69      1
886     1
247     1
Name: Industrial Category - R to U - Non HHI - Males, Length: 314, dtype: int64

[ ] Industrial_category_counts = df['Industrial Category - R to U - Non HHI - Females'].value_counts()
print(Industrial_category_counts)

0      71
6      12
2      11
32      8
30      8
..
99      1
39      1
962     1
548     1
774     1
Name: Industrial Category - R to U - Non HHI - Females, Length: 342, dtype: int64
```

CONCLUSION:

Thus the demographic analysis and create visualizations are created

The distribution of marginal workers based on age, industrial category, and sex are calculated using data aggregation and manipulation.