NAAN MUDHALVAN PROJECT PHASE 3 : DEVELOPMENT PART 1

PROJECT TITTLE: TN MARGINAL WORKERS - SOCIOECONOMIC ANALYSIS

DATA ANALYTICS

Import Libraries:

pythonCopy code

import pandas as pd

Load the Dataset:

If the dataset is in a CSV file:

pythonCopy code

df = pd.read_csv('your_dataset.csv')

If the dataset is in an Excel file:

pythonCopy code

df = pd.read_excel('your_dataset.xlsx')

For other formats, pandas provides similar functions (e.g., pd.read_sql() for databases, pd.read_json() for JSON files, etc.).

Explore the Dataset:

After loading, it's a good practice to take a look at the first few rows of the dataset using df.head() and check the data types and null values using df.info().

Handle Missing Data:

If there are missing values, you can choose to drop them or fill them with appropriate values using df.dropna() or df.fillna().

Data Cleaning and Transformation:

Perform any necessary data cleaning steps (e.g., removing duplicates, correcting data types, etc.) and transformations (e.g., feature scaling, encoding categorical variables, etc.).

Exploratory Data Analysis (EDA):

Conduct exploratory data analysis to understand the characteristics and relationships within the dataset. This may include summary statistics, visualizations, etc.

Feature Engineering:

If needed, create new features or modify existing ones to enhance the performance of your analysis or machine learning model.

Split Data (if applicable):

If you're doing predictive modeling, split the dataset into training and testing sets using techniques like cross-validation.

```
In [1]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
```

```
In [2]: main_dat = pd.read_csv("DDW_B06SC_3300_State_TAMIL_NADU-2011 (1).csv")
        ks = main_dat.copy()
```

```
In [3]: ks.head()
```

Out[3]:

	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	...	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	Inc Ca - F Fe
0	B0806SC	`33	`000	State - TAMIL NADU	Total	Total	1200828	589003	611825	221386	...	3565	11080	4019	7061	16833	4266	
1	B0806SC	`33	`000	State - TAMIL NADU	Total	`5-14	27791	14125	13666	2447	...	11	122	71	51	427	169	
2	B0806SC	`33	`000	State - TAMIL NADU	Total	15-34	514340	259560	254780	92423	...	1754	7536	2718	4818	8346	2127	
3	B0806SC	`33	`000	State - TAMIL NADU	Total	35-59	542581	251957	290624	99202	...	1619	3205	1131	2074	6591	1487	
4	B0806SC	`33	`000	State - TAMIL NADU	Total	60+	115103	62833	52270	27165	...	175	211	93	118	1457	483	

5 rows × 69 columns

5 rows × 69 columns

```
In [4]: ks.sample(5)
```

Out[4]:

	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	...	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	I (
426	B0806SC	`33	`624	District - Theni	Urban	Total	5214	3003	2211	1151	...	18	62	12	50	38	
252	B0806SC	`33	`615	District - Perambalur	Total	Total	16797	7835	8962	1741	...	12	140	36	104	166	
575	B0806SC	`33	`632	District - Coimbatore	Urban	Age not stated	13	10	3	0	...	0	0	0	0	0	
336	B0806SC	`33	`619	District - Thiruvarur	Urban	Total	4807	2585	2222	788	...	0	20	0	20	34	
465	B0806SC	`33	`626	District - Ramanathapuram	Urban	35-59	923	440	483	253	...	0	26	0	26	18	

5 rows × 69 columns

```
In [5]: ks.shape
```

Out[5]: (594, 69)

In [6]:
```
ks.columns
```

Out[6]:
```
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 - 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')
```

In [7]:
```
pd.isnull(ks).sum()
```

Out[7]:
```
Table Code                                        0
State Code                                        0
District Code                                     0
Area Name                                         0
Total/ Rural/ Urban                               0
                                                 ..
Industrial Category - R to U - HHI - Males        0
Industrial Category - R to U - HHI - Females      0
Industrial Category - R to U - Non HHI - Persons  0
Industrial Category - R to U - Non HHI - Males    0
Industrial Category - R to U - Non HHI - Females  0
Length: 69, dtype: int64
```

In [8]:
```
ks.dropna(inplace=True)
pd.isnull(ks).sum()
```

Out[8]:
```
Table Code                                        0
State Code                                        0
District Code                                     0
Area Name                                         0
Total/ Rural/ Urban                               0
                                                 ..
Industrial Category - R to U - HHI - Males        0
Industrial Category - R to U - HHI - Females      0
Industrial Category - R to U - Non HHI - Persons  0
Industrial Category - R to U - Non HHI - Males    0
Industrial Category - R to U - Non HHI - Females  0
Length: 69, dtype: int64
```

In [8]:
```
ks.dropna(inplace=True)
pd.isnull(ks).sum()
```

Out[8]:
```
Table Code                                        0
State Code                                        0
District Code                                     0
Area Name                                         0
Total/ Rural/ Urban                               0
                                                 ..
Industrial Category - R to U - HHI - Males        0
Industrial Category - R to U - HHI - Females      0
Industrial Category - R to U - Non HHI - Persons  0
Industrial Category - R to U - Non HHI - Males    0
Industrial Category - R to U - Non HHI - Females  0
Length: 69, dtype: int64
```

In [9]:
```
ks.describe()
```

Out[9]:

	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	...
count	5.940000e+02	594.000000	594.000000	594.000000	594.000000	594.000000	594.000000	594.000000	594.000000	594.000000	...
mean	1.617277e+04	7932.700337	8240.067340	2981.629630	1338.289562	1643.340067	865.117845	466.424242	398.693603	12225.616162	...
std	7.607172e+04	36864.822704	39259.545337	13909.621137	6127.047670	7808.832522	4274.458077	2298.072295	1978.682322	60458.382586	...
min	0.000000e+00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...
25%	2.872500e+02	147.250000	144.000000	27.000000	14.250000	13.000000	9.000000	5.000000	4.000000	79.250000	...
50%	2.225500e+03	1147.000000	1076.000000	430.000000	198.500000	213.000000	69.500000	35.500000	32.000000	1094.000000	...
75%	9.628500e+03	4770.500000	4887.500000	1775.250000	774.250000	946.500000	466.000000	244.250000	204.750000	6279.750000	...
max	1.200828e+06	589003.000000	611825.000000	221386.000000	99368.000000	122018.000000	64235.000000	34632.000000	29603.000000	907752.000000	...

8 rows × 63 columns

In [10]:
```
ks.nunique()
```

Out[10]:
```
Table Code                                          1
State Code                                          1
District Code                                      33
Area Name                                          33
Total/ Rural/ Urban                                 3
                                                 ...
Industrial Category - R to U - HHI - Males        120
Industrial Category - R to U - HHI - Females      187
Industrial Category - R to U - Non HHI - Persons  397
Industrial Category - R to U - Non HHI - Males    314
Industrial Category - R to U - Non HHI - Females  342
Length: 69, dtype: int64
```

In [11]:
```
ks.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 594 entries, 0 to 593
Data columns (total 69 columns):
 #   Column                                                                   Non-Null Co
unt  Dtype
---  ------                                                                   -----------
---  -----
 0   Table Code                                                               594 non-nul
l    object
 1   State Code                                                               594 non-nul
l    object
 2   District Code                                                            594 non-nul
l    object
 3   Area Name                                                                594 non-nul
l    object
 4   Total/ Rural/ Urban                                                      594 non-nul
l    object
 5   Age group                                                                594 non-nul
l    object
```

Downloads/ | Marginal_workers - Jupyter Note | www.google.com

localhost:8888/notebooks/Downloads/Marginal_workers.ipynb

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

jupyter Marginal_workers (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

In [12]:
```
ks.dtypes
```

Out[12]:
```
Table Code                                        object
State Code                                        object
District Code                                     object
Area Name                                         object
Total/ Rural/ Urban                               object
                                                   ...
Industrial Category - R to U - HHI - Males         int64
Industrial Category - R to U - HHI - Females       int64
Industrial Category - R to U - Non HHI - Persons   int64
Industrial Category - R to U - Non HHI - Males     int64
Industrial Category - R to U - Non HHI - Females   int64
Length: 69, dtype: object
```

In [19]:
```
ks = ks.drop(['Table Code'], axis=1)
```

In [20]:
```
ks = ks.drop(['State Code'], axis=1)
```

In [21]:
```
ks
```

Out[21]:

	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 - 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	Industria Categor - R to U HHI Male
0	`000	State - TAMIL NADU	Total	Total	1200828	589003	611825	221386	99368	122018	...	3565	11080	4019	7061	16833	426
1	`000	State - TAMIL NADU	Total	`5-14	27791	14125	13666	2447	1247	1200	...	11	122	71	51	427	16
2	`000	State - TAMIL NADU	Total	15-34	514340	259560	254780	92423	43892	48531	...	1754	7536	2718	4818	8346	212
3	`000	State - TAMIL NADU	Total	35-59	542581	251957	290624	99202	40691	58511	...	1619	3205	1131	2074	6591	148

```
In [12]: ks.dtypes
```

```
Out[12]: Table Code                                    object
         State Code                                    object
         District Code                                 object
         Area Name                                     object
         Total/ Rural/ Urban                           object
                                                        ...
         Industrial Category - R to U - HHI - Males     int64
         Industrial Category - R to U - HHI - Females   int64
         Industrial Category - R to U - Non HHI - Persons  int64
         Industrial Category - R to U - Non HHI - Males    int64
         Industrial Category - R to U - Non HHI - Females  int64
         Length: 69, dtype: object
```

```
In [19]: ks = ks.drop(['Table Code'], axis=1)
```

```
In [20]: ks = ks.drop(['State Code'], axis=1)
```

```
In [21]: ks
```

```
Out[21]:
```

	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 - 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
0	`000	State - TAMIL NADU	Total	Total	1200828	589003	611825	221386	99368	122018	...	3565	11080	4019	7061	16833	4268
1	`000	State - TAMIL NADU	Total	`5-14	27791	14125	13666	2447	1247	1200	...	11	122	71	51	427	16
2	`000	State - TAMIL NADU	Total	15-34	514340	259560	254780	92423	43892	48531	...	1754	7536	2718	4818	8346	2124
3	`000	State - TAMIL NADU	Total	35-59	542581	251957	290624	99202	40691	58511	...	1619	3205	1131	2074	6591	1485
4	`000	State - TAMIL NADU	Total	60+	115803	62833	52970	27185	13455	13730	...	179	211	93	118	1457	488

```
In [27]: ks
```

```
Out[27]:
```

	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 - 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
0	000	State - TAMIL NADU	Total	Total	1200828	589003	611825	221386	99368	122018	3565	11080	4019	7061	16833	4268
1	000	State - TAMIL NADU	Total	`5-14	27791	14125	13666	2447	1247	1200	11	122	71	51	427	16
2	000	State - TAMIL NADU	Total	15-34	514340	259560	254780	92423	43892	48531	1754	7536	2718	4818	8346	2124
3	000	State - TAMIL NADU	Total	35-59	542581	251957	290624	99202	40691	58511	1619	3205	1131	2074	6591	1485
4	000	State - TAMIL NADU	Total	60+	115803	62833	52970	27185	13455	13730	179	211	93	118	1457	488
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
980	633	District - Tiruppur	Urban	`5-14	272	120	143	18	6	12	0	0	0	0	0	0
990	633	District - Tiruppur	Urban	15-34	3281	1654	1631	473	238	235	20	44	15	29	62	18
991	633	District - Tiruppur	Urban	35-59	3672	1769	1903	522	247	275	33	36	12	23	36	10
992	633	District - Tiruppur	Urban	60+	690	390	297	111	50	61	0	3	0	3	10	3
993	633	District - Tiruppur	Urban	Age not stated	2	1	1	0	0	0	0	0	0	0	0	0

994 rows × 67 columns

```
In [10]: ks = ks.drop(['Table Code'], axis=1)
```

```
In [11]: ks = ks.drop(['State Code'], axis=1)
```

```
In [12]: ks
```

```
Out[12]:
```

	District Code	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 - 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
0	000	1	Total	1200828	589003	611825	221386	99368	122018	64239	3565	11080	4019	7061	16833	4268
1	000	1	`5-14	27791	14125	13666	2447	1247	1200	1710	11	122	71	51	427	16
2	000	1	15-34	514340	259560	254780	92423	43892	48531	24563	1754	7536	2718	4818	8346	2124
3	000	1	35-59	542581	251957	290624	99202	40691	58511	29092	1619	3205	1131	2074	6591	1485
4	000	1	60+	115803	62833	52970	27185	13455	13730	7230	179	211	93	118	1457	488
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
989	633	0	`5-14	272	120	143	18	6	12	0	0	0	0	0	0	0
990	633	0	15-34	3281	1654	1631	473	238	235	22	20	44	15	29	62	18
991	633	0	35-59	3672	1769	1903	522	247	275	25	33	36	12	23	36	10
992	633	0	60+	690	390	297	111	50	61	0	0	3	0	3	10	3
993	633	0	Age not stated	2	1	1	0	0	0	0	0	0	0	0	0	0

994 rows × 66 columns

```
In [11]: ks['Age group'] = ks['Age group'].replace(['Total','`5-14','15-34','35-59','60+','Age not stated'],['0','1','2','3','4','5'])
```

```
In [10]: ks
```

```
In [29]: ks
```

```
Out[29]:
```

	District Code	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 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 - 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
0	000	total	Total	1200828	589003	221386	99368	122018	64239	...	3565	11080	4019	7061	16833	4268
1	000	total	`5-14	27791	14125	2447	1247	1200	1710	...	11	122	71	51	427	16
2	000	total	15-34	514340	259560	92423	43892	48531	24563	...	1754	7536	2718	4818	8346	2124
3	000	total	35-59	542581	251957	99202	40691	58511	29092	...	1619	3205	1131	2074	6591	1485
4	000	total	60+	115803	62833	27185	13455	13730	7230	...	179	211	93	118	1457	488
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
989	633	Urban	`5-14	272	120	18	6	12	0	...	0	0	0	0	0	0
990	633	Urban	15-34	3281	1654	473	238	235	22	...	20	44	15	29	62	18
991	633	Urban	35-59	3672	1769	522	247	275	25	...	33	36	12	23	36	10
992	633	Urban	60+	690	390	111	50	61	0	...	0	3	0	3	10	3
993	633	Urban	Age not stated	2	1	0	0	0	0	...	0	0	0	0	0	0

994 rows × 66 columns

```
In [13]: ks['Total/ Rural/ Urban'] = ks['Total/ Rural/ Urban'].replace(['Urban','Total','Rural'],['0','1','2'])
```

```
In [14]: ks
```

994 rows × 66 columns

```
In [71]: ks.corr()
```

```
In [35]: ks.corr()
```

```
Out[35]:
```

	District Code	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 - N to O - Females	Industrial Category - P to Q - Persons	Industrial Category - P to Q - Males
District Code	1.000000e+00	1.703029e-16	1.561161e-16	-0.584309	-0.591554	-0.576722	-0.589599	-0.601056	-0.578626	-0.555306	...	-0.596740	-0.589169	-0.58
Total/ Rural/ Urban	1.703029e-16	1.000000e+00	5.361496e-19	0.079478	0.073996	0.084519	0.075662	0.064609	0.084081	0.106286	...	-0.017692	0.014386	0.02
Age group	1.561161e-16	5.361496e-19	1.000000e+00	-0.148080	-0.150075	-0.146007	-0.145439	-0.149113	-0.142067	-0.137326	...	-0.154607	-0.168363	-0.16
Worked for 3 months or more but less than 6 months - Persons	-5.843088e-01	7.947800e-02	-1.480796e-01	1.000000	0.999263	0.999351	0.998865	0.994883	0.998631	0.987308	...	0.886220	0.923176	0.93
Worked for 3 months or more but less than 6 months - Males	-5.915544e-01	7.399611e-02	-1.500753e-01	0.999263	1.000000	0.997232	0.999020	0.997622	0.996757	0.982657	...	0.897920	0.934548	0.94
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Industrial Category - R to U - HHI - Males	-5.967098e-01	6.429709e-02	-1.564372e-01	0.987514	0.991627	0.982327	0.988538	0.993430	0.981375	0.963460	...	0.910143	0.956272	0.96
Industrial Category - R to U - HHI - Females	-6.003483e-01	4.692903e-02	-1.569529e-01	0.983226	0.987635	0.977768	0.985253	0.991781	0.976818	0.952677	...	0.940298	0.965278	0.96
Industrial Category														

Out[41]:
```
Age group
0     99
1     99
2     99
3     99
4     99
5     99
Name: count, dtype: int64
```

In [43]:
```
ks
```

Out[43]:

	District Code	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 - 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	Indus Cate - R t I M
0	000	1	0	1200828	589003	611825	221386	99368	122018	64235	...	3565	11080	4019	7061	16833	
1	000	1	1	27791	14125	13666	2447	1247	1200	1710	...	11	122	71	51	427	
2	000	1	2	514340	259560	254780	92423	43892	48531	24863	...	1754	7536	2718	4818	8346	
3	000	1	3	542581	251957	290624	99202	40691	58511	29692	...	1619	3205	1131	2074	6591	
4	000	1	4	115103	62833	52270	27165	13465	13700	7930	...	175	211	93	118	1457	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
589	633	0	1	272	129	143	18	6	12	6	...	0	0	0	0	0	
590	633	0	2	3285	1654	1631	473	238	235	22	...	20	44	15	29	62	
591	633	0	3	3672	1769	1903	522	247	275	25	...	33	35	12	23	36	
592	633	0	4	696	399	297	111	50	61	5	...	0	3	0	3	10	
593	633	0	5	2	1	1	0	0	0	0	...	0	0	0	0	0	

594 rows × 66 columns

In [44]:
```
ks.to_csv('cleaned_data.csv', index=False)
```