320-govt-upd-1-1

March 20, 2025

1 Lagishetty Akhila

2 2211cs010320

3 group 4

4 Indra Gandhi Widow Pension Scheme FY 2023-2024

The **Indira Gandhi Widow Pension Scheme (2023-2024) ensures structured tracking of welfare benefits across regions and demographics. LGD codes uniquely identify states and districts, while scheme codes categorize programs. This enhances transparency and accountability, ensuring funds reach the right beneficiaries. Aadhaar and mobile numbers aid in DBT validation, preventing fraud. Demographic analysis (SC, ST, OBC, General) ensures fair distribution.Time-based tracking (financial year, month, last update) allows trend analysis and irregularity detection. Covering multiple districts in Jammu & Kashmir the dataset provides a comprehensive and structured approachto welfare monitoring.

[2]:
```
import pandas as pd
import numpy as np
import seaborn as sb
import matplotlib.pyplot as plt
import math
df=pd.read_csv("govt_dataset.csv");
df
```

[2]:

	fin_year	mnth	lgd_state_code	state_name	lgd_district_code	\
0	2023-2024	2	1	JAMMU AND KASHMIR	13	
1	2023-2024	2	1	JAMMU AND KASHMIR	14	
2	2023-2024	2	1	JAMMU AND KASHMIR	2	
3	2023-2024	2	1	JAMMU AND KASHMIR	4	
4	2023-2024	2	1	JAMMU AND KASHMIR	622	
…	…	…	…	…	…	
14377	2024-2025	2	9	UTTAR PRADESH	172	
14378	2024-2025	2	9	UTTAR PRADESH	173	
14379	2024-2025	2	9	UTTAR PRADESH	176	
14380	2024-2025	2	9	UTTAR PRADESH	177	
14381	2024-2025	2	9	UTTAR PRADESH	181	

```
       district_name scheme_code  total_beneficiaries  total_sc  total_st  \
0           SRINAGAR      IGNWPS                  546         0         1
1           UDHAMPUR      IGNWPS                  353        85        17
2             BADGAM      IGNWPS                  176         3         1
3               DODA      IGNWPS                  246         9        18
4             KULGAM      IGNWPS                  290        20         0
…                …           …                    …         …         …
14377  MUZAFFARNAGAR      IGNWPS                 7601      1817        13
14378       PILIBHIT      IGNWPS                 8268      1796         0
14379         RAMPUR      IGNWPS                23245      3169         0
14380     SAHARANPUR      IGNWPS                10667      3649       143
14381      SHRAVASTI      IGNWPS                 6168      1689       332

       total_gen  total_obc  total_aadhar  total_mobileno lastUpdated
0            544          1           403             127  2024-07-31
1            238         13           250             121  2024-07-31
2            170          2           150              25  2024-07-31
3            219          0           237              33  2024-07-31
4            257         13           290               0  2024-07-31
…            …           …             …               …
14377       2719       3052          3680            3177  2025-03-07
14378       1756       4716          4037            4306  2025-03-07
14379       7868      12208          9288            9096  2025-03-07
14380       2940       3935          5896            5587  2025-03-07
14381       2203       1944          1649            2063  2025-03-07

[14382 rows x 15 columns]
```

[ ]:

5 This loads necessary libraries and reads “govt_dataset.csv” into a DataFrame for analysis.

[3]:
```
df.head(10)
```

[3]:
```
    fin_year  mnth  lgd_state_code        state_name  lgd_district_code  \
0  2023-2024     2               1  JAMMU AND KASHMIR                 13
1  2023-2024     2               1  JAMMU AND KASHMIR                 14
2  2023-2024     2               1  JAMMU AND KASHMIR                  2
3  2023-2024     2               1  JAMMU AND KASHMIR                  4
4  2023-2024     2               1  JAMMU AND KASHMIR                622
5  2023-2024     2               1  JAMMU AND KASHMIR                624
6  2023-2024     2               1  JAMMU AND KASHMIR                625
7  2023-2024     2              10              BIHAR                196
8  2023-2024     2              10              BIHAR                202
```

```
9  2023-2024     2              10               BIHAR                204

  district_name scheme_code  total_beneficiaries  total_sc  total_st  \
0      SRINAGAR      IGNWPS                  546         0         1
1      UDHAMPUR      IGNWPS                  353        85        17
2        BADGAM      IGNWPS                  176         3         1
3          DODA      IGNWPS                  246         9        18
4        KULGAM      IGNWPS                  290        20         0
5         SAMBA      IGNWPS                  117        14         4
6       SHOPIAN      IGNWPS                  454         0        47
7          GAYA      IGNWPS                15236      6590        87
8      KHAGARIA      IGNWPS                15163      1623       271
9    LAKHISARAI      IGNWPS                 4116       684        34

   total_gen  total_obc  total_aadhar  total_mobileno lastUpdated
0        544          1           403             127  2024-07-31
1        238         13           250             121  2024-07-31
2        170          2           150              25  2024-07-31
3        219          0           237              33  2024-07-31
4        257         13           290               0  2024-07-31
5         98          1            83               2  2024-07-31
6        407          0           454             315  2024-07-31
7       7644        915         11236           11618  2024-07-31
8      11498       1771         10253           12412  2024-07-31
9       1588       1810          2883            3181  2024-07-31
```

Displays the first 10 rows of the DataFrame to preview the data.

[5]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
,RangeIndex: 14382 entries, 0 to 14381
,Data columns (total 15 columns):
, #   Column               Non-Null Count  Dtype
,---  ------               --------------  -----
, 0   fin_year             14382 non-null  object
, 1   mnth                 14382 non-null  int64
, 2   lgd_state_code       14382 non-null  int64
, 3   state_name           14382 non-null  object
, 4   lgd_district_code    14382 non-null  int64
, 5   district_name        14382 non-null  object
, 6   scheme_code          14382 non-null  object
, 7   total_beneficiaries  14382 non-null  int64
, 8   total_sc             14382 non-null  int64
, 9   total_st             14382 non-null  int64
, 10  total_gen            14382 non-null  int64
, 11  total_obc            14382 non-null  int64
, 12  total_aadhar         14382 non-null  int64
```

```
, 13  total_mobileno       14382 non-null  int64
, 14  lastUpdated          14382 non-null  object
,dtypes: int64(10), object(5)
,memory usage: 1.6+ MB
```

Shows the DataFrame's structure, including column names, data types, and non-null counts.

[7]:
```
print(df.isnull().sum())
```

```
fin_year               0
,mnth                  0
,lgd_state_code        0
,state_name            0
,lgd_district_code     0
,district_name         0
,scheme_code           0
,total_beneficiaries   0
,total_sc              0
,total_st              0
,total_gen             0
,total_obc             0
,total_aadhar          0
,total_mobileno        0
,lastUpdated           0
,dtype: int64
```

Displays the total number of missing (null) values in each column of the DataFrame.

[9]:
```
df['state_name'].unique()
```

[9]:
```
array(['JAMMU AND KASHMIR', 'BIHAR', 'SIKKIM', 'ARUNACHAL PRADESH',
       'NAGALAND', 'MANIPUR', 'MIZORAM', 'TRIPURA', 'MEGHALAYA', 'ASSAM',
       'WEST BENGAL', 'HIMACHAL PRADESH', 'JHARKHAND', 'ODISHA',
       'CHHATTISGARH', 'MADHYA PRADESH', 'GUJARAT', 'MAHARASHTRA',
       'ANDHRA PRADESH', 'KARNATAKA', 'PUNJAB', 'GOA', 'KERALA',
       'TAMIL NADU', 'TELANGANA', 'LADAKH',
       'THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU', 'UTTARAKHAND',
       'HARYANA', 'LAKSHADWEEP', 'CHANDIGARH', 'DELHI', 'RAJASTHAN',
       'UTTAR PRADESH', 'PUDUCHERRY', 'ANDAMAN AND NICOBAR'], dtype=object)
```

Lists all unique state names present in the "state_name" column.

[11]:
```
df['state_name'].value_counts()
```

[11]:
```
state_name
UTTAR PRADESH                                1360
MADHYA PRADESH                               1020
BIHAR                                         760
MAHARASHTRA                                   680
```

```
GUJARAT                                          660
CHHATTISGARH                                     660
ASSAM                                            660
TELANGANA                                        660
RAJASTHAN                                        660
TAMIL NADU                                       640
KARNATAKA                                        600
ODISHA                                           600
ARUNACHAL PRADESH                                500
JHARKHAND                                        480
WEST BENGAL                                      460
PUNJAB                                           440
HARYANA                                          440
ANDHRA PRADESH                                   430
JAMMU AND KASHMIR                                400
UTTARAKHAND                                      260
KERALA                                           252
HIMACHAL PRADESH                                 240
MIZORAM                                          220
MEGHALAYA                                        220
NAGALAND                                         220
DELHI                                            200
MANIPUR                                          180
TRIPURA                                          160
SIKKIM                                            80
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU      60
GOA                                               40
LADAKH                                            40
ANDAMAN AND NICOBAR                               40
LAKSHADWEEP                                       20
CHANDIGARH                                        20
PUDUCHERRY                                        20
Name: count, dtype: int64
```

Shows the count of occurrences for each unique state in the "state_name" column.

[13]:
```
union_territories = ['JAMMU AND KASHMIR', 'DELHI', 'PUDUCHERRY', 'LADAKH',␣
↪'CHANDIGARH', 'ANDAMAN AND NICOBAR ISLANDS']
union_territories
```

[13]:
```
['JAMMU AND KASHMIR',
 'DELHI',
 'PUDUCHERRY',
 'LADAKH',
 'CHANDIGARH',
 'ANDAMAN AND NICOBAR ISLANDS']
```

Creates a list of union territories for later filtering or classification.

[15]:
```
df['region_type'] = df['state_name'].apply(lambda x: 'Union Territory' if x in␣
↪union_territories else 'State')
```

Adds a new column “region_type” labeling each row as either “State” or “Union Territory”.

[17]:
```
df.head
```

[17]:
```
<bound method NDFrame.head of         fin_year  mnth  lgd_state_code
state_name  lgd_district_code  \
0       2023-2024     2               1  JAMMU AND KASHMIR                 13
1       2023-2024     2               1  JAMMU AND KASHMIR                 14
2       2023-2024     2               1  JAMMU AND KASHMIR                  2
3       2023-2024     2               1  JAMMU AND KASHMIR                  4
4       2023-2024     2               1  JAMMU AND KASHMIR                622
…             …   …             …                 …                  …
14377   2024-2025     2               9      UTTAR PRADESH                172
14378   2024-2025     2               9      UTTAR PRADESH                173
14379   2024-2025     2               9      UTTAR PRADESH                176
14380   2024-2025     2               9      UTTAR PRADESH                177
14381   2024-2025     2               9      UTTAR PRADESH                181

       district_name scheme_code  total_beneficiaries  total_sc  total_st  \
0           SRINAGAR      IGNWPS                  546         0         1
1           UDHAMPUR      IGNWPS                  353        85        17
2             BADGAM      IGNWPS                  176         3         1
3               DODA      IGNWPS                  246         9        18
4             KULGAM      IGNWPS                  290        20         0
…                …          …                    …        …        …
14377  MUZAFFARNAGAR      IGNWPS                 7601      1817        13
14378       PILIBHIT      IGNWPS                 8268      1796         0
14379         RAMPUR      IGNWPS                23245      3169         0
14380     SAHARANPUR      IGNWPS                10667      3649       143
14381      SHRAVASTI      IGNWPS                 6168      1689       332

       total_gen  total_obc  total_aadhar  total_mobileno lastUpdated  \
0            544          1           403             127  2024-07-31
1            238         13           250             121  2024-07-31
2            170          2           150              25  2024-07-31
3            219          0           237              33  2024-07-31
4            257         13           290               0  2024-07-31
…          …        …             …                …          …
14377       2719       3052          3680            3177  2025-03-07
14378       1756       4716          4037            4306  2025-03-07
14379       7868      12208          9288            9096  2025-03-07
14380       2940       3935          5896            5587  2025-03-07
14381       2203       1944          1649            2063  2025-03-07
```

```
          region_type
0      Union Territory
1      Union Territory
2      Union Territory
3      Union Territory
4      Union Territory
…                    …
14377            State
14378            State
14379            State
14380            State
14381            State

[14382 rows x 16 columns]>
```

[18]:
```
df
```

[18]:
```
        fin_year  mnth  lgd_state_code        state_name  lgd_district_code  \
0      2023-2024     2               1  JAMMU AND KASHMIR                 13
1      2023-2024     2               1  JAMMU AND KASHMIR                 14
2      2023-2024     2               1  JAMMU AND KASHMIR                  2
3      2023-2024     2               1  JAMMU AND KASHMIR                  4
4      2023-2024     2               1  JAMMU AND KASHMIR                622
…            …     …               …                 …                  …
14377  2024-2025     2               9      UTTAR PRADESH                172
14378  2024-2025     2               9      UTTAR PRADESH                173
14379  2024-2025     2               9      UTTAR PRADESH                176
14380  2024-2025     2               9      UTTAR PRADESH                177
14381  2024-2025     2               9      UTTAR PRADESH                181

      district_name scheme_code  total_beneficiaries  total_sc  total_st  \
0          SRINAGAR      IGNWPS                  546         0         1
1          UDHAMPUR      IGNWPS                  353        85        17
2            BADGAM      IGNWPS                  176         3         1
3              DODA      IGNWPS                  246         9        18
4            KULGAM      IGNWPS                  290        20         0
…               …          …                    …       …        …
14377  MUZAFFARNAGAR      IGNWPS                 7601      1817        13
14378      PILIBHIT      IGNWPS                 8268      1796         0
14379        RAMPUR      IGNWPS                23245      3169         0
14380    SAHARANPUR      IGNWPS                10667      3649       143
14381     SHRAVASTI      IGNWPS                 6168      1689       332

      total_gen  total_obc  total_aadhar  total_mobileno lastUpdated  \
0           544          1           403             127  2024-07-31
1           238         13           250             121  2024-07-31
2           170          2           150              25  2024-07-31
```

```
3            219          0           237              33  2024-07-31
4            257         13           290               0  2024-07-31
…            …           …            …                …           …
14377       2719       3052          3680            3177  2025-03-07
14378       1756       4716          4037            4306  2025-03-07
14379       7868      12208          9288            9096  2025-03-07
14380       2940       3935          5896            5587  2025-03-07
14381       2203       1944          1649            2063  2025-03-07

          region_type
0      Union Territory
1      Union Territory
2      Union Territory
3      Union Territory
4      Union Territory
…                  …
14377            State
14378            State
14379            State
14380            State
14381            State

[14382 rows x 16 columns]
```

[19]:
```
df.groupby('state_name')['total_beneficiaries'].sum()
```

[19]:
```
state_name
ANDAMAN AND NICOBAR                    7970
ANDHRA PRADESH                     11729503
ARUNACHAL PRADESH                    120200
ASSAM                               2290292
BIHAR                              12684420
CHANDIGARH                            49715
CHHATTISGARH                        4318904
DELHI                                729357
GOA                                  117940
GUJARAT                             7762285
HARYANA                             2657668
HIMACHAL PRADESH                     439480
JAMMU AND KASHMIR                    138212
JHARKHAND                           5084571
KARNATAKA                           8196100
KERALA                              7940714
LADAKH                                 8880
LAKSHADWEEP                            1500
MADHYA PRADESH                     16521580
MAHARASHTRA                         1993335
```

```
MANIPUR                                              149261
MEGHALAYA                                            164478
MIZORAM                                               46401
NAGALAND                                              84820
ODISHA                                             10415590
PUDUCHERRY                                           199180
PUNJAB                                               379420
RAJASTHAN                                           6585102
SIKKIM                                                29380
TAMIL NADU                                         10181512
TELANGANA                                           4017940
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU          98200
TRIPURA                                              346542
UTTAR PRADESH                                      20504720
UTTARAKHAND                                          477974
WEST BENGAL                                        14402441
Name: total_beneficiaries, dtype: int64
```

Sums the total beneficiaries for each state, grouped by the "state_name" column.

[21]:
```
df['total_beneficiaries'].sum()
```

[21]:
```
150875587
```

Calculates the total number of beneficiaries across the entire dataset.

[23]:
```
df.groupby('state_name')['total_beneficiaries'].sum().
  ↪sort_values(ascending=False)
```

[23]:
```
state_name
UTTAR PRADESH                                      20504720
MADHYA PRADESH                                     16521580
WEST BENGAL                                        14402441
BIHAR                                              12684420
ANDHRA PRADESH                                     11729503
ODISHA                                             10415590
TAMIL NADU                                         10181512
KARNATAKA                                           8196100
KERALA                                              7940714
GUJARAT                                             7762285
RAJASTHAN                                           6585102
JHARKHAND                                           5084571
CHHATTISGARH                                        4318904
TELANGANA                                           4017940
HARYANA                                             2657668
ASSAM                                               2290292
MAHARASHTRA                                         1993335
DELHI                                                729357
```

```
UTTARAKHAND                                    477974
HIMACHAL PRADESH                               439480
PUNJAB                                         379420
TRIPURA                                        346542
PUDUCHERRY                                     199180
MEGHALAYA                                      164478
MANIPUR                                        149261
JAMMU AND KASHMIR                              138212
ARUNACHAL PRADESH                              120200
GOA                                            117940
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU    98200
NAGALAND                                        84820
CHANDIGARH                                      49715
MIZORAM                                         46401
SIKKIM                                          29380
LADAKH                                           8880
ANDAMAN AND NICOBAR                              7970
LAKSHADWEEP                                      1500
Name: total_beneficiaries, dtype: int64
```

Displays total beneficiaries per state, sorted in descending order.

[25]:
```
df.groupby('state_name')['total_beneficiaries'].sum().
  ↪sort_values(ascending=False).head(10)
```

[25]:
```
state_name
UTTAR PRADESH     20504720
MADHYA PRADESH    16521580
WEST BENGAL       14402441
BIHAR             12684420
ANDHRA PRADESH    11729503
ODISHA            10415590
TAMIL NADU        10181512
KARNATAKA          8196100
KERALA             7940714
GUJARAT            7762285
Name: total_beneficiaries, dtype: int64
```

Shows the top 10 states with the highest total number of beneficiaries.

[27]:
```
df[df['total_mobileno'] == 0][['state_name', 'district_name',␣
  ↪'total_beneficiaries']]
```

[27]:
```
              state_name        district_name  total_beneficiaries
4      JAMMU AND KASHMIR               KULGAM                  290
17     ARUNACHAL PRADESH  UPPER DIBANG VALLEY                   78
18     ARUNACHAL PRADESH         KURUNG KUMEY                  391
21     ARUNACHAL PRADESH                TIRAP                  277
```

```
22     ARUNACHAL PRADESH            KRA DAADI                  143
…                  …                  …                    …
14255  ARUNACHAL PRADESH              SIANG                  123
14256  ARUNACHAL PRADESH      PAKKE KESSANG                   74
14257           NAGALAND               PHEK                  436
14259           NAGALAND            LONGLENG                 111
14278   HIMACHAL PRADESH              CHAMBA                2384

[1116 rows x 3 columns]
```

Lists states and districts where no beneficiaries have mobile numbers, along with their total beneficiaries.

[29]:
```
df[df['state_name']=='JAMMU AND KASHMIR']
```

[29]:
```
        fin_year  mnth  lgd_state_code         state_name  lgd_district_code  \
0      2023-2024     2               1  JAMMU AND KASHMIR                 13
1      2023-2024     2               1  JAMMU AND KASHMIR                 14
2      2023-2024     2               1  JAMMU AND KASHMIR                  2
3      2023-2024     2               1  JAMMU AND KASHMIR                  4
4      2023-2024     2               1  JAMMU AND KASHMIR                622
…          …   …             …                 …                  …
14239  2024-2025     2               1  JAMMU AND KASHMIR                 14
14240  2024-2025     2               1  JAMMU AND KASHMIR                  3
14241  2024-2025     2               1  JAMMU AND KASHMIR                622
14242  2024-2025     2               1  JAMMU AND KASHMIR                623
14243  2024-2025     2               1  JAMMU AND KASHMIR                  8

      district_name scheme_code  total_beneficiaries  total_sc  total_st  \
0          SRINAGAR      IGNWPS                  546         0         1
1          UDHAMPUR      IGNWPS                  353        85        17
2            BADGAM      IGNWPS                  176         3         1
3              DODA      IGNWPS                  246         9        18
4            KULGAM      IGNWPS                  290        20         0
…               …          …                    …       …         …
14239      UDHAMPUR      IGNWPS                  356        85        17
14240     BARAMULLA      IGNWPS                  674         0         0
14241        KULGAM      IGNWPS                  395        20         0
14242     BANDIPORA      IGNWPS                  477         0        54
14243       KUPWARA      IGNWPS                  492         5        10

       total_gen  total_obc  total_aadhar  total_mobileno lastUpdated  \
0            544          1           403             127  2024-07-31
1            238         13           250             121  2024-07-31
2            170          2           150              25  2024-07-31
3            219          0           237              33  2024-07-31
4            257         13           290               0  2024-07-31
```

```
…          …        …           …              …          …
14239        241         13           251             127  2025-03-07
14240        674          0           671             150  2025-03-07
14241        361         14           395             366  2025-03-07
14242        422          1           477             134  2025-03-07
14243        272        205           492             450  2025-03-07

          region_type
0      Union Territory
1      Union Territory
2      Union Territory
3      Union Territory
4      Union Territory
…                  …
14239  Union Territory
14240  Union Territory
14241  Union Territory
14242  Union Territory
14243  Union Territory

[400 rows x 16 columns]
```

Filters and displays all rows where the state name is “JAMMU AND KASHMIR”.

[31]:
```
df['mnth'] = df['mnth'].replace({
    1: 'Jan', 2: 'Feb', 3: 'Mar', 4: 'Apr',
    5: 'May', 6: 'Jun', 7: 'Jul', 8: 'Aug',
    9: 'Sep', 10: 'Oct', 11: 'Nov', 12: 'Dec'
})
df
```

[31]:
```
        fin_year mnth  lgd_state_code         state_name  lgd_district_code  \
0      2023-2024  Feb               1  JAMMU AND KASHMIR                 13
1      2023-2024  Feb               1  JAMMU AND KASHMIR                 14
2      2023-2024  Feb               1  JAMMU AND KASHMIR                  2
3      2023-2024  Feb               1  JAMMU AND KASHMIR                  4
4      2023-2024  Feb               1  JAMMU AND KASHMIR                622
…          …  …              …                 …                  …
14377  2024-2025  Feb               9      UTTAR PRADESH                172
14378  2024-2025  Feb               9      UTTAR PRADESH                173
14379  2024-2025  Feb               9      UTTAR PRADESH                176
14380  2024-2025  Feb               9      UTTAR PRADESH                177
14381  2024-2025  Feb               9      UTTAR PRADESH                181

       district_name scheme_code  total_beneficiaries  total_sc  total_st  \
0           SRINAGAR      IGNWPS                  546         0         1
1           UDHAMPUR      IGNWPS                  353        85        17
```

```
2              BADGAM      IGNWPS                  176         3         1
3                DODA      IGNWPS                  246         9        18
4              KULGAM      IGNWPS                  290        20         0
…                 …          …                     …         …         …
14377  MUZAFFARNAGAR      IGNWPS                 7601      1817        13
14378       PILIBHIT      IGNWPS                 8268      1796         0
14379          RAMPUR      IGNWPS                23245      3169         0
14380     SAHARANPUR      IGNWPS                10667      3649       143
14381      SHRAVASTI      IGNWPS                 6168      1689       332

       total_gen  total_obc  total_aadhar  total_mobileno lastUpdated  \
0            544          1           403             127  2024-07-31
1            238         13           250             121  2024-07-31
2            170          2           150              25  2024-07-31
3            219          0           237              33  2024-07-31
4            257         13           290               0  2024-07-31
…            …          …             …               …         …
14377       2719       3052          3680            3177  2025-03-07
14378       1756       4716          4037            4306  2025-03-07
14379       7868      12208          9288            9096  2025-03-07
14380       2940       3935          5896            5587  2025-03-07
14381       2203       1944          1649            2063  2025-03-07

          region_type
0      Union Territory
1      Union Territory
2      Union Territory
3      Union Territory
4      Union Territory
…                  …
14377            State
14378            State
14379            State
14380            State
14381            State

[14382 rows x 16 columns]
```

Replaces month numbers with their short-form names (Jan, Feb, etc.) in the “mnth” column.

[33]:
```
df.groupby('mnth')['total_beneficiaries'].sum()
```

[33]:
```
mnth
Apr    15332724
Aug     7825327
Dec     7558180
Feb    15028134
```

```
Jan     7558954
Jul    15321763
Jun    15364095
Mar     7539522
May    15379860
Nov    14985987
Oct    14500548
Sep    14480493
Name: total_beneficiaries, dtype: int64
```

Sums the total beneficiaries for each month, grouped by the "mnth" column.

[35]:
```
df.groupby('state_name')['total_aadhar'].sum().sort_values(ascending=False)
```

[35]:
```
state_name
WEST BENGAL                                  14065667
ODISHA                                       10168479
MADHYA PRADESH                                9311717
BIHAR                                         8334296
KARNATAKA                                     7985203
GUJARAT                                       7683545
TAMIL NADU                                    7587049
ANDHRA PRADESH                                7432652
RAJASTHAN                                     5535692
JHARKHAND                                     5006173
KERALA                                        4692877
UTTAR PRADESH                                 4148635
CHHATTISGARH                                  4001228
TELANGANA                                     3729388
MAHARASHTRA                                   1743080
ASSAM                                         1719339
HARYANA                                       1653813
DELHI                                          725728
PUNJAB                                         216856
UTTARAKHAND                                    188706
HIMACHAL PRADESH                               164740
TRIPURA                                        151914
JAMMU AND KASHMIR                              112615
MANIPUR                                        106508
GOA                                            103736
MEGHALAYA                                       96931
ARUNACHAL PRADESH                               76300
CHANDIGARH                                      49248
MIZORAM                                         34335
PUDUCHERRY                                      26632
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU    19068
NAGALAND                                         7890
```

```
ANDAMAN AND NICOBAR                                  4806
LADAKH                                               1931
LAKSHADWEEP                                          1500
SIKKIM                                                  0
Name: total_aadhar, dtype: int64
```

Shows the total Aadhaar-linked beneficiaries per state, sorted in descending order.

[37]:
```
df.groupby('district_name')['total_beneficiaries'].mean()
```

[37]:
```
district_name
24 PARGANAS NORTH    74552.70
24 PARGANAS SOUTH    45113.00
ADILABAD             13699.00
AGAR MALWA            5769.00
AGRA                  4577.00
                        …
YADADRI               5540.00
YADGIR                1075.70
YAMUNANAGAR           8802.90
YAVATMAL              2044.05
ZUNHEBOTO              434.00
Name: total_beneficiaries, Length: 721, dtype: float64
```

Calculates the average number of beneficiaries for each district.

[39]:
```
df.groupby('state_name')['lastUpdated'].max()
```

[39]:
```
state_name
ANDAMAN AND NICOBAR                  2025-03-07
ANDHRA PRADESH                       2025-03-07
ARUNACHAL PRADESH                    2025-03-07
ASSAM                                2025-03-07
BIHAR                                2025-03-07
CHANDIGARH                           2025-03-07
CHHATTISGARH                         2025-03-07
DELHI                                2025-03-07
GOA                                  2025-03-07
GUJARAT                              2025-03-07
HARYANA                              2025-03-07
HIMACHAL PRADESH                     2025-03-07
JAMMU AND KASHMIR                    2025-03-07
JHARKHAND                            2025-03-07
KARNATAKA                            2025-03-07
KERALA                               2025-03-07
LADAKH                               2025-03-07
LAKSHADWEEP                          2025-03-07
MADHYA PRADESH                       2025-03-07
```

```
MAHARASHTRA                                          2025-03-07
MANIPUR                                              2025-03-07
MEGHALAYA                                            2025-03-07
MIZORAM                                              2025-03-07
NAGALAND                                             2025-03-07
ODISHA                                               2025-03-07
PUDUCHERRY                                           2025-03-07
PUNJAB                                               2025-03-07
RAJASTHAN                                            2025-03-07
SIKKIM                                               2025-03-07
TAMIL NADU                                           2025-03-07
TELANGANA                                            2025-03-07
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU         2025-03-07
TRIPURA                                              2025-03-07
UTTAR PRADESH                                        2025-03-07
UTTARAKHAND                                          2025-03-07
WEST BENGAL                                          2025-03-07
Name: lastUpdated, dtype: object
```

Shows the most recent update date for each state.

[41]:
```
df[df['total_mobileno'] == 0]['state_name'].unique()
```

[41]:
```
array(['JAMMU AND KASHMIR', 'ARUNACHAL PRADESH', 'NAGALAND',
       'HIMACHAL PRADESH', 'LADAKH', 'LAKSHADWEEP', 'CHANDIGARH',
       'SIKKIM', 'MANIPUR', 'PUDUCHERRY',
       'THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU', 'TRIPURA'],
      dtype=object)
```

Lists the unique states where no beneficiaries have mobile numbers.

[43]:
```
df.
  ↪groupby('state_name')[['total_beneficiaries','total_sc','total_st','total_gen','total_obc']]
  ↪sum()
```

[43]:

state_name	total_beneficiaries	total_sc	\
ANDAMAN AND NICOBAR	7970	0	
ANDHRA PRADESH	11729503	2187771	
ARUNACHAL PRADESH	120200	0	
ASSAM	2290292	128180	
BIHAR	12684420	2059880	
CHANDIGARH	49715	0	
CHHATTISGARH	4318904	346288	
DELHI	729357	16317	
GOA	117940	2080	
GUJARAT	7762285	342774	
HARYANA	2657668	586454	

HIMACHAL PRADESH	439480	128600
JAMMU AND KASHMIR	138212	9577
JHARKHAND	5084571	710744
KARNATAKA	8196100	741149
KERALA	7940714	347804
LADAKH	8880	0
LAKSHADWEEP	1500	0
MADHYA PRADESH	16521580	3171540
MAHARASHTRA	1993335	102123
MANIPUR	149261	3449
MEGHALAYA	164478	1075
MIZORAM	46401	40
NAGALAND	84820	20
ODISHA	10415590	1229465
PUDUCHERRY	199180	30040
PUNJAB	379420	23220
RAJASTHAN	6585102	1538505
SIKKIM	29380	2420
TAMIL NADU	10181512	2472065
TELANGANA	4017940	842420
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU	98200	3860
TRIPURA	346542	62202
UTTAR PRADESH	20504720	5720240
UTTARAKHAND	477974	97223
WEST BENGAL	14402441	3977488

state_name	total_st	total_gen	total_obc
ANDAMAN AND NICOBAR	2270	5700	0
ANDHRA PRADESH	891352	2618222	6032158
ARUNACHAL PRADESH	119080	1100	20
ASSAM	249472	1669274	243366
BIHAR	199880	7953940	2470720
CHANDIGARH	0	49715	0
CHHATTISGARH	1156282	1638101	1178233
DELHI	163	709303	3574
GOA	30380	61400	24080
GUJARAT	1045214	4496532	1876012
HARYANA	0	1552919	518295
HIMACHAL PRADESH	24180	230260	56440
JAMMU AND KASHMIR	8091	112759	7785
JHARKHAND	1407118	1264247	1702462
KARNATAKA	329900	1944622	5180429
KERALA	145506	7447404	0
LADAKH	8858	22	0
LAKSHADWEEP	660	840	0
MADHYA PRADESH	4830140	1702800	6817100

MAHARASHTRA	293757	1427562	169893
MANIPUR	34430	77740	33642
MEGHALAYA	145642	16874	887
MIZORAM	43685	2676	0
NAGALAND	84480	120	200
ODISHA	2304825	5022703	1858597
PUDUCHERRY	740	160920	7480
PUNJAB	160	350620	5420
RAJASTHAN	1574271	1242608	2229718
SIKKIM	12140	8540	6280
TAMIL NADU	105123	2628970	4975354
TELANGANA	545160	277280	2353080
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU	42460	15320	36560
TRIPURA	88032	138674	57634
UTTAR PRADESH	115640	4700240	9968600
UTTARAKHAND	15240	324606	40905
WEST BENGAL	1027654	9120798	256592

Sums the total beneficiaries and their category-wise breakdown (SC, ST, Gen, OBC) for each state.

[45]:
```
df.groupby('fin_year')['total_beneficiaries'].sum().sort_index()
```

[45]:
```
fin_year
2023-2024    90896320
2024-2025    59979267
Name: total_beneficiaries, dtype: int64
```

Shows the total beneficiaries for each financial year, sorted in chronological order.

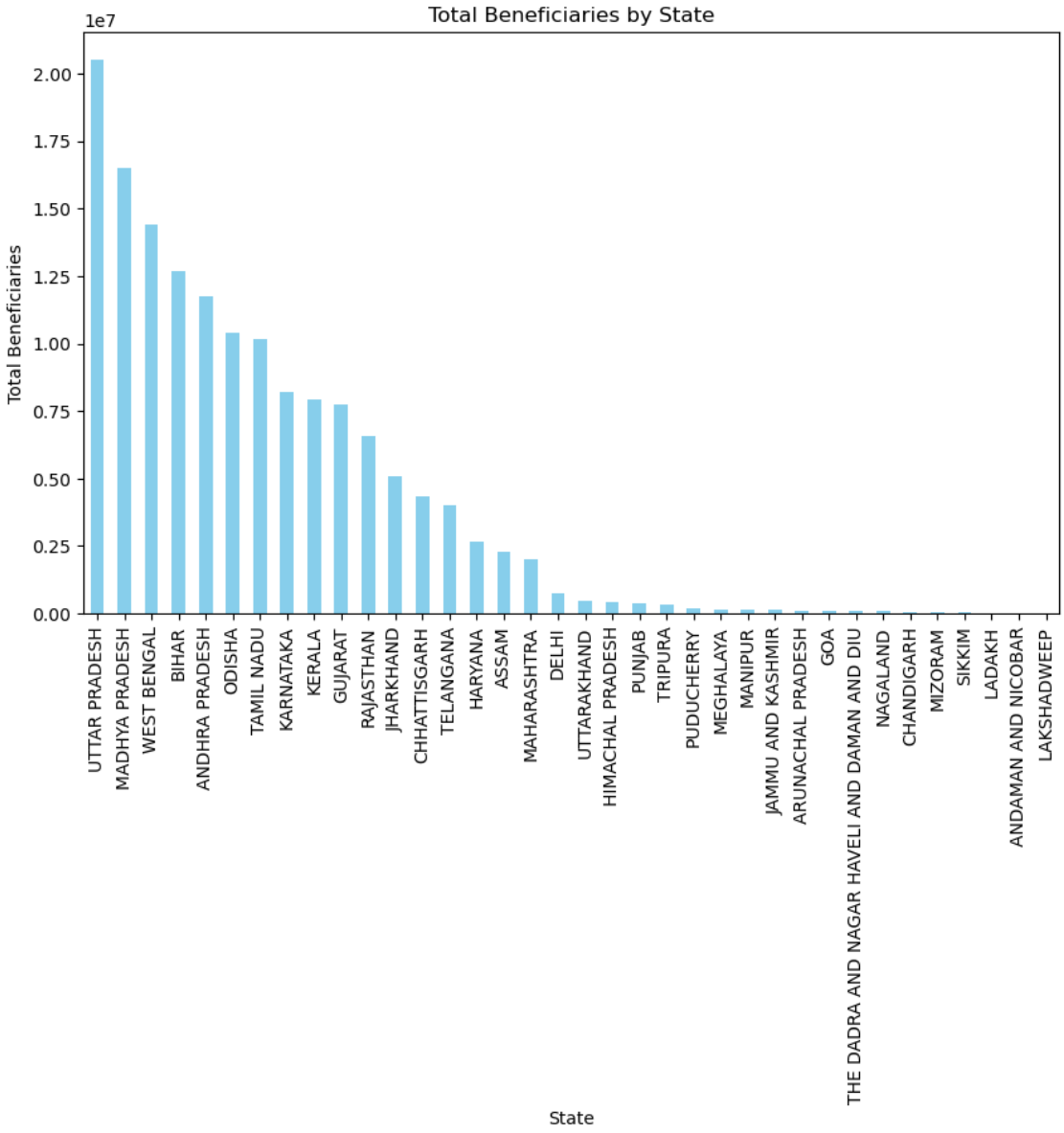
[47]:
```
df.groupby('scheme_code')['total_beneficiaries'].sum().
 ↪sort_values(ascending=False).head(5)
```

[47]:
```
scheme_code
IGNWPS    150875587
Name: total_beneficiaries, dtype: int64
```

Displays the top 5 schemes with the highest total number of beneficiaries.

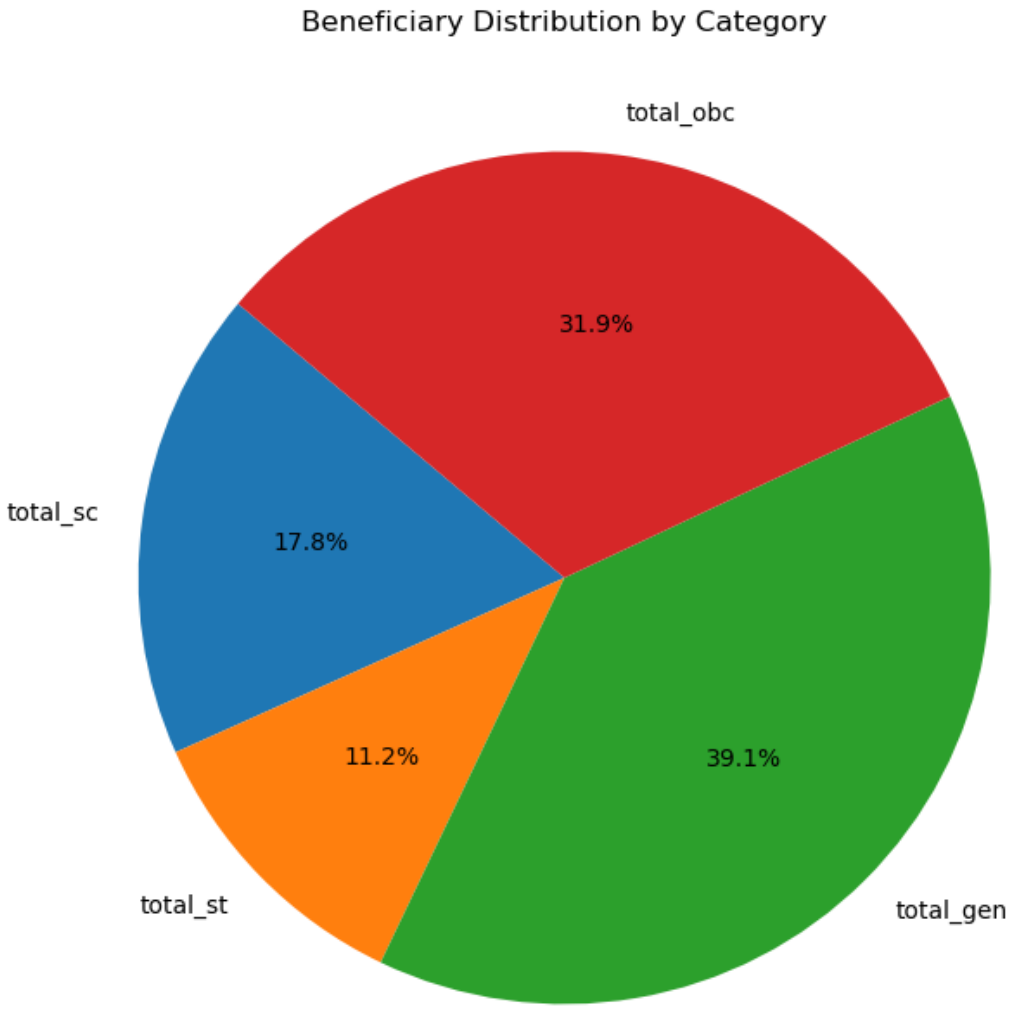
[49]:
```
import matplotlib.pyplot as plt

statewise = df.groupby('state_name')['total_beneficiaries'].sum().
 ↪sort_values(ascending=False)
statewise.plot(kind='bar', figsize=(10, 6), title='Total Beneficiaries by␣
 ↪State', color='skyblue')
plt.xlabel('State')
plt.ylabel('Total Beneficiaries')
plt.show()
```



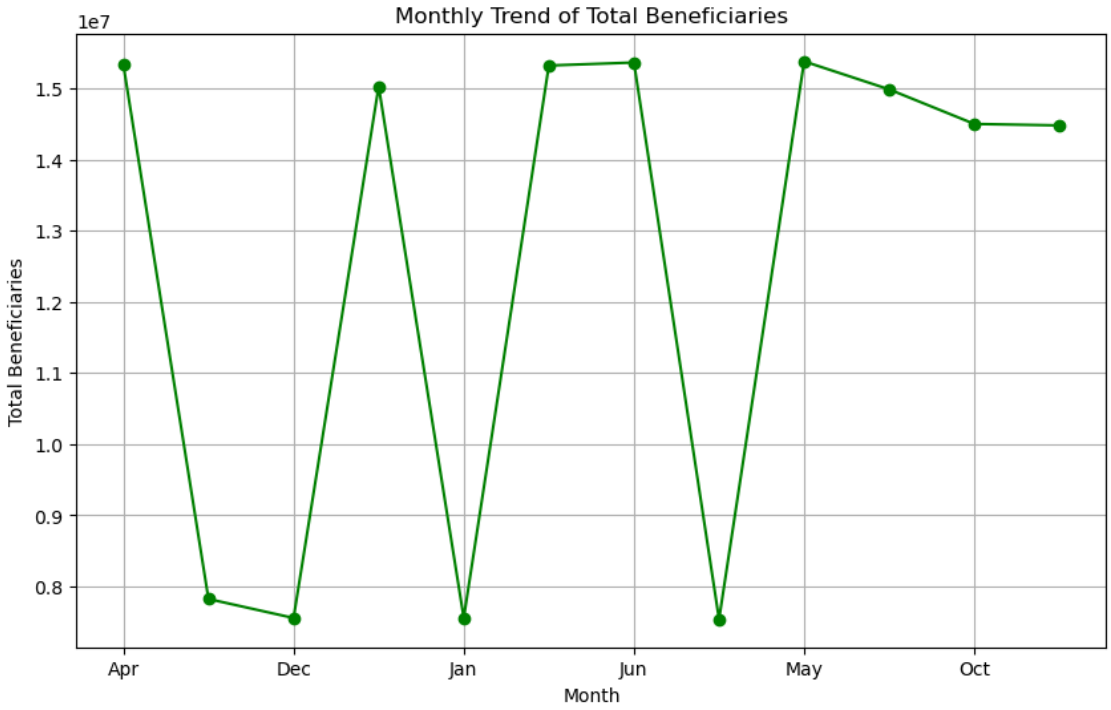
Creates a bar chart showing total beneficiaries by state, sorted in descending order.

[51]:
```
category_totals = df[['total_sc', 'total_st', 'total_gen', 'total_obc']].sum()
category_totals.plot(kind='pie', autopct='%1.1f%%', startangle=140, figsize=(8,␣
↪8), title='Beneficiary Distribution by Category')
plt.show()
```



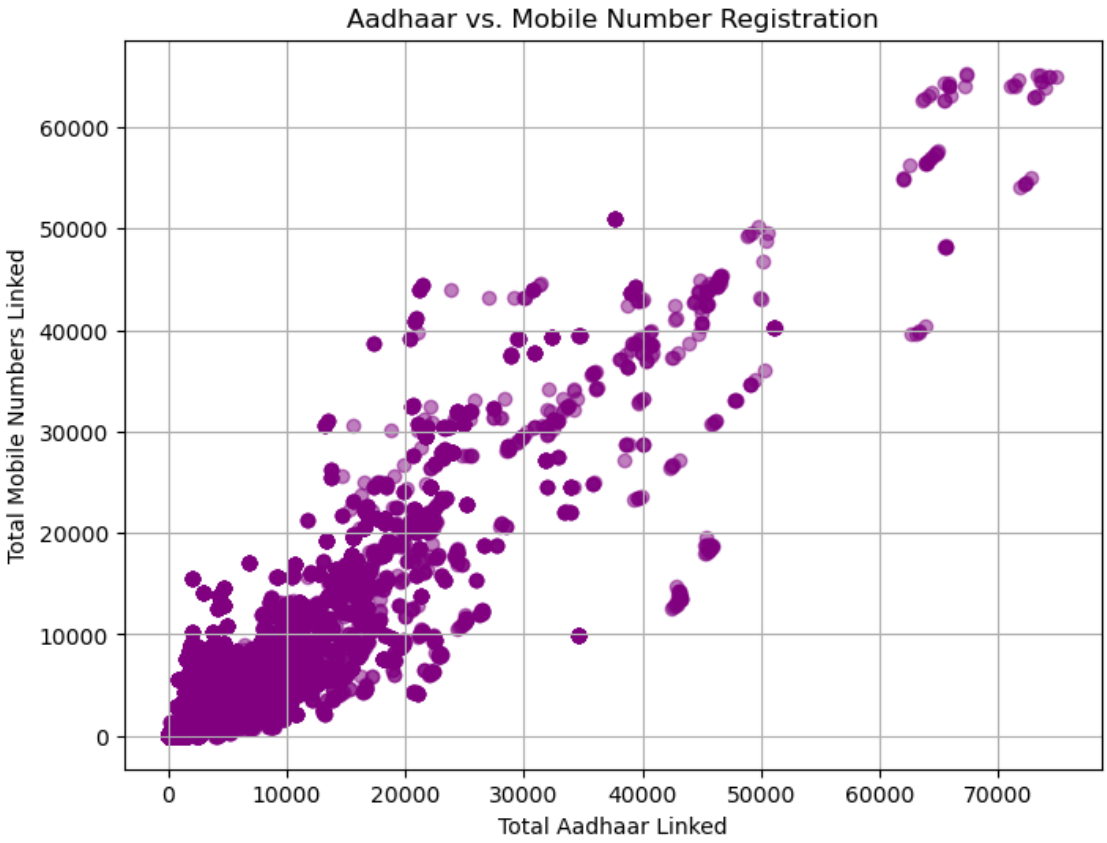
Creates a pie chart showing the percentage distribution of beneficiaries across different categories (SC, ST, Gen, OBC).

[53]:
```
monthly_trend = df.groupby('mnth')['total_beneficiaries'].sum()
monthly_trend.plot(kind='line', marker='o', figsize=(10, 6), title='Monthly␣
↪Trend of Total Beneficiaries', color='green')
plt.xlabel('Month')
plt.ylabel('Total Beneficiaries')
plt.grid(True)
plt.show()
```



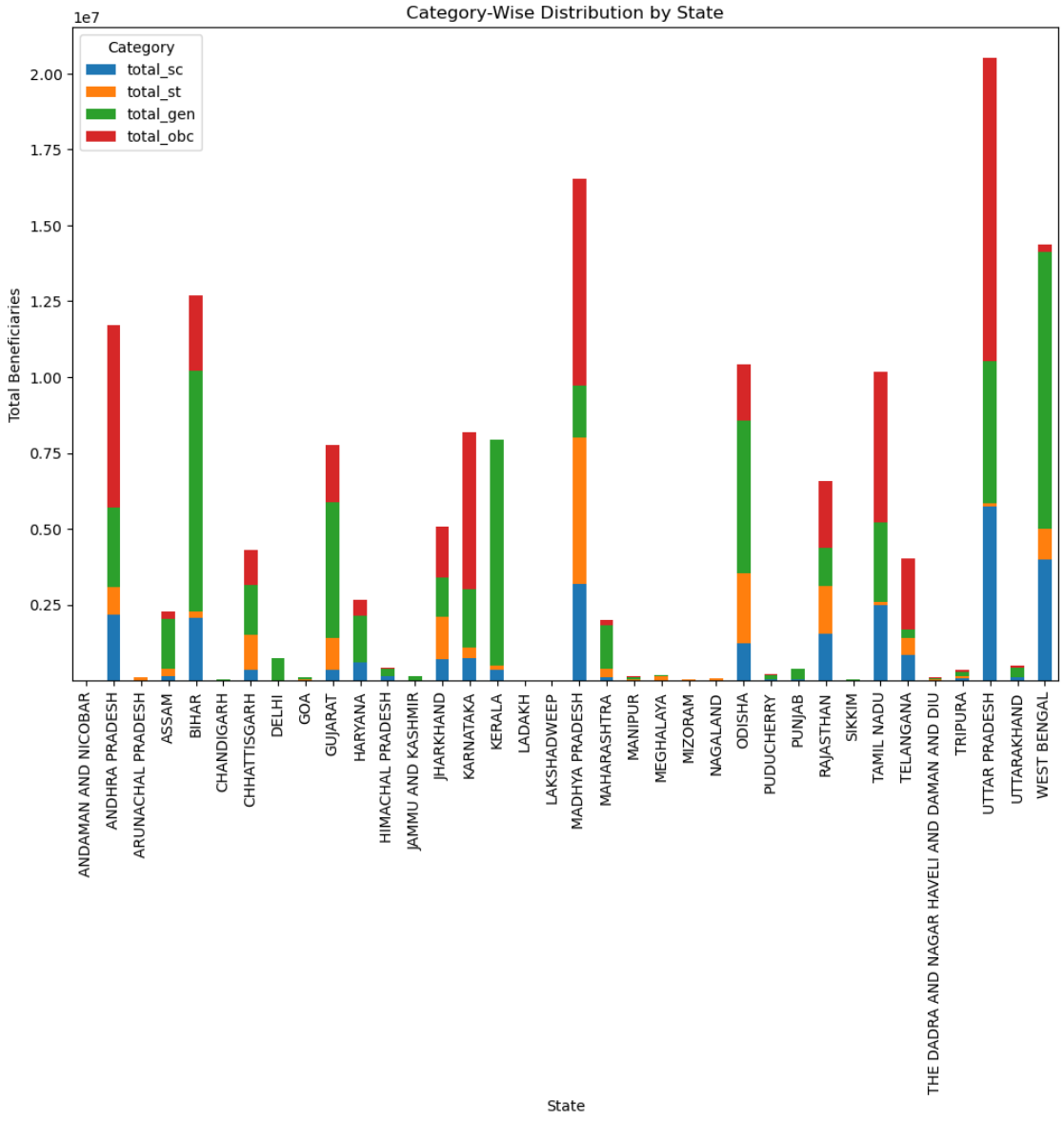
Plots a line chart showing the monthly trend of total beneficiaries, with markers and grid for clarity.

[55]:
```
plt.figure(figsize=(8, 6))
plt.scatter(df['total_aadhar'], df['total_mobileno'], alpha=0.5, color='purple')
plt.xlabel('Total Aadhaar Linked')
plt.ylabel('Total Mobile Numbers Linked')
plt.title('Aadhaar vs. Mobile Number Registration')
plt.grid(True)
plt.show()
```



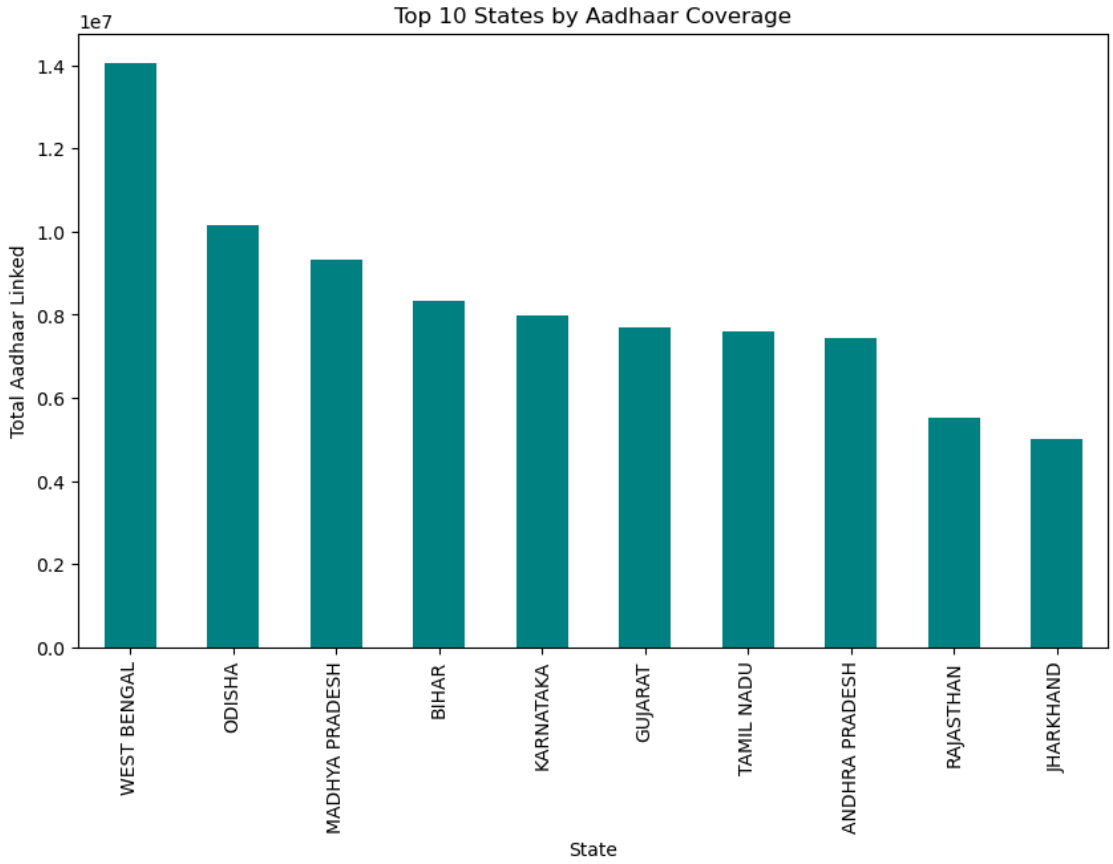
Creates a scatter plot showing the relationship between Aadhaar-linked and mobile-linked beneficiaries.

[57]:
```
statewise_category = df.groupby('state_name')[['total_sc', 'total_st',␣
↪'total_gen', 'total_obc']].sum()
statewise_category.plot(kind='bar', stacked=True, figsize=(12, 8),␣
↪title='Category-Wise Distribution by State')
plt.xlabel('State')
plt.ylabel('Total Beneficiaries')
plt.legend(title='Category')
plt.show()
```



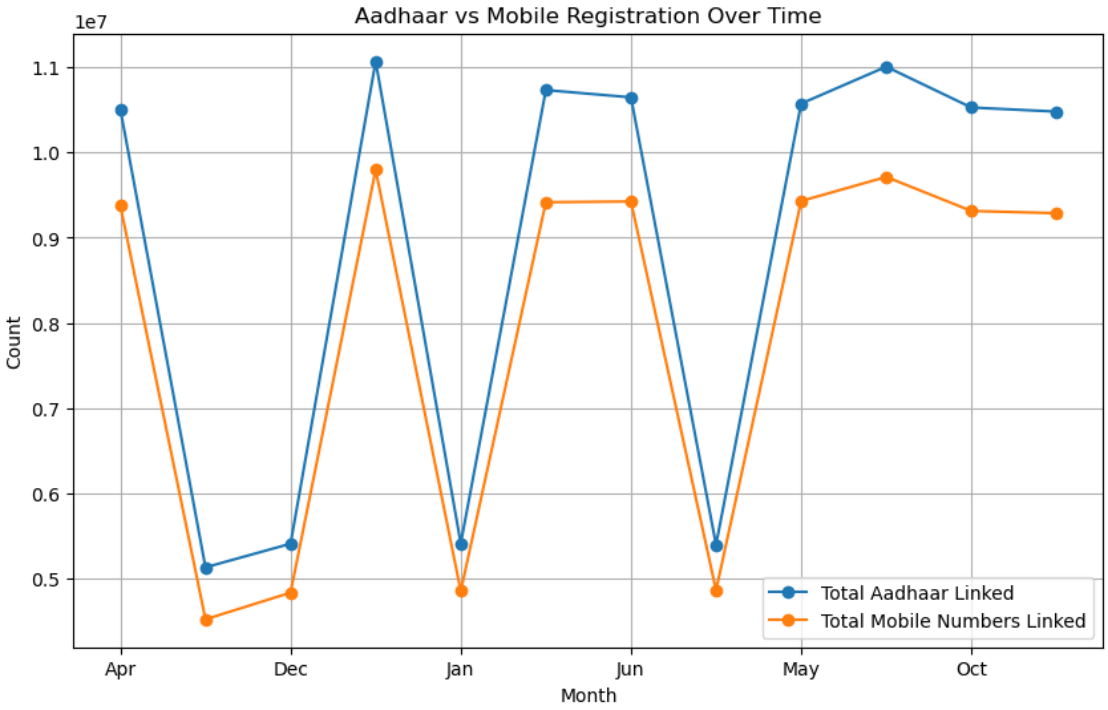
Creates a stacked bar chart showing the category-wise distribution of beneficiaries for each state.

[59]:
```
top_aadhar = df.groupby('state_name')['total_aadhar'].sum().
 ↪sort_values(ascending=False).head(10)
top_aadhar.plot(kind='bar', color='teal', figsize=(10, 6), title='Top 10 States␣
 ↪by Aadhaar Coverage')
plt.xlabel('State')
plt.ylabel('Total Aadhaar Linked')
plt.show()
```



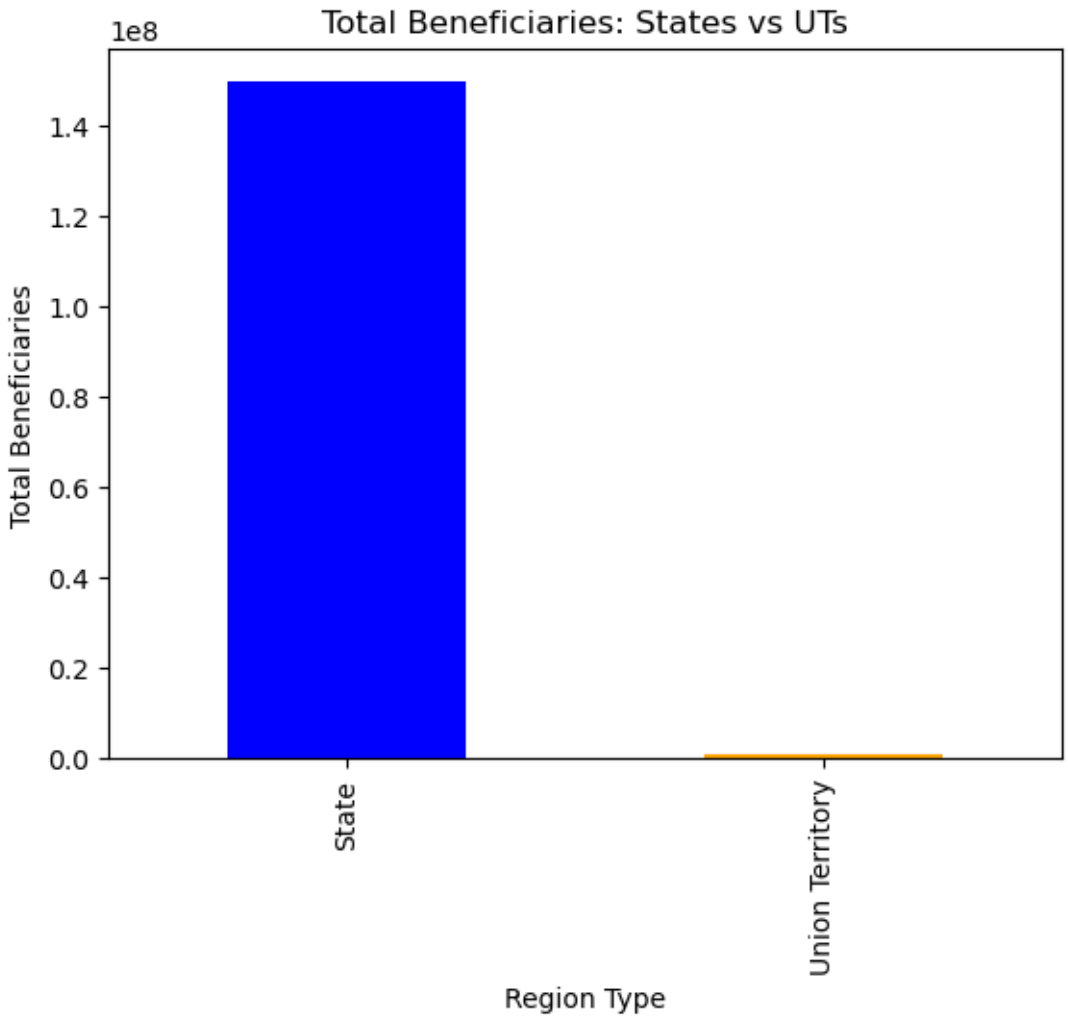
Displays a bar chart of the top 10 states with the highest total Aadhaar-linked beneficiaries.

[61]:
```
monthly_totals = df.groupby('mnth')[['total_aadhar', 'total_mobileno']].sum()
monthly_totals.plot(kind='line', marker='o', figsize=(10, 6), title='Aadhaar vs␣
↪Mobile Registration Over Time')
plt.xlabel('Month')
plt.ylabel('Count')
plt.legend(['Total Aadhaar Linked', 'Total Mobile Numbers Linked'])
plt.grid(True)
plt.show()
```



Plots a line chart comparing Aadhaar and mobile registrations over the months, showing trends over time.

[63]:
```
state_ut_totals = df.groupby('region_type')['total_beneficiaries'].sum()  #␣
  ↪Assuming 'region_type' distinguishes State/UT
state_ut_totals.plot(kind='bar', color=['blue', 'orange'], title='Total␣
  ↪Beneficiaries: States vs UTs')
plt.xlabel('Region Type')
plt.ylabel('Total Beneficiaries')
plt.show()
```

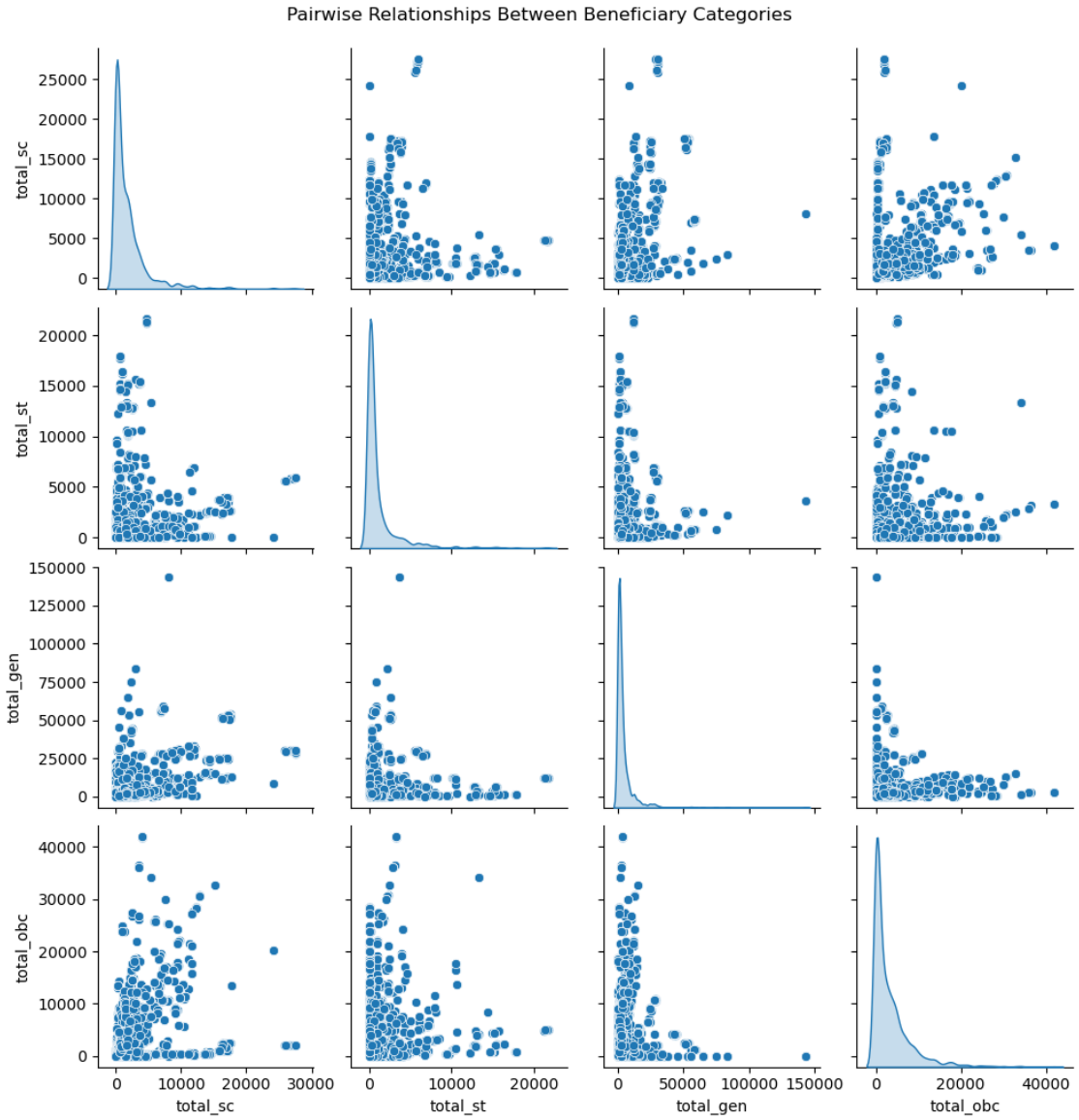


Creates a bar chart comparing total beneficiaries between states and union territories.

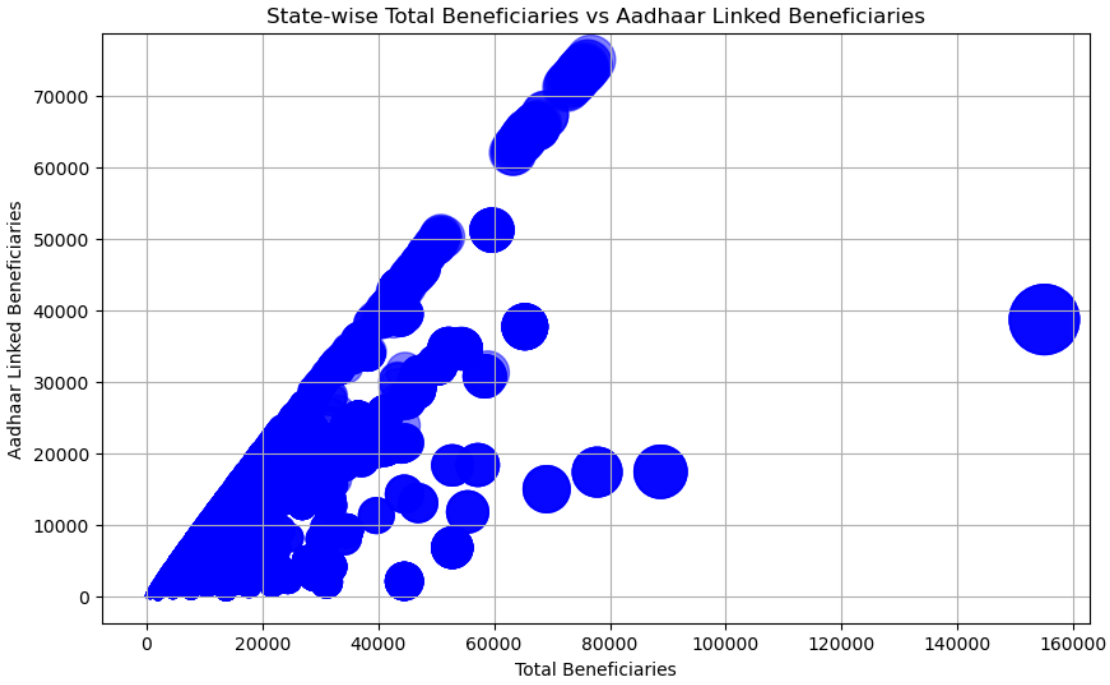
[70]:
```
import seaborn as sns
sns.pairplot(df[['total_sc', 'total_st', 'total_gen', 'total_obc']],␣
 ↪diag_kind='kde', palette='husl')
plt.suptitle("Pairwise Relationships Between Beneficiary Categories", y=1.02)
plt.show()
```

```
C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1513: UserWarning:
Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=vector, **plot_kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1513:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=vector, **plot_kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1513:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
```

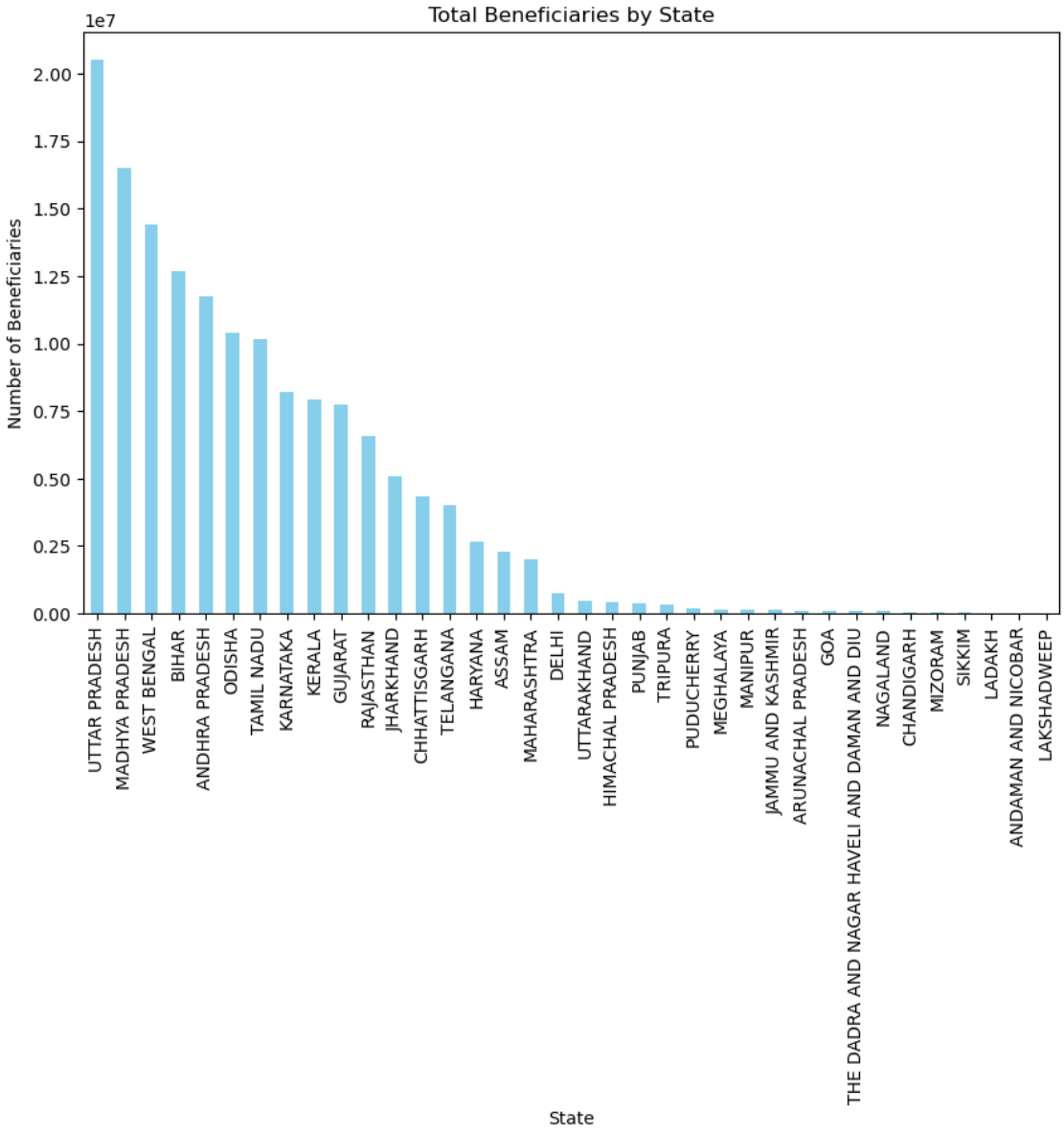
```
,  func(x=vector, **plot_kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1513:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=vector, **plot_kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
,C:\Users\Dell\anaconda3\Lib\site-packages\seaborn\axisgrid.py:1615:
UserWarning: Ignoring `palette` because no `hue` variable has been assigned.
,  func(x=x, y=y, **kwargs)
```



[66]:
```
plt.figure(figsize=(10, 6))
plt.scatter(df['total_beneficiaries'], df['total_aadhar'],␣
  ↪s=df['total_beneficiaries'] / 100, alpha=0.5, color='blue')
plt.title("State-wise Total Beneficiaries vs Aadhaar Linked Beneficiaries")
plt.xlabel("Total Beneficiaries")
plt.ylabel("Aadhaar Linked Beneficiaries")
plt.grid(True)
plt.show()
```



[72]:
```
import matplotlib.pyplot as plt
state_beneficiaries = df.groupby('state_name')['total_beneficiaries'].sum()
plt.figure(figsize=(10, 6))
state_beneficiaries.sort_values(ascending=False).plot(kind='bar',␣
 ↪color='skyblue')
plt.title("Total Beneficiaries by State")
plt.xlabel("State")
plt.ylabel("Number of Beneficiaries")
plt.xticks(rotation=90)
plt.show()
```


Total Beneficiaries by State
1e7
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Number of Beneficiaries
UTTAR PRADESH
MADHYA PRADESH
WEST BENGAL
BIHAR
ANDHRA PRADESH
ODISHA
TAMIL NADU
KARNATAKA
KERALA
GUJARAT
RAJASTHAN
JHARKHAND
CHHATTISGARH
TELANGANA
HARYANA
ASSAM
MAHARASHTRA
DELHI
UTTARAKHAND
HIMACHAL PRADESH
PUNJAB
TRIPURA
PUDUCHERRY
MEGHALAYA
MANIPUR
JAMMU AND KASHMIR
ARUNACHAL PRADESH
GOA
THE DADRA AND NAGAR HAVELI AND DAMAN AND DIU
NAGALAND
CHANDIGARH
MIZORAM
SIKKIM
LADAKH
ANDAMAN AND NICOBAR
LAKSHADWEEP
State

[ ]:

6 **Confirmed 14,382 rows, 15 columns, no missing values.**

7 **Listed the complete set of states, confirming coverage across multiple regions.**

8 **Showed which states had more records, indicating regions with higher data granularity or reporting frequency.**

9 **Highlighted the leading states in terms of outreach and those with lower coverage, pointing to potential gaps.**

10 **Showed that certain states had significantly higher Aadhaar penetration, indicating better digital integration.**

11 **Identified districts with the highest outreach, potentially reflecting higher population density or better program implementation**

12 **Revealed monthly patterns — useful for detecting seasonal impacts or administrative trends.**

13 **Confirmed when each state's data was last refreshed, helping to identify stale data sources.**

14 **Created new column name region type as we divided it into union teritories and states.**

[ ]: