UIDAI Data Hackathon 2026

Digital India Readiness Index: State-wise Analysis of Aadhaar Ecosystem Health

Analyst: Meet Wadekar

Dataset Period: Multi-source UIDAI data across 10 datasets

Total States/UTs Analyzed: 36

1. Problem Statement and Approach

1.1 Problem Statement

Measuring State-wise Digital Service Delivery Readiness: A Multi-dimensional Analysis of Aadhaar Ecosystem Health for Inclusive Welfare Distribution

The Unique Identification Authority of India (UIDAI) has deployed Aadhaar as the cornerstone of India's digital infrastructure, enabling a wide range of citizen services including Public Distribution System (PDS), MGNREGS wage payments, Direct Benefit Transfer (DBT), and MSME integration.

Core Challenge: Despite widespread Aadhaar enrollment (>99% nationally), significant disparities exist in the utilization and integration of Aadhaar across states, leading to:

 * Digital exclusion of vulnerable populations from welfare schemes
 * Inefficient subsidy delivery due to incomplete beneficiary seeding
 * Regional imbalances, particularly in North-Eastern states
 * Data quality gaps that hinder evidence-based policy making

1.2 Hypothesis

Primary Hypothesis: States with higher Aadhaar integration across PDS, MGNREGS, and MSME ecosystems demonstrate stronger digital readiness and more efficient welfare delivery.

Secondary Hypotheses:

 * North-Eastern states lag significantly in Aadhaar-based service delivery due to infrastructure constraints
 * States with 100% ration card seeding will show correspondingly high ABPS coverage in MGNREGS
 * MSME density correlates positively with overall digital readiness

1.3 Analytical Approach

The analysis employs a Multi-dimensional Index approach to quantify digital readiness:

Dimension	Weight	Rationale
Aadhaar Coverage	20%	Foundation metric
PDS Readiness	35%	Highest citizen impact
MGNREGS ABPS	30%	Employment guarantee critical
MSME Density	15%	Economic formalization

2. Datasets Used

2.1 Primary Datasets

Dataset File	Description
RS_Session_254_AU_1356.csv	PDS metrics and FPS automation
RS_Session_246_AU2800.csv	Ration cards with deletions
RS_Session_254_AU_1540.1.ii_.c	MSME data by state
RS_Session_260_AU_1546_C.csv	MGNREGS ABPS eligibility
rs_session-241_au2785_1.1.csv	Aadhaar vs 2011 census
rs_session-241_au3451_1.1.csv	DBT scheme beneficiaries
rs_session243_au721_1.1.csv	Deleted ration cards
rs_session_239_AU1492_1.1.csv	Transgender Aadhaar data
session_244_AU85_1.1_1.csv	Alternative seeding data
Registered_device_april-may_25	Biometric device performance

2.2 Key Columns Used

PDS Metrics (RS_Session_254_AU_1356.csv):

  * State/UT - Name of state or union territory
  * % Aadhaar Seeding with Ration Cards - Percentage of ration cards linked to Aadhaar
  * % Aadhaar Seeding with Beneficiaries - Percentage of beneficiaries linked
  * % FPS Automation - Fair Price Shop automation percentage

MGNREGS Data (RS_Session_260_AU_1546_C.csv):

  * State/UT - Name of state or union territory
  * Numbers of Active Workers (In Lakh) - Total active MGNREGS workers
  * Number of Workers Eligible for ABPS (In Lakh) - Workers eligible for Aadhaar-Based Payment

MSME Data (RS_Session_254_AU_1540.1.ii_.csv):

  * State/UT - Name of state
  * Micro/Small/Medium Manufacturing/Services - Enterprise counts by category
  * Total - Total MSMEs registered

Aadhaar Generation (rs_session-241_au2785_1.1.csv):

  * State/UT - Name of state
  * Population as per 2011 census - Census population
  * Aadhaar generated - Total Aadhaar numbers generated
  * % Aadhaar generated - Coverage percentage

3. Methodology

3.1 Data Loading and Preprocessing

All datasets were loaded using pandas with appropriate encoding handling. The preprocessing pipeline included:

Data Loading Code:
```
import pandas as pd
import numpy as np
import os

DATA_PATH = 'data'

# Load primary datasets
df_pds_metrics = pd.read_csv(os.path.join(DATA_PATH, 'RS_Session_254_AU_1356.csv'))
df_ration_cards = pd.read_csv(os.path.join(DATA_PATH, 'RS_Session_246_AU2800.csv'))
df_msme = pd.read_csv(os.path.join(DATA_PATH, 'RS_Session_254_AU_1540.1.ii_.csv'))
df_mgnregs = pd.read_csv(os.path.join(DATA_PATH, 'RS_Session_260_AU_1546_C.csv'))
df_aadhaar_gen = pd.read_csv(os.path.join(DATA_PATH, 'rs_session-241_au2785_1.1.csv'))
```

3.2 Data Cleaning and Standardization

State names were standardized across all datasets to enable accurate merging:

State Name Standardization:
```
def standardize_state_name(name):
    if pd.isna(name) or str(name).strip() in ['Total', 'Grand Total']:
        return None
    name = str(name).strip()
    replacements = {
        'A & N Islands': 'Andaman and Nicobar Islands',
        'Andaman & Nicobar': 'Andaman and Nicobar Islands',
        'Dadra & Nagar Haveli': 'Dadra and Nagar Haveli',
        'Daman & Diu': 'Daman and Diu',
        'Jammu & Kashmir': 'Jammu and Kashmir',
    }
    for old, new in replacements.items():
        if old.lower() == name.lower():
            return new
    return name

# Apply standardization
df_pds_clean['State'] = df_pds_clean['State'].apply(standardize_state_name)
df_pds_clean = df_pds_clean[df_pds_clean['State'].notna()]
```

3.3 Feature Engineering

Additional metrics were calculated from raw data:

Feature Engineering:
```
# Calculate ABPS Coverage
df_mgnregs_clean['ABPS_Coverage'] = (
    df_mgnregs_clean['ABPS_Eligible_Lakh'] /
    df_mgnregs_clean['Active_Workers_Lakh'] * 100
).round(2)

# Calculate MSME Density (per 10,000 population)
master_df['MSME_Density'] = (
    master_df['Total_MSMEs'] / master_df['Population_2011'] * 10000
).round(2)

# Cap Aadhaar coverage at 100%
master_df['Aadhaar_Coverage_Capped'] = master_df['Aadhaar_Percentage'].clip(upper=100)
```

3.4 Index Calculation

The Digital Readiness Index was calculated using min-max normalization and weighted aggregation:

Index Calculation Code:

```python
def normalize_score(series, higher_is_better=True):
    min_val = series.min()
    max_val = series.max()
    if max_val == min_val:
        return pd.Series([50] * len(series), index=series.index)
    if higher_is_better:
        return ((series - min_val) / (max_val - min_val) * 100).round(2)
    return ((max_val - series) / (max_val - min_val) * 100).round(2)


# Define weights
WEIGHTS = {
    'Score_Aadhaar_Coverage': 0.20,
    'Score_PDS_Readiness': 0.35,
    'Score_MGNREGS_ABPS': 0.30,
    'Score_MSME_Density': 0.15
}


# Calculate composite index
master_df['Digital_Readiness_Index'] = (
    master_df['Score_Aadhaar_Coverage'] * WEIGHTS['Score_Aadhaar_Coverage'] +
    master_df['Score_PDS_Readiness'] * WEIGHTS['Score_PDS_Readiness'] +
    master_df['Score_MGNREGS_ABPS'] * WEIGHTS['Score_MGNREGS_ABPS'] +
    master_df['Score_MSME_Density'] * WEIGHTS['Score_MSME_Density']
).round(2)
```

4. Data Analysis and Visualisation

4.1 Key Findings

Coverage Statistics:

Metric	Value
Total States/UTs Analyzed	36
Average Digital Readiness Inde	70.9
Median Digital Readiness Index	78.1
Standard Deviation	24.3

MGNREGS ABPS Coverage:

Metric	Value
Total Active Workers	1,496.9 Lakh
ABPS Eligible Workers	1,043.3 Lakh
National ABPS Coverage	69.7%
Workers NOT Covered	453.6 Lakh

4.2 Top Performing States

Rank	State	DRI Score	Key Strength
1	Tamil Nadu	94.5	Full PDS + High MSME
2	Andhra Pradesh	91.1	100% coverage all metrics
3	Telangana	91.0	Excellent MGNREGS
4	Maharashtra	89.3	High MSME + PDS
5	Puducherry	86.7	100% FPS automation

4.3 Critical Findings

Finding 1: PDS Digital Exclusion (HIGH PRIORITY)

  * Assam and Meghalaya have 0% Aadhaar-ration card seeding
  * Impact: ~40+ lakh ration cards without digital linkage
  * Recommendation: Emergency enrollment drives in NE states

Finding 2: MGNREGS ABPS Gap (HIGH PRIORITY)

  * 453.6 lakh workers NOT eligible for ABPS direct payment
  * Only 69.7% of active workers are ABPS-ready
  * Recommendation: Expedite Aadhaar seeding campaigns

Finding 3: North-Eastern States Lagging (HIGH PRIORITY)

  * NE States average readiness: ~40 vs national average: 70.9
  * Affected: Assam, Meghalaya, Nagaland, Manipur, Mizoram, Arunachal
  * Recommendation: Door-to-door enrollment programs

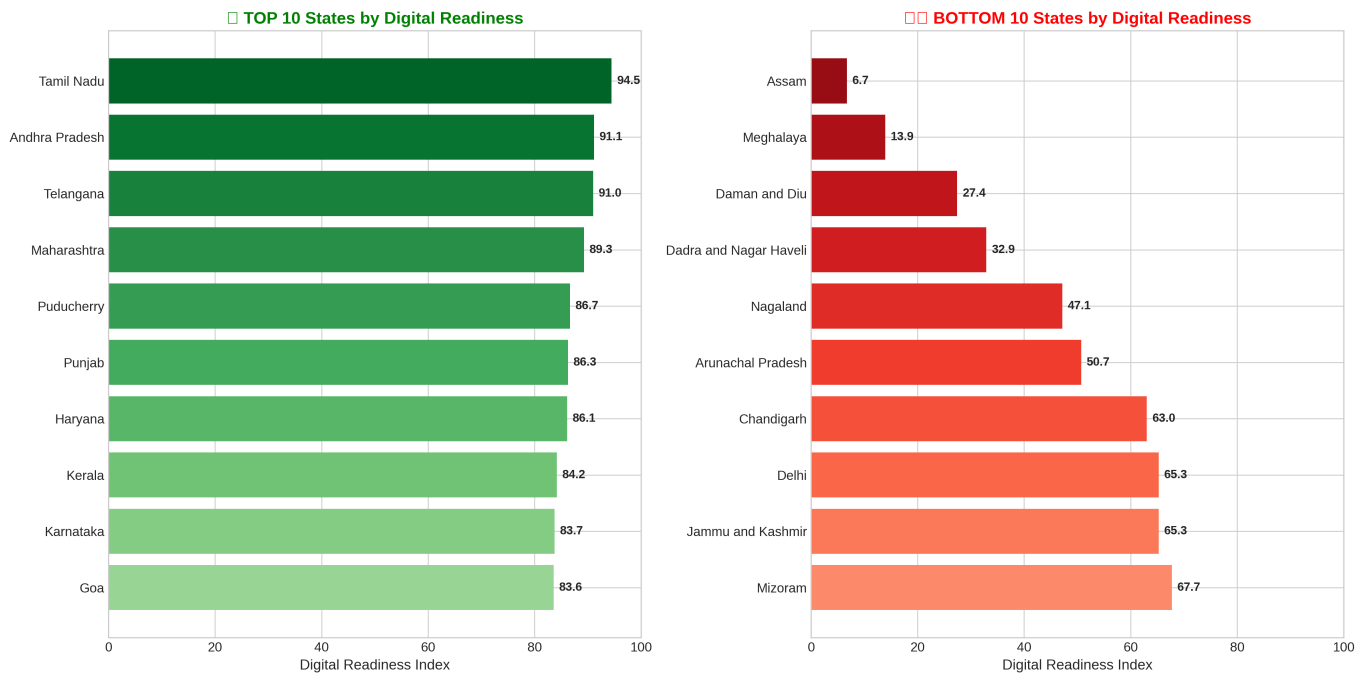## 4.4 Visualisations

State Rankings - Top 10 and Bottom 10:

Figure 1: Top 10 and Bottom 10 States by Digital Readiness Index

Digital Readiness Heatmap Matrix:

▢ Digital Readiness Score Matrix by State

State	Aadhaar Coverage	PDS Readiness	MGNREGS ABPS Dimension	MSME Density	Overall Index
Tamil Nadu	91	100	94	86	94
Andhra Pradesh	100	100	100	41	91
Telangana	100	100	95	50	91
Maharashtra	98	96	70	100	89
Puducherry	100	97	83	51	87
Punjab	100	100	82	45	86
Haryana	100	100	81	46	86
Kerala	100	99	89	21	84
Karnataka	96	100	81	35	84
Goa	100	98	79	37	84
Madhya Pradesh	95	99	59	74	82
Rajasthan	89	99	76	49	82
Himachal Pradesh	100	100	84	10	82
Tripura	99	96	88	6	80
Uttarakhand	95	100	75	23	80
Chhattisgarh	100	99	76	13	79
Jharkhand	100	93	74	26	79
Gujarat	90	98	46	81	78
Sikkim	97	97	79	6	78
Odisha	87	99	72	15	76
Bihar	76	93	70	49	76
Manipur	68	93	59	77	75
Uttar Pradesh	83	100	61	24	73
West Bengal	86	84	83	13	73
Lakshadweep	100	99	58	5	73
Andaman and Nicobar Islands	100	98	0	96	69
Mizoram	48	94	77	14	68
Jammu and Kashmir	68	97	58	0	65
Delhi	100	100	0	69	65
Chandigarh	100	93	0	69	63
Arunachal Pradesh	64	65	49	4	51
Nagaland	54	86	21	1	47
Dadra and Nagar Haveli	100	0	0	86	33
Daman and Diu	83	0	0	72	27
Meghalaya	3	35	3	0	14
Assam	0	0	22	0	7

Figure 2: Multi-dimensional Score Matrix by State

Radar Chart - Top 3 vs Bottom 3 States:
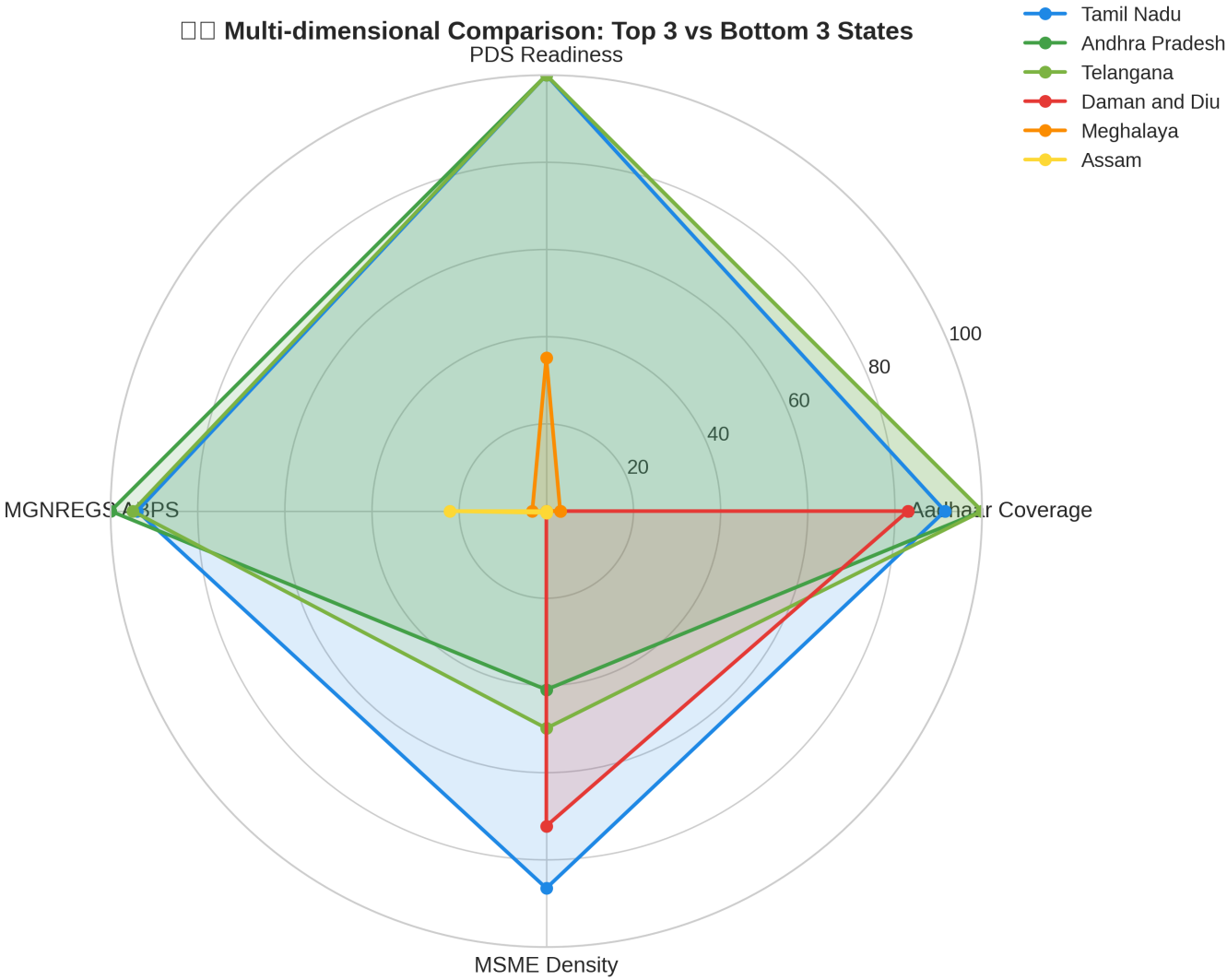
Figure 3: Multi-dimensional Comparison

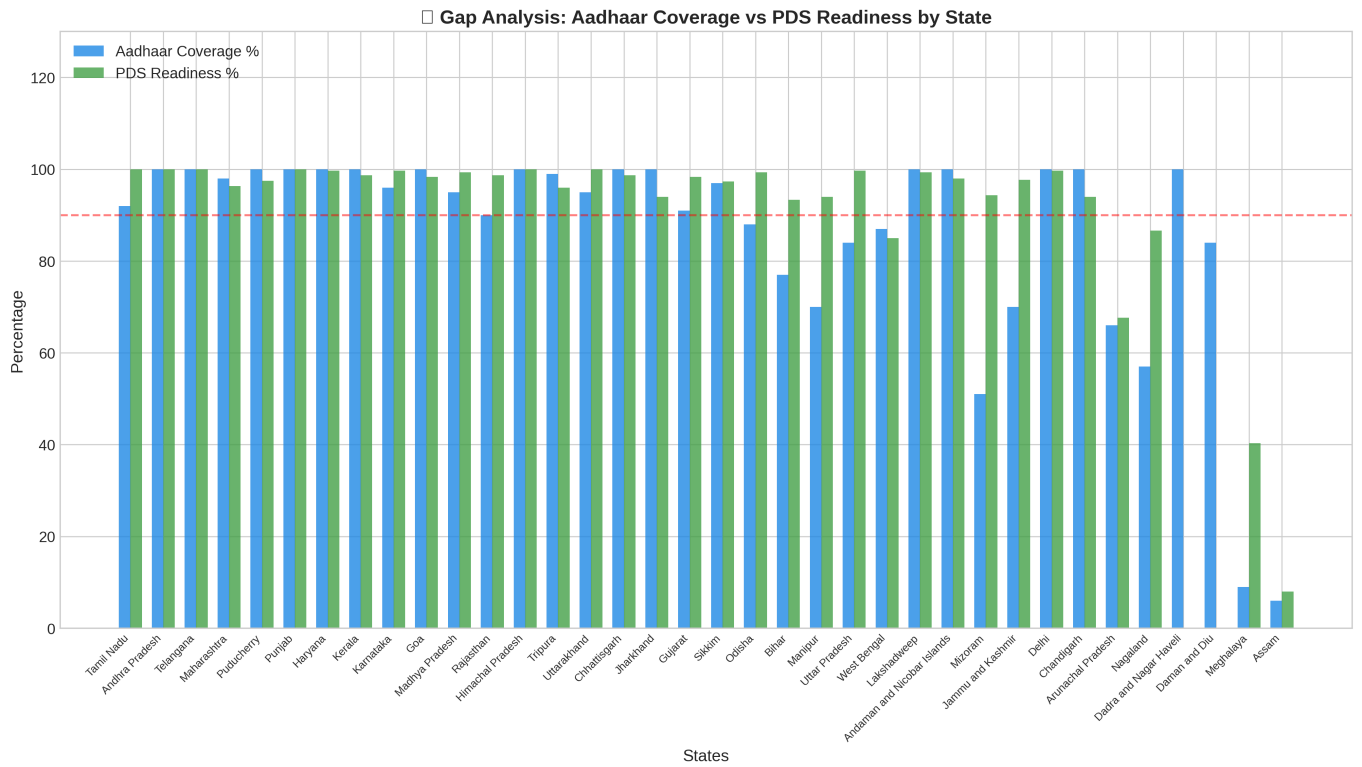Gap Analysis - Aadhaar Coverage vs PDS Readiness:

Figure 4: Gap Analysis by State

Correlation Matrix:

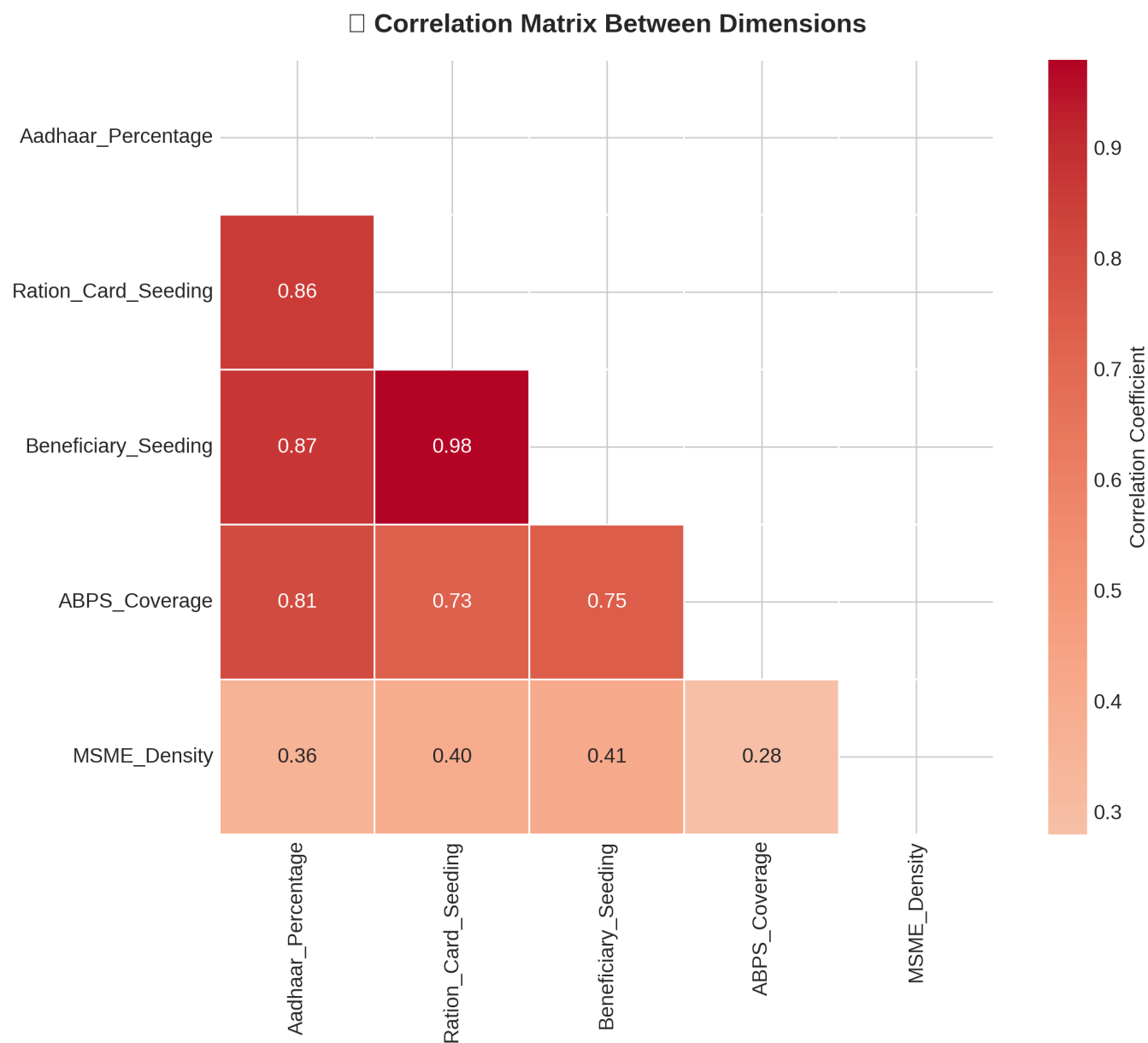*Figure 5: Correlation Between Dimensions*
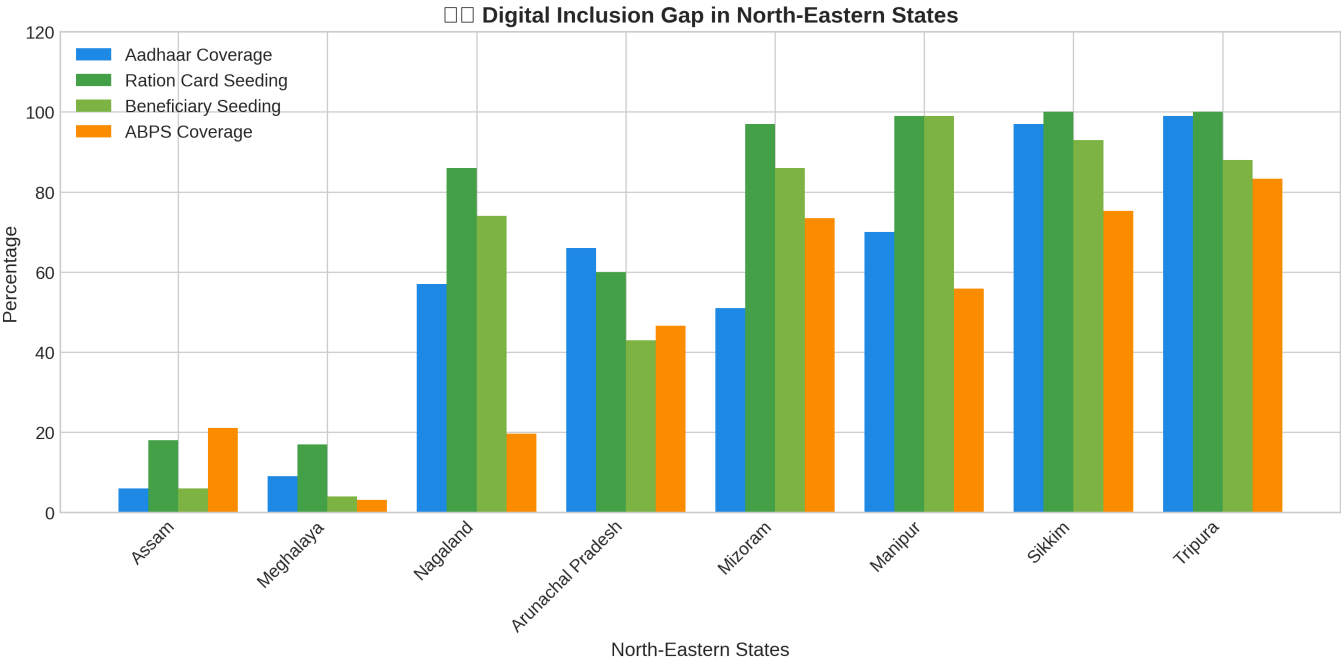
NE States Analysis:

Figure 6: Digital Inclusion Gap in North-Eastern States

MGNREGS ABPS Gap Analysis:

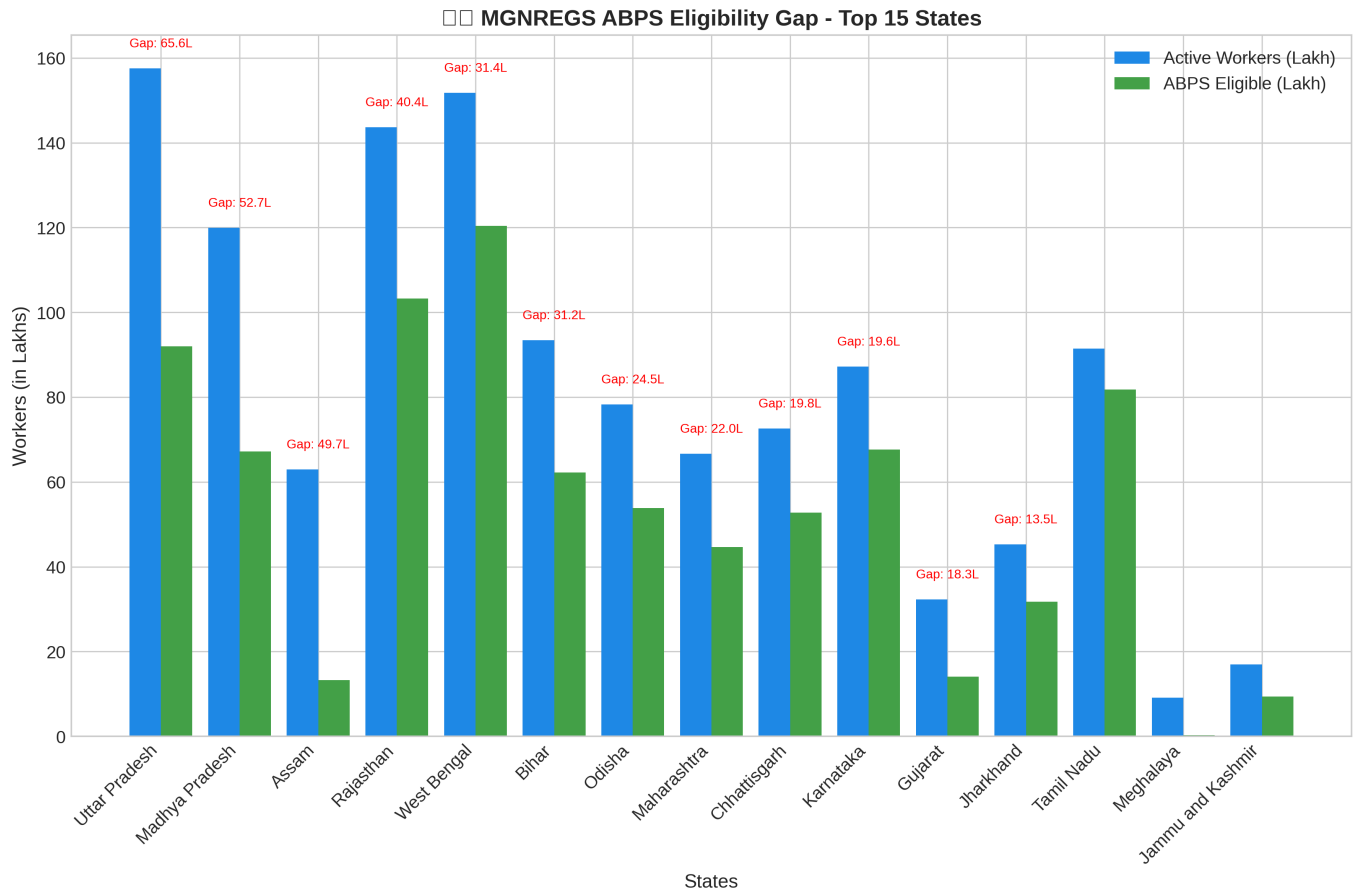**□□ MGNREGS ABPS Eligibility Gap - Top 15 States**

Figure 7: MGNREGS ABPS Eligibility Gap

4.5 Visualization Code

State Rankings Code:

```python
# State Rankings Visualization
fig, axes = plt.subplots(1, 2, figsize=(16, 8))

top_10 = master_df.head(10).sort_values('Digital_Readiness_Index')
colors_top = plt.cm.Greens(np.linspace(0.4, 0.9, 10))
axes[0].barh(top_10['State'], top_10['Digital_Readiness_Index'], color=colors_top)
axes[0].set_xlabel('Digital Readiness Index', fontsize=12)
axes[0].set_title('TOP 10 States', fontsize=14, fontweight='bold', color='green')

bottom_10 = master_df.tail(10).sort_values('Digital_Readiness_Index', ascending=False)
colors_bottom = plt.cm.Reds(np.linspace(0.4, 0.9, 10))
axes[1].barh(bottom_10['State'], bottom_10['Digital_Readiness_Index'], color=colors_bottom)
axes[1].set_xlabel('Digital Readiness Index', fontsize=12)
axes[1].set_title('BOTTOM 10 States', fontsize=14, fontweight='bold', color='red')

plt.savefig('assets/state_rankings.png', dpi=300, bbox_inches='tight')
```

Heatmap Code:

```python
# Heatmap Matrix
heatmap_data = master_df.set_index('State')[[
    'Score_Aadhaar_Coverage', 'Score_PDS_Readiness',
    'Score_MGNREGS_ABPS', 'Score_MSME_Density', 'Digital_Readiness_Index'
]]

fig, ax = plt.subplots(figsize=(12, 16))
sns.heatmap(heatmap_data, annot=True, fmt='.0f', cmap='RdYlGn',
            linewidths=0.5, ax=ax, vmin=0, vmax=100)
ax.set_title('Digital Readiness Score Matrix by State', fontsize=16)
plt.savefig('assets/heatmap_matrix.png', dpi=300, bbox_inches='tight')
```

Radar Chart Code:

```python
# Radar Chart for Top vs Bottom States
from math import pi

categories = ['Aadhaar Coverage', 'PDS Readiness', 'MGNREGS ABPS', 'MSME Density']
N = len(categories)
angles = [n / float(N) * 2 * pi for n in range(N)]
angles += angles[:1]

fig, ax = plt.subplots(figsize=(10, 10), subplot_kw=dict(polar=True))

for state in compare_states:
    state_data = master_df[master_df['State'] == state].iloc[0]
    values = [state_data['Score_Aadhaar_Coverage'],
              state_data['Score_PDS_Readiness'],
              state_data['Score_MGNREGS_ABPS'],
              state_data['Score_MSME_Density']]
    values += values[:1]
    ax.plot(angles, values, 'o-', linewidth=2, label=state)
    ax.fill(angles, values, alpha=0.15)

ax.set_xticks(angles[:-1])
ax.set_xticklabels(categories)
ax.legend(loc='upper right')
plt.savefig('assets/radar_chart.png', dpi=300, bbox_inches='tight')
```

5. Strategic Recommendations

Immediate Actions (0-3 months)

* Launch emergency Aadhaar seeding drives in Assam and Meghalaya
* Deploy mobile enrollment units to remote NE areas
* Fast-track ABPS integration for 453+ lakh unlinked workers
* Establish regional monitoring dashboards for NE states

Short-term Actions (3-6 months)

* State-specific intervention programs for bottom 10 states
* Training programs for FPS operators in low-automation states
* Awareness campaigns in vernacular languages
* Integration with other welfare schemes (DBT)

Medium-term Actions (6-12 months)

* Target 95% ration card seeding nationally
* Achieve 90% MGNREGS ABPS coverage
* Reduce regional disparity in Digital Readiness Index to under 20 points
* Quarterly progress reviews with state governments