Data Analytics for Monitoring & Evaluation of Government Schemes

Presented by: Thamizhvaanan I
Application No: SSMDA2024-1087

Case Study: Puducherry Welfare Schemes – Predictive & Preventive Insights

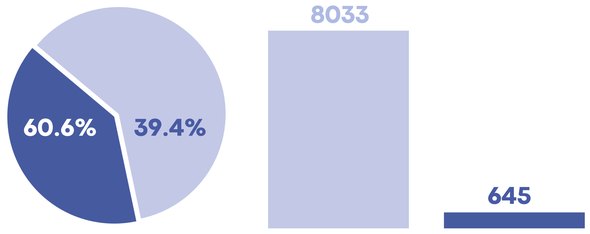
IMPACTS

- PM-KISAN Monitoring - 645 missed payments flagged (7.43% gap)
- Pension Fraud Detection - 108 duplicates, ₹53.16L risk detected
- Pension Forecasting - +27% growth, ₹367 Cr required by 2030
- RGSSS Data Validation - 1,090 real entries, no anomalies

01 PM-KISAN DBT MONITORING (2024)

- Total Registered Farmers: 8,678
- Beneficiaries Received DBT: 8,033
- Missed Payments: 645 (7.43% gap)
- Women Beneficiaries: 3,168 (39.4%)

Insight: Delivery gap requires field investigation, especially among women & smallholder farmers.



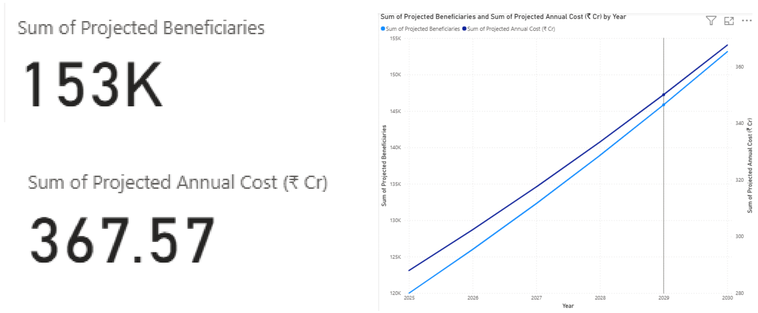
02 DUPLICATE DETECTION – OLD AGE PENSION (Simulated Dataset)

- Dataset: 5,000+ simulated entries
- Method: Fuzzy Matching (Name + DOB + Taluk)
- Duplicates Flagged: 108
- Estimated Annual Leakage: ₹53.16 Lakhs
- Real-World Reference: Delhi Pension Scam (2023), Assam Widow Scheme Fraud (2022)

03 FORECASTING DEMAND – OLD AGE PENSION (2025–2030)

- Insight: Pension demand will rise steadily. Need for strategic forecasting and proactive budget allocation.
- Assumptions: 5% annual growth, 80% coverage, seniors (2024): 1.55 lakh

Year	Projected Beneficiaries	Avg. Monthly Payout (₹)	Projected Annual Cost (₹ Cr)
2025	120000	2000	288
2026	126000	2000	302.4
2027	132300	2000	317.52
2028	138915	2000	333.4
2029	145860	2000	350.06
2030	153153	2000	367.57



Projected Rise in Beneficiaries & Cost

04 REAL-WORLD DATA VALIDATION – RGSSS (2024–2025)

- Official PDFs Processed: 4 (Oct 2024 – Apr 2025)
- Entries Parsed & Cleaned: 1,090
- Claimants Standardized: Name, Relationship, Address
- Duplicate Check Logic: Name ≥ 90%, Address ≥ 85%
- Duplicates Found: 0 — Data is clean and consistent!
- Proof: Structured CSV exported for audit
- Tools: Python (fitz, regex, pandas), Excel

TECH STACK USED

Python (Pandas, FuzzyWuzzy, Forecasting), Power BI, Excel, Streamlit

DATA SOURCES

Census, NSSO, PM-KISAN DBT, RGSSS PDF Orders

Scan to view the real RGSSS data validation, featured in my portfolio.

SUMMARY

This case study showcases how data analytics improved the monitoring and evaluation of Puducherry's welfare schemes. Key insights include a 7.43% DBT delivery gap, ₹53.16L potential pension fraud detected, a 27% projected rise in pension demand by 2030, and clean validation of 1,090 RGSSS records. Using Python, Excel, and Power BI, the project highlights how data can drive transparency, prevent leakage, and support smarter governance.