



Government of Madhya Pradesh



Madhya Pradesh State Electronics Development Corporation

(Personal Project)

Data Analytics Dashboard
for
Government Scheme Performance and MSME Impact
in
Madhya Pradesh

Exploratory Data Analysis & Analytics Report

Prepared By:

Ritik Pratap Singh Patel
Graduate Engineer Trainee
MPSEDC

Technologies Used:

Python | Power BI | Data Analytics

Date: December 2025

Contents

1. Introduction	3
2. Objectives of the Study	4
3. Data Sources.....	5
3.1 Open Government Data Platform (data.gov.in)	5
3.2 Direct Benefit Transfer (DBT) Scheme-wise Beneficiary Data	5
3.3 MSME Project Approval and Employment Data	5
3.4 Madhya Pradesh Administrative Hierarchy Dataset	6
3.5 Data Coverage and Availability Limitations	6
4. Data Understanding	7
4.1 DBT Scheme-wise Beneficiary Data	7
4.2 MSME Units, Investment, and Employment Data.....	7
4.3 Madhya Pradesh Administrative Hierarchy Data	8
4.4 Dimension Tables for Analytical Modeling.....	8
5. Data Cleaning and Preprocessing.....	9
5.1 Issues Identified in the Raw Data	9
5.2 Data Cleaning and Preprocessing Steps Performed	9
5.3 Importance of Data Cleaning	10
6. Feature Engineering	11
6.1 Fund Utilization Percentage	11
6.2 Year–Month Aggregation	11
6.3 District-wise Distribution Logic	12
6.4 MSME Growth and Trend Extension Features	12
6.5 Aggregated Performance Metrics	12
6.6 Time Dimension Features	13
7. Synthetic Data Generation and Assumptions	14
7.1 Need for Synthetic Data Generation	14
7.2 Baseline Data Used for Simulation	14
7.3 Growth Assumptions.....	14
7.4 COVID-19 Impact Modeling	15
7.5 Post-COVID Recovery Trends.....	15
7.6 Transparency and Analytical Integrity.....	15
8. Exploratory Data Analysis (EDA).....	16
8.1 Overall Beneficiary and Fund Utilization Trends	16

8.2 Fund Utilization Efficiency Over Time	16
8.3 District-wise Performance Comparison	16
8.4 District-Level Trends Over Time	17
8.5 MSME Units, Investment, and Employment Trends	17
8.6 Relationship Between MSME Investment and Employment	17
9. Key Insights	18
9.1 Consistent Growth in Beneficiary Coverage.....	18
9.2 High Levels of Fund Utilization.....	18
9.3 Impact of COVID-19 on Scheme Performance	18
9.4 District-Level Performance Variation	18
9.5 Relationship Between MSME Investment and Employment	18
10. Dashboard Overview.....	20
10.1 Executive Overview Dashboard	20
10.2 District Performance Dashboard.....	20
10.3 MSME & Employment Impact Dashboard	20
11. Limitations of the Study	21
12. Future Scope & Recommendations	22
13. Conclusion.....	23

1. Introduction

The effective implementation of government welfare schemes and economic development programs plays a crucial role in improving social well-being and driving inclusive growth. In a state like Madhya Pradesh, where multiple welfare schemes and MSME initiatives operate simultaneously across diverse districts, monitoring performance through traditional reporting methods becomes challenging.

Madhya Pradesh State Electronics Development Corporation (MPSEDC) plays an important role in supporting digital governance and data-driven decision-making for state departments. As government programs expand in scale and complexity, there is an increasing need for analytical tools that can consolidate data from multiple sources, identify trends, and provide actionable insights for administrators and policymakers.

This project focuses on developing a centralized **Data Analytics Dashboard** to evaluate the performance of selected government welfare schemes and to analyze the impact of MSME-related investments on employment generation. The dashboard has been designed using Power BI and supported by a structured analytics pipeline developed in Python.

The analysis emphasizes three critical aspects:

- Beneficiary reach and growth trends over time
- Fund allocation and utilization efficiency
- District-level performance comparison and MSME employment impact

Due to limited public availability of detailed scheme-wise datasets, the project combines official open government data with statistically simulated data based on realistic assumptions. This approach enables demonstration of end-to-end analytics workflows while maintaining transparency regarding data limitations.

The outcome of this project is a decision-support dashboard that provides stakeholders with a clear, visual, and data-driven understanding of scheme performance and MSME impact across Madhya Pradesh.

2. Objectives of the Study

The primary objective of this study is to design and develop a data-driven analytics framework that enables effective monitoring and evaluation of government welfare schemes and MSME-related initiatives in Madhya Pradesh. The study aims to support informed decision-making through structured data analysis and interactive visual dashboards.

The specific objectives of the study are as follows:

1. To analyze the overall beneficiary reach of selected government welfare schemes and identify year-wise trends in scheme coverage.
2. To evaluate fund allocation and utilization patterns in order to assess financial efficiency and identify potential gaps in implementation.
3. To perform district-level performance analysis by comparing beneficiary distribution and fund utilization across districts.
4. To examine the impact of MSME-related investments on employment generation and enterprise growth over time.
5. To develop an interactive Power BI dashboard that presents insights in a clear, intuitive, and decision-oriented manner for administrators and policymakers.
6. To demonstrate an end-to-end analytics workflow, including data collection, cleaning, feature engineering, exploratory analysis, and visualization.
7. To highlight the role of data analytics in improving transparency, accountability, and performance monitoring within government programs.

Together, these objectives provide a structured framework for evaluating both social welfare delivery and economic development initiatives in the state. By combining beneficiary analysis, financial performance assessment, and MSME impact evaluation, the study aims to present a holistic view of program effectiveness. The defined objectives also ensure that the analysis remains focused, measurable, and aligned with the broader goals of data-driven governance.

3. Data Sources

The analysis presented in this report is based exclusively on official and publicly available government datasets. All data sources were selected to ensure authenticity, reliability, and relevance to government scheme monitoring and MSME performance analysis. The datasets used in this study were obtained from recognized government platforms and departmental releases.

3.1 Open Government Data Platform (data.gov.in)

The primary source of scheme-related data was the Open Government Data (OGD) Platform of India (data.gov.in). This platform provides access to a wide range of datasets published by various ministries and departments of the Government of India. Scheme-wise and state-level beneficiary datasets related to Direct Benefit Transfer (DBT) programs were sourced from this platform.

These datasets form the foundation for analysing beneficiary trends and overall scheme reach over time.

3.2 Direct Benefit Transfer (DBT) Scheme-wise Beneficiary Data

DBT scheme-wise beneficiary data was used to analyze the reach and performance of selected welfare schemes. The available datasets included information on:

- Scheme names
- Reporting periods (month and year)
- Total number of beneficiaries

Due to the limited availability of detailed scheme-wise time-series data for recent years, the DBT data primarily covered an earlier time period. This dataset served as the baseline for beneficiary analysis and trend identification.

3.3 MSME Project Approval and Employment Data

To assess the economic impact of MSME-related initiatives, MSME project approval datasets were utilized. These datasets contained year-wise information on:

- Number of MSME units approved
- Total investment amounts
- Employment generated

The MSME data was used to analyze trends in enterprise growth, investment patterns, and employment generation across multiple years, providing an economic perspective alongside welfare scheme analysis.

3.4 Madhya Pradesh Administrative Hierarchy Dataset

An administrative hierarchy dataset for Madhya Pradesh was used to obtain district-level information. This dataset enabled geographic mapping and district-wise aggregation of scheme performance indicators. It provided standardized district identifiers and names, which were essential for consistent regional analysis and visualization.

3.5 Data Coverage and Availability Limitations

While the datasets used in this study are official and reliable, certain limitations were observed:

- Detailed scheme-wise beneficiary data was not available for all years in the public domain.
- MSME datasets were available only up to a specific time period and required extension for recent years.
- Real-time or near real-time data access was not available through public sources.

To address these limitations, statistically simulated data was generated for selected years based on historical trends and realistic assumptions. All such assumptions are explicitly documented in later sections of this report to maintain transparency and analytical integrity.

4. Data Understanding

This section provides an overview of the datasets used in the study, explaining their structure, key attributes, and level of detail. Understanding the nature and granularity of each dataset was essential to ensure correct data modeling, aggregation, and interpretation during analysis.

4.1 DBT Scheme-wise Beneficiary Data

The DBT scheme-wise beneficiary dataset contains information related to the number of beneficiaries covered under specific government welfare schemes. The dataset includes scheme names, reporting periods, and beneficiary counts.

Key fields used:

- Scheme Name
- Reporting Period (Month and Year)
- Total Number of Beneficiaries

Granularity:

- Monthly and yearly (depending on data availability)

This dataset was used to analyze beneficiary trends over time and to assess the reach of welfare schemes. Due to limited availability of recent scheme-wise data, this dataset primarily served as a baseline for beneficiary analysis.

4.2 MSME Units, Investment, and Employment Data

The MSME dataset provides year-wise information related to approved MSME projects. It captures the scale of enterprise activity and its contribution to employment generation.

Key fields used:

- Year
- Number of MSME Units Approved
- Total Investment Amount (₹ in lakh)
- Employment Generated

Granularity:

- Yearly

This dataset was used to study long-term trends in MSME growth, investment behaviour, and employment impact, enabling economic impact assessment alongside welfare scheme analysis.

4.3 Madhya Pradesh Administrative Hierarchy Data

The administrative hierarchy dataset contains geographic information related to districts within Madhya Pradesh. It was used to enable district-level analysis and visualization.

Key fields used:

- District Identifier
- District Name
- Administrative Hierarchy Attributes

Granularity:

- District level

This dataset ensured consistent geographic mapping across all analyses and supported district-wise comparison of scheme performance indicators.

4.4 Dimension Tables for Analytical Modeling

To support structured analysis and efficient data modeling, additional dimension tables were created as part of the analytics pipeline.

These included:

- **Time dimensions** (year and month-level tables) to support temporal analysis
- **Scheme dimension** to standardize scheme-related attributes
- **District dimension** to maintain geographic consistency

Granularity:

- Yearly and monthly, aligned with corresponding fact tables

These dimension tables enabled the creation of a star-schema-based data model, ensuring accurate filtering and aggregation across different analytical views.

Section Summary

Understanding the structure and granularity of each dataset helped define appropriate aggregation levels and relationships within the data model. This step ensured that subsequent analysis and visualizations were based on correctly interpreted data rather than assumptions, forming a reliable foundation for the exploratory data analysis and dashboard development.

5. Data Cleaning and Preprocessing

Raw datasets obtained from public government sources often contain inconsistencies, formatting issues, and gaps that can impact analytical accuracy if not addressed carefully. Before performing any analysis or visualization, all datasets used in this study underwent systematic data cleaning and preprocessing to ensure reliability and consistency.

5.1 Issues Identified in the Raw Data

During initial data inspection, several issues were observed across the raw datasets:

- **Inconsistent date formats:**
Reporting periods were represented in multiple formats such as financial years, abbreviated month-year combinations, and text-based period labels, making direct time-based analysis unreliable.
- **Limited and uneven time coverage:**
Scheme-wise beneficiary data was available only for specific years, while MSME datasets covered a different time range. This resulted in gaps when attempting year-on-year comparisons.
- **Non-standard numerical representations:**
Numeric fields such as beneficiary counts, investment amounts, and employment figures contained commas, mixed data types, and occasional text values.
- **Missing or incomplete records:**
Certain records contained missing values, particularly in beneficiary counts and investment-related fields, which could distort aggregate calculations.
- **Geographic inconsistencies:**
District names appeared with variations in spelling or formatting across datasets, preventing direct geographic joins.

5.2 Data Cleaning and Preprocessing Steps Performed

To address the identified issues, the following preprocessing steps were applied:

- **Standardization of date fields:**
All time-related attributes were converted into standardized year and month formats. Separate month-level and year-level time dimensions were created to align with the granularity of different fact tables.
- **Handling missing values:**
Records with critical missing values, such as beneficiary counts or investment figures, were either excluded from analysis or addressed through controlled assumptions where appropriate.
- **Numeric data normalization:**
All numerical fields were cleaned by removing non-numeric characters and converting them into consistent numeric data types to support accurate aggregation and calculation.

- **Filtering relevant time periods:**
Analysis was restricted to relevant reporting years to maintain consistency across datasets and to focus on periods meaningful for performance evaluation.
- **Geographic standardization:**
A centralized district dimension table was created to ensure consistent district identifiers and names across all datasets, enabling reliable district-level analysis.
- **Data aggregation:**
Raw records were aggregated to appropriate levels (monthly or yearly) based on analytical requirements, ensuring alignment between datasets of differing granularity.

5.3 Importance of Data Cleaning

These preprocessing steps were essential to ensure that subsequent analysis and visualizations reflected true trends rather than data artifacts. Proper data cleaning enabled:

- Accurate year-on-year and district-wise comparisons
- Reliable calculation of key performance indicators such as fund utilization percentage
- Consistent filtering and slicing across dashboard visuals
- A stable and scalable data model suitable for decision-support analytics

Without these steps, the analysis would risk producing misleading insights and unreliable conclusions.

Section Summary

Effective data cleaning and preprocessing formed the foundation of this study by transforming raw, inconsistent datasets into a structured and analysis-ready format. This ensured that all insights derived from the dashboard were based on accurate, comparable, and well-understood data.

6. Feature Engineering

Feature engineering refers to the process of creating new variables and derived metrics from raw data in order to enhance analytical depth and interpretability. In this study, multiple features were engineered to support meaningful comparisons, trend analysis, and performance evaluation across schemes, districts, and time periods.

These engineered features played a critical role in transforming raw datasets into actionable insights.

6.1 Fund Utilization Percentage

One of the most important derived metrics created during the analysis is **Fund Utilization Percentage**, which measures the efficiency of fund usage under government schemes.

Derived metric:

- $\text{Fund Utilization \%} = (\text{Funds Utilized} / \text{Funds Allocated}) \times 100$

Why it matters:

- Indicates financial efficiency of scheme implementation
- Helps identify underutilization or execution bottlenecks
- Enables comparison across districts and years regardless of absolute budget size

This metric was used extensively across executive and district-level dashboards to highlight performance variations.

6.2 Year–Month Aggregation

To support time-based analysis, a **Year–Month aggregation feature** was created by combining year and month attributes into a single temporal identifier.

Derived feature:

- Year–Month (formatted as YYYY–MM)

Why it matters:

- Enables consistent monthly trend analysis
- Aligns datasets with different time granularities
- Supports correct time-based filtering in Power BI

This aggregation was essential for modeling monthly beneficiary trends while avoiding ambiguity in time relationships.

6.3 District-wise Distribution Logic

Detailed district-level beneficiary data was not uniformly available across all datasets. To enable district-level analysis, a **controlled distribution logic** was applied.

Approach used:

- Total scheme beneficiaries were proportionally distributed across districts
- Distribution weights were applied to maintain realistic variations
- Aggregate totals were preserved to avoid distortion at state level

Why it matters:

- Enables district-wise performance comparison
- Maintains consistency with official aggregate figures
- Allows regional analysis without fabricating unrealistic data

All assumptions related to this logic are transparently documented to preserve analytical integrity.

6.4 MSME Growth and Trend Extension Features

The MSME dataset was available only up to a limited time period. To analyze recent trends, additional features were engineered to extend the dataset.

Derived features included:

- Year-on-year MSME unit growth
- Investment trend extension
- Employment growth patterns

Special considerations:

- COVID-19 impact was explicitly modeled as a dip
- Post-COVID recovery trends were incorporated
- Growth rates were applied cumulatively rather than uniformly

Why it matters:

- Enables long-term trend visualization
- Supports economic impact analysis
- Prevents misleading flat or duplicated data patterns

6.5 Aggregated Performance Metrics

Several aggregation-level features were created to support executive-level insights:

Examples:

- Total Beneficiaries (yearly and cumulative)
- Total Funds Allocated and Utilized
- Total MSME Units Approved
- Total MSME Investment and Employment

Why it matters:

- Enables high-level performance monitoring
- Simplifies complex datasets into decision-friendly indicators
- Supports KPI-based dashboard design

These aggregated metrics form the backbone of the Executive Overview dashboard.

6.6 Time Dimension Features

To ensure correct data modeling and filtering, separate time-related features were engineered:

- Year-level dimension for yearly fact tables
- Month-level dimension for monthly fact tables

Why it matters:

- Prevents many-to-many relationship issues
- Ensures accurate slicing and filtering
- Supports scalability of the data model

This design choice ensured analytical correctness across multiple fact tables with different granularities.

Section Summary

The feature engineering process significantly enhanced the analytical value of the datasets by introducing meaningful metrics, resolving granularity mismatches, and enabling consistent comparisons across dimensions. These engineered features transformed raw government data into a structured analytical framework capable of supporting performance evaluation and decision-making.

7. Synthetic Data Generation and Assumptions

Publicly available government datasets often have limitations in terms of coverage, granularity, and continuity across years. In this study, synthetic data generation was selectively employed to address such limitations while preserving analytical continuity and interpretability. This section clearly documents the rationale, methodology, and assumptions used for generating synthetic data to maintain transparency and analytical integrity.

7.1 Need for Synthetic Data Generation

During data exploration, it was observed that detailed scheme-wise beneficiary data and MSME datasets were available only for limited time periods in the public domain. Specifically:

- Scheme-wise DBT beneficiary data was available primarily for early reporting periods.
- MSME project approval and employment data was available up to the financial year 2013–14.

However, meaningful performance evaluation and trend analysis require continuous time-series data across recent years. To enable year-on-year comparisons, trend visualization, and dashboard functionality for recent periods, synthetic data was generated for selected years where official data was not publicly accessible.

Synthetic data was used strictly to **extend trends**, not to replace or contradict officially available data.

7.2 Baseline Data Used for Simulation

All synthetic data generation was anchored to official baseline datasets to ensure realism and consistency.

- **DBT Scheme Baseline:**
Official scheme-wise DBT beneficiary data from the year 2013 was used as the reference point for beneficiary trend extension.
- **MSME Baseline:**
Official MSME units, investment, and employment data from the period 2004–2014 was used as the baseline for extending MSME-related trends into subsequent years.

Using real historical data as the foundation ensured that simulated values followed realistic patterns observed in official records.

7.3 Growth Assumptions

To extend the datasets beyond available years, controlled growth assumptions were applied based on typical trends observed in government program expansion and economic activity.

Key assumptions included:

- Gradual year-on-year growth in beneficiary counts and MSME activity during normal periods.

- Growth rates were applied incrementally rather than as flat repetitions to avoid artificial plateaus.
- Randomized but bounded variation was introduced to prevent uniform or unrealistic patterns.

These assumptions were intentionally conservative to avoid overstating performance or impact.

7.4 COVID-19 Impact Modeling

The COVID-19 pandemic had a significant impact on government operations, economic activity, and program implementation across the country. To reflect this disruption:

- The year 2020 was modelled with a **temporary dip** in beneficiary growth, MSME investment, and employment generation.
- Reduced fund utilization efficiency was incorporated to reflect operational challenges during the pandemic period.

This modeling ensures that the dashboards realistically reflect the disruption observed during this period rather than presenting an uninterrupted growth curve.

7.5 Post-COVID Recovery Trends

Following the pandemic-related slowdown, recovery trends were incorporated for subsequent years.

Assumptions included:

- Gradual recovery in beneficiary coverage and MSME activity during 2021.
- Accelerated but controlled growth during 2022, reflecting economic normalization and policy interventions.

These recovery patterns were applied consistently across relevant datasets to maintain coherence in trend analysis.

7.6 Transparency and Analytical Integrity

All synthetic data generated as part of this study:

- Is clearly documented and distinguishable from official data.
- Preserves aggregate consistency with baseline datasets.
- Is used solely for analytical demonstration and decision-support visualization.

The purpose of synthetic data generation in this project is to enable end-to-end analytics workflows and performance monitoring demonstrations, not to represent official or audited figures.

8. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted to identify trends, patterns, and variations in government scheme performance and MSME-related economic impact. The analysis focuses on beneficiary growth, fund utilization efficiency, district-level performance differences, and MSME investment-employment relationships. Visual analytics were used extensively to support interpretation and decision-oriented insights.

8.1 Overall Beneficiary and Fund Utilization Trends

The executive-level analysis reveals a clear upward trend in total beneficiaries over the observed period. The yearly trend indicates steady growth in beneficiary coverage, reflecting expansion and improved reach of welfare schemes over time.

A noticeable dip is observed in the year **2020**, corresponding to the COVID-19 pandemic. During this period:

- Growth in beneficiaries slowed
- Fund utilization showed a temporary decline
- Operational disruptions affected implementation efficiency

Post-2020, the trend demonstrates recovery and acceleration, indicating restoration of administrative capacity and increased program uptake. Fund utilization closely follows beneficiary growth, suggesting a strong alignment between financial allocation and actual scheme delivery.

Dashboard Reference: Executive Overview – Yearly Beneficiaries and Funds Trend

8.2 Fund Utilization Efficiency Over Time

Analysis of fund utilization percentage highlights consistently high utilization levels across most years, generally exceeding 90%. This indicates effective deployment of allocated funds and relatively low levels of underspending.

However, a slight decline in utilization is observed during the pandemic period, reinforcing the impact of external disruptions on execution efficiency. The recovery in utilization percentage in subsequent years reflects stabilization of operations and improved financial management.

Dashboard Reference: Executive Overview – Fund Utilization % by Year

8.3 District-wise Performance Comparison

District-level analysis reveals significant variation in beneficiary coverage and fund utilization across districts. While some districts consistently show higher beneficiary counts and strong utilization efficiency, others lag behind, indicating uneven scheme penetration.

Top-performing districts demonstrate:

- Higher beneficiary coverage

- Efficient fund utilization
- Consistent year-on-year growth

Conversely, lower-performing districts may require targeted administrative attention, improved outreach mechanisms, or operational support to enhance scheme effectiveness.

The district-wise comparison highlights the importance of geographic disaggregation in identifying localized performance gaps that are not visible in aggregate state-level analysis.

Dashboard Reference: District Performance – Top Districts by Beneficiaries and Utilization Table

8.4 District-Level Trends Over Time

Trend analysis at the district level shows that most districts follow a similar growth trajectory, with a synchronized dip around 2020. This indicates that external factors, rather than district-specific issues, were the primary cause of the slowdown during that period.

Post-pandemic recovery trends vary slightly across districts, suggesting differences in the speed of administrative recovery and scheme uptake. Such variations provide an opportunity for comparative analysis and knowledge transfer between districts.

Dashboard Reference: District Performance – Beneficiaries by Year and District

8.5 MSME Units, Investment, and Employment Trends

The MSME-focused analysis highlights a strong positive relationship between MSME investment and employment generation. Over time, increases in investment levels are associated with corresponding growth in employment, underscoring the role of MSME initiatives in economic development.

A clear slowdown in MSME activity is observed during 2020, reflected by reduced investment and lower employment generation. This aligns with broader economic disruptions caused by the pandemic.

Subsequent years show recovery and renewed growth, indicating resilience in the MSME sector and the effectiveness of supportive measures introduced during the post-COVID period.

Dashboard Reference: MSME & Employment Impact – Investment and Employment Trends

8.6 Relationship Between MSME Investment and Employment

Scatter plot analysis reveals a strong correlation between MSME investment and employment generation. Years with higher investment levels tend to correspond with greater employment outcomes, reinforcing the importance of sustained financial support for MSME initiatives.

The size variation in MSME units across years further illustrates how enterprise creation contributes to employment outcomes, providing a multi-dimensional view of economic impact.

Dashboard Reference: MSME & Employment Impact – Investment vs Employment Scatter Plot

9. Key Insights

This section summarizes the most important findings derived from the exploratory data analysis and dashboard visualizations. The insights are presented in a non-technical manner to support understanding by decision-makers and administrative stakeholders.

9.1 Consistent Growth in Beneficiary Coverage

The analysis indicates a steady increase in total beneficiaries over the observed period, demonstrating expanded reach and improved penetration of government welfare schemes. This trend suggests progressive enhancement in scheme awareness, accessibility, and implementation effectiveness across the state.

9.2 High Levels of Fund Utilization

Fund utilization levels remained consistently high across most years, generally exceeding 90 percent. This reflects efficient financial execution and limited underspending. The alignment between fund allocation and utilization indicates that allocated resources are largely translating into on-ground delivery.

9.3 Impact of COVID-19 on Scheme Performance

A temporary slowdown is clearly visible during the year 2020, coinciding with the COVID-19 pandemic. This period shows:

- Reduced growth in beneficiaries
- Slight decline in fund utilization efficiency
- Lower MSME investment and employment generation

The recovery observed in subsequent years highlights resilience in administrative systems and the effectiveness of post-pandemic recovery measures.

9.4 District-Level Performance Variation

Significant variation exists in scheme performance across districts. While several districts consistently demonstrate higher beneficiary coverage and strong utilization efficiency, others lag behind. These differences highlight the importance of district-level monitoring to identify areas requiring targeted administrative support or improved outreach.

9.5 Relationship Between MSME Investment and Employment

The MSME analysis reveals a strong positive relationship between investment levels and employment generation. Years with higher MSME investment correspond to increased employment

outcomes, reinforcing the role of MSME initiatives as a key driver of economic development and job creation.

Section Summary

Overall, the insights emphasize effective scheme execution, efficient financial utilization, and meaningful economic impact through MSME initiatives, while also identifying opportunities for targeted district-level improvements.

10. Dashboard Overview

This section provides a brief overview of the Power BI dashboard developed as part of this study, outlining the purpose of each page and how stakeholders can use it for decision-making.

10.1 Executive Overview Dashboard

The Executive Overview page presents a consolidated view of scheme performance at the state level. It includes key performance indicators such as total beneficiaries, total funds allocated and utilized, utilization percentage, and year-wise trends.

Purpose:

- Enable senior officials to quickly assess overall performance
- Identify broad trends and disruptions (such as the COVID-19 impact)
- Monitor financial efficiency at a high level

This page is intended for rapid, high-level review and strategic oversight.

10.2 District Performance Dashboard

The District Performance page focuses on geographic comparison of scheme implementation. It highlights district-wise beneficiary coverage, fund utilization efficiency, and comparative rankings.

Purpose:

- Identify high-performing and underperforming districts
- Support targeted administrative interventions
- Enable comparative analysis across regions

This page is particularly useful for district-level monitoring and operational planning.

10.3 MSME & Employment Impact Dashboard

The MSME & Employment Impact page analyzes economic outcomes related to MSME initiatives. It visualizes trends in MSME units approved, investment levels, and employment generation over time.

Purpose:

- Assess the economic impact of MSME-related initiatives
- Understand the relationship between investment and employment
- Monitor recovery patterns following economic disruptions

This page supports economic planning and policy evaluation related to enterprise development.

11. Limitations of the Study

While the analysis provides meaningful insights into government scheme performance and MSME impact, certain limitations must be acknowledged. These limitations arise primarily from data availability constraints and the scope of publicly accessible datasets.

- **Limited availability of detailed scheme-wise data:**
Publicly available datasets did not provide continuous, detailed scheme-wise beneficiary data for all recent years. As a result, in-depth comparative analysis across multiple schemes could not be performed.
- **Use of simulated data for continuity:**
To enable time-series analysis and dashboard functionality for recent years, synthetic data was generated based on historical trends and realistic assumptions. While this approach supports analytical demonstration, the simulated values do not represent official audited figures.
- **Lack of real-time or live data feeds:**
The analysis is based on static datasets obtained from open government sources. Real-time monitoring of scheme performance and MSME activity was not possible due to the absence of live data integration.
- **Limited scope of economic indicators:**
The MSME analysis focuses primarily on units approved, investment, and employment generation. Other economic indicators such as sector-wise distribution, credit flow, or survival rates were outside the scope of this study.

These limitations do not undermine the analytical approach but instead highlight opportunities for enhancement with improved data access and system integration.

12. Future Scope & Recommendations

The analytics framework and dashboard developed in this study can be significantly enhanced with expanded data access and advanced analytical capabilities. The following recommendations outline potential directions for future development:

- **Integration with real-time databases:**
Connecting the dashboard to live data sources such as departmental MIS systems or DBT platforms would enable near real-time monitoring of scheme performance and fund utilization.
- **Expansion to additional welfare schemes:**
Incorporating data from a broader range of government schemes would allow comparative analysis across programs and improve the comprehensiveness of performance evaluation.
- **Predictive analytics and forecasting:**
Machine learning models could be introduced to forecast beneficiary demand, fund requirements, and MSME employment trends, supporting proactive planning and resource allocation.
- **Grievance and feedback analytics integration:**
Integrating grievance redressal and beneficiary feedback data could provide qualitative insights into scheme effectiveness and citizen satisfaction.
- **Enhanced geographic and demographic analysis:**
Future iterations could include block-level analysis and demographic segmentation to further refine targeting and impact assessment.

These enhancements would transform the dashboard from a descriptive analytics tool into a predictive and prescriptive decision-support system.

13. Conclusion

This study demonstrates the practical application of data analytics in evaluating government welfare scheme performance and assessing the economic impact of MSME initiatives in Madhya Pradesh. By consolidating data from multiple official sources and applying structured analytical techniques, the project illustrates how complex administrative data can be transformed into meaningful, decision-support insights.

The developed Power BI dashboard provides a clear and interactive view of beneficiary reach, fund utilization efficiency, district-level performance variations, and MSME-related employment outcomes. Through visual analytics, the dashboard enables stakeholders to move beyond static reporting and engage with data dynamically, supporting evidence-based governance.

Despite limitations in publicly available data, the use of transparent and well-documented assumptions ensured continuity in analysis while maintaining analytical integrity. The project highlights how even with constrained data environments, robust analytical frameworks can be designed to support monitoring, evaluation, and strategic planning.

For MPSEDC, this project reinforces the value of data-driven approaches in strengthening digital governance initiatives. The analytics framework and dashboard can serve as a foundation for future integrations with real-time systems, expanded scheme coverage, and advanced analytical capabilities.

Overall, the study underscores the role of data analytics as a critical enabler for informed decision-making, operational efficiency, and improved public service delivery.