



STANDARD OPERATING PROCEDURE FOR DATA-LED REMOTE AUDIT (DLRA) OF REGISTRATION AND STAMPS DEPARTMENT

COMPTROLLER AND AUDITOR GENERAL OF INDIA

CONTENTS

Foreword.....	3
Acknowledgement.....	4
CHAPTER 1: Executive Summary.....	5
CHAPTER 2: Background and Scope.....	7
CHAPTER 3: Data-Led Remote Audit Methodology.....	13
CHAPTER 4: Data Acquisition and Management Framework.....	19
CHAPTER 5: Data Analytics Framework.....	23
CHAPTER 6: Root Cause Analysis and System Audit.....	28
CHAPTER 7: Revenue Impact Quantification Methodology.....	30
CHAPTER 8: Audit Observations and Reporting Framework.....	31
CHAPTER 9: Continuous Audit and Quarterly Risk Monitoring Framework.....	33
CHAPTER 10: Key Performance Indicators.....	34
CHAPTER 11: Benefits and Way Forward reflected in	
○ Case Study of State of Uttar Pradesh.....	35
○ Case Study of Telangana.....	72
ANNEXURE.....	87

FOREWORD



The Registration and Stamps Department plays a vital role in State revenue administration through the collection of stamp duty, registration fees, and other related levies, while also providing legal recognition to property transactions. The increasing digitisation of registration processes has resulted in the availability of extensive electronic data, creating new opportunities for technology-enabled and data-driven auditing.

The Indian Audit and Accounts Department has been progressively adopting data analytics, automation, and remote auditing techniques to enhance audit coverage, efficiency, and effectiveness. Experience gained from such audits has demonstrated the value of analytics-based approaches in identifying revenue leakages, compliance deficiencies, systemic weaknesses as well as providing enhanced assurance in respect of underlying transactions.

This Standard Operating Procedure (SOP) on Data-Led Remote Audit (DLRA) of Registration and Stamps Departments provides a structured framework for conducting risk-based, evidence-driven, and technology-enabled audits. The SOP aims to promote uniformity in audit practices, strengthen the use of data analytics, and support effective assessment of revenue administration and internal controls.

I hope that this SOP will serve as a useful reference for audit offices and contribute towards strengthening accountability, transparency, and sound public financial management. The continued refinement of this framework through field experience, feedback and innovation will further strengthen the Department's ability to provide independent and credible assurance in an increasingly digital governance environment.

Date:

Place: New Delhi

(Shri K. Sanjay Murthy)
Comptroller and Auditor General of India

ACKNOWLEDGEMENT

The Indian Audit and Accounts Department (IA&AD) acknowledges the contributions of all officers and staff whose dedication, innovation, and professional commitment have contributed to the development of this Standard Operating Procedure (SOP) on Data-Led Remote Audit of the Registration and Stamps Department.

IA&AD places on record its appreciation for the pioneering efforts of the Office of the Principal Accountant General (Audit)-II, Uttar Pradesh, and Lucknow. The conceptualisation, development, and successful implementation of the Data-Led Remote Audit framework were carried out under the leadership of Shri Rajiv Kumar Pandey, Principal Accountant General (Audit)-II, Uttar Pradesh, with supervision by Shri Akash Chaudhary, Deputy Accountant General. The contributions of Shri D. P. Singh, Senior Audit Officer, Shri Durgesh Awasthi, Assistant Audit Officer, Shri Shubham Yadav, Assistant Audit Officer and Shri Kamlesh Kumar Singh, Senior Auditor, are gratefully acknowledged.

IA&AD also acknowledges the valuable contribution of the team from the Office of the Accountant General (Audit), Telangana, comprising Shri C. Sachin, Senior Audit Officer, Shri Bh. V. Ramana Murthy, Assistant Audit Officer and Smt. Ch. Geetanjali, Assistant Audit Officer, under the guidance of Smt. P. Madhavi, Principal Accountant General (Audit), Telangana. Their inputs and support have enriched the framework and facilitated its wider applicability across audit offices.

CHAPTER -1

EXECUTIVE SUMMARY

The Registration and Stamps Department is a major contributor to State Government revenues through the collection of stamp duty, registration fees, and other related charges. With the widespread adoption of digital registration platforms such as the National Generic Document Registration System (NGDRS) and various State-specific applications, substantial volumes of transaction data are now available in electronic form. This transformation presents significant opportunities for modernising audit methodologies and enhancing audit effectiveness.

The Standard Operating Procedure (SOP) for Data-Led Remote Audit (DLRA) has been developed to provide a structured framework for conducting technology-enabled, risk-based audits of Registration and Stamps Departments. The SOP seeks to institutionalise the use of data analytics, digital records, artificial intelligence, and remote audit techniques to improve audit coverage, strengthen assurance and identify systemic weaknesses affecting revenue administration.

Unlike traditional audits that rely heavily on physical inspections and sample-based examination of records, the DLRA framework promotes full-population analysis of digital transactions. By leveraging analytical tools and automated risk assessment techniques, auditors can identify high-risk transactions, detect undervaluation of properties, uncover short levy of stamp duty and registration fees and evaluate the effectiveness of system controls embedded within registration applications.

The SOP adopts a risk-based audit philosophy that focuses audit attention on areas and transactions presenting higher potential for revenue leakage and non-compliance. Audit planning is guided by data-driven risk assessment covering transaction volumes, revenue trends, urban development patterns, previous audit findings, citizen complaints, and other risk indicators. Analytical procedures are then applied to the complete population of transactions to generate audit leads and identify anomalies requiring detailed examination.

The framework emphasises the detection of individual irregularities along with the identification of their root causes. Audit observations are expected to assess whether deficiencies arise from software design weaknesses, inadequate internal controls, poor data quality, ineffective supervision, gaps in

policy, or administrative shortcomings. This approach enables audit recommendations to address systemic issues and promote sustainable improvements in governance and revenue administration.

The SOP further establishes comprehensive procedures for data acquisition, validation, standardisation, analytics, evidence gathering, and revenue impact quantification, reporting and follow-up. It encourages the use of advanced technologies such as Optical Character Recognition (OCR), Artificial Intelligence (AI), Geographic Information Systems (GIS), and automated audit algorithms to enhance audit efficiency and analytical capability.

An important feature of the DLRA framework is the introduction of continuous audit and periodic risk monitoring mechanisms. By utilising real-time or near real-time data, audit offices can move beyond periodic inspections towards a proactive model of assurance that enables early identification of emerging risks and in time reporting.

The implementation of this SOP is expected to significantly enhance audit coverage, improve detection of revenue leakages, strengthen internal controls, facilitate evidence-based decision-making, and support the development of a sustainable technology-enabled audit ecosystem. Ultimately, the framework aims to contribute to improved revenue administration, greater transparency, stronger governance, and enhanced public confidence in the functioning of Registration and Stamps Departments.

The SOP is intended to serve as a practical guide for audit offices across the Indian Audit and Accounts Department and to promote a uniform, consistent, and effective approach to Data-Led Remote Audits throughout the country.

The SOP identifies difference in the system architecture that may be present in different States & Union Territories and incorporates the efforts in DLRA in the States of Telangana and Uttar Pradesh.

This SOP prescribes a common audit framework. Individual State Audit Offices shall adapt the audit checks, datasets, legal references, valuation mechanisms, software architecture and risk indicators to suit State-specific Registration systems.

CHAPTER 2

BACKGROUND AND SCOPE

BACKGROUND

The registration and stamping of instruments relating to immovable property form an integral part of the revenue administration framework of the State. The registration process is primarily governed by the Indian Registration Act, 1908, which prescribes the documents that are compulsorily registrable, the procedures for registration, and the maintenance and preservation of records. Registration serves to provide legal validity, public notice of transactions, and protection of property rights.

The levy and collection of stamp duty are governed by the Indian Stamp Act, 1899, together with State-specific amendments, rules, notifications, and valuation guidelines issued by State Governments. The assessment of stamp duty and registration fees is largely based on notified market values, circle rates, ready reckoner rates, or guideline values prescribed for different categories of properties and transactions. The regulatory framework is further supplemented by departmental instructions, valuation manuals, and policies governing exemptions, concessions, and special categories of instruments.

In recent years, most States have adopted digital registration platforms and integrated them with e-stamping systems, land records databases, and other applications containing details of relevant information and transactions. These developments have improved efficiency and transparency as well as increased the complexity of revenue administration. This has led to identification of new challenges relating to valuation, compliance, system controls, and data integrity. Consequently, there is a need for change in audit approaches capable of examining both transaction-level compliance and the effectiveness of the digital systems and controls established under the legal and regulatory framework governing registration and stamp duty administration.

SCOPE

The scope of Data-Led Remote Audit shall include examination of all digital systems, processes, databases, records, and transactions maintained by the Registration and Stamps Department that have a bearing on government revenue, legal compliance, operational efficiency, and system effectiveness.

Figure-1: Pictorial presentation of the Scope of Data-Led Remote Audit (DLRA) of Registration and Stamps Department



A. COVERAGE OF OFFICES AND ORGANISATIONAL UNITS

The audit may cover:

1. Office of the Inspector General of Registration.
2. Registration Directorates.

3. District Registrar Offices.
4. Sub-Registrar Offices (SROs).
5. Any other offices involved in registration and stamp duty administration.

B. COVERAGE OF DIGITAL SYSTEMS AND APPLICATIONS

The audit may include examination of:

1. National Generic Document Registration System (NGDRS).
2. State-specific Registration and Stamps applications.
3. Centralised registration databases.
4. Property valuation and market value determination systems.
5. E-stamping and stamp duty management systems.
6. Digital document repositories and archival systems.
7. Revenue accounting and management information systems.
8. Workflows including approval management systems.
9. Citizen service portals and online registration platforms.

C. COVERAGE OF REGISTRABLE INSTRUMENTS

The audit may encompass all categories of instruments liable to stamp duty and registration fees, including:

1. Sale deeds.
2. Gift deeds.
3. Mortgage deeds.
4. Lease deeds.
5. Release deeds.
6. Settlement deeds.
7. Partition deeds.
8. Exchange deeds.
9. Development agreements.
10. Powers of Attorney.
11. Conveyance deeds.
12. Partnership and dissolution deeds.
13. Trust deeds.
14. Any other instrument notified under the Registration Act, Stamp Act, or applicable State laws.

D. PERIOD OF COVERAGE

The audit may cover:

1. Current year transactions.
2. Historical transactions within the prescribed audit period.
3. Real-time or near real-time transaction data, wherever continuous auditing mechanisms are implemented.

E. COVERAGE OF REVENUE AND COMPLIANCE ASPECTS

The audit may examine:

1. Assessment and collection of stamp duty.
2. Assessment and collection of registration fees.
3. Exemptions, concessions, and rebates as may be applicable.
4. Under-valuation and revenue leakage risks.
5. Compliance with statutory provisions, rules, notifications, and government orders.
6. Timeliness and completeness of revenue remittances.

F. COVERAGE OF SYSTEM CONTROLS AND GOVERNANCE

The audit may assess:

1. Application controls.
2. User access management controls.
3. Role-based access controls.
4. Workflow and approval controls.
5. Change management controls.
6. Audit trail mechanisms.
7. Data security and confidentiality controls.
8. Backup and disaster recovery arrangements.
9. System monitoring and exception management controls.

G. COVERAGE OF DATA MANAGEMENT PROCESSES

The audit may include examination of:

1. Master data creation and maintenance.
2. Property valuation master data.
3. Guideline value databases.
4. Data quality and integrity controls.
5. Data validation mechanisms.
6. Duplicate record controls.
7. Data migration and archival processes.

H. COVERAGE OF BUSINESS PROCESSES

The audit may examine:

1. Registration workflow processes.
2. Document scrutiny and approval processes.
3. Valuation and assessment procedures.
4. Mutation and land-record integration processes.
5. Revenue accounting and reconciliation procedures.
6. Generation and monitoring of management reports.
7. Citizen Service delivery processes.
8. Compliance to provisions of Right to Service Act & Rules thereof.

I. USE OF EXTERNAL DATA SOURCES

Where feasible, audit may utilise external databases and independent data sources for validation, reconciliation, benchmarking, and risk assessment, including:

1. Land Records and Revenue Department databases.
2. Municipal and Urban Local Body records.
3. Development Authority databases.
4. Property tax databases.
5. Geographic Information Systems (GIS).
6. Survey and settlement records.
7. E-stamping databases.
8. Court and legal databases, where relevant.
9. Other Government databases relevant to property transactions and revenue administration.

J. COVERAGE OF ANALYTICS-BASED AUDIT EXAMINATION

The audit may utilise data analytics to:

1. Identify instances of under-valuation.
2. Detect short levy or non-levy of stamp duty and registration fees.
3. Identify duplicate, anomalous, or suspicious transactions.
4. Detect unusual patterns, trends, and outliers.
5. Assess compliance with prescribed business rules.
6. Quantify potential revenue leakage.

K. SYSTEM-GENERATED OUTPUTS AND REPORTING

The audit may examine:

1. System-generated reports.
 2. Exception reports.
 3. Revenue collection reports.
 4. Valuation reports.
 5. Management Information System (MIS) reports.
 6. Dashboard analytics and monitoring reports.
 7. Audit logs and transaction histories.
-

CHAPTER-3

DATA-LED REMOTE AUDIT METHODOLOGY

The Data-Led Remote Audit (DLRA) framework shall be implemented through the following sequential steps:

STEP 1: STRATEGIC UNDERSTANDING OF THE DOMAIN

1. Obtain an overview of the Registration and Stamps Department's organisational structure, business processes, digital systems, and revenue streams.
2. Identify key applications, databases, and sources of audit-relevant information.
3. Understand applicable legal provisions, rules, notifications, and valuation frameworks governing registration and stamp duty administration.

STEP 2: RISK ASSESSMENT AND AUDIT PLANNING

1. Conduct a strategic assessment of the registration ecosystem.
2. Identify areas with higher revenue, compliance, operational, or systemic risks.
3. Develop a risk profile of districts, regions, and Sub-Registrar Offices (SROs) using factors such as:
 - Revenue collection trends.
 - Transaction volumes.
 - Urbanisation and market activity.
 - Concentration of high-value transactions.
 - Previous audit findings.
 - Citizen complaints and grievances.
 - Emerging risks identified through data analytics.
4. Prioritise audit units based on risk scores and materiality considerations.

STEP 3: SELECTION OF AUDIT UNITS

1. Select districts, registration offices, or transaction categories based on the risk assessment.
2. Document the rationale for selection.
3. Determine the audit period and scope of coverage.

Based on the risk assessment, audit units may be classified as:

Risk Category	Description
High Risk	Significant revenue exposure and elevated probability of irregularities
Medium Risk	Moderate revenue exposure and control risks
Low Risk	Relatively lower risk profile and stable compliance history

STEP 4: DATA ACQUISITION AND PREPARATION

1. Obtain data from relevant registration systems and databases.
2. Collect supporting master data, valuation tables, guideline values, and workflow records.
3. Verify data completeness, accuracy, consistency, and integrity.
4. Standardise and cleanse data for analytical processing.
5. Document data limitations, if any.

STEP 5: FULL-POPULATION TRANSACTION ANALYSIS

1. Analyse the complete universe of available transactions within the selected audit scope.
2. Minimise dependence on statistical or judgmental sampling wherever full-population analysis is feasible.
3. Apply data analytics techniques to identify:
 - Revenue leakage risks.
 - Under-valuation of properties.
 - Short levy of stamp duty.
 - Incorrect registration fees.
 - Duplicate or unusual transactions.
 - Outlier transactions.
 - Pattern deviations and anomalies.
 - Deviations from prescribed timelines.

STEP 6: RISK INDICATOR TESTING AND LEAD GENERATION

1. Apply predefined risk indicators and audit algorithms.
2. Screen all transactions against established risk parameters.
3. Generate exception reports and audit leads.

4. Rank identified exceptions based on risk and materiality.

STEP 7: DETAILED EXAMINATION OF FLAGGED TRANSACTIONS

1. Subject flagged transactions to detailed scrutiny.
2. Examine:
 - Digital records.
 - Registered documents.
 - Scanned images.
 - Approval records.
 - Supporting evidence.
3. Validate findings through corroborative records and external databases wherever necessary.
4. Quantify revenue implications arising from identified irregularities.

STEP 8: VALIDATION THROUGH EXTERNAL DATA SOURCES

1. Cross-verify audit findings using:
 - Land records.
 - Municipal records.
 - Development Authority databases.
 - Geographic Information Systems (GIS).
 - Property tax databases.
 - Other government databases.
2. Assess consistency between departmental records and external sources.

STEP 9: ROOT CAUSE ANALYSIS

1. Analyse the underlying causes of significant audit observations.
2. Determine whether deficiencies arose due to:
 - Software design weaknesses.
 - Inadequate system controls.
 - Deficient business processes.
 - Weak supervisory mechanisms.
 - Inadequate training and capacity building.
 - Poor master data management.
 - Policy or regulatory gaps.
3. Assess the systemic impact of identified deficiencies.

STEP 10: ASSESSMENT OF SYSTEMIC RISKS AND CONTROL ENVIRONMENT

1. Evaluate the effectiveness of:
 - Application controls.

- Workflow controls.
 - Validation mechanisms.
 - User access controls.
 - Monitoring and exception management systems.
2. Identify systemic vulnerabilities contributing to recurring irregularities.

STEP 11: FORMULATION OF AUDIT FINDINGS AND RECOMMENDATIONS

1. Develop evidence-based audit observations.
2. Quantify financial and revenue impacts wherever possible.
3. Highlight systemic issues and governance concerns.
4. Recommend corrective, preventive, and long-term remedial measures.
5. Prioritise recommendations according to risk and impact.

STEP 12: REPORTING AND COMMUNICATION

1. Prepare audit reports supported by analytical evidence.
2. Present significant findings, risk assessments, and revenue implications.
3. Communicate systemic weaknesses and recommended reforms.
4. Share dashboards, exception reports, and management information as appropriate.

STEP 13: FOLLOW-UP AND CONTINUOUS MONITORING

1. Monitor implementation of audit recommendations.
2. Track corrective actions taken by the Department.
3. Update risk indicators and audit algorithms based on audit outcomes.
4. Facilitate continuous auditing through automated monitoring and periodic analytics.

STEP 14: CONTINUOUS IMPROVEMENT OF THE DLRA FRAMEWORK

1. Review effectiveness of audit methodologies and analytical models.
2. Refine risk indicators and audit algorithms based on emerging risks.
3. Incorporate advancements in data analytics, artificial intelligence, and automation tools.
4. Strengthen the continuous, technology-enabled audit framework for enhanced assurance and wider audit coverage.

EXPECTED OUTCOMES

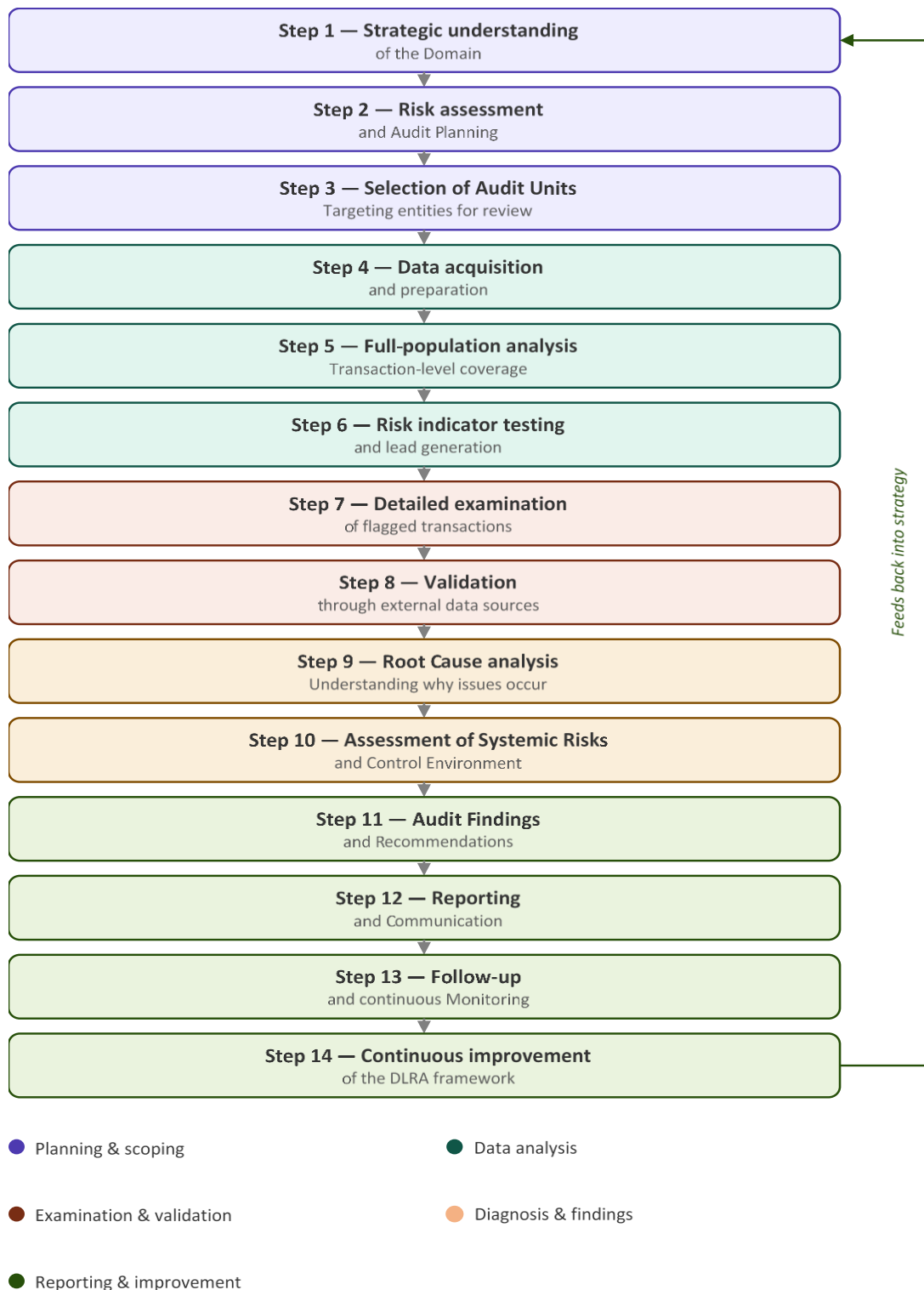
The DLRA framework aims to:

1. Enhance audit coverage through full-population analysis.
2. Improve detection of revenue leakages and compliance deficiencies.

3. Strengthen governance and internal controls.
4. Promote evidence-based decision-making.
5. Facilitate timely corrective action.
6. Establish a sustainable and continuous technology-enabled audit model.
7. Improve revenue administration and stakeholder confidence in the Registration and Stamps Department.

While Data-Led Remote Audit shall constitute the primary audit methodology, Audit Offices may undertake selective field verification wherever considered necessary for verification of original records, validation of audit evidence, examination of internal controls or confirmation of audit findings.

FIGURE: 2 PICTORIAL PRESENTATION OF SEQUENTIAL STEPS



CHAPTER -4

DATA ACQUISITION AND MANAGEMENT FRAMEWORK

I. IDENTIFICATION OF DATA REQUIREMENTS

Before commencement of audit, the Audit Office may prepare a **Data Requirement Matrix** specifying:

1. Datasets required for audit.
2. Source systems from which data is to be obtained.
3. Period of coverage.
4. Data fields required.
5. Frequency of data extraction.
6. Responsible departmental officer for providing data.

The Data Requirement Matrix may be communicated to the Department at least four weeks before the scheduled commencement of audit.

II. PROCUREMENT OF DATA

The Audit Team shall obtain data through one or more of the following methods:

1. Read-only access to NGDRS or State Registration System databases.
2. Access to departmental dashboards and Management Information Systems (MIS).
3. Extraction of data by the Department and submission to Audit.
4. Secure transfer through departmental servers, APIs, or encrypted storage devices.
5. Access to document repositories containing scanned registration documents.

Where direct access to database is available, preference should be given to automated extraction of data to reduce dependence on manually prepared reports.

III. MANDATORY DATASETS

The following minimum datasets shall be obtained for the audit period:

Dataset	Key Fields
Registration Transactions	Registration Number, Registration Date, Sub-Registrar Office Code, Instrument Type

Valuation Data	Market Value, Circle Rate, Guideline Value, Property Area
Revenue Data	Stamp Duty Payable, Stamp Duty Paid, Registration Fee Payable, Registration Fee Paid
Property Details	Survey Number, Property Type, Village/Ward, GIS Coordinates (where available)
Party Details	Executant, Claimant, PAN, Aadhaar Masked ID (where permitted), Mobile Number
Document Repository	Scanned Deeds, Attachments, Supporting Documents
User Activity Logs	User ID, Date-Time Stamp, Transaction History

IV. COLLECTION OF SUPPORTING DATA

To facilitate validation and risk analysis, the Audit Team shall obtain the following supporting datasets:

1. Circle Rate/Guideline Value Registers.
2. Land Records Database.
3. Mutation Records.
4. Development Authority Approvals.
5. Property Tax Databases.
6. GIS and Geo-Spatial Data.
7. Notifications relating to stamp duty exemptions and concessions.
8. Government Orders, Circulars and Instructions.
9. Master Tables used in the registration software.

V. DATA VALIDATION

Immediately after receipt of data, the Audit Team shall perform the following checks:

COMPLETENESS CHECKS

- Verify that all SROs and audit periods are covered.
- Check for missing transaction records.
- Reconcile transaction counts with departmental Management Information System reports.

ACCURACY CHECKS

- Compare sample records with source documents.
- Verify revenue figures against departmental accounts.
- Validate key fields such as market value, stamp duty and registration fee.

INTEGRITY CHECKS

- Identify duplicate records.
- Verify unique registration numbers.
- Check consistency between linked tables and databases.

A Data Validation Report shall be prepared before commencement of detailed analysis.

VI. DATA STORAGE AND SECURITY

The Audit Office shall ensure that:

1. All datasets are stored on secured departmental servers.
2. Access is restricted to authorised audit personnel.
3. Data transfers are encrypted.
4. Audit logs of data access are maintained.
5. Backup copies are maintained in secure locations.
6. Personal information is masked wherever feasible.
7. Data is not shared outside the audit process without approval of the competent authority.

VII. CREATION OF AUDIT DATA REPOSITORY

After validation, all datasets shall be consolidated into a structured Audit Data Repository containing:

1. Transaction Data.
2. Master Data.
3. Reference Data.
4. External Validation Data.
5. Audit Exception Data.

The repository shall serve as the single source of data for analytics, risk assessment and audit evidence generation.

VII. RESPONSIBILITY MATRIX

Activity	Responsibility
Issue Data Requirement Matrix	Audit Office
Provide Transaction Data	Registration Department
Provide Supporting Datasets	Concerned Department/Agency
Validate Data	Audit Team

Secure Storage of Data	Audit Office
Data Analytics and Risk Assessment	Audit Team
Preservation and Disposal of Data	Audit Office

IX. DELIVERABLES

Before commencing detailed audit examination, the Audit Team shall ensure availability of:

1. Data Requirement Matrix.
 2. Data Receipt Register.
 3. Data Validation Report.
 4. Audit Data Repository.
 5. List of Missing Data and Data Gaps.
 6. Certification of Data Security Compliance.
-

CHAPTER 5

DATA ANALYTICS FRAMEWORK

The data-led remote audit shall be supported by the use of analytical tools and techniques for examination of digital records, identification of risk indicators, and generation of audit leads. Audit teams may utilise the following tools and technologies:

- * **SQL (Structured Query Language)** for extraction, querying, and validation of data from departmental databases.
- * **Python** and **R** for data analytics, statistical analysis, automation, text mining, and risk modelling.
- * **KNIME (Konstanz Information Miner), IDEA (Interactive Data Extraction and Analysis), ACL (Access Control List)** and similar audit analytics tools for workflow-based analysis, exception testing, and full-population transaction examination.
- * **Microsoft excel** and **Power Query** for data cleaning, standardisation, reconciliation, and preliminary analysis.
- * **OCR (Optical Character Recognition)** tools for extraction of data from scanned deeds, notifications, valuation schedules, and supporting documents.
- * **Artificial Intelligence** and **Machine Learning** tools for anomaly detection, pattern recognition, classification of instruments, and predictive risk assessment, wherever appropriate.
- * **GIS** and **geo-spatial** tools for location-based validation of properties and assessment of valuation risks.
- * **Power Bi, Tableau**, and similar visualisation platforms for dashboards, trend analysis, heat maps, geographical mapping, and presentation of audit findings.

The office of Accountant General (Audit), Kerala has suggested that **OCR-based tools** should be positioned as instruments for risk identification and productivity enhancement rather than as primary sources of audit evidence. Since OCR technology is still evolving and its accuracy varies depending upon document quality, handwritten entries and regional languages, the SOP may provide for manual validation of selected outputs, periodic accuracy assessment and confidence-based review mechanisms before audit conclusions are drawn.

Incorporation of these safeguards would enhance the reliability and credibility of audit findings generated through OCR-based analysis.

The analytical procedures may include:

- * Trend analysis and comparative analysis.
- * Outlier and anomaly detection.
- * Risk scoring and prioritisation.
- * Pattern recognition and clustering.
- * Revenue leakage analysis.
- * Detection of undervaluation and misclassification.
- * Duplicate and suspicious transaction analysis.
- * Compliance and exception testing.

Audit offices shall maintain a repository of standard audit algorithms, analytical scripts, queries, dashboards, and risk models for recurring audit tests. Such repositories shall be periodically reviewed and updated to address changes in legislation, business processes, technology platforms, and emerging audit risks.

ILLUSTRATIVE AUDIT LEADS AND RISK INDICATORS

The analytical methodology may be utilised to generate audit leads across a wide range of revenue, compliance, valuation, and system-control risks. The following categories of audit leads are illustrative and may be expanded or modified based on State-specific legal provisions, registration systems, and emerging risk patterns.

A. INSTRUMENT CLASSIFICATION RISKS

1. Mortgage deeds with potential of being registered as security bonds, Memoranda of Deposit of title deeds, agreements, or other instruments carrying lower duty.
2. Lease deeds with attendant risk of being misclassified under categories attracting lower stamp duty.
3. Incorrect classification of development agreements, powers of attorney, release deeds, and settlement deeds.
4. Inappropriate assessment in respect of instruments containing composite transactions.

B. PROPERTY VALUATION RISKS

1. Residential or commercial properties valued at agricultural rates.
2. Developed land valued as undeveloped land.
3. Highway, link-road, or road-facing properties valued at ordinary agricultural rates.
4. Properties located within municipal limits or urban development zones valued using lower rural rates.
5. Transactions where declared market value is substantially lower than guideline value or comparable transactions.

C. CONCEALMENT AND MISREPRESENTATION RISKS

1. Suppression of residential or commercial land use.
2. Concealment of road frontage, corner plots, or locational advantages.
3. Non-disclosure of development activities affecting market value.
4. Incomplete or misleading property descriptions affecting duty assessment.
5. Potential violations of statutory provisions requiring full and true disclosure of transaction particulars.

D. STAMP DUTY AND REGISTRATION FEE RISKS

1. Short levy of stamp duty.
2. Incorrect levy of registration fees.

3. Non-levy or short levy of surcharge, cess, or additional stamp duty.
4. Incorrect application of exemptions and concessions.
5. Errors in lease rent calculations, premium valuation, or consideration assessment.

E. DATA QUALITY AND MASTER DATA RISKS

1. Missing or invalid PAN, Aadhaar, mobile numbers, or property identifiers.
2. Duplicate registrations or duplicate party records.
3. Incomplete property descriptions and geographical information.
4. Inconsistencies between registration records and master databases.
5. Deficiencies in valuation master tables and reference data.

F. SYSTEM CONTROL AND SOFTWARE RISKS

1. Bypass of automated valuation modules.
2. Weak input validation controls.
3. Absence of mandatory field checks.
4. Deficiencies in audit trails and exception logging.
5. Inadequate integration with land records, GIS, e-stamping, and external databases.

G. REVENUE LEAKAGE AND FRAUD INDICATORS

1. Unusual concentration of transactions involving common parties.
2. Repeated use of exemptions by related entities.
3. Significant deviations from prevailing market trends.
4. Multiple transactions involving the same property within a short period.
5. Patterns indicative of undervaluation, concealment, or collusive behaviour.

H. GOVERNANCE AND COMPLIANCE RISKS

1. Persistent irregularities across multiple registration offices.
2. Non-utilisation of available software controls and validation tools.

3. Delayed updating of guideline values and master data.
 4. Non-compliance with departmental instructions and statutory provisions.
 5. Recurring exceptions indicating systemic control weaknesses
-

CHAPTER 6

ROOT CAUSE ANALYSIS AND SYSTEM AUDIT

One of the fundamental distinctions between a Data-Led Remote Audit and a conventional transaction-based audit is the emphasis on identifying the underlying causes of irregularities rather than merely reporting individual instances of short levy or non-compliance. Accordingly, every significant audit finding shall be subjected to a structured root cause analysis.

The objective of root cause analysis shall be to determine whether the identified irregularity arose due to deficiencies in software design, weaknesses in internal controls, gaps in master data management, inadequate supervision, lack of training, ambiguities in legal provisions, or deliberate non-compliance by stakeholders.

Audit teams shall categorise the causes of irregularities under the following broad heads:

SOFTWARE AND SYSTEM DEFICIENCIES

Software-related deficiencies may include:

- Incomplete master data.
- Failure to synchronise valuation databases with notified circle rates.
- Weak validation controls.
- Absence of automated alerts.
- Inadequate audit trails.
- Deficiencies in accounting modules.
- Improper integration with land records and external databases.

The audit shall assess whether the software architecture itself creates opportunities for undervaluation, misclassification, concealment, or incorrect duty assessment.

ADMINISTRATIVE DEFICIENCIES

Administrative causes may include:

- Inadequate scrutiny by registering authorities.
- Failure to utilise available software tools.
- Lack of supervisory review.
- Non-compliance with departmental instructions.
- Deficiencies in monitoring mechanisms.

POLICY AND REGULATORY DEFICIENCIES

Certain irregularities may arise due to:

- Ambiguous legal provisions.
- Outdated notifications.
- Contradictory valuation guidelines.
- Lack of clarity regarding treatment of specific instruments.

Audit observations shall clearly distinguish between transaction-level failures and systemic weaknesses so that corrective measures can be directed appropriately.

Where full-population analysis is not feasible owing to data limitations or where additional assurance is considered necessary, Audit Offices may supplement analytics with random sampling or judgemental selection.

Professional judgement shall continue to play an important role in determining additional audit checks, emerging risk areas and transaction selection beyond automated analytical outputs.

CHAPTER-7

REVENUE IMPACT QUANTIFICATION METHODOLOGY

A major strength of the Data-Led Remote Audit framework is its ability to quantify the aggregate financial implications of identified irregularities with consistency. Unlike traditional audits that often focus on isolated cases, DLRA seeks to estimate the broader revenue impact arising from systemic deficiencies.

Audit teams shall adopt a standard methodology for quantification of:

- Short levy of stamp duty.
- Short levy of registration fees.
- Non-accountal of surcharge or cess.
- Additional stamp duty losses.
- Incorrect exemptions and concessions.
- Revenue implications of software deficiencies.

The quantification process shall be transparent, reproducible, and supported by documentary evidence. Calculations shall be based on applicable legal provisions, notified valuation rates, government orders, and departmental instructions.

Where systemic deficiencies affect large populations of transactions, audit may estimate the total revenue impact across all affected transactions after appropriate validation and documentation.

Revenue implications shall be presented at:

- Transaction level.
- Sub-Registrar Office level.
- District level.
- State level.

Aggregation of findings at the State level shall enable policymakers to appreciate the implications of systemic weaknesses.

CHAPTER - 8

AUDIT OBSERVATIONS AND REPORTING FRAMEWORK

- * Audit observations shall be issued through the OIOS (One IA&AD One System) or any other approved digital platform.
- * Responses of the auditee units shall be obtained through OIOS and examined before finalisation of audit findings.
- * Entry and Exit Conferences may be conducted through virtual mode and suitably documented.
- * Audit findings shall be supported by analytical evidence gathered during the audit process.
- * The Inspection Report shall highlight both individual irregularities and systemic deficiencies.
- * Similar audit observations identified across multiple offices may be consolidated and analysed at the departmental level.
- * Significant cases involving substantial revenue implications shall be specifically highlighted.
- * Financial impact of recurring irregularities may be aggregated at the State level for reporting purposes.
- * Departmental responses received through OIOS or other approved channels shall be duly considered while finalising the Report.
- * The Inspection Report shall identify process gaps, control weaknesses, and areas requiring corrective action.
- * Recommendations shall focus on recovery of revenue, strengthening of internal controls, and improvement of compliance mechanisms.
- * Where deficiencies arise from system limitations, recommendations may include enhancement of software validations and application controls.
- * The final Report shall facilitate timely corrective action and promote accountability within the Department.
- * Audit reporting should contribute to improved governance, transparency, and effectiveness of revenue administration.

Issuance of Inspection Report

In compliance with the strategic direction issued vide SMU letter No. 489/88-SMU/2025 dated 31.12.2025, which envisages a transition from “unit-centric compliance” to “systemic risk assurance”,

the findings emerging from data analytics, audit memos, and field verification shall ordinarily be consolidated into a single Departmental Inspection Report (DIR) for the Registration and Stamps Department. However, in cases involving specific local administrative structures or processes, the Head of Department, with the prior approval of the SMU (Strategic Management Unit) at Headquarters, may consider issuing Inspection Reports at a level below the Administrative Head of the Department in the State Government. In such cases, significant audit findings and systemic concerns shall simultaneously be brought to the notice of the Administrative Head of the Department through a Departmental Appreciation Note or Management Letter.

CHAPTER-9

CONTINUOUS AUDIT AND QUARTERLY RISK MONITORING FRAMEWORK

The availability of real-time transaction data creates opportunities for moving beyond periodic audits towards a continuous audit model.

Under this framework, audit offices may establish quarterly or half-yearly risk monitoring cycles whereby predefined analytical scripts are executed against updated departmental datasets.

The objectives of continuous audit shall include:

- Early identification of emerging risks.
- Monitoring of recurring irregularities.
- Tracking implementation of audit recommendations.
- Identification of new patterns suggestive of potential to bypass prescribed controls or instructions.
- Periodic assessment of effectiveness of control in IT Applications.

Continuous monitoring shall enable audit offices to provide timely assurance rather than relying solely on post-facto observations.

A State-level risk dashboard may be developed for monitoring:

- Revenue trends.
- High-risk SROs.
- Repeat observations.
- Software vulnerabilities.
- Implementation status of audit recommendations.

This approach shall gradually transform audit from a retrospective activity into a proactive governance tool.

CHAPTER-10

KEY PERFORMANCE INDICATORS

To assess the effectiveness of Data-Led Remote Audit, measurable performance indicators shall be established.

These may include:

COVERAGE INDICATORS

- Number of SROs covered.
- Percentage of State revenue covered.
- Number of transactions analysed.

EFFICIENCY INDICATORS

- Reduction in audit cycle time.
- Reduction in travel expenditure.
- Number of transactions analysed per auditor.

EFFECTIVENESS INDICATORS

- Revenue impact identified.
- Number of systemic deficiencies detected.
- Number of control weaknesses identified in respect of involved IT Applications.
- Number of recommendations accepted by Government.

OUTCOME INDICATORS

- Revenue recovered.
- Modifications implemented in respect of IT Applications.
- Policy changes initiated.
- Reduction in recurring irregularities.

These indicators shall facilitate periodic evaluation and continuous improvement of the DLRA framework.

CHAPTER -11

BENEFITS AND WAY FORWARD REFLECTED IN CASE STUDY OF STATE OF UTTAR PRADESH

OFFICE OF THE PRINCIPAL ACCOUNTANT GENERAL (AUDIT) II, UTTAR PRADESH

DATA-LED REMOTE AUDIT

Stamp and Registration Department, Government of Uttar Pradesh

Concept Paper . June 2026 . PAG (Audit) II, Lucknow

Rs.30.81 Crore

Total Revenue Leakage

8

SROs Audited

10

Finding Categories

~2,50,000+

Transactions Screened

DISTRICTS

Ghaziabad & Gautam Buddh Nagar, UP

METHODOLOGY

Data-Led Remote Audit

SROs EXAMINED

SRO-I, II, III (Sadar GZB), Modinagar,
SRO-GN, SRO-Dadri, SR-I & SR-II Noida (GB Nagar)

TOOL CHAIN

Python - KNIME - OCR - IGRS Portal

Concept Paper . June 2026 . PAG (Audit) II, Lucknow

1. SUMMARY

The Office of the **Principal Accountant General (Audit) II, Uttar Pradesh, Lucknow**, has conducted a **Data-Led Remote Audit (DLRA) of the Stamp and Registration Department of the Government of Uttar Pradesh**. This part of the SoP documents the rationale, audit selection framework, methodology, digital infrastructure leveraged, analytical techniques deployed, and consolidated audit findings from eight **Sub-Registrar Offices (SRO)** — SRO-I (Sadar), SRO-II (Sadar), SRO–III (Sadar) and SRO- Modinagar in Ghaziabad district, and SRO-Greater Noida, SRO-Dadri, SRO-I, Noida, SRO-II, Noida in Gautam Buddha Nagar district. The audit leverages the IGRS (**Inspector General Registration and Stamps**¹) portal and a dedicated AG Office Dashboard that provides secure, read-only access to transactional data across all Sub-Registrar Offices in Uttar Pradesh.

Key Revenue Impact — Eight SROs (Ongoing Audit)

Total quantified revenue impact so far: approximately **₹ 30.81 crore** across ten categories of irregularities, covering eight Sub-Registrar Offices in Ghaziabad and Gautam Buddha Nagar districts.

This is an ongoing audit. As more Sub-Registrar Offices are brought under the Data-Led Remote Audit (DLRA) framework, the cumulative revenue impact will rise substantially.

Findings span misclassification of mortgage instruments, systematic property undervaluation, irregular lease rent computation, and software design gaps in the PRERNA registration system. The systemic nature of these findings means identical irregularities are likely present across many more SROs statewide.

¹ **Inspector General Registration and Stamps (IGRS) Portal** is a comprehensive digital platform of the Registration and Stamps Department, Government of Uttar Pradesh, designed to facilitate online property registration, stamp duty management, citizen services, and grievance redressal through a single integrated interface.

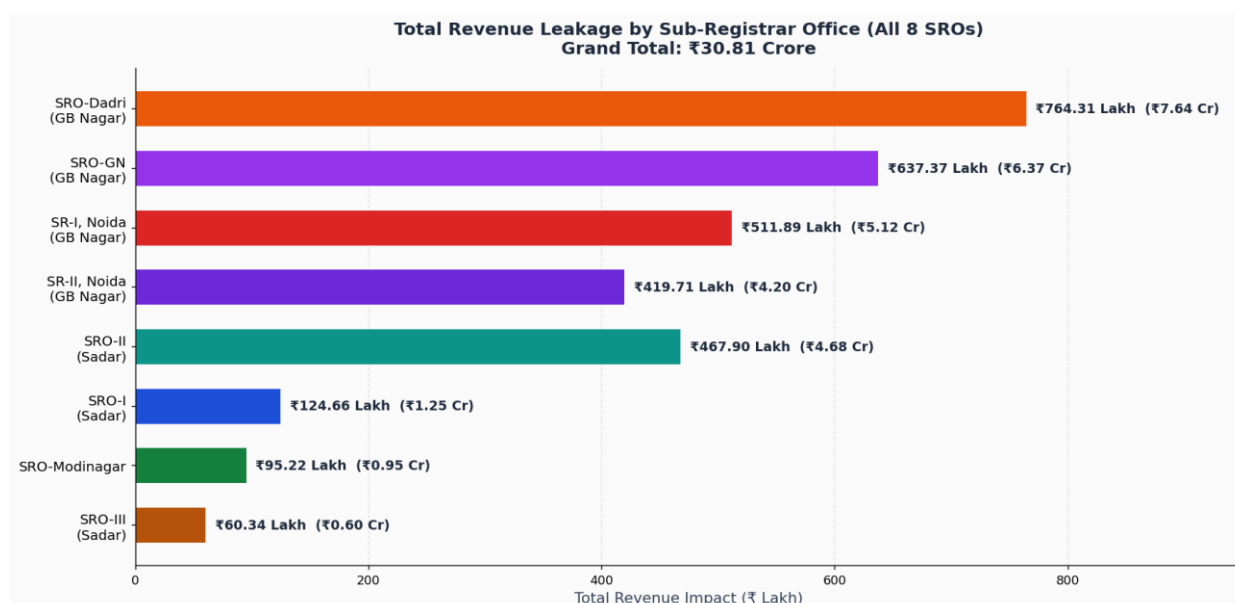


Figure 3: Total Revenue Leakage by Sub-Registrar Office

2. BACKGROUND AND CONTEXT

2.1 THE REGISTRATION SYSTEM

The system of registration of immovable property transactions in India includes Sale Deeds, Gift Deeds, Settlement Deeds, Development Agreement-cum-General Power of Attorney, Mortgages, Leases, and Release Deeds. The registration process involves document presentation, market value determination, duty calculation, admission of execution, and secure digital storage.

2.2 DIGITAL TRANSFORMATION IN UTTAR PRADESH

Uttar Pradesh has computerized its registration operations through the IGRS portal. The PRERNA software powers the actual registration workflow at SRO level, facilitating online document presentation, automated duty computation, and digital storage. The system was designed to achieve on-the-spot registration, improved monitoring, standardized document language, system transparency, auto-valuation of property, electronic document storage, auto-generation of mutation notices and single-window service delivery. The registered deeds are stored digitally and are accessible remotely through the IGRS portal, forming the technological foundation for this Data-Led Remote Audit.

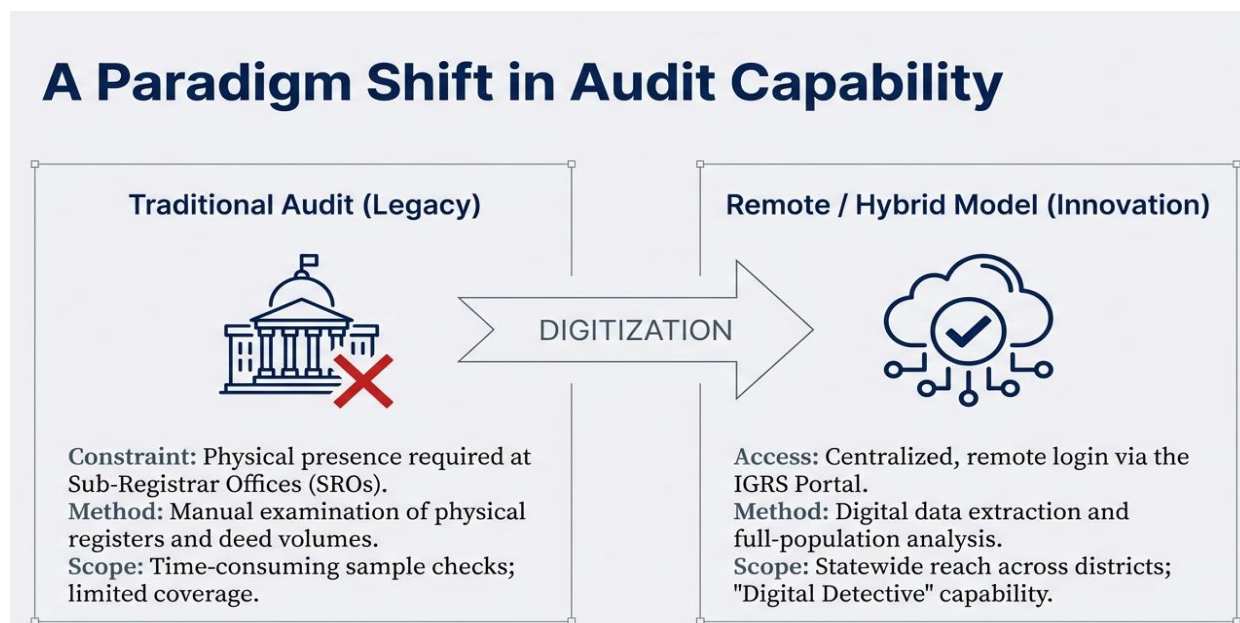
2.3 SECURING OF SYSTEM ACCESS

This audit became possible after efforts to secure the creation of a dedicated AG Office Dashboard within the IGRS (Inspector General Registration and Stamps) portal of the Stamp and Registration Department, Government of Uttar Pradesh. This dashboard provides secure, read-only access to Fehrist-2 (Deed Index), Syaha (Fee Register), and scanned registered deed copies across all SROs in Uttar Pradesh, thereby, enabling remote audit without any requirement for physical presence at the auditee office.

3. DATA-LED REMOTE AUDIT FRAMEWORK

Traditionally, audit of the Stamp and Registration Department required on-site visits to SROs for physical examination of registers and documents which is time-consuming, constrained to small samples, and dependent on manual production

of records. The DLRA framework responds to CAG's mandate by demonstrating that remote audit is not only feasible but delivers superior outcomes in terms of observations generated, revenues involved and audit efficiency achieved.



3.1 AUDIT SELECTION AND EXECUTION METHODOLOGY

The design principle of this audit framework is the **audit approach** which focuses on achieving both manageable scope and comprehensive coverage:

STAGE 1 — SRO SELECTION (Revenue-Based Sampling)
- All Sub-Registrar Offices in the district are ranked by total revenue collected. - Top revenue-generating SROs are selected for audit, thereby, maximising revenue-risk coverage. - Rationale: <i>Khasra (Araji/khasra/gata indicate the particular number of a land holding in a locality)</i> categorisations vary across Sub-Registrar Offices (SROs), and the same Khasra number may appear in multiple offices with different land-use classifications. Since applicable stamp duty rates differ by locality (e.g., 5% versus 7%), accurate assessment requires deed-level scrutiny and office-wise analysis of transaction data. - Stage 1 selection ensures audit resources target entities with highest revenue collection.
STAGE 2 — FULL UNIVERSE ANALYSIS (Within Each Selected SRO)
- All transactions registered during the audit period are downloaded from the IGRS portal. - Full-population algorithmic screening is applied.

- Data cleaning, integrity checks, and cross-referencing with digitised *Mulyankan Soochi*² is carried out.
- Python/KNIME scripts automatically flag potential irregularities, which are subsequently verified through examination of the scanned deeds.
- 100% population coverage where every transaction in the audit period is algorithmically screened.

Why full-state population analysis is not feasible in a single pass:

- *Khasra* categorisations (residential, near-Abadi, near national highway) vary from one SRO to another. A *Khasra* declared residential under SRO-I may be agricultural under SRO-III. The same *Khasra* number exists across multiple SROs with different land-use classifications.
- Stamp duty rates are not uniform statewide. Where some localities attract 5% and others 7% (2% additional stamp duty), concessional rates/rebate for women purchasers further complicate application of uniform rules across the full population.
- Revenue shortfall relating to Findings 2, 3, and 4³ requires cross-referencing each *Khasra* against the District Magistrate's *Mulyankan Soochi*, which is specific to each district/registration jurisdiction.
- Misclassification of Security Bond (Finding 1⁴) requires reading the actual deed text which is required to be done office by office as instrument characteristics vary.
- Stage 1 selection is therefore based on total revenue collected by each SRO where high-revenue offices are prioritised as they represent the greatest risk of revenue leakage in absolute terms.
- For each selected SRO (Stage 2), all transactions recorded during the audit period are subjected to full-population algorithmic screening, eliminating the need for any additional sampling.

Khasra classifications in the *Mulyankan Soochi* are jurisdiction-specific and vary across districts and Sub-Registrar Offices (SROs). The same *Khasra* number may exist in multiple villages falling under different SROs and may be assigned different land-use categories, resulting in different valuation rates and stamp duty implications. Further, the applicable stamp duty rates vary across localities, with some areas attracting a duty rate of 5% and others 7% including an additional Stamp Duty of 2% where applicable. Concessional rates available for women executants/transferees add another layer of complexity to the valuation process.

Accordingly, the *Mulyankan Soochi* must be digitised and mapped separately for each district/SRO jurisdiction before analytical scripts can be executed. Consequently, a single-pass statewide analysis is not feasible without first creating district-specific reference datasets that capture the local valuation classifications and duty structures applicable to each SRO.

² Mulyankan Soochi is the district-wise valuation schedule notified by the District Magistrate under the Uttar Pradesh Stamp (Valuation of Property) Rules, 1997, as applicable in Uttar Pradesh, specifying minimum rates for various categories of land and properties to facilitate determination of market value and levy of stamp duty. It is publically available on IGRS portal of Stamp and Registration Department, UP.

³ Reference Para No.6

Finding 2 – Residential Properties Registered at Agricultural Rates

Finding 3 – Potential Misdeclaration under Section 27

Finding 4 – Properties Located Adjacent to Abadi Areas

⁴ Finding 1 – Misclassification of Security Bonds

3.2 AUDIT PROCESS

I	Data Cleaning & Preparation • Download Deeds Details Index-II (<i>Fehrist-2</i>) & Fee Register (<i>Syaha</i>) ⁵from IGRS portal. • Merge datasets using Registration Number as primary key (imperfect match rate — Deeds dataset is primary) • No intermediate data transformation — Raw data used as extracted • Check for null values, duplicates, data-type errors • Extract & digitise <i>Mulyankan Soochi</i> via OCR + Python
II	Algorithmic Screening • Apply pre-defined Python/KNIME scripts for each Finding category. • Filter Security Bond documents and use OCR-based text analysis to identify and flag records containing references to “Mortgage Deed” • Cross-reference <i>Khasra</i> numbers against digitised <i>Mulyankan Soochi</i> • Group transactions by village and <i>Khasras</i> to detect mixed residential/agricultural patterns • Flag 2% Additional Stamp Duty column missing from <i>Syaha</i> and auto-generated by PRERNA • Python scripts compute revenue shortfall for each flagged transaction
III	Deed Scrutiny & Observations • Manual Examination of scanned deeds retrieved from IGRS portal • Cross-verify flagged transactions with circle rate documents • Confirm misclassification of instrument type (Security Bond vs. Mortgage) • Draft audit observations with document references

3.3 ADVANTAGES OF THE DATA-LED REMOTE AUDIT APPROACH

Direct Read-Only Access Secure access to the IGRS portal's AG Office Dashboard enables remote examination of records, thereby eliminating the need for physical visits.	Full-Universe SRO Coverage Within each selected SRO, the entire population of transactions is subjected to audit scrutiny, providing comprehensive assurance for the period under review without reliance on sampling techniques.
--	--

⁵ *Syaha* is maintained in format no. 13 as per rule 211 under the stamp manual. It contains the collected amount of stamp duty, additional stamp duty, registration fee etc.

Speed of Mathematical Checks

Algorithmic scripts verify stamp duty calculations and categorisation errors across thousands of transactions in minutes.

Documentary Evidence

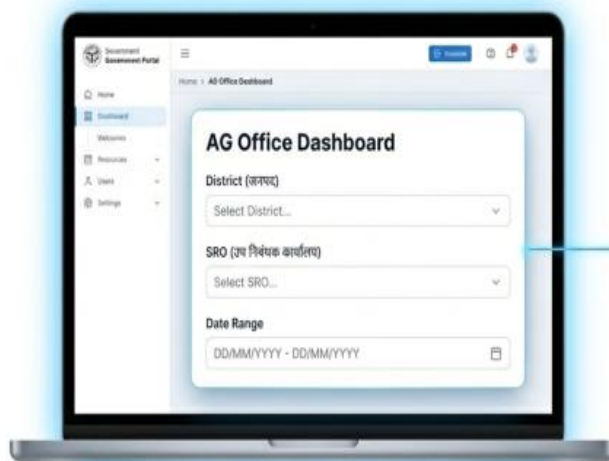
Scanned deeds available on the IGRS portal constitute reliable digital evidence, eliminating the need for physical deeds for evidentiary purposes.

4. SYSTEM ACCESS AND DATA ARCHITECTURE IN UTTAR PRADESH

4.1 THE IGRS PORTAL AND AG OFFICE MODULE

The Office obtained secure login access to the "AG Office Dashboard" within the IGRS portal, enabling remote retrieval of registration records from any Sub-Registrar Office (SRO) through configurable district, office, and date-range parameters. This facility significantly streamlined data collection and facilitated comprehensive audit coverage without the logistical constraints associated with field visits.

The Ecosystem: The AG Office Dashboard



Secure login module enabling 100% statewide data access without field visits.

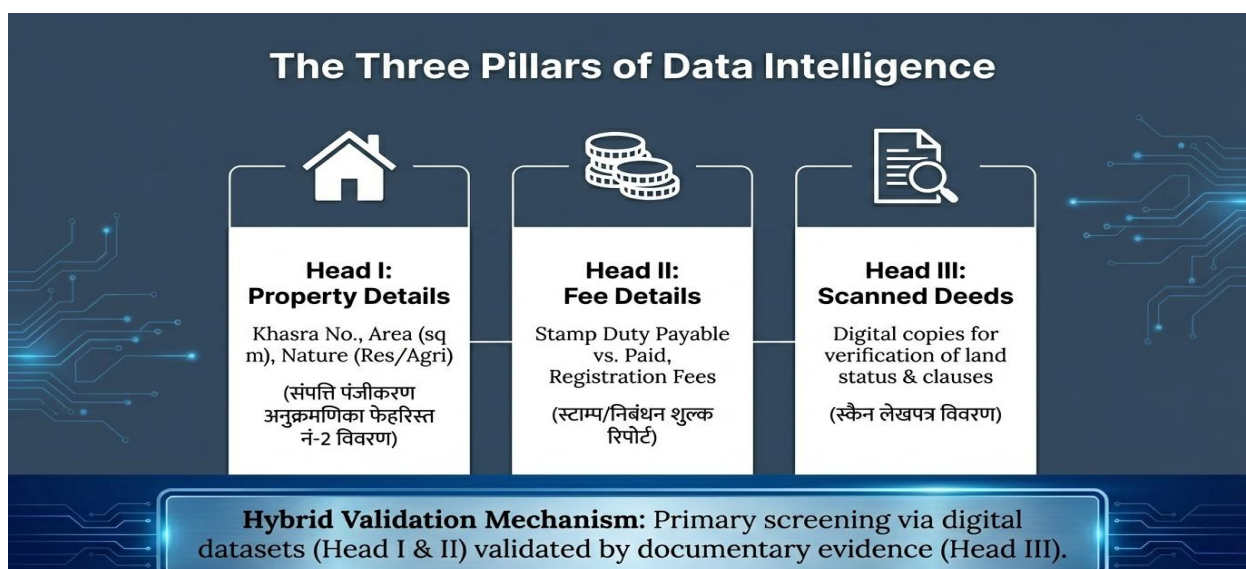
Data extraction via IGRS processed using **KNIME**, **Python-based analytics**, and **AI tools**.

4.2 DATA ARCHITECTURE — THREE FUNCTIONAL HEADS

HEAD I: Deed Index (Fehrist No. 2 / अनुक्रमणिका पंजिका)	HEAD II: Fee Register (Syaha / स्टाम्प/निबंधन शुल्क रिपोर्ट)
क्र. सं. Serial No.	स्टाम्प शुल्क Stamp Duty Collected
मोहल्ला/गाँव Locality / Village	2% Additional Stamp Duty
वार्ड/परगना Ward / Pargana	रसीद नम्बर Receipt Number
संपत्ति का विवरण Property Details (Khasra, boundaries, area)	REGISTRATION ID (cross-ref. key)
अभिलेख प्रकार Nature of Deed	लेख/प्रतिलेखों की किस्म Nature of Deed & Presenter Name
प्रथम पक्ष First Party (Executant/Seller)	विनिमय मूल्यांकन Exchange Valuation
द्वितीय पक्ष Second Party (Claimant/Purchaser)	निबंधन शुल्क Registration Fee
निष्पादन तिथि Date of Execution	अन्य प्राप्ति Other Receipts (incl. certification)
पंजीकरण तिथि Date of Registration	कुल योग Grand Total
बाज़ारी मूल्य Market Value	चालान विवरण Challan Details (No. & Date)
प्रतिफल Consideration Value	दिनांक Date of Return of Original Document
बही नं-1 की क्रम सं. Registration Number (Primary Key)	
स्कैन प्रपत्र Scanned Deed Availability	

HEAD III: Scanned Copies of Registered Deeds (स्कैन लेखपत्र विवरण)

This module enables retrieval of scanned deed copies using Registration Number and year. It facilitates verification of the nature of rights transferred, clauses impacting valuation, built-up vs. vacant land classification, applicability of concessional rates, and special conditions affecting duty liability. Scanned deeds serve as digital documentary evidence strengthening audit conclusions without requiring physical inspection of deed volumes.



4.2.1 FOUR-LAYER DATA FRAMEWORK

All data required for this audit is available through the IGRS portal (Layers 1, 2, and 4) or through one-time OCR digitisation of the Mulyankan Soochi (Layer 3). This single-source availability is what makes the full-remote model feasible.

Data Layer	Type of Data	Source	Audit Use
LAYER 1 Fehrist-2 (Deed Index)	Deed-level data: Serial No., Village/Mohalla, Pargana, Description of Property (Khasra, area, boundaries - Col 3), Type of Deed (Col 4), Parties and Permanent Account Number (PAN) (Col 5/6), Dates, Market Value (Col 11),	IGRS Portal -> AG Dashboard -> Fehrist-2 download	The primary dataset comprises 16 columns, with Columns 3 and 4 containing the most analytically valuable information.

Data Layer	Type of Data	Source	Audit Use
	Consideration (Col 12), Registration No. (Col 14).		
LAYER 2 Syaha (Fee Register)	Col 1: Stamp Duty paid. Col 2: 2% Additional Stamp Duty (universally blank = systemic finding F8). Col 3: Registration ID Col 7: Registration Fee. Col 10: Total receipts. Col 13: Challan details.	IGRS Portal -> AG Dashboard -> Syaha download (CSV).	Cross-referenced with Fehrist-2 via Registration ID to verify actual duty paid.
LAYER 3 Mulyankan Soochi (District Magistrate's Circle Rate List)	Khasra number, village, land-use category (Abadi/Residential, Near-Abadi, Agricultural, National Highway/State Highway/Link Road, Commercial), circle rate per sq. metre by category.	PDF Mulyankan Soochi obtained from the District Magistrate Office → OCR extraction of data → Manual quality checks and corrections → Conversion into a structured CSV format	Cross-referenced with Col 3 of Fehrist-2 to detect category mismatches. Underpins Findings Finding 2 – Residential Properties Registered at Agricultural Rates Finding 3 – Potential Misdeclaration under Section 27 Finding 4 – Properties Located Adjacent to Abadi Areas Finding 5 – Properties Adjacent to National Highways/Link Roads
LAYER 4 Scanned Deed Copies	Full text of the registered instrument, including deed clauses, IDC/GDA ⁶ references, property boundary descriptions, land-use declarations, and premium and lease rent terms, enabling detailed content-based audit analysis.	IGRS Portal -> Scanned Deed module (search by Reg. No. + Year).	Required for Security Bond Misclassification – OCR-based identification of Security Bond and GDA-related clauses. Misdeclaration under Section 27 – Detection of omissions or deficiencies in property boundary descriptions. Short Levy of Lease Rent and Premium – Verification of premium and lease rent provisions contained in registered instruments.

⁶ References to Development Authority approvals (e.g., Ghaziabad Development Authority) and Infrastructure Development Charges (IDC).

4.2.2 FEHRIST-2 COLUMN REFERENCE

The *Fehrist-2* (Deed Index) downloaded from the IGRS portal contains 16 columns, with the following fields being the most critical for data-driven audit analysis:

Index	Column Name (Hindi)	Type	Audit Use
0	S.No. (Serial No.)	Int	Row identifier
1	Mohalla/Gaon (Village/Locality)	Text	Group by village for Mulyankan Soochi cross-reference
2	Ward/Pargana (Pargana)	Text	Geographic classification
3	Sampatti ka Vivaran (Property Description)	Long Text	PRIMARY FIELD -- Khasra extraction (regex), area in hectares, boundary disclosure
4	Abhilekh Prakar (Deed Type)	Text	Filter for Security Bond, Patta >30 yr, Vikreya Patra + Krishi, etc.
5	Pratham Paksh (First Party)	Text	PAN and mobile number extraction and validation
6	Dvitiya Paksh (Second Party)	Text	PAN cross-check
7	Nishpadan Tithi (Execution Date)	Date	Applicable rate list version
11	Bazari Mulya (Market Value)	Numeric	Stamp duty base -- higher of Market Value/consideration
12	Pratifal (Consideration)	Numeric	Stamp duty base -- higher of Market Value /consideration

Index	Column Name (Hindi)	Type	Audit Use
14	Bahi Kram No. (Registration No.)	Text	For downloading the scanned deeds.

4.2.3 OCR DIGITISATION OF MULYANKAN SOOCHI

OCR Workflow -- One-Time per District

Step 1: Obtain Mulyankan Soochi PDF from the District Collector's office or IGRS/Revenue portal.

Step 2: Apply OCR (Tesseract) to extract tabular content.

Step 3: Clean in Python -- correct numeral misrecognition, standardise Khasra format (remove trailing spaces).

Step 4: Save as CSV: Village_Name | Khasra_Number | Category (Abadi/Near-Abadi/Agri/NH-Link-Road) | Circle_Rate_per_SqM.

Step 5: Spot-check 20-30 entries against printed Soochi.

4.3 DATA EXTRACTION AND CONSOLIDATION

Data under Heads I and II is downloaded in Excel format for selected SROs and time periods. The two datasets are merged using the Registration Number as the primary key. However, the match rate between the Registration Dataset (Fehrist-2) and the Fee Register (Syaha) is not absolute — there are known gaps arising from differences in how the two systems record identifiers. As a result, the consolidated merged dataset serves primarily as a supplementary cross-reference tool. The core automated data analysis is conducted directly on the **Registration Dataset (Fehrist-2)** files with the **Fee Register (Syaha)** consulted on a case-by-case basis when specific financial verification is required to support an observation. Raw data is used as extracted, without intermediate data transformation, to preserve the integrity of original records.

5. ANALYTICAL TOOLS AND TECHNOLOGY STACK

Tool	Primary Use	Findings Covered
Python (Jupyter Notebook)	Khasra number extraction using regular expressions (Regex); cross-verification with Mulyankan Soochi data; computation of	Finding 1 – Misclassification of Security Bonds Finding 2 – Residential Properties Registered at Agricultural Rates Finding 3 – Potential Misdeclaration under Section 27 Finding 4 – Properties Located Adjacent to Abadi Areas Finding 5 – Properties Abutting National Highways/Link Roads Finding 8 – Non/Short Levy of Additional 2% Stamp Duty Finding 9 – PAN Validation and Compliance Checks

Tool	Primary Use	Findings Covered
	valuation shortfalls; PAN validation; and OCR-based text analysis to identify Security Bond documents and GDA-related clauses.	Finding 10 – Verification of PRERNA Auto-Valuation Assessments
KNIME Analytics Platform	Visual workflow illustrating data loading, data cleansing, deed-type filtering, and generation of flagged transactions for detailed audit examination.	All findings -- pre-processing layer
OCR	(a) Digitising Mulyankan Soochi PDF -> structured CSV. (b) Extracting text from scanned deeds for Security Bond clause detection.	F1 – Misclassification of Security Bonds F2 – Residential Properties Registered at Agricultural Rates F3 – Potential Misdeclaration under Section 27 F4 – Properties Located Adjacent to Abadi Areas F5 – Properties Abutting National Highways/Link Roads
Microsoft Excel / CSV	IGRS data download format. Initial inspection and pre-processing before Python/KNIME.	Data extraction -- all findings
IGRS Portal (AG Office Dashboard)	Secure read-only access to Fehrist-2, Syaha, and scanned deeds across all SROs in Uttar Pradesh.	All findings -- primary data source
GPT / LLM (Non-Data Mode)	Code generation: Python scripts, regex patterns, SQL logic. No audit data shared externally.	Code development only

5.1 LLM-ASSISTED SCRIPT GENERATION

AI assistants can significantly reduce the effort required to write **Python** and Structured Query Language (**SQL**) code. Users can simply describe the analytical task in plain language without sharing actual data and the LLM can generate the required code. To do so effectively, the model should be provided with: (a) the column names, (b) a single illustrative row with fictitious values demonstrating the data format, (c) a description of the analysis to be performed, and (d) the target execution environment (e.g., Jupyter Notebook using Python).

5.1.1 EXAMPLE PROMPT STRUCTURE FOR KHASRA CROSS-REFERENCE SCRIPT

Platform: Python (Jupyter Notebook)

I have an Excel file (Fehrist-2) with column index 3 = property description in Hindi.

Sample format: 'कृषि भूमि, खाता नं० 100 के खसरा नं० 234मि०, 0.25 हैक्टेयर, पूर्व- सड़क'

I also have khasra_aabadi.csv with columns: Village, Khasra_Number

Task: Extract Khasra number from column 3, match against khasra_aabadi.csv,

flag rows where Khasra is listed as Aabadi (residential) but deed says 'कृषि'.

Compute shortfall = (area_sqm * 14400 - max(MV, consideration)) * 0.08

Output flagged rows with Khasra, area, expected basis, actual basis, shortfall.

6. REVENUE AUDIT CHECK FRAMEWORK

The following table sets out the Revenue Audit Check Framework for data-led remote audit of the Stamp and Registration Department. It translates each legal provision governing stamp duty levy into a specific, testable audit check that can be answered from digital data alone — without physical presence at the SRO.

Finding	Audit Question & Legal Basis	Data Fields	Analytical Method
Finding 1 Misclassification of Mortgage deed as Security Bonds	Whether a mortgage instrument misclassified as a Security Bond to avoid higher duty? Sec. 2(17) ISA 1899; Art. 40, Sch. IA; State Notif. 25.05.2001 (Rs.5L cap).	Col 4: deed type filter 'Jamanatname'/'BOND'. Layer 4 scanned deed: OCR for GDA/IDC clause.	KNIME filter -> OCR -> Python regex for 'Mortgage Deed'/'IDC'/'GDA demand'.

Finding	Audit Question & Legal Basis	Data Fields	Analytical Method
Finding 2 Residential Properties Registered at Agricultural Rates	Whether a Khasra officially notified as residential (Abadi) in District Magistrate's Mulyankan Soochi valued at agricultural rate? ISA 1899; UP Val. Rules 1997; IG Cir. 4302/2018.	Col 3: Khasra No. regex + area extraction. Col 4: Sale Deed + Krishi. Cross-ref: khasra_aabadi_list.csv.	Python: extract Khasra, check vs aabadi list, convert hectares->sqm, compute.
Finding 3 Potential Misdeclaration under Section 27	Whether facts affecting stamp duty chargeability (residential land use, boundaries) deliberately omitted? Section 27, ISA 1899; IG Cir. 4302/2018.	Col 3: 'Krishi', residential, institutional written in property description in same khasra. Col 1+extracted Khasra: groupby. Layer 4: boundary omission.	Group by village+Khasra; flag mixed deed-type Khasras (agri + residential); download scanned deed; verify boundary omission; compute shortfall.
Finding 4 Properties Located Adjacent to Abadi Areas	Whether Near-Abadi designated land valued at the lower general agricultural rate? ISA 1899; UP Val. Rules 1997 (Near-Abadi column in Mulyankan Soochi).	Col 3: Khasra. Cross-ref: Near-Abadi category in Mulyankan Soochi CSV.	Same script as F2, using Near-Abadi Khasra list and Near-Abadi circle rate.
Finding 5 Properties situated adjacent to National	Whether highway-proximity agricultural land valued at the lower general agricultural rate? ISA 1899; UP Val.	Col 3: Khasra. Cross-ref: NH/Link Road column in Mulyankan Soochi CSV.	Same script as F2, using NH/Link Khasra list and highway-proximity circle rate.

Finding	Audit Question & Legal Basis	Data Fields	Analytical Method
Highways/Link Roads	Rules 1997 (Highway column in Mulyankan Soochi).		
Finding 6/7 Incorrect Assessment of Lease Rent	Whether the consideration value in lease deeds executed by the Greater Noida Industrial Development Authority (GNIDA) for lease terms exceeding 30 years has been correctly computed by incorporating the applicable lease rent component (Commercial: 37.5% of premium; Residential: 15% of premium), as required under Article 35(c) of the Indian Stamp Act, 1899.	Col 4: Patta Vilekh (>30 yr). Layer 4: premium + rent clause from deed.	Filter deed type; download scanned deed; identify premium; compute shortfall
Finding 8 Irregularities in Accounting of Additional Stamp Duty (2%)	Whether the 2% urban cess being correctly accounted for in the PRERNA Syaha? Sec. 39 UP Urban Planning Act 1973; GO 03.09.2024.	Syaha Col 1: 2% column. Syaha Col 0: stamp duty > 0.	Load Syaha; filter stamp duty > Rs.100; check if Col 1 is universally blank -- if yes, systemic PRERNA failure confirmed.

Finding	Audit Question & Legal Basis	Data Fields	Analytical Method
Finding 9 Non-Validation or Incorrect Capture of PAN Details	Whether PAN and mobile numbers of the parties are correctly captured and PAN has been mandatorily furnished in accordance with the Income Tax Act provisions applicable to immovable property transfer.	Col 5/6: First/Second Party PAN field.	Python regex: flag absent or malformed PAN [A-Z]{5}[0-9]{4}[A-Z]; count and report.
Finding 10 Deviation from PRERNA Auto-Valuation Benchmarks	Are properties being valued via PRERNA's auto-valuation module? PRERNA design mandate.	Col 3: check for boundary terms and land-use descriptors.	Flag records where Col 3 lacks East/West/North/South and activity descriptors -- indicates PRERNA auto-val bypass.

7. ANALYTICAL METHODOLOGY

7.1 OVERVIEW

The audit adopts a full-population analysis methodology within each selected SRO — every transaction registered during the audit period is subjected to automated algorithmic screening. Findings from data analytics are then validated through examination of scanned deeds on the IGRS portal. The analytical toolchain includes KNIME, Python (for scripting, text parsing, and computation), OCR (Optical Character Recognition for scanned document text extraction), Microsoft Excel (data sorting) etc.

7.2 METHODOLOGY BY FINDING CATEGORY

A. Misclassification of Mortgage Deeds as Security Bonds / Deposit of Title Deeds

- 1 Import Registration datasets into KNIME analytics software.

2	Apply a targeted filter on the 'Nature of Deed' (अभिलेख प्रकार) column to isolate all transactions registered as 'Security Bonds' or 'Deposit of Title Deed Memoranda'.
3	Download scanned copies of flagged instruments from the IGRS portal.
4	Process scanned deeds through OCR to extract machine-readable text.
5	Run a Python script on OCR-extracted text to identify trigger phrases (e.g., 'Mortgage Deed', 'IDC', 'Internal Development Charges', 'land lien') indicating potential misclassification of instruments chargeable as Mortgage Deeds under Article 40 of Schedule IA of the Indian Stamp Act, 1899
6	Manually scrutinise all flagged instruments to validate misclassification, review GDA demand notices, and compute the correct stamp duty liability using the applicable government circle rates for the mortgaged property.

B. Residential Land Valued at Agriculture Rate (Declared Abadi Khasras)

1	Digitize the District Magistrate's Mulyankan Soochi (circle rate and Khasra classification list) using OCR and Python — convert the structured Khasra category data into a searchable Excel reference file. This Soochi explicitly lists each Khasra within a revenue village as residential (Abadi), near-Abadi, on Link Road, on State Highway, or general agriculture.
2	Import registration datasets and the digitized Khasra list into a computational notebook.
3	Develop a custom Python script to parse the unstructured 'Property Details' (संपत्ति का विवरण) column, extract the exact Khasra numbers embedded in the text for each transaction.
4	Electronically cross-reference extracted Khasra numbers against the Mulyankan Soochi. The script automatically flags transactions where the Khasra is officially declared as residential (Abadi) in the rate list but has been valued using the lower general agriculture rate.
5	The Python script also computes the revenue shortfall for each flagged transaction — applying the applicable residential circle rate to the area transferred, comparing against the stamp duty actually paid, and calculating the difference as the revenue loss.

```

# Core shortfall computation script (actual code used in audit)
import pandas as pd, re

RES_CIRCLE_RATE = 14400 # Rs/sqm from Mulyankan Soochi (residential category)
TOTAL_RATE = 0.08 # 7% stamp duty + 1% registration fee

res_set = set(pd.read_csv('khasra_aabadi_list.csv')['Khasra_Number'].astype(str).str.strip())
df = pd.read_excel('fehris2_SRO.xlsx', header=1)

def extract_khasra(desc):
    m = re.search(r'Khasra[\s\-\]*([0-9/]+)', str(desc))
    return re.sub(r'^[0-9/]', '', m.group(1)) if m else None

def extract_area_sqm(desc):
    m = re.search(r'([\d\.\.]+)\s*hectare', str(desc))
    return float(m.group(1)) * 10000 if m else None # 1 hectare = 10,000 sqm

results = []
for _, row in df.iterrows():
    desc = str(row.iloc[3])
    khasra = extract_khasra(desc)
    if khasra not in res_set: continue
    area_sqm = extract_area_sqm(desc)
    if not area_sqm: continue
    mv = float(str(row.iloc[11]).replace(',', '')) if str(row.iloc[11]).replace(',', '').replace('.', '').isdigit() else 0
    con = float(str(row.iloc[12]).replace(',', '')) if str(row.iloc[12]).replace(',', '').replace('.', '').isdigit() else 0
    expected = max(area_sqm * RES_CIRCLE_RATE, con)
    shortfall = (expected - max(mv, con)) * TOTAL_RATE
    if shortfall > 0:
        results.append({'Reg_No': row.iloc[14], 'Khasra': khasra,
                        'Area_SqM': round(area_sqm, 2), 'Shortfall': round(shortfall, 2)})

out = pd.DataFrame(results)
print(f'Cases: {len(out)} | Total Shortfall: Rs. {out.Shortfall.sum():.2f}')

```

C. Violation of Section 27 — Concealment of Residential Character

1

Deploy a Python script to segregate transaction data by revenue village and further subdivide by individual Khasra numbers.

2	Flag Khasras that simultaneously exhibit a mix of transaction types — some registered as 'agricultural' and others as 'residential plots', 'houses', or 'commercial' — within the same Khasra during the audit period. This mixed pattern is a strong data-driven indicator of potential concealment.
3	For flagged Khasras, cross-reference nearby registrations in the same Khasra to confirm that residential colonies, developed plots, or constructed buildings pre-existed the questioned agricultural transaction.
4	Download and manually examine scanned deeds for flagged instruments to confirm that boundaries, road access, and surrounding residential activity were deliberately omitted from the instrument — constituting a violation of Section 27 of the Indian Stamp Act, 1899.
5	Note: these instruments also tend to be registered in online mode via PRERNA software, which provides tools for potential cross-checking; the failure to utilise these tools by the registering authority forms part of the observation.

D. Agriculture Land near ABADI Valued at General Agriculture Rate

1	Using the same Python cross-referencing workflow as in (B) above, extend the Mulyankan Soochi matching to specifically identify Khasras categorised as 'near Abadi' (आबादी के निकट) in the District Magistrate's rate list.
2	The Mulyankan Soochi prescribes a separate, higher column rate for 'near Abadi' agricultural land — the script compares the rate actually applied in the transaction against this applicable column.
3	Transactions where the 'near Abadi' Khasra was valued using the lower 'general agriculture' column are automatically flagged.
4	The Python script calculates the revenue shortfall for each case — applying the applicable 'near Abadi' circle rate to the transferred area, computing the correct stamp duty and registration fee payable, and subtracting the amounts actually collected.
5	Output includes a case-wise tabulation of revenue shortfall, ready for inclusion in the Inspection Report.

E. Non-Accountal of 2% Additional Stamp Duty	
1	Download the Fee Register (Syaha) auto-generated by PRERNA software for the selected SRO.
2	Identify transactions in the Syaha where the '2%' column (representing the Additional Stamp Duty under Section 39 of the UP Urban Planning and Development Act, 1973) is blank or zero, despite the transaction being in a developed area where the Additional Stamp duty is applicable.
3	Compare against the base stamp duty column to confirm the 2% was collected from parties but not separately accounted for in the Syaha.
4	This finding indicates a software-level failure in the PRERNA accounting module rather than an individual officer error.

F. Irregular Calculation of Lease Rent on Lease Deeds Exceeding 30 Years (GNIDA Leases)	
1	Filter the registration dataset to identify all instruments classified as 'Patta Vilekh' (पट्टा विलेख) or lease deeds with a term exceeding 30 years, which are potentially chargeable under Article 35(c).
2	Within this filtered subset, apply a further filter to isolate leases registered under the Greater Noida Industrial Development Authority (GNIDA) framework, identifiable from the 'Executant/Claimant' party details showing GNIDA as lessor.
3	Download and manually examine the scanned deeds of flagged instruments from the IGRS portal — no OCR-based text extraction is required at this stage. The examination focuses on reading the premium amount, the applicable lease rent formula (annual rate and escalation clause), and the one-time lease rent option terms as stated in the lease deed.
4	For commercial plots: calculate correct consideration amount as 37.5% of total premium (15-year equivalent at 2.5% per annum). For residential plots: calculate correct consideration as 15% of total premium (15-year equivalent at 1% per annum). Apply the applicable stamp duty rate (as per market value / consideration, whichever is higher) to the correctly computed consideration.
5	Compare the stamp duty and registration fee actually levied (as reflected in the Syaha) against the correctly computed amount. Compute the revenue shortfall for each case and tabulate SRO-wise.

G. Agricultural Land on State Highway / National Highway / Link Road Valued at General Agriculture Rate

- 1 Using the digitised Mulyankan Soochi, identify Khasras designated under higher-value columns such as 'National Highway', 'State Highway', or 'Link Road' in the rate list.
- 2 Cross-reference against the Deeds dataset to flag cases where these Khasras were valued using the general agriculture column.
- 3 Python script computes the revenue shortfall based on the difference between the applicable road-proximity rate and the rate actually applied.

H. Incorrect PAN / Aadhaar Number and Mobile Number in Registration Records

- 1 Extract the buyer and seller PAN (Permanent Account Number) and mobile number fields from the Deeds Details (Fehrist-2) auto-generated by PRERNA software.
- 2 Apply validation checks: PAN format validation (10-character alphanumeric, specific positional rules), mobile number format validation (10 digits, valid network prefixes).
- 3 Flag records where PAN numbers appear in incorrect formats, are repeated across unrelated parties, or mobile numbers appear clearly erroneous.
- 4 Under the Income Tax Act, both buyer and seller are mandated to provide valid PAN for immovable property transfers. Incorrect PAN entries adversely affect tax liability determination including TDS on high-value transactions.
- 5 This finding highlights a systemic data-quality issue in PRERNA's input validation and warrants a software-level correction requiring mandatory format checking at the time of document presentation.

I. Properties Not Being Valued through PRERNA Software Auto-Valuation

- 1 Examine the 'Property Details' (संपत्ति का विवरण) column in the Fehrist-2 dataset for completeness of boundary (Chouhaddi) details and surrounding land-use description.

2	PRERNA software's auto-valuation function requires specific fields — road name, road width, park proximity, surrounding activity type (agricultural/residential/commercial) — to compute market value automatically.
3	Flag records where these mandatory descriptive fields are blank or generic, indicating the registering authority bypassed the software's auto-valuation capability and accepted the party's declared value without system-level cross-checking.
4	This defeats the core purpose of PRERNA's adoption — accurate, objective property valuation — and creates the conditions for systematic undervaluation.

8. AUDIT FINDINGS — IRREGULARITIES IDENTIFIED

Audit of registration records of eight Sub-Registrar Offices across Ghaziabad and Gautam Buddha Nagar districts, conducted through secure read-only access to the IGRS portal, revealed nine categories of revenue irregularities. The findings are systemic in nature and indicate institutional gaps in valuation practices, instrument classification, lease rent computation, and software design that likely exist across multiple SROs in the State. Individual case details are maintained in the working papers; this section presents the nature, legal basis, and aggregated quantum of each finding category.

8.1 SUMMARY OF FINDINGS ACROSS ALL SAMPLED SROs

		SRO-I	SRO-II	SRO-III	Modinagar	SRO-GN	Dadri	SRO-I Noida	SRO-II Noida	Total (₹L)	SRO-I	SRO-II	SRO-III	Mng.	SRO-GN	Dadri	Noida	SRO-II Noida	Total Cases
F1	Security Bond in Place of Mortgage Deed	21.90	109.17	35.83	4.56	0	0	0	0	171.46	2	11	5	1	0	0	0	0	19
F2	Deposit of Title Deed in Place of Mortgage Deed	0	5.63	0	0	0	0	0	0.10	5.73	0	3	0	0	0	0	0	1	4
F3	Residential Land Valued at Agriculture Rate	33.75	12.72	3.95	0	0	0	0	0	50.42	5	3	1	0	0	0	0	0	9
F4	Violation of Section 27 — Concealment of	23.20	333.58	20.56	88.93	299.66	179.26	446.59	338.66	1730.44	7	41	15	17	66	58	52	38	294

		SRO-I	SRO-II	SRO-III	Modinagar	SRO-GN	Dadri	SRO-I Noida	SRO-II Noida	Total (₹L)	SRO-I	SRO-II	SRO-III	Mng.	SRO-GN	Dadri	Noida	SRO-II Noida	Total Cases
F5	Residential Character																		
	Agriculture Land near ABADI Valued at General Agriculture Rate	45.81	3.17	0	0	0	0	0	0	48.98	13	2	0	0	0	0	0	0	15
F6	Agriculture Land on NH/State Highway/Link Road at General Rate	0	3.63	0	1.73	4.54	46.74	0	0	56.64	0	4	0	3	14	9	0	0	30
F7	Lease Rent Irregularity (GNIDA Leases Exceeding 30 Years)	0	0	0	0	333.17	538.31	65.30	80.95	1017.73	0	0	0	0	14	8	17	14	53
F8	Non-Accountal of 2% Additional Stamp Duty	Yes	Yes	Yes	Yes	Yes	No	No	No	—	—	—	—	—	—	—	—	—	Mixed
F9	Properties Not Valued via PRERNA Auto-Valuation Module	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	—	—	—	—	—	—	—	—	—	Selected
F10	Incorrect PAN / Mobile Number in PRERNA Records	No	Yes	Yes	No	No	Yes	No	No	—	—	—	—	—	—	—	—	—	Selected
	GRAND TOTAL	124.66	467.90	60.34	95.22	637.37	764.31	511.89	419.71	3081.40	27	64	21	21	94	75	69	53	424

8.2 VISUAL SUMMARY — REVENUE IMPACT ANALYSIS

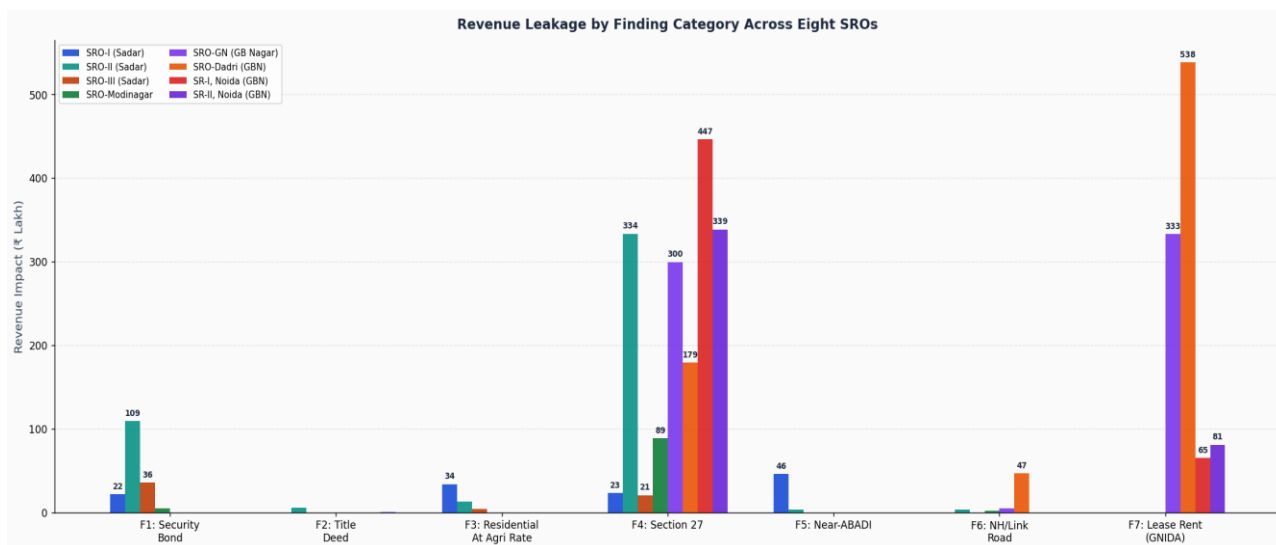


Figure 4: Revenue Impact by Finding Category Across eight SROs (₹ Lakh)

Share of Total Revenue Leakage by Finding Category (All 8 SROs)

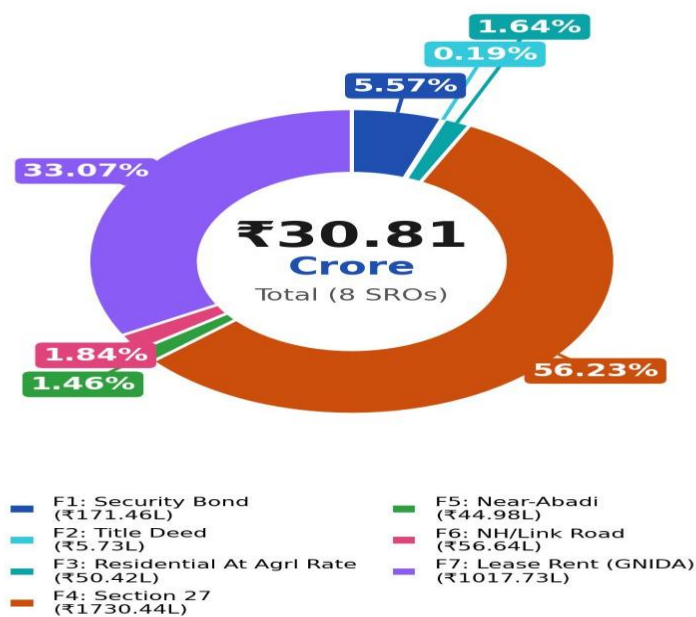


Figure 5: Share of Total Revenue Leakage by Finding Category

8.3 DESCRIPTION OF INDIVIDUAL FINDING CATEGORIES

Finding 1 — Security Bond in Place of Mortgage Deed

Legal Basis: **Section 2(17) of the Indian Stamp Act, 1899; Article 40 of Schedule IB.**

Nature: Developers mortgaged project land in favour of the Ghaziabad Development Authority (GDA) as security for payment of Internal Development Charges (IDC), but executed the instruments as Security Bonds under Article 57, attracting nominal stamp duty of ₹100–₹5000. Since the developers created a charge on their own immovable property to secure a liability borne by them, the instruments are classifiable as Mortgage Deeds under Section 2(17) and chargeable under Article 40. Although the State Government Notification dated 25.05.2001 caps stamp duty on mortgage deeds at ₹5 lakh, the instruments remain liable to stamp duty up to the prescribed cap along with applicable registration fees. Similar instances were observed in memoranda relating to Deposit of Title Deeds under Section 58F of the Transfer of Property Act. The issue was found across all SROs registering GDA-related instruments.

Quantum: SRO-I: ₹21.90 lakh (2 cases); SRO-II: ₹109.17 lakh (11 cases); SRO-III: ₹35.83 lakh (5 cases); Modinagar: ₹4.56 lakh (1 case). **Total: ₹171.46 lakh (19 cases).**

Finding 1B — Deposit of Title Deed in Place of Mortgage Deed

Legal Basis: **Section 58(f) of the Transfer of Property Act, 1882; Article 40 of Schedule IB, Indian Stamp Act, 1899.**

Nature: Memoranda of deposit of title deeds, where a debtor deposits original title documents with a creditor to create an equitable mortgage, are liable to stamp duty under Article 40 as Mortgage Deeds. Registering these instruments at a nominal stamp duty rate results in revenue loss.

Quantum: SRO-II: ₹5.63 lakh (3 cases); SRO-II, Noida: ₹0.10 lakh (1 case). **Total: ₹5.73 lakh (4 cases).**

Finding 2 — Residential Land Valued at Agriculture Rate

Legal Basis: **Indian Stamp Act, 1899 (as amended for UP); UP (Property Valuation) Rules, 1997; Inspector General Registration Circular No. 4302/2018 dated 27.11.2018.**

Nature: Properties situated in Khasra numbers classified as residential (Abadi) in the District Magistrate's Circle Rate List were registered and valued at the lower rates applicable to general agricultural land. The Circle Rate List specifically prescribes Khasra-wise land-use classifications, with residential land attracting substantially higher valuation rates than agricultural land. However, the PRERNA software master data does not incorporate

these Khasra-specific classifications, enabling selection of incorrect land categories without generating system validation alerts, resulting in undervaluation and consequent short levy of stamp duty and registration fees.

Quantum: SRO-I: ₹33.75 lakh (5 cases); SRO-II: ₹12.72 lakh (3 cases); SRO-III: ₹3.95 lakh (1 case). **Total: ₹50.42 lakh (9 cases).**

Finding 3 – Violation of Section 27: Concealment of Residential Character

Legal Basis: Section 27 of the Indian Stamp Act, 1899, which requires full and true disclosure of all facts and circumstances affecting the chargeability of stamp duty.

Nature: In numerous instruments across the sampled SROs, land was declared and valued as agricultural despite evidence indicating a residential character. Analysis revealed that the same Khasra numbers either contained or were contiguous to developed residential colonies, plotted housing schemes, constructed buildings, and residential transactions registered at higher residential rates. Cross-analysis of registration records further disclosed a recurring pattern of both residential and agricultural transactions being registered within the same Khasra, suggesting misdeclaration of land use and consequent undervaluation. Despite the availability of verification tools and land-use indicators within the PRERNA registration platform, these indicators were not effectively utilised during the registration process. This constituted the highest-value category of audit findings across the sampled SROs.

Quantum: SRO-I: ₹23.20 lakh (7 cases); SRO-II: ₹333.58 lakh (41 cases); SRO-III: ₹20.56 lakh (15 cases); Modinagar: ₹88.93 lakh (17 cases); SRO-GN (GB Nagar): ₹299.66 lakh (66 cases); SRO-Dadri (GB Nagar): ₹179.26 lakh (58 cases); SRO-I, Noida (GB Nagar): ₹446.59 lakh (52 cases); SRO-II, Noida (GB Nagar): ₹338.66 lakh (38 cases). **Total: ₹1,730.44 lakh (294 cases) — the single largest finding category.**

Finding 4 – Agricultural Land Near Abadi Valued at General Agricultural Rate

Legal Basis: Indian Stamp Act, 1899 (as amended in Uttar Pradesh); Uttar Pradesh Stamp (Valuation of Property) Rules, 1997; Inspector General of Registration Circular No. 4302/2018.

Nature: Several properties situated in Khasra numbers classified as 'Near Abadi' (आबादी के निकट) in the District Magistrate's Circle Rate List were valued at the lower rates applicable to general agricultural land. Under the prescribed valuation framework, land falling in the 'Near Abadi' category attracts a higher minimum valuation owing to its greater development potential and proximity to residential settlements. However, in the identified cases, the registering authorities applied the general agricultural land rates instead of the prescribed 'Near Abadi' rates, resulting in undervaluation and consequent short levy of stamp duty and registration fees. The issue was facilitated by the absence of Khasra-wise 'Near Abadi' classifications in the PRERNA software master data, which allowed incorrect category selection without generating validation alerts.

Quantum: SRO-I: ₹45.81 lakh (13 cases); SRO-II: ₹3.17 lakh (2 cases). **Total: ₹48.98 lakh (15 cases).**

Finding 5 – Agricultural Land Abutting State/National Highways or Link Roads Valued at General Agricultural Rate

Legal Basis: Indian Stamp Act, 1899 (as amended in Uttar Pradesh); Uttar Pradesh Stamp (Valuation of Property) Rules, 1997.

Nature: The District Magistrate's Circle Rate List prescribes separate and higher valuation rates for agricultural land situated along National Highways, State Highways, and Link Roads, recognising its superior accessibility and higher development potential. Audit scrutiny revealed instances where such land was valued at the lower rates applicable to general agricultural land instead of the prescribed road-frontage rates. The incorrect application of valuation categories resulted in undervaluation of the properties and consequent short levy of stamp duty and registration fees.

Quantum: SRO-II: ₹3.63 lakh (4 cases); Modinagar: ₹1.73 lakh (3 cases); SRO-GN (GB Nagar): ₹4.54 lakh (14 cases); SRO-Dadri (GB Nagar): ₹46.74 lakh (9 cases). Total: ₹56.64 lakh (30 cases).

Finding 6 – Irregular Computation of Lease Rent in Lease Deeds Exceeding 30 Years

Legal Basis: Article 35(c) of Schedule I-B of the Indian Stamp Act, 1899 (as applicable in Uttar Pradesh); lease conditions prescribed by the New Okhla Industrial Development Authority (NOIDA) & Greater Noida Industrial Development Authority (GNIDA).

Nature: NOIDA & GNIDA allots residential, commercial, industrial, and institutional plots on long-term leases, generally for a period of 90 years. The lease conditions prescribe annual lease rent at specified percentages of the premium, with periodic enhancements, and also provide an option for payment of one-time lease rent calculated as a multiple of the annual lease rent. Under Article 35(c), lease deeds executed for a term exceeding 30 years are chargeable to stamp duty in the same manner as conveyances, based on the market value or consideration, whichever is higher. Audit scrutiny of lease deeds registered in Gautam Buddha Nagar district revealed that the consideration amount adopted for stamp duty purposes did not correctly incorporate the applicable lease rent components prescribed under the lease terms. As a result, the consideration value was understated, leading to short assessment of stamp duty and registration fees. The issue was predominantly observed in registrations involving NOIDA & GNIDA lease transactions.

Quantum: SRO-GN: ₹333.17 lakh (14 lease deeds); SRO-Dadri: ₹538.31 lakh (8 lease deeds); SRO-I, Noida: ₹65.30 lakh (17 lease deeds); SRO-II, Noida: ₹80.95 lakh (14 lease deeds).

Combined Total (SRO-GN + SRO-Dadri + SRO-I, Noida + SRO-II, Noida): ₹1,017.73 lakh (53 lease deeds). This finding is concentrated in SRO-GN, SRO-Dadri, SRO-I, Noida and SRO-II, Noida given the concentration of GNIDA, NOIDA lease transactions in Gautam Buddha Nagar district.

Finding 7 – Non- Accountal of Additional Stamp Duty

Legal Basis: Section 39 of the Uttar Pradesh Urban Planning and Development Act, 1973; Government Order No. 16/2024/1356/94-1-2024-312(35)/2023 dated 03.09.2024.

Nature: In respect of properties situated within notified development areas, an additional stamp duty of 2 per cent is leviable on the consideration amount and is collected at the time of registration. Audit scrutiny revealed that, although the additional duty was being collected from the parties, the Syaha (Fee Register) auto-generated through the PRERNA software did not separately capture or account for this amount in the prescribed column. Consequently, the additional stamp duty collected could not be distinctly identified, reconciled, or distributed in accordance with the revised accounting and distribution mechanism prescribed by the Government Order dated 03.09.2024. The issue arises from a deficiency in the system design and accounting module of the PRERNA software and represents a systemic control weakness rather than an error attributable to individual registering officers.

Finding 8 — Incorrect PAN / Mobile Number in Registration Records

Legal Basis: Income Tax Act provisions mandating PAN for immovable property transfers; PRERNA system design requirements.

Nature: Sample examination of the Fehrist-2 (Deed Index) revealed instances of incorrect, malformed, or repeated PAN numbers and mobile numbers entered for buyers and sellers in the PRERNA system's auto-generated register. This adversely affects tax liability assessment including TDS on high-value transactions and source-wise tax deduction. PRERNA lacks mandatory format-validation checks at the point of document presentation, allowing incorrect data to be recorded without system alerts. The Department should implement mandatory PAN format validation in PRERNA.

Finding 9 – Non-Utilisation of PRERNA's Auto-Valuation Module

Legal Basis: PRERNA software valuation framework; District Magistrate Circle Rate List and prescribed property valuation requirements.

Nature: The PRERNA system incorporates an auto-valuation module designed to determine the applicable circle rate and market value based on key property attributes, including road name, road width, proximity to parks or public amenities, and surrounding land-use characteristics. Effective functioning of the module requires these details to be accurately captured in the property description at the time of registration. Audit scrutiny of Fehrist-2 records revealed that a substantial number of registered instruments lacked these mandatory descriptive particulars in the Property Details column, indicating that the auto-valuation functionality was either not being utilised or was being bypassed during registration. The absence of critical valuation parameters undermines the system's ability to automatically determine the correct valuation category, weakens embedded revenue

safeguards, and creates conditions conducive to undervaluation and short levy of stamp duty and registration fees across multiple transaction categories.

9. SYSTEMIC ISSUES AND ROOT CAUSE ANALYSIS

The audit findings reveal structural gaps rather than isolated officer-level failures. The root causes are identified below:

PRERNA Software — Master Data Gaps

7.1

Systemic Issue: A key control weakness identified is the lack of integration between PRERNA's valuation master data and the District Magistrate's Mulyankan Soochi (Circle Rate List). Khasra-specific classifications prescribed in the official rate lists are not fully embedded in the system, preventing automated validation of the applicable valuation category. As a result, transactions may be processed at lower rates without generating exception alerts, contributing to the undervaluation issues observed in Findings 2, 3, 4 and 5.

PRERNA Software — Accounting Module (2% Additional Stamp Duty)

7.2

Systemic Issue: The omission of the 2 per cent additional stamp duty from the relevant column of the auto-generated Syaha register indicates a deficiency in PRERNA's accounting module. Although the amount is collected during registration, the system does not separately capture, classify, or account for it in the prescribed records (fee register). This creates a risk of incorrect accounting, reconciliation, and distribution of receipts under the revised framework prescribed by the Government Order of September 2024.

PRERNA Software — Data Validation Gaps

7.3

The absence of mandatory format for validation for PAN and mobile numbers in PRERNA's input screens allows incorrect taxpayer identification data to be recorded, adversely impacting TDS compliance and revenue tracking.

Registering Authority — Non-Utilisation of Software Capabilities

7.4

Systemic/Control Issue: PRERNA incorporates validation and auto-valuation tools designed to assess land characteristics and compute market value based on property particulars and locational attributes. Audit scrutiny indicated that these functionalities were not consistently utilised during the registration process, with property classifications often being accepted on the basis of party declarations without adequate

verification against available system data. This weakened the effectiveness of built-in valuation controls and points to deficiencies in both operational compliance and supervisory oversight.

Absence of Instrument Classification Verification

7.5

Systemic/Control Issue: No effective supervisory or system-based mechanism exists to verify whether instruments executed as Security Bonds in developer–GDA transactions are correctly classified for stamp duty purposes. The recurrence of similar cases across multiple SROs indicates a broader control deficiency, resulting in the risk of misclassification of Mortgage Deeds and consequent short levy of stamp duty and registration fees.

Registering Authority — Incorrect Lease Rent Computation (GNIDA Leases)

7.6

Systemic/Control Issue: In lease deeds executed under the GNIDA allotment framework, the consideration amount was not consistently computed by incorporating all lease rent components prescribed under the applicable lease conditions. As Article 35(c) requires stamp duty to be levied on the market value or consideration, whichever is higher, in respect of leases exceeding 30 years, any understatement of consideration directly results in short levy of stamp duty and registration fees. The recurring nature of the issue indicates inadequate verification of lease deed valuations and insufficient understanding or application of GNIDA-specific lease rent provisions during the registration process.

10. RECOMMENDATIONS

Synchronisation of PRERNA Master Data with Circle Rate Lists:

R1

Establish a formal mechanism to periodically update PRERNA's valuation master database in line with revisions to the District Magistrate's Mulyankan Soochi (Circle Rate List). The system should generate validation alerts wherever the declared land category differs from the prescribed Khasra-wise classification.

R2

Automated Validation of Khasra-Based Classifications:

Enhance PRERNA to automatically cross-reference Khasra numbers with the digitised Circle Rate List and require supervisory approval for transactions involving category mismatches or valuation exceptions.

R3

Rectification of 2% Additional Stamp Duty Accounting Module in PRERNA:

Direct the PRERNA development team to rectify the Syaha accounting module to ensure that the additional 2 per cent stamp duty is separately captured, accounted for, and distributed in accordance with Government Order dated 03.09.2024.

R4

Mandatory Use of Auto-Valuation Features:

Make entry of complete property particulars, including Chouhaddi (boundaries), road details, and surrounding land-use characteristics, mandatory in PRERNA to enable effective functioning of the auto-valuation module.

R5

PAN and Mobile Number Validation Controls:

Introduce system-based validation checks for PAN and mobile numbers at the document presentation stage and prevent submission of instruments containing invalid or incomplete details.

R6

Guidance on Classification of Security Instruments:

Issue departmental instructions for registering authorities on the classification of developer–development authority security instruments, with specific reference to Section 2(17) and Article 40 of the Indian Stamp Act, 1899

R7

Recovery of Short-Levied Revenue:

Initiate action for recovery of short-levied stamp duty and registration fees identified during audit and furnish an action-taken report to the audit office within a prescribed timeframe..

Standardisation of Lease Rent Computation for Long-Term Leases (Development Authority Leases):

- R8** Issue detailed guidelines for assessment of lease deeds exceeding 30 years, requiring inclusion of all applicable lease rent components, escalation provisions, and one-time lease rent payments while determining consideration under Article 35(c) of Schedule I-A of the Indian Stamp Act, 1899

11. DEMONSTRATED EFFICIENCY GAINS AND INSTITUTIONAL IMPACT

The shift to Data-Led Remote Audit methodology has delivered measurable improvements across three dimensions: staff time, travel expenditure, and audit output quality. The following section presents a comparative assessment of the DLRA approach against the traditional on-site audit model, based on the experience of auditing eight Sub-Registrar Offices across Ghaziabad and Gautam Buddha Nagar districts.

DLRA Institutional Impact Summary



Figure 6: Key Efficiency Gains — DLRA vs Traditional Audit

11.1 PERFORMANCE COMPARISON: TRADITIONAL AUDIT VS DATA-LED REMOTE AUDIT

The table below consolidates the comparative performance of the Data-Led Remote Audit (DLRA) approach against the traditional on-site audit model across four dimensions: revenue impact, observation output, time and cost efficiency, and methodology and coverage. All DLRA figures reflect the cumulative position across the eight Sub-Registrar Offices audited so far under this ongoing exercise.

PERFORMANCE METRIC	TRADITIONAL AUDIT	DATA-LED REMOTE AUDIT (DLRA)	JUMP
A. Revenue Impact			
Total Revenue Covered (% of total stamp revenue audited)	0.9% of total stamp revenue	18.71% of total revenue already covered	20.8x more

PERFORMANCE METRIC	TRADITIONAL AUDIT	DATA-LED REMOTE AUDIT (DLRA)	JUMP
Total Money Value of Audit Observations	₹3.14 Crore (traditional audit cycle)	₹30.81 Crore (ongoing — will rise further)	9.8× higher
Average Money Value per SRO Office Audited	₹31.43 Lakh per SRO	₹3.85 Crore per SRO	12.2× per SRO
Total Number of Cases (Audit Observations)	39 cases (traditional audit cycle)	424 cases (across 8 SROs)	10.9× more
B. Observation Output			
Average Observations Issued per SRO	4 observations per SRO	53 observations per SRO (range: 21–94)	13× more obs.
Types of Audit Observations / Findings	2 types (mainly Section 27 concealment)	10 types (comprehensive coverage across all categories)	5× broader
C. Time & Cost Efficiency			
Audit Duration per SRO	5–7 working days (incl. travel + on-site examination of registers)	~3 working days (from headquarters only, no travel required)	40–57% faster
Man-Day Savings (Staff Time per SRO)	Full cycle: travel, examination, document requisition	~3 days/SRO — same team can cover significantly more SROs per planning cycle	Major Saving
TA/DA & Travel Expenditure	Full TA/DA per SRO visit (travel, accommodation, transport)	Zero — 100% eliminated (entire audit conducted from headquarters)	100% Saved
D. Methodology & Coverage			
Sampling Approach (within each SRO)	Manual sampling of physical registers — small sample, high risk of missed cases	100% population coverage — every transaction algorithmically screened, no sampling bias	Full universe

PERFORMANCE METRIC	TRADITIONAL AUDIT	DATA-LED REMOTE AUDIT (DLRA)	JUMP
Documentary Evidence	Physical registers & deeds at SRO premises — requires physical presence	Scanned deeds on IGRS portal — digital evidence, secure remote access	Full remote

11.2 REDUCTION IN AUDIT DURATION

Under the traditional audit model, examination of a Sub-Registrar Office typically required **5–7 working days**, including travel and on-site scrutiny. Under the Data-Led Remote Audit model, the same examination can be completed from headquarters in approximately **3 working days**, resulting in a **40–57% reduction in audit cycle time per SRO**.

11.3 ELIMINATION OF TRAVEL / TA-DA EXPENDITURE

The audit is conducted remotely through the IGRS portal's read-only dashboard, eliminating the need for field visits and associated TA/DA expenditure, thereby enabling coverage of more SROs within the same budget.

11.4 SIGNIFICANT INCREASE IN AUDIT OBSERVATIONS

By replacing manual sampling with full-population algorithmic screening, the Data-Led Remote Audit model increased audit output from about **1-7 observations per SRO** to **21–94 observations per SRO**, representing **improvement in audit detection capability**, with correspondingly higher revenue identification.

DLRA Efficiency Gains: Before vs After Comparison

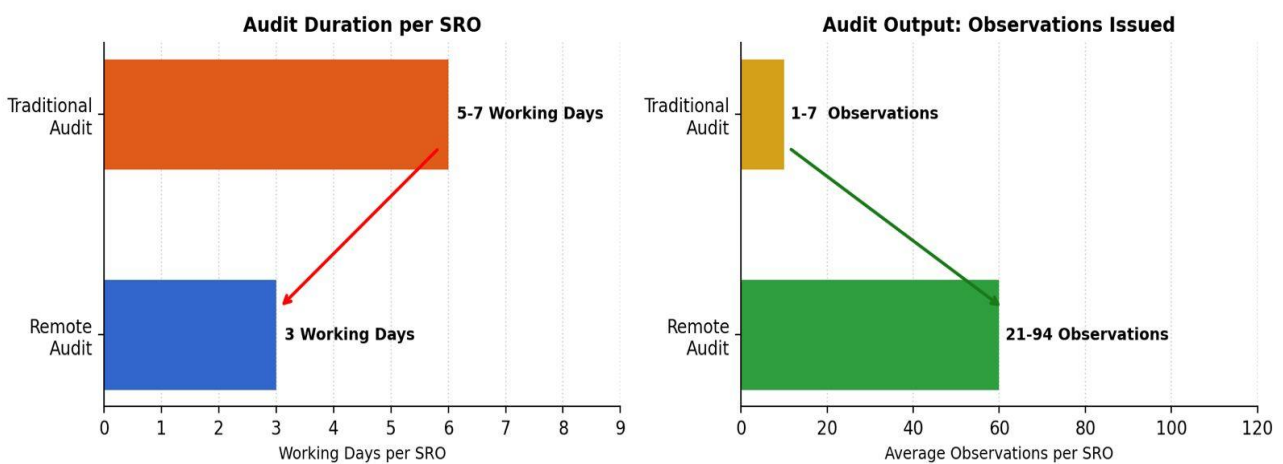


Figure 7: Before vs After — Audit Duration and Observations per SRO

11.5 IGRS PORTAL — DATA ACCESS LIMITATION (PRE-2025 DOCUMENTS)

The IGRS portal currently provides three categories of access: (1) download of Deed Index (Fehrist-2), (2) download of Fee Register (Syaha), and (3) retrieval of scanned deed copies. However, it has been observed that documents registered prior to 2025 are not consistently available for download in the scanned deed section. This limitation affects the ability to examine scanned deeds in older observation cases that require manual instrument verification.

12. WAY FORWARD — SCALING THE DLRA MODEL

THIS IS AN ONGOING AUDIT

The ₹30.81 crore revenue impact reported in this concept paper covers eight Sub-Registrar Offices audited so far. As audit coverage is progressively extended to more SROs across Uttar Pradesh's 75 districts, the cumulative money value of identified irregularities will rise substantially. The systemic nature of the findings — rooted in software gaps and valuation non-compliance that are common across all SROs — means that every additional SRO brought under audit will add significantly to the revenue impact.

12.1 EXTENSION TO ADDITIONAL SROs

The exercise across eight SROs in Ghaziabad and Gautam Buddha Nagar districts has validated the analytical framework, toolchain, and observation methodology. The same scripts, with SRO-specific Mulyankan Soochi inputs, can be deployed for other high-revenue SROs across UP's 75 districts. Revenue-based selection (Stage 1) will ensure audit resources are directed where revenue risk is highest.

12.2 CONTINUOUS AUDIT CYCLE

The IGRS portal's real-time data access opens the possibility of a continuous or quarterly audit model. Pre-defined risk scripts can be run periodically to generate fresh outliers and flag new irregularity patterns as they emerge.

12.3 AI AND OCR AT SCALE

As OCR and AI capabilities mature, automation of deed recital analysis will reduce manual verification time for Findings 1, 4, and 10 — the categories currently requiring most manual review. Investment in AI-assisted document analysis will further enhance the depth and speed of audit scrutiny.

13. CONCLUSION

The Data-Led Remote Audit of the Stamp and Registration Department across eight Sub-Registrar Offices in Ghaziabad and Gautam Buddha Nagar districts represents a significant methodological advance in receipts audit. By leveraging the IGRS portal's digital infrastructure through a audit framework — revenue-based SRO selection followed by full-universe algorithmic analysis within each selected office — the audit has identified revenue leakage of approximately **₹30.81 crore** across eight Sub-Registrar Offices.

The findings are systemic, not episodic: they reflect structural gaps in PRERNA's master data, accounting module, and validation design. Remediation through software upgrades and data synchronisation will have a

far more durable impact than case-by-case recovery. The Office proposes to extend the DLRA framework progressively to all high-revenue SROs in Uttar Pradesh, establishing a continuous, risk-based assurance framework for stamp duty and registration fee realisation across the State.

₹30.81 Crore	8 Offices	10	Python/KNIME/OCR
Total Impact	SROs Covered	Types Finding Categories	Tool Chain

BENEFITS AND WAY FORWARD REFLECTED IN CASE STUDY OF TELANGANA

Issued by	Office of PAG (Audit), Telangana
Intended for	All Offices of the Accountant General (Audit) — Registration & Stamps Audit
Scope	Registration and Stamps Departments across States

1. DATA-LED REMOTE AUDIT FRAMEWORK

1.1 OVERALL APPROACH

Stage	Phase	Key Activities
Preliminary	Data Preparation	Obtain data dump; integrity checks; convert to audit-friendly format; clean and standardise
Stage I	Analytical Screening	Apply pre-defined offline tool checks and logical queries to identify outliers and risk indicators across the full dataset
Stage II	Document Scrutiny	Download and examine selected documents using OCR-based offline tools; confirm or rule out audit suspicion
Stage III	Reporting	Issue observations through OIOS; collate departmental responses; prepare Inspection Report identifying systemic flaws

1.2 ADVANTAGES

Direct Access: Audit teams obtain read-only access to the State registration portal or receive data dumps from the Department, without physically visiting any SRO.

State-wide Analysis: Offline analytical tools enable examination of large volumes of records to identify outliers and short levy of stamp duty across all SROs simultaneously, giving a holistic picture of systemic lapses.

Speed of Mathematical Checks: Rigid mathematical rules such as stamp duty computation and market value comparison can be verified across the entire dataset at once which is not feasible in manual audit.

1.3 CHALLENGES

Volume of Data: Transaction data is obtainable in bulk, but scanned documents must typically be downloaded individually from the registration portal, which is time-consuming.

Data Confidentiality: The data obtained contains personal asset information including details of public figures. Strict data handling protocols must be followed; data must not leave the office network.

Vetting of Recitals: Data analysis identifies suspected transactions but cannot by itself confirm an irregularity. Each flagged document requires its recitals to be examined manually. The office-developed OCR tools assist significantly in this step.

2. PRELIMINARY STAGE — DATA PREPARATION

2.1 OBTAINING DATA

A formal data requisition shall be addressed to the Commissioner, Registration and Stamps of the State, specifying:

- The period for which data is sought;
- The format required — machine-readable formats such as CSV or Excel are preferred;
- The list of mandatory fields (Section 2.2);
- A clear deadline for compliance and the audit team contact details.

The Department shall simultaneously be requested to provide read-only access to the State registration portal for direct retrieval of scanned deed copies. Non-compliance within the stipulated timeline shall be escalated through the appropriate channel.

2.2 MANDATORY DATA FIELDS

The following fields, specific to Telangana State, may be included in the data dump to enable the audit checks. Field names may vary by State, therefore, the offices should map these to the equivalent fields in the local registration system:

S.No.	Field	Purpose in Audit
1	SRO Name and Code	Unit-level stratification and grouping of findings
2	Document Number and Date	Primary identifier; duplicate detection; date-range validation
3	Nature of Deed	Filtering by deed type for targeted checks

S.No.	Field	Purpose in Audit
4	Market Value (Government Rate)	Undervaluation detection; stamp duty base computation
5	Consideration Value	Stamp duty base; higher of Market Value or consideration value
6	Stamp Duty Paid	Shortfall computation
7	Registration Fees Paid	Fee shortfall computation
8	Boundaries on all Four Sides	Completeness of recital; concealment detection
9	Area (with unit — Sq.Ft / Sq.Yds / Acres as applicable in the State)	Valuation computation; unit consistency check
10	Village / Locality; Sub-District; District	Geographic cross-reference with market value guidelines
11	Survey / Plot / Door Number	Land category matching; chain transaction detection
12	Transaction ID	Unique identifier across systems
13	Executant (Seller / Donor) Name	Relationship check for gift deeds
14	Claimant (Buyer / Donee) Name	Relationship check; family concession eligibility
15	Stated Relationship (for Gift Deeds)	Family member concession verification
16	Number of Pages	Document completeness; OCR workload estimation
17	Link SRO Code and Link Document Number	Prior transaction reference — for previous-value comparison check

2.3 INTEGRITY CHECK

Before proceeding to analytical screening, validate the data using MS Excel or IDEA:

- **Duplicate Documents:** Check for duplicate document numbers within the same SRO and year.
- **Null Values:** Identify blank entries in mandatory fields — Stamp Duty, Market Value, Consideration, Nature of Deed.
- **Data Type Compliance:** Verify that numeric fields (Stamp Duty, Consideration, Area) do not contain text entries.
- **Date Range:** Confirm all registration dates fall within the period for which data was requested.
- **Consistency:** Where data has been received in tranches, cross-check record counts across tranches for omissions.

2.4 CONVERSION TO AUDIT-FRIENDLY FORMAT AND DATA CLEANING

Convert the data to a format suitable for the tools available. MS Excel with Power Query or IDEA is sufficient for most checks at the volumes typically involved in registration audits. Data cleaning shall address:

- Normalising the Nature of Deed field — standardise variations in capitalisation and spelling;
- Removing commas and special characters from Consideration and Stamp Duty columns so they are treated as numbers;
- Standardising area to a single unit for computational checks;
- Standardising locality / village name spellings to match the market value reference dataset used in the State.

The cleaned dataset shall be saved as a separate master file. The original data as received must be retained unchanged as evidence.

3. IDENTIFICATION OF RISK INDICATORS

Before running targeted checks, stratify the data by SRO, deed type, consideration value, and geographic location to identify SROs with anomalous stamp duty collections, deed types with disproportionately low average consideration, and geographic clusters of zero or concessional stamp duty transactions.

The following risk indicators shall be specifically flagged for Stage I analytical checks:

Sl.No.	Risk Indicator	Possible Irregularity
1	Short Levy of Stamp Duty Due to Adoption of Lower of Market Value and Consideration Value Instead of the Higher Applicable Value.	Short levy — the legal requirement is to charge on the higher of the two
2	Previously registered value of a property is higher than the current transaction value	Possible suppression of consideration; higher of the two values is required to be adopted
3	Gift deeds where the stated relationship falls outside the definition of 'relative' under Section 56(2) of Income Tax, Act	Short Levy of Stamp Duty Due to Ineligible Availment of Family Member Stamp Duty Concession.
4	Transfer of Development Rights (TDR) instruments registered without stamp duty or at nominal value.	Short levy — TDR transfers attract conveyance duty on consideration / market value.

Sl.No.	Risk Indicator	Possible Irregularity
5	Zero or concessional stamp duty on deeds that do not qualify for the claimed exemption	Incorrect exemption; possible collusion at registration
6	Deposit of Title Deed (MSME / equitable mortgage) registered at lower duty rate	Under-stamped mortgage instrument; revenue loss
7	Misclassification trends — instruments registered under a lower-duty deed type	Higher-duty instrument wrongly classified to reduce stamp duty liability

Note on Scope

Detailed step-by-step methodology is provided in Section 4 for checks where data and documents have been received and tested. Remaining checks are listed in Annexure A as 'To be detailed in the next version'. Methodology for those checks will be documented as audit experience grows.

Risk Indicator 5 (TDR) is currently documented specifically in the context of Municipal Corporation-issued TDR certificates (see Section 4.2). Offices in States where TDR schemes are operated by different bodies should adapt the terminology and legal references accordingly.

4. STAGE I — ANALYTICAL METHODOLOGY

Stage I applies pre-defined checks to the cleaned dataset to identify transactions with a high probability of revenue irregularity. Checks are conducted using offline JavaScript-based tools developed for this audit, in combination with MS Excel and IDEA. No audit data is shared with any external system at this stage.

4.1 GIFT DEED — FAMILY MEMBER CONCESSION CHECK

Legal Basis

Concessional stamp duty on gift deeds is admissible only where the transfer is made in favour of a 'relative' as defined under Section 56(2) of the Income Tax Act, 1961. The definition covers the spouse, siblings, lineal ascendants and descendants, and their respective spouses. Relationships such as nephews, nieces, cousins, friends, and other unrelated persons do not qualify for the concession. Accordingly, where a gift deed has been registered at a concessional rate despite the stated relationship falling outside the prescribed definition, the differential stamp duty is recoverable.

How the Check is Performed

1. **Identify potentially concessional gift deeds** by filtering the registration dataset for instruments registered at concessional family-member rates or where the stamp duty rate is lower than the standard rate applicable to gift deeds.
2. **Download deed copies** from the registration portal using the office-developed Document Downloader tool:
 - Prepare an input Excel file containing the following fields: District, Sub-Registrar Office, Book Number, Document Number, and Registration Year.
 - Upload the Excel file into the tool, which automatically populates the portal search fields and downloads the corresponding documents in batch mode.
 - In cases where the portal returns an error, use the tool's retry functionality to reattempt the download without re-entering document details.
3. **Process the downloaded PDFs** through the office-developed Gift Deed Audit Tool, which analyses the deed content and identifies cases requiring further scrutiny based on the relationship declared between the parties.

```
// For each uploaded deed PDF:
// 1. Extract all pages as images using PDF.js
// 2. Run Tesseract.js OCR on each page (English + regional language packs)
// 3. Combine OCR text from all pages into a single string per document

// Relationship extraction — search English text patterns:
const eligibleRelatives = [
  /\b(son|daughter|wife|husband|mother|father|brother|sister|spouse)\b/i,
  /\b(grandson|granddaughter|grandfather|grandmother)\b/i
];
const ineligibleRelatives = [
  /\b(nephew|niece|cousin|friend)\b/i // explicitly excluded
];

// Regional language patterns should be configured per State
// (Example: Telugu patterns for son/daughter/wife are pre-loaded in the tool)

// Verdict logic:
if (ineligibleRelationshipFound) {
  verdict = 'FLAG FOR SCRUTINY';
  action = 'Recover differential stamp duty';
} else if (eligibleRelationshipFound) {
  verdict = 'ELIGIBLE';
  action = 'No action required';
} else {
  verdict = 'RELATIONSHIP NOT DETECTED — MANUAL REVIEW';
}
```

4. **Manually scrutinise** all deeds flagged for review to verify the relationship stated in the recitals and determine whether it qualifies as a 'relative' under Section 56(2) of the Income Tax Act, 1961.
5. **Recalculate the applicable stamp duty** by applying the non-concessional rate on the market value or consideration, whichever is higher, and determine the short levy by deducting the stamp duty already paid.
6. **Aggregate and quantify** the short levy of stamp duty SRO-wise for reporting and recovery purposes.

Tools Used — Gift Deed Check

Document Downloader Tool: Automated batch download of deed copies from the State registration portal. Handles form-filling, retry on errors, and per-document status tracking.

Gift Deed Audit Tool: Offline HTML tool using Tesseract.js OCR (English + regional language). Batch processes deed PDFs. Applies Section 56(2) of Income Tax Act, relationship validation. Exports audit CSV with verdict and action column.

Both tools run entirely in the browser on the auditor's local machine. No data is transmitted externally.

4.2 TRANSFER OF DEVELOPMENT RIGHTS (TDR) — STAMP DUTY CHECK

Background — Municipal TDR Schemes

Transferable Development Right (TDR) certificates are issued by Urban Development Authorities and Municipal Bodies to landowners whose land is acquired or surrendered for public purposes such as roads, parks, and other public amenities. These certificates confer transferable development rights and possess market value. Audit scrutiny focuses on cases where transfers of TDR certificates are registered as gifts or other nominal instruments attracting lower stamp duty, potentially resulting in undervaluation and short levy of stamp duty.

State-Specific Context — Municipal Corporation TDRs

This methodology has been developed and tested in the context of TDR certificates issued by the Greater Hyderabad Municipal Corporation (GHMC) and the Hyderabad Metropolitan Development Authority (HMDA), which represent the predominant category of TDR transactions in urban areas of the pilot State.

In other States, TDRs may be issued by the respective Municipal Corporation, Development Authority, or Town Planning body. The instrument name, the issuing authority, and the applicable Government Orders will differ — offices should identify the relevant legal framework and GO references for their State before applying this check.

A significant number of Gift Deeds in favour of Municipal bodies / Local Bodies involving TDR issuances are registered either without stamp duty or at a nominal value, rather than on the market value of the extent of land for which Development Rights are conveyed.

Legal Basis

Article 20 of Schedule I-A of the Indian Stamp Act, governing conveyance deeds; the provisions of the Greater Hyderabad Municipal Corporation (GHMC) Act, 1955; the Hyderabad Metropolitan Development Authority (HMDA) Master Plan provisions relating to Transferable Development Rights (TDR); and the Government Orders, notifications, and guidelines

issued from time to time by the Municipal Administration and Urban Development (MA&UD) Department regulating the issuance, transfer, and utilisation of TDR.

How the Check is performed

1. Filter the dataset for all instruments where the Nature of Deed is 'Gift in favour of GHMC/Local Body' and download the PDFs and scanned images using the Tool
2. Apply a keyword search like 'TDR', 'Transfer of Development Rights', 'Development Rights', 'Transferable Development Right', or 'TDR Certificate' across the text of downloaded deed PDFs using the office-developed PDF and Image Word Finder tool. The core logic of the tool is:

```
// For each uploaded file (PDF or image):
// Step 1 — Attempt text extraction from PDF text layer via PDF.js:
const pdfText = await extractTextFromPDF(file);

// Step 2 — If text layer is empty or too short (scanned deed), fall back to OCR:
if (pdfText.trim().length < 50) {
  const ocrText = await runTesseract(pageImage); // Tesseract.js, eng + regional lang
}

// Step 3 — Search for TDR-related keywords:
const tdrKeywords = [
  'TDR',
  'Transfer of Development Rights',
  'Transferable Development Right',
  'TDR Certificate',
  'Development Right Certificate',
  'GHMC TDR', // configure per State — replace with local issuing body
  //add relevant GO numbers or scheme names as needed
];
const found = tdrKeywords.some(kw =>
  combinedText.toLowerCase().includes(kw.toLowerCase())
);

// Output per file:
// File name | Pages searched | TDR keyword found (Yes/No) | Matching page numbers
// Export as CSV for audit working paper
```

- This step identifies cases where TDR is present in the deed recitals.
3. Check the TDR certificate number cited in the deed. Cross-reference with the TDR register (if available) to confirm the conveyance of TDR Rights.
 - For each TDR instrument identified, flag all cases where Stamp duty is zero or nil; or
 - Stamp duty has been paid on only the nominal registration value rather than the market value of the extent for which Development Rights are conveyed
 4. Download and manually examine the recitals of each flagged instrument to confirm by crosschecking with Government records

5. Compute the revenue shortfall: applicable stamp duty for conveyance × market value of the extent of TDR minus stamp duty actually paid.
6. Aggregate shortfall SRO-wise. TDR transactions tend to be geographically concentrated in SROs covering the jurisdiction of the issuing Municipal Body

Tool Used — TDR Check

PDF and Image Word Finder: Offline HTML tool that batch-searches TDR-related keywords across uploaded deed PDFs and scanned image files. Applies OCR where the PDF has no embedded text layer. Returns a file-wise list showing which documents contain TDR keywords, the matching page numbers, and the surrounding text context.

The tool's keyword list is configurable — auditors should add the name of the local issuing body, relevant GO numbers, and scheme-specific terms for their State to narrow results appropriately.

4.3 STAMP DUTY CALCULATED ON LOWER OF MARKET VALUE / CONSIDERATION

Legal Basis

Under the Indian Stamp Act, stamp duty on conveyance instruments is mandatorily calculated on the higher of: (a) the consideration stated in the deed, or (b) the Government-notified market value for the area. Calculating Stamp Duty on the lower of the two values constitutes a violation of the Act and results in short levy.

How the Check is performed

1. Compute the higher of the **Market Value** and **Consideration Amount** for each sale deed or conveyance instrument.
2. Calculate the **expected stamp duty** by applying the applicable duty rate to the higher of the two values.
3. Compare the expected stamp duty with the **stamp duty actually paid** and flag cases where the duty paid is lower than the amount payable, excluding minor rounding differences.
4. Analyse flagged records SRO-wise to identify patterns indicative of systemic under-assessment.
5. Perform manual scrutiny of high-value cases by examining the registered deed and supporting details to confirm the short levy and quantify the revenue implication.

4.4 PREVIOUSLY REGISTERED VALUE HIGHER THAN CURRENT TRANSACTION VALUE

Legal Basis

As per the orders issued by the Commissioner and Inspector General of Registration, Telangana, where a higher value has been previously adopted for a property, such value is required to be considered for subsequent transactions relating to the same property. Accordingly, where the consideration value or market value declared in a subsequent sale deed is lower than the value adopted in an earlier transaction of the same property, stamp duty and registration fees are required to be levied on the previously adopted higher value.

How the Check is Performed — Using IDEA and the Link Document Table

The Computer-Aided Administration of Registration Department (CARD) system maintains transaction history through a Link Document Table, which records previous transactions associated with a registered property.

1. Import the **Link Document Table** into IDEA as a separate dataset along with the main registration dataset.
2. The table contains the following key fields:

Field	Description
Document Number (Link)	Document number of the previous transaction relating to the property
Registration Date (Link)	Registration date of the previous transaction
SRO (Link)	Sub-Registrar Office where the previous transaction was registered
Document Number (Main)	Document number of the current transaction
Registration Date (Main)	Registration date of the current transaction
SRO (Main)	Sub-Registrar Office where the current transaction was registered

3. In IDEA, join the main transaction dataset with the Link Document Table using the common property identifier (SRO Code + Registration Year + Document Number).
4. Extract all records where the taxable value adopted in the linked previous document exceeds the taxable value adopted in the current document, and rank the results in descending order of value difference.
5. Apply additional filters to retain cases where the extent of land, plinth area, and other key property attributes are identical or substantially similar in both transactions.
6. For flagged cases, obtain and examine both the current deed and the linked previous deed from the CARD portal. Compare the property description, extent, nature of transfer, consideration value, and taxable value adopted in each transaction.
7. Where the same property, or substantially the same property, has been valued lower in the current transaction than in the linked previous transaction without adequate justification, quantify the differential duty and raise an audit observation.

Tool — IDEA Link Document Table Approach

This check is performed entirely within IDEA using its table joining capability. The Link Document table is a standard component of the data dump from the State registration system. No additional offline tools are needed for this check beyond IDEA.

Caution: The Link Document table may contain multiple prior transactions per current document (where a property has changed hands more than once). Use the most recent prior registration as the primary comparator, or compare against the highest value across all prior transactions, whichever is more appropriate.

This check is particularly effective for SROs covering areas with rapidly appreciating property values, as the gap between prior and current values is likely to be more pronounced in such areas.

4.5 MSME — DEPOSIT OF TITLE DEED / EQUITABLE MORTGAGE

Legal Basis

Under Section 58(f) of the Transfer of Property Act, 1882, a mortgage by deposit of title deeds is created when a borrower deposits title documents with a creditor as security for a loan without executing a formal mortgage deed. Instruments evidencing such transactions are chargeable to stamp duty as Memoranda of Deposit of Title Deeds (MoDTD). In certain cases, concessional stamp duty is available to enterprises classified as **Micro** or **Small** under the Udyam Registration framework. Enterprises classified as **Medium** are not eligible for the concession. Incorrect grant of the concession or under-stamping of the instrument results in short levy of stamp duty.

How the Check is Performed

1. Filter the registration dataset to identify all instruments classified as **Deposit of Title Deeds / Memorandum of Deposit of Title Deeds (MoDTD)** and download the corresponding deed PDFs and scanned documents using the office-developed Document Downloader tool.
2. Load the downloaded documents into the office-developed **MSME Title Deed Checker Tool**.
3. Search the deed documents and supporting uploads for keywords such as “**MSME**,” “**Udyam**,” “**Micro**,” “**Small**,” and “**Medium**” to identify cases where concessional duty may have been claimed.
4. The core logic of the tool is:

```
// For each uploaded deed PDF:
// Step 1 — OCR all pages (Tesseract.js, English + Telugu)
const fullText = await ocrAllPages(pdfFile);

// Step 2 — Extract Udyam Registration Number
const udyamPattern = /UDYAM-[A-Z]{2}-\d{2}-\d{7}/g;
const udyamNumbers = fullText.match(udyamPattern) || [];

// Step 3 — Determine MSME category from Udyam certificate text
// (Udyam certificate is typically the first page of the uploaded file set)
const microSmallPattern = /\b(micro|small)\b/i;
const isMicroOrSmall = microSmallPattern.test(udyamCertificateText);

// Step 4 — Verdict
if (isDeposit && !isMicroOrSmall) {
  verdict = 'FLAG — Concession not applicable; recover differential duty';
} else if (isDeposit && isMicroOrSmall) {
  verdict = 'Concession appears valid — verify year of Udyam cert';
} else {
  verdict = 'Hypothecation of movables — stamp duty check not applicable here';
}

// Export: Doc No. | Udyam No. | MSME Category | Verdict
```

- The tool verifies whether the instrument is a Memorandum of Deposit of Title Deeds (MoDTD) and whether an Udyam Registration Certificate has been uploaded as part of the registration record.
 - Udyam Registration Numbers identified during OCR processing are extracted and exported to an Excel file for further verification.
 - The exported data can be processed through a separate utility to validate the Udyam Registration details and obtain the corresponding Udyam Certificates.
 - The key audit check is whether the enterprise was classified as **Micro** or **Small** under the Udyam Registration framework on the date of registration. Enterprises classified as **Medium** are not eligible for the concessional stamp duty rate.
5. Compute the differential stamp duty payable in cases where the concession was incorrectly availed or the instrument was otherwise under-stamped, and aggregate the revenue impact SRO-wise.

Tools Used — MSME Check

MSME Title Deed Checker: Offline HTML tool using Tesseract.js OCR. Batch processes deed PDFs. Extracts Udyam Registration Numbers. Checks MSME enterprise category (Micro/Small vs. Medium).

Udyam Certificate Downloader (companion tool): Processes the exported Udyam numbers and downloads the corresponding Udyam certificates from the MSME Ministry portal for verification of enterprise category and registration year.

Both tools run entirely in the browser on the auditor's local machine.

5. STAGE II — OBSERVATION, RESPONSE AND REPORTING

5.1 ISSUE OF AUDIT OBSERVATIONS

1. After document scrutiny, audit observations shall be issued through the CAG Connect mode in OIOS. Responses from the Department shall be obtained through the same platform.
2. Entry and Exit Conferences shall be conducted in virtual mode.
3. Each audit observation must include:
 - The legal provision under which the short levy or irregularity has arisen;
 - Whether the irregularity is systemic or isolated;
 - Number of cases and total short levy or revenue impact;
 - Sample case details with document references, computed shortfall, and evidence from the deed;
 - Recommended action — recovery, software correction, or issue of circular instructions.

5.2 RESPONSE OF THE DEPARTMENT

Since DLRA generates observations covering many SROs simultaneously, the overall departmental response addressing the finding category as a whole shall be obtained and considered before finalising the report, rather than processing each case individually.

5.3 INSPECTION REPORT

The final Inspection Report shall:

1. Identify systemic flaws and not merely list individual errors. Where a finding reflects a software or process gap causing the same error across all SROs, it shall be presented as a systemic issue requiring a system-level fix.
2. Aggregate total short levy at the State level by deed type and SRO, for communication to the Heads of the Department.
3. Include significant isolated cases where the money value is high, even if the pattern is not widespread.

6. OFFLINE AUDIT TOOLS DEVELOPED BY PAG(AUDIT) TELANGANA

The following tools have been developed to support the DLRA of Registration and Stamps Departments. All are browser-based, offline HTML applications — they require no software installation and transmit no data to any external system. They run entirely on the auditor's local machine and can be adopted by any State AG office.

Tool	Function	SOP Section
Document Downloader Tool	Automated batch download of certified deed copies from the State registration portal. Handles portal form-filling, retry on errors.	Sections 4.1, 4.2, 4.5,
Gift Deed Audit Tool	Offline OCR tool (Tesseract.js — English + regional language) that extracts party names, father names, and stated relationships from gift deed PDFs and images. Validates relationships under Section 56(2) of Income Tax Act with verdict and action column. Batch processes multiple documents.	Section 4.1
PDF and Image Word Finder	Batch keyword search across uploaded deed PDFs and scanned image files. Searches for TDR-related and other configurable keywords. Applies OCR where the PDF has no embedded text layer. Returns file-wise results with matching page numbers and surrounding text context.	Section 4.2
MSME Title Deed Checker	OCR-based tool that reads MoDTD deed PDFs and Udyam certificates together. Extracts Udyam Registration Numbers, checks the MSME enterprise category (Micro/Small vs. Medium)	Section 4.5
Udyam Certificate Downloader	Companion tool to the MSME Checker. Processes the exported list of Udyam numbers and downloads the	Section 4.5

Tool	Function	SOP Section
	corresponding Udyam certificates from the MSME Ministry portal for verification.	

Security and Data Handling

All tools operate entirely offline on the auditor's local machine. No deed content, party names, property details, or document numbers are transmitted to any external system or cloud service.

The Tesseract.js OCR engine embedded in the tools processes documents locally within the browser. Any AI assistance used in developing these tools was limited to generating the tool source code — no actual audit data was shared with any AI system at any stage.

Regional language support: The tools currently include English and Telugu OCR language packs. Offices in other States should load the appropriate Tesseract language pack for their State's regional language before use.

7. WAY FORWARD

7.1 EXPANDING AUDIT COVERAGE

- As the full transaction dataset and document access are progressively obtained from the Department, all risk indicator checks listed in Section 5 shall be applied across the complete data, covering all SROs and deed types.
- The detailed methodology sections of this SOP shall be expanded to cover all remaining checks as they are operationalised. Remaining checks are listed as 'To be detailed in the next version' in Annexure A.

7.2 ADOPTION BY OTHER STATE OFFICES

- Other State AG offices adopting this SOP should begin by mapping their State's registration system field names to the mandatory data fields in Section 2.2.
- The Document Downloader tool should be configured to match the specific portal structure and form layout of the State's registration portal before use.
- Regional language OCR packs should be loaded into the Gift Deed Audit Tool and MSME Checker for the State's primary script. Contact the lead office for assistance with tool configuration.
- State-specific legal references, the applicable stamp duty rates schedule, market value notification mechanism, and Departmental instructions should be followed.

Scaling the Audit Tools

- Additional OCR-based tools will be developed for check types such as undervaluation detection, boundary omission flagging, and mortgage deed scrutiny, as audit experience grows.
- All tools and check scripts shall be documented and maintained as a reusable repository for future audit cycles, enabling faster deployment in subsequent years.

ANNEXURE A — RISK INDICATOR TO AUDIT CHECK CROSS-REFERENCE

Sl. No.	Risk Indicator	SOP Section	Tool / Method	Status
1	Stamp duty on lower of MV / Consideration	Section 4.3	Excel computed column	Documented
2	Previous registered value higher than current transaction value	Section 4.4	IDEA — Link Document table join	Documented
3	Gift deed — wrong family concession	Section 4.1	Gift Deed Audit Tool (OCR)	Documented
4	TDR without stamp duty	Section 4.2	PDF and Image Word Finder	Documented
5	Zero / concessional duty — wrong exemption	To be added	Filter by stamp duty = 0 + deed type check	Next version
6	MSME equitable mortgage — under-stamped	Section 4.5	MSME Title Deed Checker (OCR)	Documented
7	Misclassification of deed type	To be added	Keyword filter on Nature of Deed	Next version

ANNEXURE

The following comparison statement captures the common themes and the distinguishing features.

Table 1: Similarities in Inputs Received

Sl. No.	Common Observation	States indicating similar views
1	Registration Department has largely shifted to computerized registration systems (NGDRS/State Portals/Open PEARL/SAMPDA/IGRS etc.) though the extent of access varies.	Andhra Pradesh, Kerala, Madhya Pradesh, Assam, Punjab
2	Audit has requested transaction data from the Registration Department for undertaking data-led audit.	Andhra Pradesh, Kerala, Madhya Pradesh, Assam, Punjab, Manipur
3	Read-only/system access remains the most important requirement for successful remote audit.	Kerala, Madhya Pradesh, Assam, Andhra Pradesh
4	Availability of scanned registered documents is considered essential for verification of document recitals.	Andhra Pradesh, Kerala, Madhya Pradesh
5	Most offices consider the Draft SOP comprehensive and suitable for adoption with only minor improvements.	Madhya Pradesh, Punjab, Kerala
6	Hybrid/Data-led audit is viewed as the future audit methodology replacing conventional local audit.	Andhra Pradesh, Kerala, Madhya Pradesh, Assam
7	Delay in receiving complete datasets and departmental cooperation are common constraints.	Kerala, Madhya Pradesh, Assam, Andhra Pradesh
8	Digital infrastructure and data analytics are expected to improve audit efficiency and coverage.	Andhra Pradesh, Kerala, Madhya Pradesh, Punjab

Table 2: Differences in Inputs Received

State	Distinct Features / Suggestions
Andhra Pradesh	Most advanced implementation. Successfully conducted Remote Audit of 18 SROs/DROs. Login access to IGRS available. AI-based OCR Proof of Concept developed. Suggested risk-weighted sampling methodology, clearer risk scoring framework and caution against relying solely on outlier analysis.
Kerala	Complete Open PEARL transaction database and Fair Value database available. Hybrid pilot audit completed. Awaiting Government approval for read-only login. Suggested that seamless system access is more critical than mere availability of data.
Madhya Pradesh	Hybrid audit operational since 2018-19 through SAMPDA Portal with read-only access. Suggested inclusion of timelines for departmental replies and improvement of server stability.
Assam	VPN access provided but table-level data not accessible due to technical issues. Data-led audit planned for 2026-27. Highlighted coordination with NIC and data security concerns.
Punjab	Transaction data available (2021-26) through NGDRS but no read-only access. No prior experience of Data-Led Audit. Proposed adopting Telangana SOP without modification.
Manipur	No data obtained yet. Conventional physical audit continues. Computerisation still evolving. No access to registration systems or digital records.
Meghalaya	Registration records are completely manual and not digitised. Data-led audit presently not feasible.
Goa	No Data-Led Audit planned during 2026-27 and therefore no comments on SOP.
Chhattisgarh	Forwarded detailed comments separately, indicating active consideration of the Draft SOP.

Overall Analytical Observations

1. Wide variation in digital maturity

The responses reveal substantial differences among States in terms of digitisation, system access and operational readiness. While Andhra Pradesh, Kerala and Madhya Pradesh have already moved towards hybrid or remote audit, several States are still dependent on conventional local audit due to limited digitisation or lack of system access.

2. System access is the critical enabler

Almost all States recognise that obtaining read-only access to the Registration Department's application is more valuable than receiving periodic data dumps. Direct access enables retrieval of documents, reports and supporting records without repeated requisitions.

3. Data availability alone is insufficient

Many offices have received transaction data but continue to face challenges because scanned documents, market value databases, encumbrance records and other supporting information are either unavailable or only partially accessible.

4. Need for standardised infrastructure

Responses consistently indicate the need for secure digital workspaces, reliable server connectivity, OCR capability, AI-assisted document analysis and structured data extraction tools to effectively implement Data-Led Audit.

5. Hybrid audit is emerging as the preferred model

States with practical experience favour a hybrid model wherein data analytics and preliminary scrutiny are undertaken remotely, while selective physical verification is conducted wherever required.

6. Importance of risk-based sampling

Andhra Pradesh has specifically highlighted that audit observations often arise from issues evident only upon scrutiny of document recitals. Accordingly, the SOP should balance data-driven risk indicators with targeted document-based verification, ensuring that significant issues not reflected in structured data are not overlooked.

7. Need for uniform national audit practices

The responses indicate that there is presently no uniform approach across IA&AD for auditing Registration Departments. A standard SOP would help establish consistent methodologies for data acquisition, analytics, document verification, risk assessment and reporting.

Conclusion

The feedback received from the States broadly endorses the Draft SOP and reflects a positive acceptance of Data-Led Audit as the future audit methodology for the Registration & Stamps Department. While the level of preparedness differs considerably across States, there is a common consensus that the success of Data-Led Audit depends primarily on seamless access to departmental systems, availability of complete digital records, adequate technological infrastructure, and adoption of a structured risk-based hybrid audit approach.