PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science

CSE DATA SCIENCE

Semester : VIII   Subject : AIFB   Academic Year: 2024-2025

Selecting the Sample to Investigate:
Selecting the right sample is a critical step when performing audits, compliance checks, or fraud investigations in banking and financial services. The goal is to efficiently identify high-risk records while ensuring coverage and relevance.

1. Purpose of Sampling
- Identify irregularities or fraudulent transactions
- Verify compliance with regulatory or internal policies
- Reduce workload by examining a representative subset
- Evaluate control effectiveness or risk exposure

2. Sampling Methods

◆ a. Random Sampling
- Equal chance for each transaction to be selected
- Used for general audits or when no prior risk indicators exist

Use case: Reviewing 1,000 transactions from a pool of 100,000 to test control effectiveness.

◆ b. Stratified Sampling
- Divides the dataset into strata (e.g., by transaction amount, region)
- Sample is taken from each stratum proportionally or based on risk

Use case: More samples from high-value transaction strata.

◆ c. Judgmental (Risk-Based) Sampling
- Samples selected based on expertise, suspicion, or risk indicators
- Focuses on unusual patterns or known fraud triggers

Use case: Investigating transactions above ₹10 lakh done at odd hours from new devices.

◆ d. Systematic Sampling
- Every *n-th* item is selected after a random start
- Useful for evenly distributed data sets

Use case: Every 100th transaction in a ledger is audited.

◆ e. Monetary Unit Sampling (MUS)
- Larger-value transactions have a higher chance of being selected
- Great for detecting material misstatements or fraud

Use case: Targeting high-impact errors in loan disbursements or large wire transfers.

3. Criteria to Select the Sample

Criteria	Example
Amount threshold	Transactions > ₹1,00,000
Time of transaction	Odd hours or during non-banking hours
Geographic anomaly	Foreign IP or new login location
Device change	New device ID or browser fingerprint
Customer risk score	Customers with high KYC risk or flagged accounts

Semester : VIII   Subject : AIFB   Academic Year: 2024-2025

Criteria	Example
Frequency anomaly	Unusual transaction volume in short periods
Known red flags	Accounts on a fraud watchlist

4. Example Scenario

Investigation Goal: Check for payroll fraud in a corporate banking system.
Sampling Method Used: Judgmental sampling + stratified sampling.
Steps:
1. Filter for all payroll transactions above ₹5,00,000.
2. Stratify by department (HR, Finance, IT).
3. Select 5 highest-value transactions per department.
4. Cross-verify bank account holders with employee master list.

Outcome: Detected 2 non-employee accounts receiving payroll credits.

📊 Summary Table

Sampling Method	Best For	Risk Level Addressed
Random	General compliance checks	Low/Unknown
Stratified	Complex datasets (e.g., regions)	Medium
Judgmental/Risk-Based	Suspicious or red-flag scenarios	High
Systematic	Routine periodic reviews	Low
Monetary Unit	Financial materiality checks	High