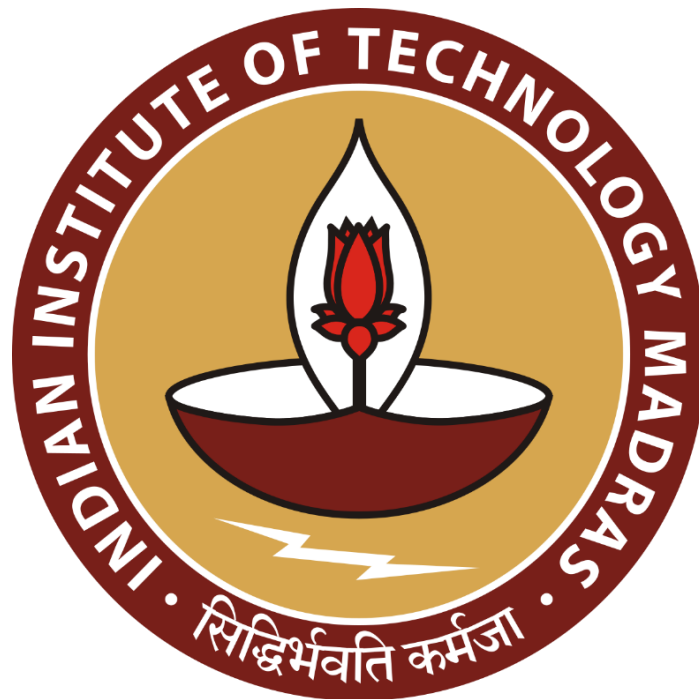# Business Analytics Solution for Investigating Third-Party Motor Insurance Claims: A Case Study of a Private Investigation Firm

Final report for the BDM capstone Project

Submitted by

Name : MUNEESHWARI N
Roll number: 24ds3000101

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 600036

Table of Contents

Title	Page No

Declaration Statement

I am working on a Project Title " Business Analytics Solution for Investigating Third-Party Motor Insurance Claims: A Case Study of a Private Investigation Firm". I extend my appreciation to Right View Investigators Company, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analyzed to assure its reliability.
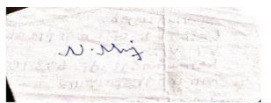
Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively, and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate

Name:MUNEESHWARI N

Date:06.07.2025

1.Executive Summary

This final report presents a business analytics solution designed to uncover operational insights from third-party motor insurance claim investigations carried out by **Right View Investigators**, a private investigation firm based in Madurai. The objective of this study was to analyze structured case records and **Right to Information** (RTI) filings to assess turnaround time (TAT), claim fee variations, and investigation status trends.

In the midterm submission, metadata and descriptive statistics were presented, focusing on case timelines, location tags, vehicle types, claim outcomes, and submission status. For the final report, deeper analytics were conducted, addressing three major objectives: optimizing the lead investigation-to-submission pipeline, identifying red flags for fraudulent or weak cases, and designing a dashboard-ready structure to monitor KPIs such as turnaround time, case closure rate, and geographical claim distribution.

Rather than providing a technical solution, the project focuses on generating actionable insights that support better decision-making, resource allocation, and client servicing. The cleaned dataset comprised **1079** records with variables such as claim status, assigned dates, investigation types, and fee amounts.

Using tools like Google Sheets, Excel, and Python (Pandas), descriptive statistics and visual analyses were performed to identify bottlenecks and operational inefficiencies. Key findings reveal that while **62%** of cases are successfully submitted, a significant portion remains pending, delaying revenue realization. TAT values ranged widely, indicating inconsistencies in investigation timelines.

The most frequent claim fees hovered around **₹1800–₹3600**. Based on these findings, strategic recommendations are proposed to streamline pending case follow-up, prioritize high-volume clients, and optimize fee collection processes. This data-driven approach empowers the organization to enhance efficiency, improve financial oversight, and strengthen stakeholder confidence.

2. Detailed Explanation of Analysis Process

The analysis process for this project was structured into clear, iterative stages to ensure data consistency, extract meaningful insights, and support business-oriented recommendations. Below is a comprehensive walkthrough of each step, covering cleaning, transformation, visualization, and interpretation:

2.1 Data Collection and Preparation

Dataset Title: *Combine RTI and Case Status Trial.xlsx*

Overview:
This dataset tracks insurance claim investigations and RTI (Right to Information) filings submitted to agencies like Police, RTO, and Hospitals. It captures both the operational process and the current case status of each investigation.

Dataset Summary:

- **Total Records:** 1079
- **Main Entities:** Insurance companies, claimants, investigators
- **Key Process:** From assignment → investigation → report submission → fee processing

Key Variables:

VARIABLE	TYPE	CATEGORY	DESCRIPTION	EXAMPLE VALUE	TAGS
Claim Number	CATEGORICAL	Identifier	Unique ID assigned to each claim	TP123456	Claim ID, Key Identifier
Company	CATEGORICAL	Client Tag	Insurance company that assigned the investigation	CHOLA	Insurer, Client Name
Assigned In	TEMPORAL	Timeline	Date when the firm received investigation assignment	2024-02-15	Assignment Date, Start
Submitted Date	TEMPORAL	Timeline	Date report was submitted to the insurer (final stage)	2024-03-12	Report Submission, End
Dispatch Date	TEMPORAL	Timeline	Date when the physical report was dispatched	2024-03-15	Dispatch, Logistics
Received Date	TEMPORAL	Financial	Date fee payment was recorded from insurer	2024-04-05	Fee Receipt, Payment Date
Status	CATEGORICAL	Case Status	Current stage of claim (Submitted = Completed)	Submitted	Progress, Closure Status
Fee Status	CATEGORICAL	Financial	Indicates whether fee has been received or is pending	Received / Pending	Payment Status, Revenue
TAT (Days)	NUMERIC	Performance Metric	Turnaround time in days (Submitted Date - Assigned In)	25	Duration, Efficiency

Figure: 1

Operational Fields:

- **Claim Number:** Unique ID for each claim
- **Company:** Insurance provider
- **Assigned In/Date:** Assignment channel and date
- **Submitted/Dispatch/Received Date:** Tracking report lifecycle
- **Status:** Investigation stage (e.g., Assigned, Pending, Submitted)
- **TAT:** Turnaround time (in days)

Investigation Fields:

- **Type:** Nature of investigation
- **Police Station / RTI / FIR:** Law enforcement interaction
- **Claimant / Insured / Driver Location:** Parties' addresses
- **Reporter Name:** Person compiling the final report

Administrative Fields:

- **Recipient Mail / Mail Links / Drive Links**
- **Remarks / Mail Sent Status:** Communication and notes

Workflow Overview:

Cases follow a path: **Assigned → Pending → Completed → Submitted → Dispatched**
(With a possible early exit at **Withdrawn**)

Figure: 2
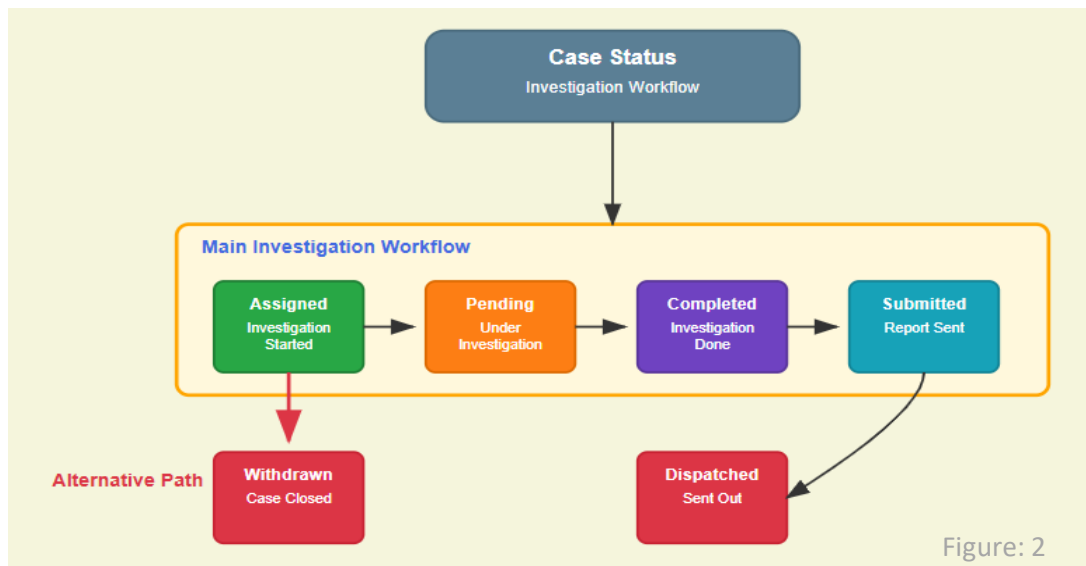
2.2 Data Cleaning Steps

1. **Standardising Column Names**
   Column headers were cleaned to maintain consistent naming conventions (e.g., *Assigned In* renamed to *Assigned IN*, *CLAIMANT* renamed to *CLAIMANT Location*).

2. **Duplicate Removal**
   Duplicate rows were identified and removed to ensure each claim record is unique, resulting in a reduced row count from 1082 to 1079.

3. **Imputation of Missing Values**
   Missing dates in *Submitted Date* were imputed using available related date columns (*Dispatch Date*) to create new columns:
   - *Submitted Date Imputed*
   - *Submitted Date Final*

4. **Date Format Standardisation**
   All date columns were formatted uniformly to ensure accurate date difference calculations for TAT.

2.3 Toolset and Technology Used

Google Sheets: For data cleaning, formula-based imputation, descriptive statistics calculation, and visualisations.

Python (Pandas): For dataset comparison and deeper validation of cleaning steps and column structures (during this documentation stage).

Excel: For local exploratory checks and pivot-based summaries.

2.4 Analytical Techniques Applied

1. **Descriptive Statistics**

   Using Google Sheets built-in functions, summary statistics (mean, median, mode, minimum, maximum, and standard deviation) were calculated for:

   - Turn Around Time (TAT)
   - Claim Amount

2. **DataVisualisation**

Created meaningful, labeled charts:

- ○ Histograms for TAT and Amount
- ○ Bar charts for Top 5 Companies by case volume
- ○ Pie chart for Case Status distribution

3. **Data Interpretation**

- ○ Each visual output was interpreted in terms of operational significance (e.g., client priority, workflow delays).
- ○ Focused on identifying skewness, outliers, and typical patterns to aid in decision-making.

4. **Comparative Validation**

- ○ Two versions of the Excel file were compared using Python (combine trial.xlsx vs midterm version) to ensure structural consistency and completeness.

2.5 Justification of Methodology

- **Spreadsheet-first approach** was selected due to the firm's current reliance on Excel formats.
- **Python scripts** provided a scalable and repeatable backup validation layer, aligning with capstone expectations.For checking if the charts created using excel is valid.
- **Charts and summaries** were chosen for clarity in stakeholder communication, especially non-technical management

3. Results and Findings

1.Descriptive Statistics:

Statistical Summary

Statistic	TAT	Amount
Mean	61.29128015	3191.919598
Median	44.5	3600
Mode	22	1800
Minimum	2	450
Maximum	589	6250
Standard Deviation	49.90496575	1328.738812

Figure: 3

◈ **Turnaround Time (TAT)**

The average Turnaround Time (TAT) is 61.3 days, with a median of 44.5 days and a mode of 22 days. This indicates that while most claims are commonly resolved in 22 days, 50% are closed within 44.5 days. The higher mean compared to the median reflects a **right-skewed** distribution, driven by a subset of significantly delayed cases extending up to 589 days. With a minimum of just 2 days and a standard deviation of 49.9 days, the data reveals moderate variability. These delayed outliers present a key area for operational improvement to reduce overall claim processing times and enhance service efficiency.

◈ **ClaimAmount**

The average investigation fee stands at ₹3192, with a median of ₹3600 and a mode of ₹1800. The median exceeding the mean suggests a **left-skewed** distribution, where most fees cluster near ₹3600, but a number of lower-value claims drag the average down. Claim amounts range from ₹450 to ₹6250, indicating substantial variance depending on claim complexity. A standard deviation of ₹1328.7 further supports moderate dispersion in fees. These findings are critical for refining billing strategies, forecasting revenue, and ensuring fair compensation for investigative effort.

2.Distribution of TAT for Claims:

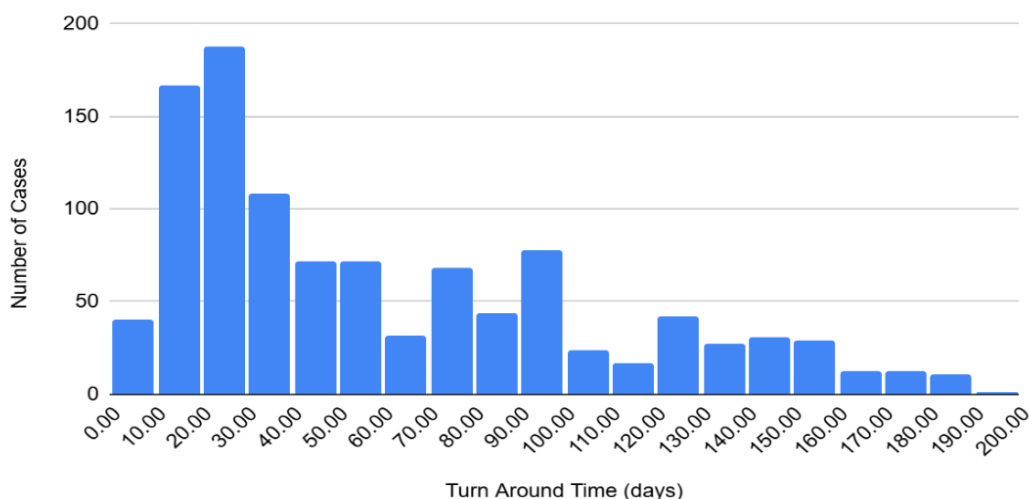**Distribution of TAT for Claims**

Figure: 4

Interpretation:

The histogram shows that most claims have a Turn Around Time (TAT) between 10 to 50 days, with frequency peaking around 20 days. A smaller number of cases extend beyond 100 days, with the longest TAT recorded as 196 days. This indicates that the process is generally efficient, but focusing on delayed claims can help further reduce average turnaround time and improve service level commitments.

3.Distribution of Claim Amounts:
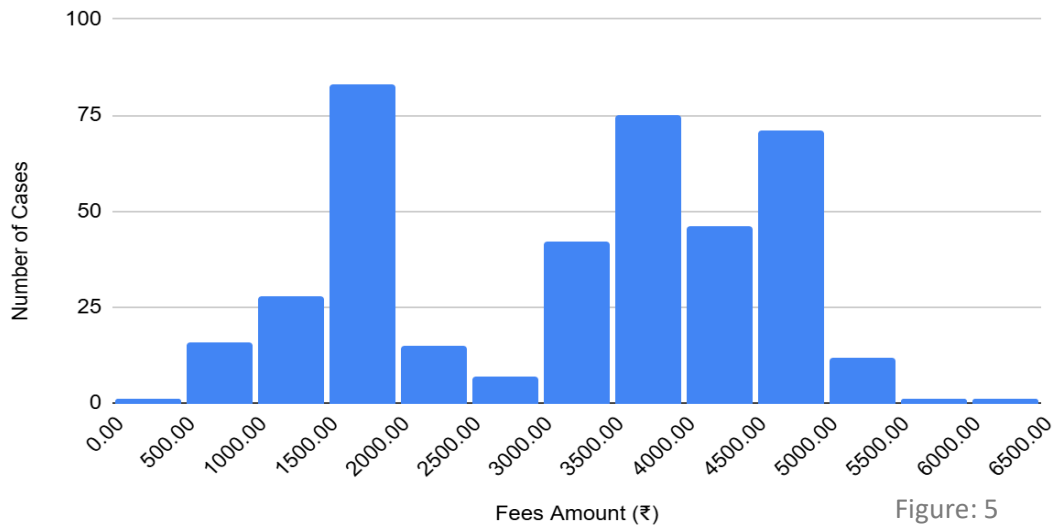
Distribution of Claim Amounts

Figure: 5

Interpretation:

The histogram indicates that most claim amounts range between ₹1500 and ₹5000, with a peak concentration in the ₹2000–₹2500 bracket. Very few claims exceed ₹6000, suggesting that the majority of payouts fall within the low to mid payout range, making overall risk exposure moderate. This helps in planning financial reserves for typical claim settlements.

4.Number of Cases per Company:

Number of Cases per Company
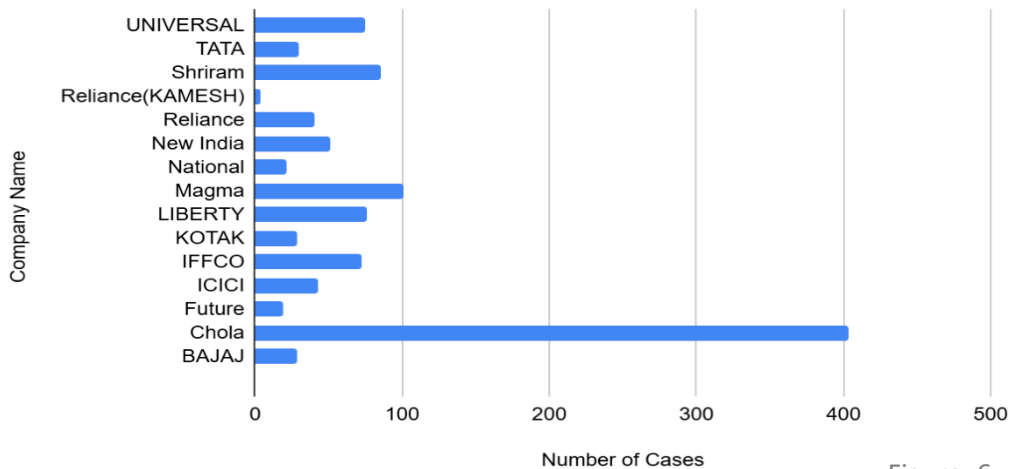
Figure: 6

Interpretation:

The bar chart shows that Chola and Magma are the top contributors to total case volume, followed by companies such as Shriram and Liberty. This implies that operational resources and process improvements can be prioritised for these high-volume clients, ensuring service quality and efficient turnaround for major stakeholders.

5.Case Status Distribution:

Case Status Distribution
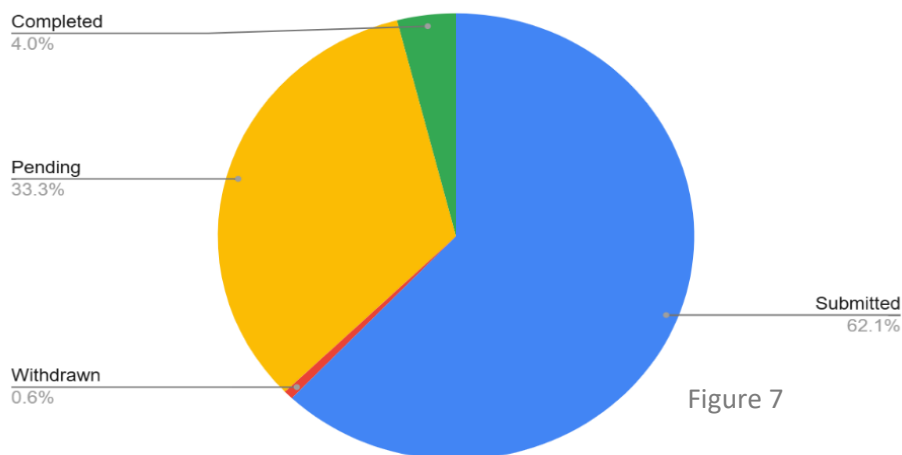
Figure 7

Interpretation:

The pie chart indicates that 62% of cases are Submitted, suggesting that the majority of investigations reach the final report submission stage. 33% remain Pending, highlighting a substantial proportion of ongoing cases, which can be targeted for process acceleration. 4% are marked Completed, and only 0.6% are Withdrawn, indicating a high completion and submission rate with minimal case dropouts. Overall, the chart demonstrates effective operational closure while pointing out areas for workflow optimisation to reduce pending backlogs.

6.Monthly Submission Trend
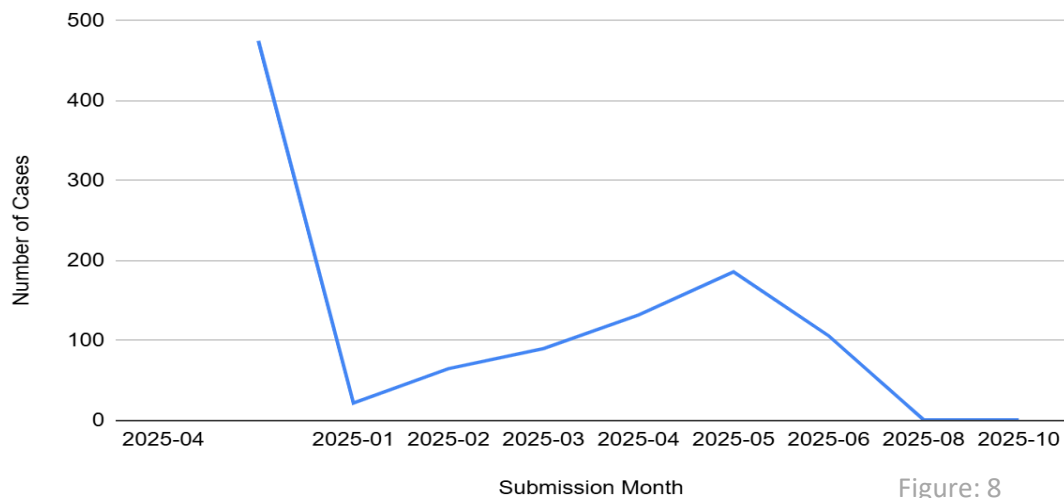
Monthly Submission Trend

Figure: 8

Interpretation:

The line chart shows fluctuating case submission volumes over the months in 2025. April had the highest submission count with **475 cases**, likely due to batch entries or a reporting backlog. May and June also saw significant activity, indicating improved operational throughput during those months. Submissions dipped in August and October, possibly due to pending case accumulation or staff availability.Overall, the trend highlights seasonality and workflow gaps that can be optimized for consistent case handling.

7. Investigation Type vs Claim Status

Investigation Type vs Claim Status
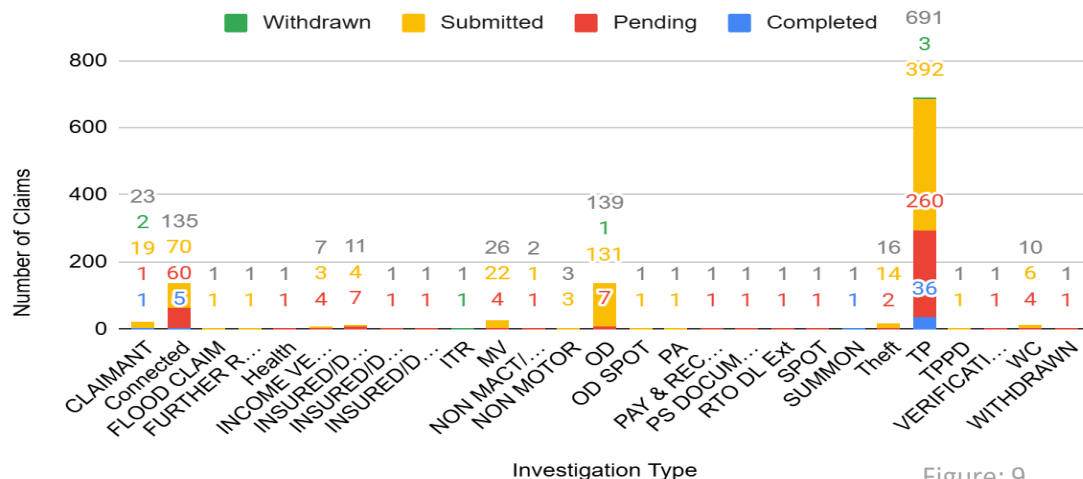
Figure: 9

The chart shows that **TP (Third Party)** investigations dominate with **691 total cases**, of which **392 are submitted**, **260 pending**, and only **36 completed**, indicating a backlog in closures.**OD (Own Damage)** follows with **139 cases**, largely submitted.Other types like **Connected** and **Claimant** have moderate volume but high pending counts.Types such as **Verification, Theft**, and **ITR** appear in smaller volumes.The overall distribution reveals operational pressure concentrated on a few high-volume case types, especially TP, suggesting a need for **focused follow-up**.

8.Turnaround Time vs Claim Amount

Relationship Between Turnaround Time and Fees Amount (₹)

Figure: 10
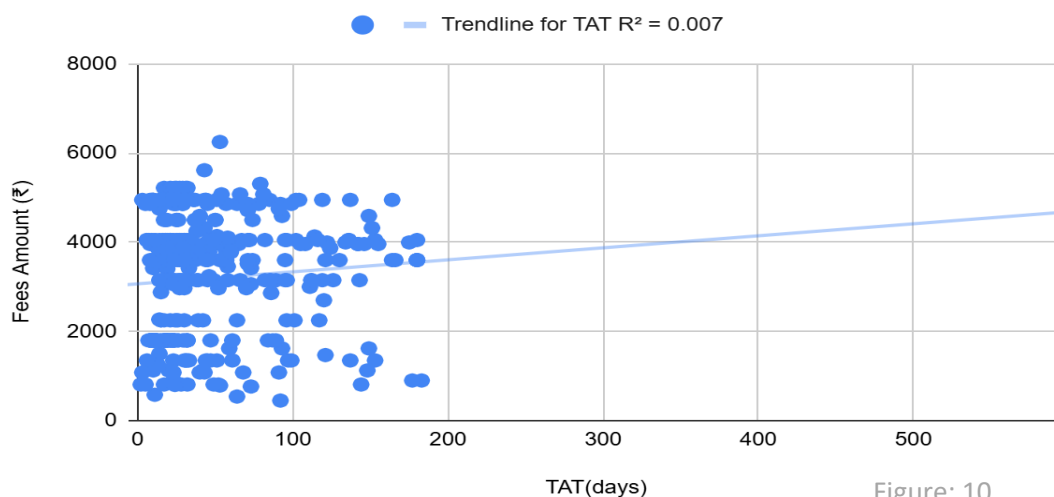
Interpretation:

- The scatter plot shows the relationship between Turnaround Time (TAT) and Fees Amount (₹) for claim cases.
- The data points are widely scattered, with no clear linear pattern, and most TAT values fall below 150 days.
- The trendline has a very low R^2 value of 0.007, indicating **negligible correlation** between TAT and fees.
- This suggests that longer or shorter claim durations do **not significantly influence** the investigation fee.
- Operationally, fee structures appear consistent regardless of delays, implying pricing is not time-dependent.
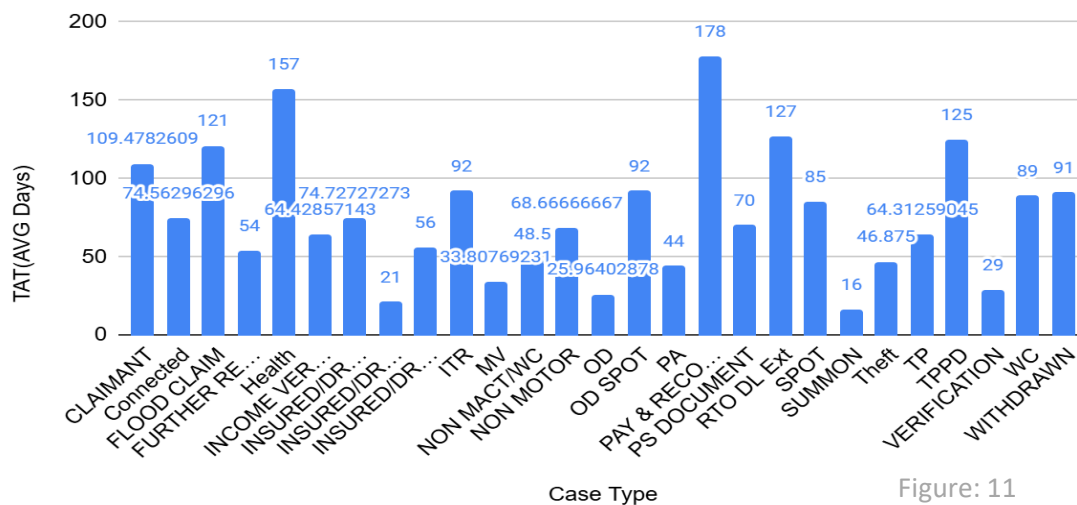
9. Distribution of TAT(AVG) for Case Type

Figure: 11

Interepretation:

The chart shows average turnaround time (TAT) for different investigation case types. **PAY & RECOVER** and **Health** cases have the highest TAT, indicating complex processing. **SUMMON**, **OD**, and **MV** have the lowest TAT, suggesting quicker resolution workflows. Duplicate entries like **INSURED/DRIV** should be cleaned for accurate insights. The chart helps identify delay areas and optimize case handling efficiency.

10. Monthly Claims Status Distribution Heatmap

Claims Heat Map: Count by Submission Month and Status

Status	Jan	Feb	Mar	Apr	May	Jun	Aug	Oct	Total
Completed	-	-	-	43	-	-	-	-	43
Pending	-	-	-	358	-	1	-	-	359
Submitted	65	1	105	67	90	22	133	186	669
Withdrawn	-	-	-	7	-	-	-	-	7
Grand Total	65	1	105	475	90	23	133	186	1078

Legend

🟩 Completed 🟨 Pending 🟦 Submitted 🟥 Withdrawn

Color intensity represents the count of claims (darker = higher count)

Figure: 1

- ◆ **April Surge**: April recorded the highest activity with **475 cases**, including **358 Pending**, **67 Submitted**, and **43 Completed**. This suggests a reporting push or backlog clearance.
- ☑ **High Submissions in October and August**: October (186) and August (133) had peak "Submitted" cases, reflecting possible SLA-driven completion or seasonal trends.
- 🗠 **Low Activity in Jan, Feb, and Jun**: Very few cases recorded in January (65), February (1), and June (23), indicating operational lulls or delayed data entry.
- 🔍 **Underuse of 'Completed' Status**: Only 43 cases are marked as "Completed", all in April. Suggests most cases may skip this stage and go straight from Pending to Submitted.
- ⬚ **Withdrawals Rare**: Just 7 Withdrawn cases were recorded — all in April — indicating that almost all investigations continue to completion.
- 📅 **Seasonality Insight**: The heat map exposes monthly processing patterns and workload clusters. This can help plan resource allocation and improve submission consistency.

4. Interpretation of Results and Recommendations

The descriptive analysis and visual exploration conducted in this project revealed several insights into operational patterns, client behavior, and workflow efficiency. Below is a breakdown of key interpretations and the corresponding strategic recommendations for Right View Investigators:

4.1 Key Observations

1. **Turnaround Time (TAT):**
   - The average TAT is 61.3 days, but the distribution is skewed with a few cases stretching over 100 days (max: 589 days).
   - Most claims are completed within 10–50 days, peaking around 22 days (mode).
   - Longer durations suggest process delays, missing data, or external dependency in specific types like TP and Pay & Recover.
2. **Claim Amounts:**
   - Majority of claims fall between ₹1800–₹3600 with outliers reaching ₹6250.
   - There is **no strong correlation** between TAT and amount, indicating fee is fixed per case type, not investigation time.
   - A few claim types (e.g., Health, Theft) may justify higher fee ranges but need validation from operational guidelines.
3. **Client Distribution:**
   - Companies like Chola and Magma account for the highest number of cases.

- These top 5 clients drive the firm's workload and thus need special dashboards, SLA tracking, and dedicated teams.

4. **Status Distribution:**
   - 62% of cases are marked as Submitted, showing good progress.
   - However, 33% remain Pending, signaling workflow congestion or missed follow-ups.
   - Very few are marked as Completed or Withdrawn, which may indicate underutilization of these closure categories.

5. **Investigation Type vs Status:**
   - TP claims dominate the dataset but also have the highest number of pending cases.
   - Verification and Summon-based cases tend to have quicker resolution.
   - Connected and ITR cases show irregular patterns and may need deeper operational review.

6. **Monthly Submissions:**
   - Submission trend is seasonal. High peaks in April–June and dips in August–October suggest batch processing or investigator workload variation.

4.2 Recommendations

Based on the analytical insights gathered, the following actionable recommendations are proposed to streamline operations and enhance decision-making at Right View Investigators:

Operational Efficiency Improvements

- 📌 **Prioritize Backlogged Cases:**
  Implement automated reminders or dashboards to target long-pending investigations, particularly in **Third-Party (TP)** and **Health** categories, to improve closure rates.

- 🕐 **Standardize Investigation Timelines:**
  Define service-level expectations per case type (e.g., TP, OD, Connected) based on historical TAT. Use alerts for cases exceeding the threshold to prompt early escalation.

- 🖼 **Monitor Monthly Trends:**
  Use seasonal submission patterns identified in the 2025 timeline to proactively **plan investigator availability** and avoid resource crunch during peak periods.

Client-Centric Strategy

- 💼 **Client-Based Workload Allocation:**
  Assign **dedicated investigation teams** or leads to high-volume clients such as Chola and Magma. This ensures faster turnaround, tailored servicing, and stronger client relationships.
- 📊 **Client-Specific Dashboards:**
  Build interactive dashboards to track key metrics (submission rates, fee received, average TAT) for each top client, helping management conduct monthly reviews.

Financial Optimization

- 🕸️ **Standardize Fee Policies:**
  Revisit billing structures to ensure that complex and high-effort cases like **Pay & Recover** or **Health** are adequately billed, preventing undercharging and aligning compensation with effort.

Data Integrity & Reporting Accuracy

- ⬜ **Duplicate Data Cleaning:**
  Fix naming inconsistencies such as "INSURED/DRIV" duplicates. Introduce **data validation steps** during entry to ensure standard field usage and avoid misclassification.
- 🗓️ **Improve Submission Date Tracking:**
  Mandate consistent recording of the **Submitted Date** for every case to ensure accurate TAT analysis and timely billing follow-ups.

Limitations and Future Scope

This project focused exclusively on structured data from 1,079 third-party motor claim investigations handled by Right View Investigators. While it provided valuable insights into turnaround times, client behavior, and workflow inefficiencies, certain aspects remained beyond the current scope.

For instance, qualitative data such as investigator remarks, client communications, or RTI response narratives were not analyzed. Additionally, real-time updates or live tracking dashboards were not implemented due to resource constraints.

In future extensions of such analytical efforts, integrating Natural Language Processing (NLP) to process unstructured comments, and developing automated case prioritization tools could help improve decision-making. Expanding the timeline to include multi-year data would also allow better seasonality analysis and forecasting.

✅ Conclusion

The analysis confirms that Right View Investigators has strong operational potential with high case submission rates. However, pending backlogs, inconsistent timelines, and fixed fee structures highlight the need for improved process control and client engagement.

By implementing these recommendations, the firm can enhance service delivery, reduce revenue leakage, and position itself as a more data-driven investigation partner for leading insurers across India.