Insurance Claims Analysis Report

By Rabab Sehalet

Table of Contents

- Executive Summary

- Introduction

- Data Overview

- Geographical Analysis of Claims

- Claims Fraud Analysis

- Client Segmentation Analysis

- Dashboard Insights

- Conclusion

- Appendices

1. Executive Summary

This analysis examines the company's financial performance, focusing on claims fraud, client segmentation, and geographical trends. Despite generating over $10 million in revenue, high claim payouts exceeding $8 million are significantly impacting profitability, leaving net gains lower than expected.

The first dashboard presents an overview of the company's financials and the geographical distribution of claims, highlighting the imbalance between gains and payouts, as well as regions with high claim frequencies and fraud risks. This reveals areas where resources may need better allocation.

Key patterns in the claims portfolio show the financial strain from fraud and the increasing severity of certain claims. To improve profitability, strategic adjustments in claims management, fraud detection, and client targeting are essential. This report highlights:

- The high cost of claims relative to profits
- Potential fraud hotspots and areas for improved claims management
- Client segmentation to identify high-risk clients and optimize resources
- Geographical analysis to pinpoint regions with frequent claims and fraud risks

The analysis aims to support decisions that could reduce claim-related expenses and enhance operational efficiency.

2. Introduction

This analysis focuses on the business performance of a company managing vehicle, property, and injury claims across three U.S. states: Indiana, Ohio, and Illinois. The company has been operating in these states since 1990, and the goal of this analysis is to assess its financial health, with an emphasis on understanding the balance between gains and losses due to claims.

The company has been successful in generating a significant number of leads and subscriptions across all three states. However, despite the strong influx of new business, several factors are negatively impacting profitability. Specifically, fraud claims, which have been identified as a growing concern, are eroding the gains generated by the company's operations.

By studying each state individually, this analysis delves into:

- The financial performance of each state from 1990 to 2015, with a focus on gains from new leads, subscriptions, and claims processing.

- The challenges posed by fraudulent claims, which have a substantial effect on the company's bottom line.

- A detailed comparison of vehicle, property, and injury claims, identifying trends, losses, and opportunities for improvement.

The following sections will break down these insights further, offering recommendations for managing risks, improving claims processing efficiency, and enhancing profitability in the future.

3. Methodology

Data Sourcing and Cleaning

The dataset was sourced from **Kaggle**, covering **customer information, policy details, and claims data** for Indiana, Ohio, and Illinois. Data was cleaned using **Python (Pandas, NumPy)** with the following steps:

- **Missing Values**: Columns with >50% missing values were removed; remaining gaps were filled using median (numerical) or mode (categorical).

- **Data Type Assignment**: Dates were formatted correctly, and categorical/numerical data types were assigned.

- **Exploratory Data Analysis (EDA)**: **Matplotlib** was used to visualize distributions, detect outliers, and analyze trends.

Machine Learning Models

1. Fraud Prediction Model

- Built using **historical claim data**, policy details, and customer demographics.

- Evaluated using **accuracy, precision, recall, and F1-score** to assess fraud risk.

2. High Claim Amount Prediction Model

- Predicts **high-risk claims** based on severity, vehicle involvement, and policy type.

- Aims to improve claim assessment and reduce financial risk.

Data Visualization & Insights

- **Tableau** dashboards were created to illustrate **geographical trends, fraud analysis, and claim segmentation** insights.
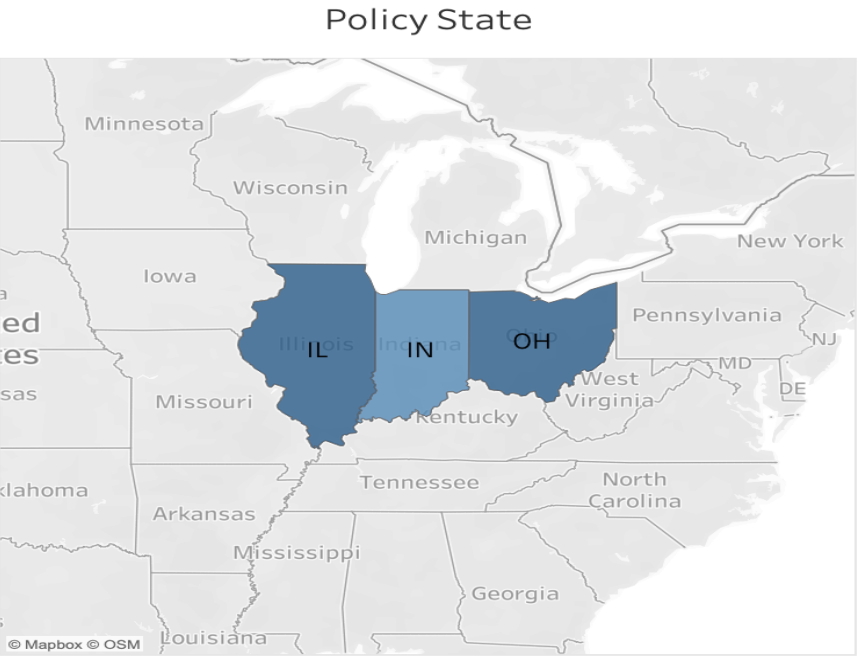
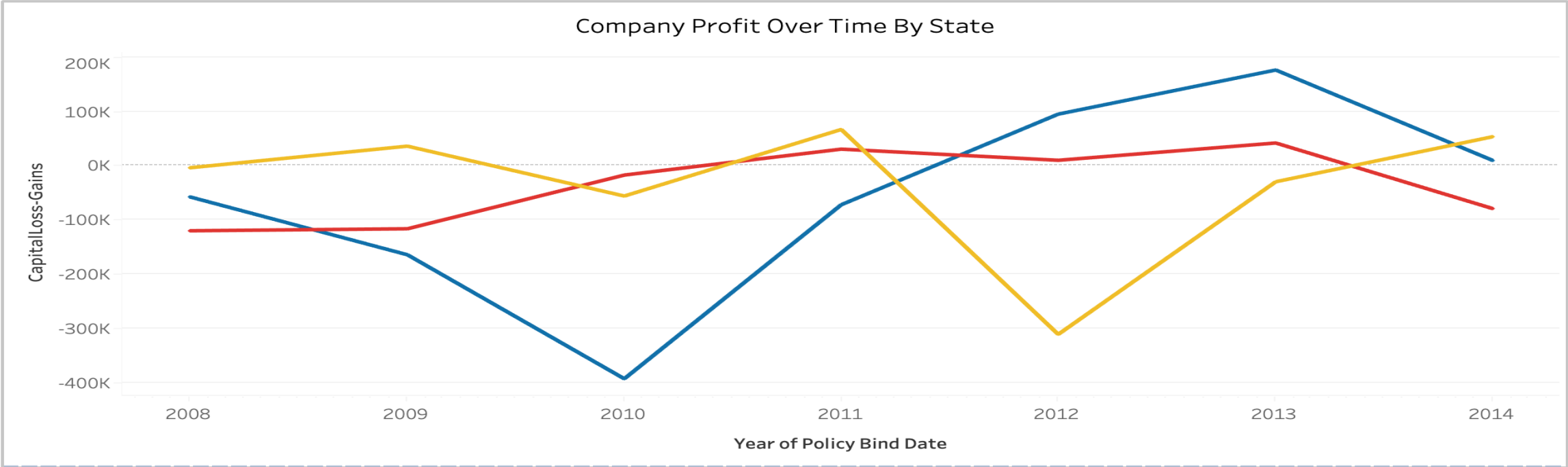# 4. Regional Claims Performance & Financial Impact

In the Financial Performance Overview dashboard, I presented a time series chart tracking the company's profit and loss trends since 1990. The data shows significant fluctuations, with losses dominating after 2008. Before then, the company saw annual gains of about $500k, with losses around $400k, leading to a balanced profit/loss situation. Post-2008, gains peaked at $300k, but profits steadily declined due to rising claim payouts.

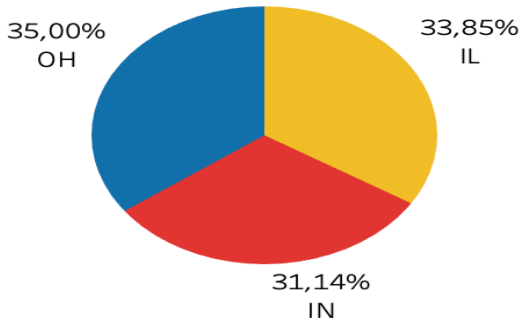Analyzing the three states Ohio, Illinois, and Indiana reveals these patterns:

- **Ohio:** Ohio has incurred the highest losses, with claim payouts exceeding $18 million. Prior to 2011, the state consistently caused major losses, but from 2011 onward, it generated profits, contributing a positive $214k since 1990.

- **Illinois:** Illinois had minimal impact initially but suffered significant losses of $340k in 2012 and further losses in 2013. Overall, Illinois has caused over $1.1 million in losses, though there were profitable years after these setbacks.

- **Indiana:** Indiana's claim payouts totaled over $16 million, contributing a loss of $2-3 million. Similar to Ohio, Indiana began contributing positively to profits after 2011.

From 1990 to 2015, Ohio has kept the profit chart slightly positive with a $214k contribution. In contrast, Illinois and Indiana have caused substantial losses, with Illinois totaling over $1.1 million and Indiana $2-3 million.
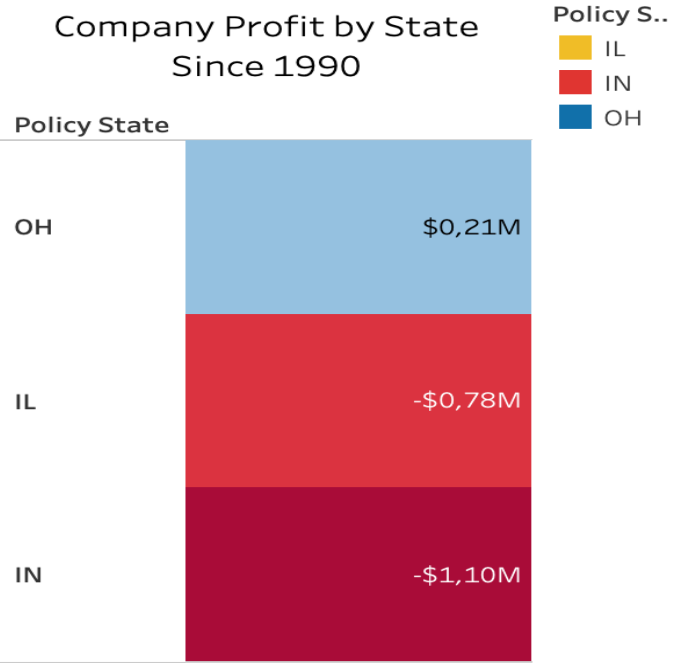
While it's not recommended to eliminate any state based on these figures alone, further analysis into fraud trends, customer segmentation, and controlling claim payouts could offer valuable insights for improving profitability across these regions.

Company Profit Over Time By State

Y-axis: CapitalLoss-Gains — 200K, 100K, 0K, -100K, -200K, -300K, -400K

X-axis: Year of Policy Bind Date — 2008, 2009, 2010, 2011, 2012, 2013, 2014

Policy State

Minnesota, Wisconsin, Michigan, New York, Iowa, Pennsylvania, Illinois (IL), Indiana (IN), OH, Missouri, Kentucky, West Virginia, MD, DE, NJ, Tennessee, North Carolina, Oklahoma, Arkansas, Mississippi, Georgia, Louisiana

© Mapbox © OSM

Percentage of Claims by State

35,00% OH

33,85% IL

31,14% IN

Company Profit by State Since 1990

Policy State

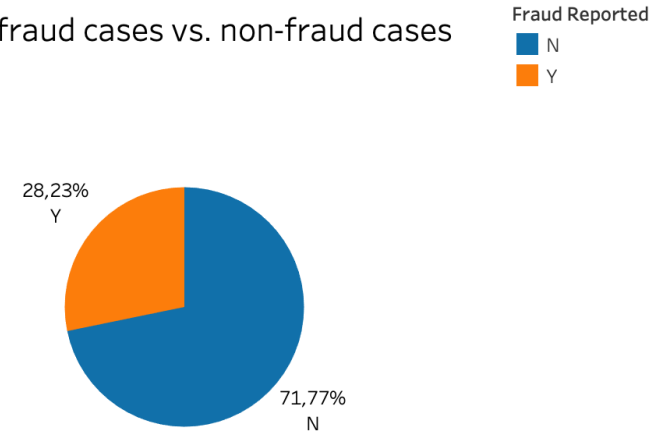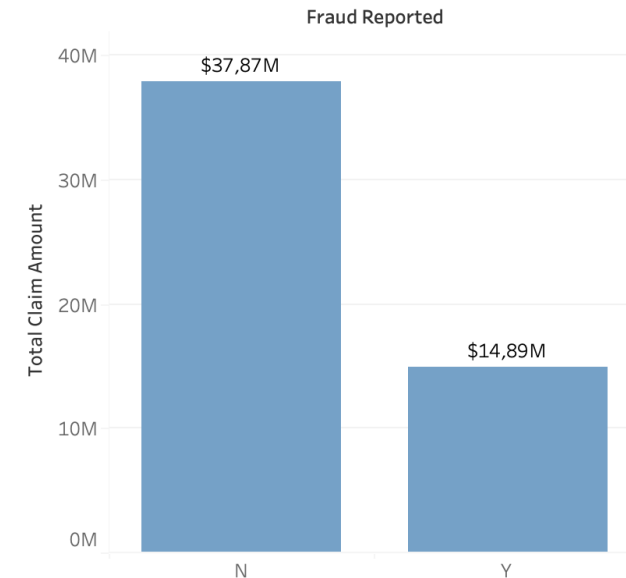| Policy State	
OH	$0,21M
IL	-$0,78M
IN	-$1,10M

Policy S..
- IL
- IN
- OH

5. Claims Fraud Analysis

•The analysis of fraud claims highlights their significant impact on the company's profitability. Nearly 30% of total claim payouts, or around $15 million, are fraudulent, which substantially reduces profits. Addressing this issue should be a top priority for improving financial performance.

•Fraudulent claims arise from various sources, including different cities, incident types, and claim categories, such as property, injury, and vehicle claims. The involvement of authorities in the claims process does not guarantee legitimacy, as fraud can still occur even with police or other officials involved.

•To reduce the financial impact of fraud, the company should implement strategies for detecting and preventing fraudulent activities. This could include stricter checks, predictive modeling, and setting limits on claim payouts. By identifying patterns and flagging high-risk claims, the company can mitigate losses and protect its profitability.

Total fraud cases vs. non-fraud cases

Fraud Reported
■ N
■ Y

28,23%
Y

71,77%
N

Fraud status and claim payout

Fraud Reported

40M

$37,87M

30M

Total Claim Amount

20M

$14,89M

10M

0M

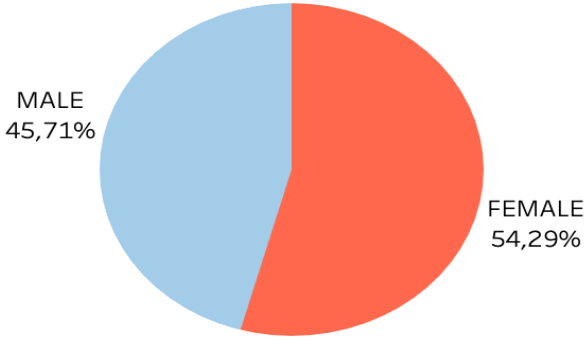N   Y

6. Client Segmentation Analysis

The client segmentation analysis offers valuable insights into the demographics and occupations most likely to file claims. Key findings include:

1. **Gender:** Claims are almost evenly split between male and female clients, with women filing 55% of claims, indicating a slightly higher likelihood of claims among female clients.

2. **Age:** Clients aged 25 to 50 are most likely to file claims, representing a significant portion of claim activity and an important group to monitor for future trends.

3. **Occupational Impact:** The top five occupations contributing to claims payouts are:

   -Machine Operators and Inspectors: Over $5 million in claims.

   -Professional Specialties (doctors, engineers, etc.), Executive and Managerial Roles,Transport and Moving Jobs : Over $4 million on each in claims.
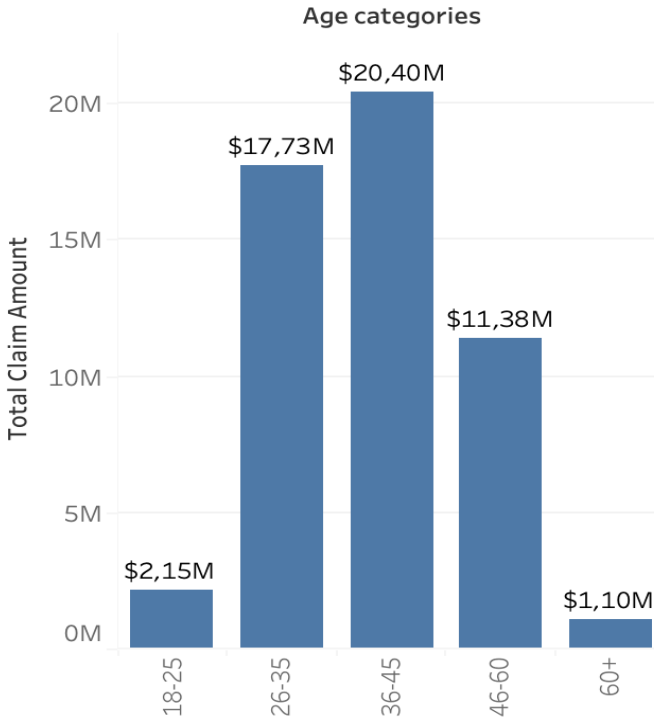
These job categories are associated with higher claims payouts, suggesting certain industries or professions are more prone to higher claims.

4. **Loyal Clients:** Clients with over 350 months of subscription are the least likely to file claims, possibly due to their familiarity with the claims process or a more stable relationship with the company.
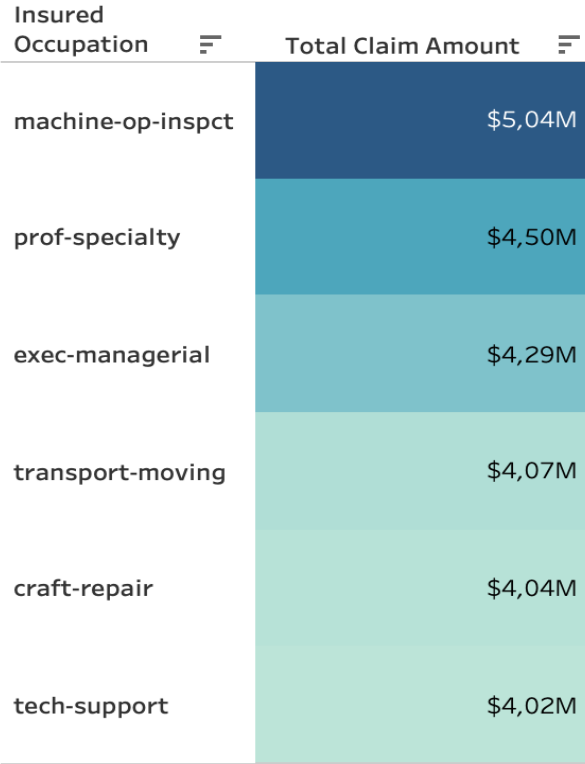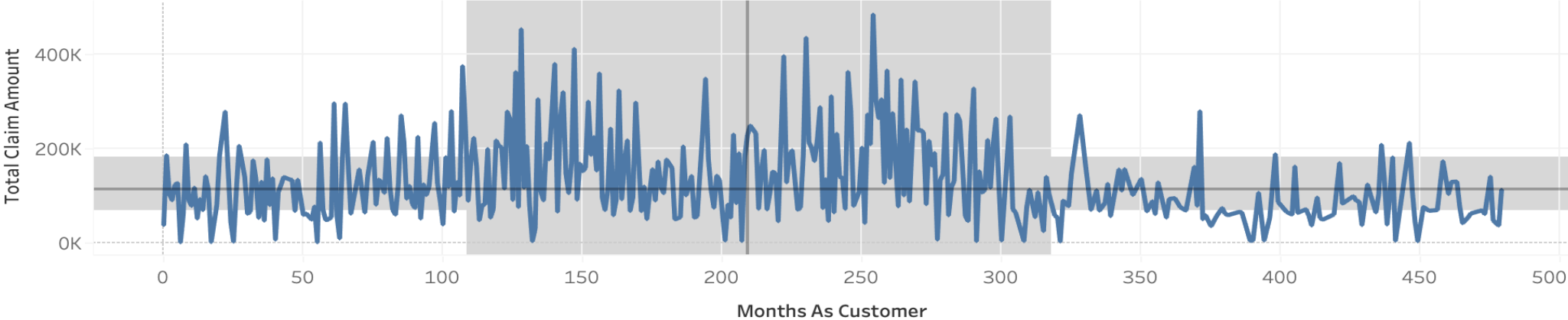
Claim Distribution By Gender

MALE
45,71%

FEMALE
54,29%

Claim Payout Amount by Age Group

Age categories

$2,15M — 18-25
$17,73M — 26-35
$20,40M — 36-45
$11,38M — 46-60
$1,10M — 60+

Total Claim Amount

Jobs Leading to High Claim Payouts

Insured Occupation	Total Claim Amount
machine-op-inspct	$5,04M
prof-specialty	$4,50M
exec-managerial	$4,29M
transport-moving	$4,07M
craft-repair	$4,04M
tech-support	$4,02M

Average Claim Payout By Months As Customer

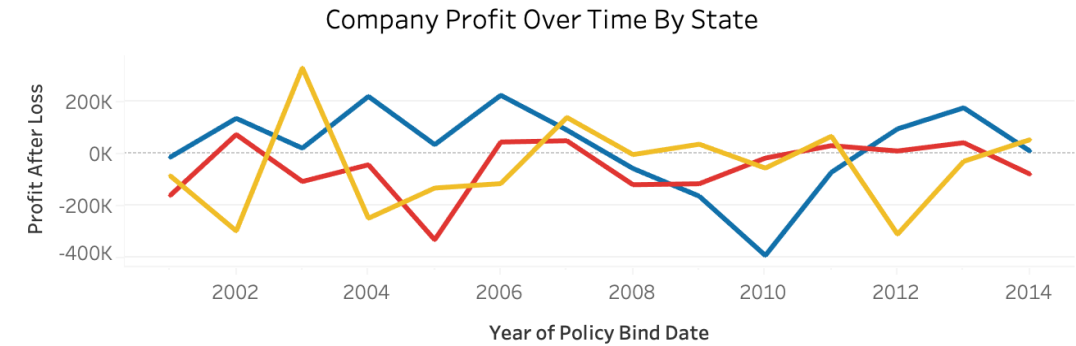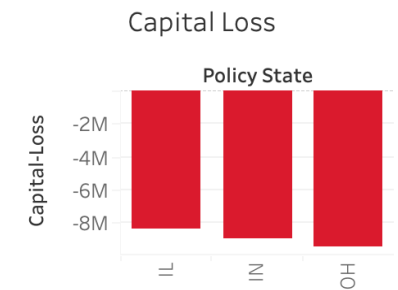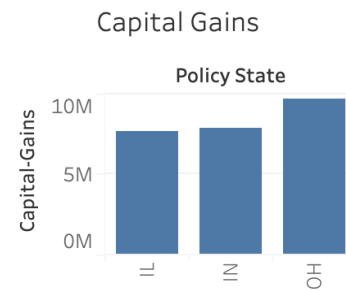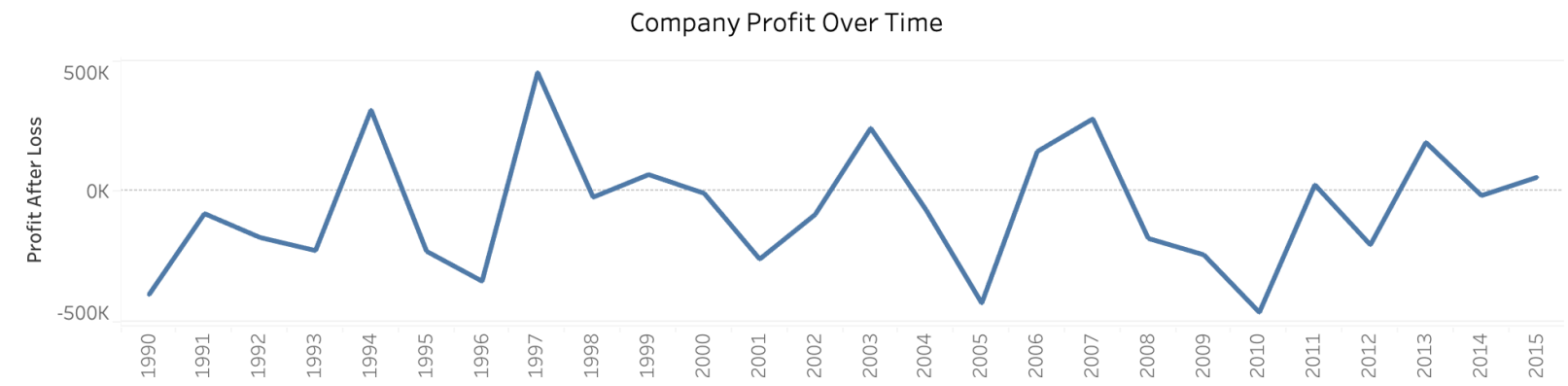Total Claim Amount

Months As Customer

7. Dashboard Insights

A **Tableau dashboards** were created to provide a **visual breakdown** of:

I- Financial Trends Over Time:
Overall Performance with Insights for Ohio, Illinois, and Indiana
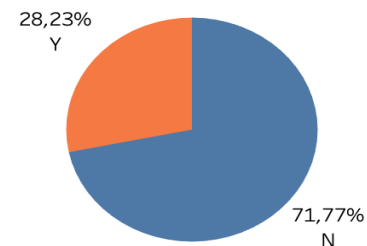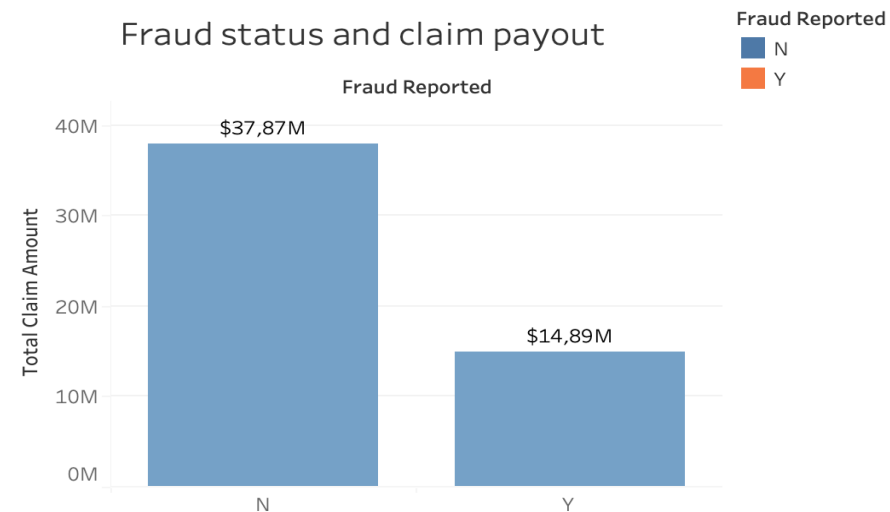
Financial Performance Overview

II - Geographical Distribution of Claims and Fraud Hotspots

Fraud Insights: Case Distribution, Severity & Risk Patterns

Total fraud cases vs. non-fraud cases
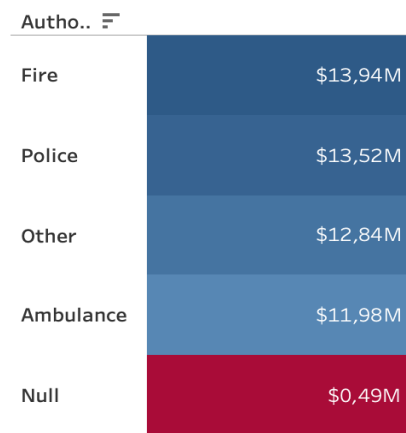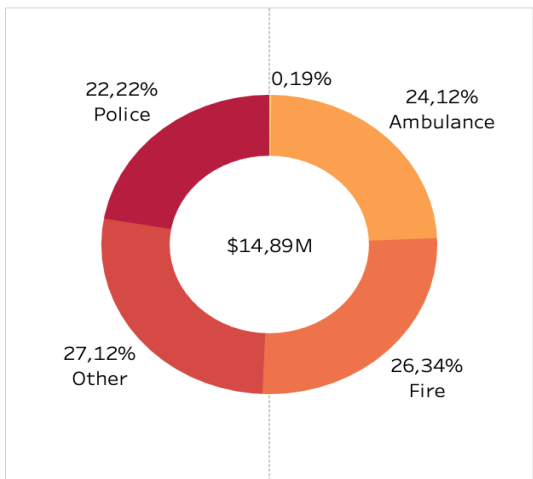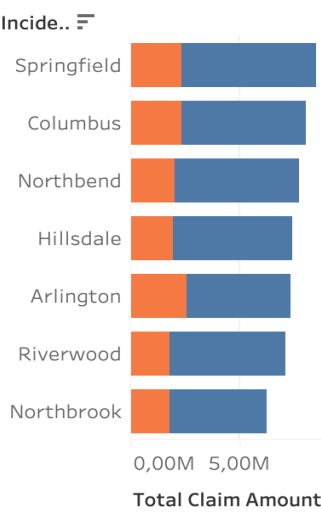
28,23%
Y

71,77%
N

Fraud status and claim payout

Fraud Reported

Fraud Reported
- N
- Y

$37,87M

$14,89M

Total Claim Amount

40M

30M

20M

10M

0M

N   Y

Authorities Contacted After Claim Occurrence

Autho..

Fire	$13,94M
Police	$13,52M
Other	$12,84M
Ambulance	$11,98M
Null	$0,49M

Fraud Percentage Despite Involvement of Authorities

22,22%
Police

0,19%

24,12%
Ambulance

$14,89M

27,12%
Other

26,34%
Fire

Fraud By City

Incide..

Springfield
Columbus
Northbend
Hillsdale
Arlington
Riverwood
Northbrook

0,00M   5,00M

Total Claim Amount

Incident Severity vs Fraud Cases

Incident Severity

Total Claim Amount

15M

10M

5M

0M

Major Damage   Total Loss   Minor Damage   Trivial Damage
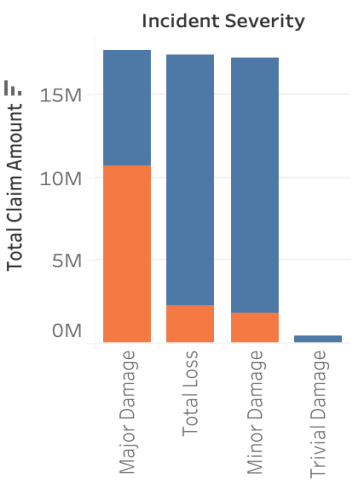
III - Client demographics and segmentation insights

Claim Distribution By Gender
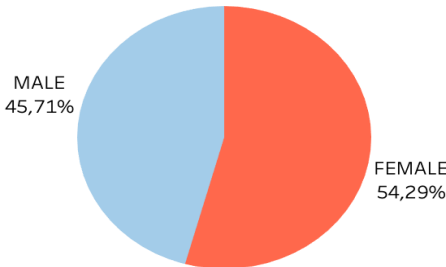
MALE
45,71%

FEMALE
54,29%

Claim Payout Amount by Age Group

Age categories

$20,40M

$17,73M

$11,38M

$2,15M

$1,10M

Total Claim Amount

18-25 26-35 36-45 46-60 60+

Jobs Leading to High Claim Payouts

Insured Occupation	Total Claim Amount
machine-op-inspct	$5,04M
prof-specialty	$4,50M
exec-managerial	$4,29M
transport-moving	$4,07M
craft-repair	$4,04M
tech-support	$4,02M

Average Claim Payout By Months As Customer

Total Claim Amount

400K

200K

0K

0   50   100 150 200 250 300 350 400 450 500
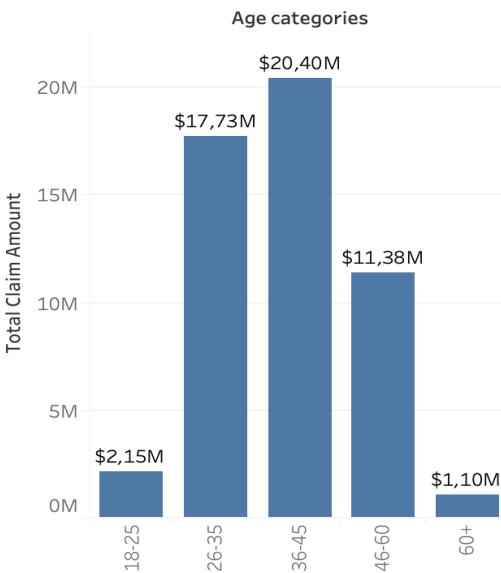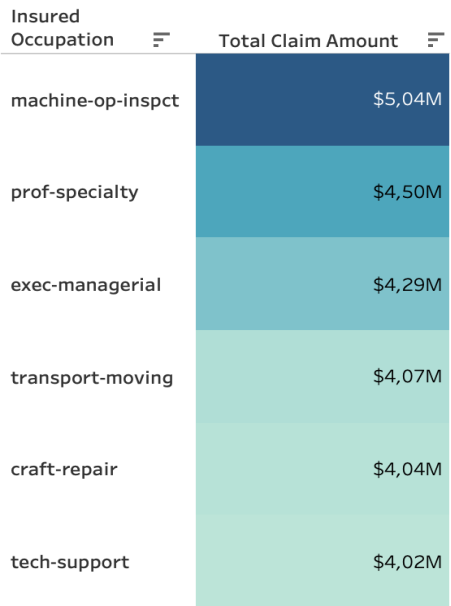
Months As Customer
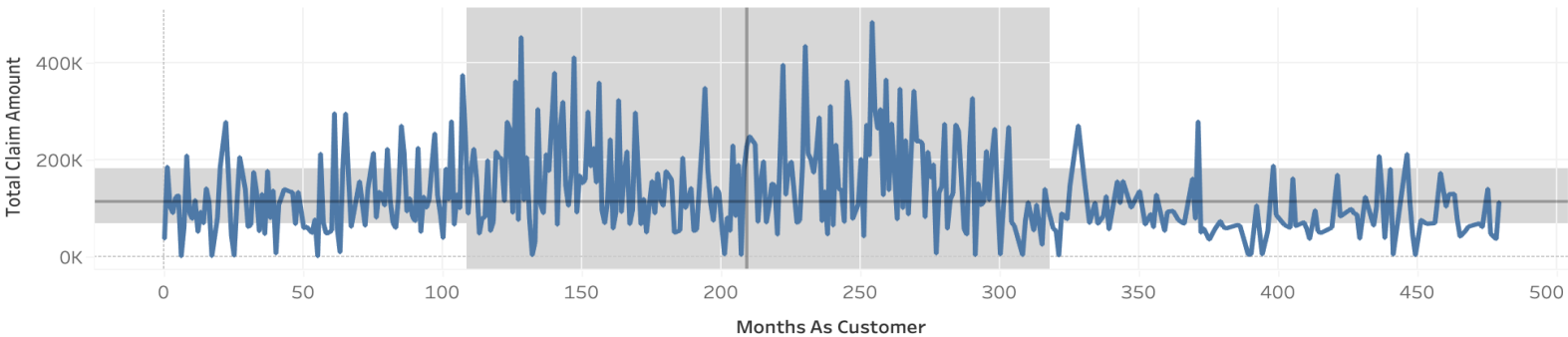
8. Conclusion & Recommendations

The analysis reveals several key insights into the company's claims performance and client behavior across the three states: **Indiana**, **Ohio**, and **Illinois**. The company has faced ongoing financial challenges, with a significant portion of its gains offset by large claim payouts. **Fraudulent claims** account for nearly **30%** of the total payout, amounting to approximately **$15 million**, which has been a major drain on profitability. Geographically, Ohio has contributed a relatively small profit of **$214k**, while Indiana and Illinois have experienced substantial losses of **$2-3 million** and over **$1.1 million**, respectively.

Client segmentation analysis shows that **women** between the ages of **25-50** are most likely to file claims. Additionally, certain occupations such as **machine operators**, **executive roles**, and **transport workers** contribute to higher claim payouts. Interestingly, long-term clients (those subscribed for over **350 months**) are the least likely to file claims, indicating that loyalty correlates with lower claims activity.

Recommendations

Based on these findings, several actionable recommendations can be made:

- **Enhance Fraud Detection and Prevention**: Since fraudulent claims represent a significant portion of the claims payout, the company should prioritize improving its fraud detection methods. Implementing predictive models using machine learning could help identify high-risk claims and reduce losses from fraud.

- **Refine Client Outreach and Segmentation**: The company can tailor its outreach efforts by focusing on the client segments that are more likely to file claims, such as women aged 25-50 and employees in higher-risk occupations. Offering personalized prevention tips and incentives could help reduce claim frequency in these groups.

- **Geographic Focus on High-Risk Areas**: While the overall geographical performance is mixed, particular attention should be given to Ohio, Indiana, and Illinois, as these states have demonstrated fluctuating financial performance. A more targeted strategy, such as focusing on areas with higher claim payouts, could help reduce losses and stabilize profits.

- **Client Retention and Loyalty Programs**: Given that long-term clients show a lower tendency to file claims, the company could develop a stronger client loyalty program to encourage long-term subscriptions and reduce the number of claims filed.

Next *Steps*

To build on this analysis, the following next steps are recommended:

- **Further Refine Client Segmentation**: More detailed segmentation analysis could provide insights into specific client behaviors, enabling the development of targeted interventions that reduce claims.

- **Leverage Machine Learning for Fraud Detection**: The company could explore the use of more sophisticated machine learning models to predict fraud based on historical data, enhancing the accuracy and effectiveness of fraud prevention efforts.

- **Evaluate Claims Processing Efficiency**: Further analysis could focus on streamlining the claims processing system to ensure that legitimate claims are handled swiftly while fraud is detected more efficiently.

- **Expand Geographic Analysis**: A deeper dive into additional geographic regions could provide a more comprehensive understanding of the company's performance across different markets, revealing untapped opportunities for growth or areas that require more attention.

9. Appendices

Company Profit Comparison
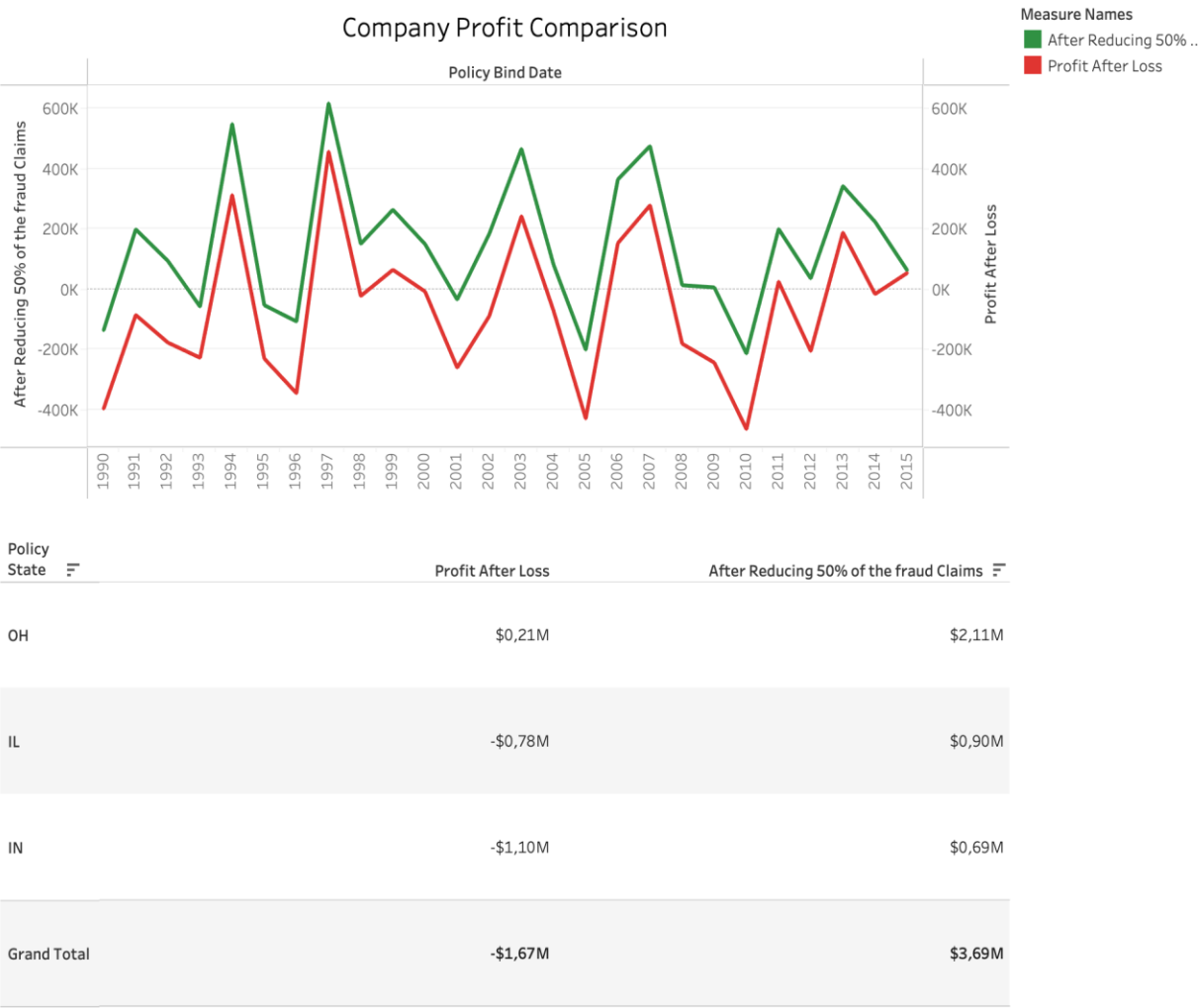
In this section, I calculated the company's profit from 1990 to 2015, factoring in a 50% reduction in fraudulent claims. The analysis shows a significant improvement: by cutting fraud, the company can maintain a consistent monthly profit, with the maximum claim payout reduced to $200k (from $500k). As a result, monthly profits exceed $400k, demonstrating that reducing fraud eliminates losses and ensures long-term profitability, even with claims being processed.

Policy State	Profit After Loss	After Reducing 50% of the fraud Claims
OH	$0,21M	$2,11M
IL	-$0,78M	$0,90M
IN	-$1,10M	$0,69M
Grand Total	-$1,67M	$3,69M