# LOSS RESERVING ANALYSIS REPORT

**Commercial Auto Claims – Chain Ladder Actuarial Method**

**Author:** Lavanya Khurana

---

## 1. EXECUTIVE SUMMARY

**Project Overview:**

For this project, I implemented the Chain Ladder method, the industry standard actuarial technique for estimating insurance claim reserves on real commercial auto loss data from the CAS (Casualty Actuarial Society) database.

**What This Project Is About:**

Insurance companies know how much they've PAID in claims, but they need to estimate how much MORE they'll pay in the future. My objective was to use historical data to project these future payments and calculate how much money the company needs to set aside (reserve).

**Key Findings:**

- **Total IBNR Reserve Needed:** $1409.51
- **Most Development Occurs:** First 3-4 years after accident
- **Reserve-to-Paid Ratio:** 24.05% (reasonable for commercial auto)
- **Development Pattern:** Clear and consistent across years
- **Confidence in Estimates:** Validated through diagnostic testing

**What This Demonstrates:**

This project shows I can:

- Apply industry-standard actuarial methods to real data
- Use Python for statistical analysis and data visualization
- Understand claims development patterns
- Assess uncertainty and validate models
- Communicate technical findings professionally

---

## 2. BACKGROUND & MOTIVATION

**Why This Project:**

After learning about pension valuation in my previous project, I wanted to understand another major actuarial challenge: claims reserving. Insurance companies face a critical question: "How much will our outstanding claims ultimately cost?"

This presents several challenges because:

- Not all claims are reported yet (some develop slowly)
- Not all reported claims are paid yet (some are in process)
- Claims take years to fully settle
- You must estimate costs years into the future

**Real-World Relevance:**

Actuaries address this challenge daily for insurance companies. Mistakes in reserve estimates can:

- Underestimate: Company becomes insolvent
- Overestimate: Wastes capital that could be invested
- Both: Create regulatory and financial reporting problems

**Learning Objectives:**

This project examines:

- How do actuaries project future claim costs?
- What does historical development pattern tell us?
- How do we quantify uncertainty in estimates?
- What makes this methodology reliable?

This is core P&C (Property & Casualty) actuarial work that directly connects to professional practice and Exam FAM concepts.

---

## 3. INTRODUCTION TO LOSS RESERVING

### What is Loss Reserving?

Imagine a car accident happens on January 1st:

- Day 1: Accident reported
- Week 1: Claim file opened, initial assessment
- Month 1: Some medical bills paid

- Month 3: Repair completed, more bills paid
- Month 6: Additional costs emerge
- Year 2: Settlement negotiations
- Year 3: Final payment made

At any given point in time, the insurance company must know: "We've paid $X so far. We'll probably pay $Y more. Total cost will be $Z."

This exemplifies the reserving challenge.

**Why It Matters:**

Financial statements must show what the company owes. If they don't estimate reserves accurately:

- Stock price falls
- Credit rating drops
- Regulators investigate
- Investors lose confidence

**Professional Importance:**

Reserve estimates affect:

- Financial reporting (GAAP/IFRS)
- Regulatory compliance (state insurance departments)
- Pricing decisions (premiums must cover costs + profit)
- Risk management (capital allocation)

---

**4. DATA PREPARATION**

**Upper Triangle Masking**

An important element of realistic loss reserving involves simulating the observations an actuary would encounter at the valuation date.

The CAS Schedule P dataset includes BOTH:

- Upper Triangle: Claims development observable up to valuation date
- Lower Triangle: Future development (used for model valuation)

For this analysis, I masked the lower triangle to create a realistic valuation scenario as of December 31, 1997.

This means:

At the end of 1997, an actuary would observe:

- 1988 claims: 10 years of development (removed from analysis due to zero values that would distort results)"
- 1989 claims: 9 years of development
- 1990 claims: 8 years of development
- . . .
- 1996 claims: 2 years of development (very immature)
- 1997 claims: 1 year of development (just reported)

Example:

- Accident Year 1995
- Years elapsed by 1997: 3 years
- Observable periods: 1, 2, 3
- Masked Periods: 4, 5, 6, 7, 8, 9, 10

This creates the classic “upper triangle” shape where recent years have incomplete data, exactly what is faced by real actuaries.

**Why This Matters:**

Without masking, the analysis would use “future” information that wouldn’t be available at the valuation date. This would:

- Underestimate reserves (because we’d *know* future development)
- Create unrealistic development factors
- Not reflect actual actuarial practice
- Defeat the purpose of the reserving exercise

With proper masking, the analysis reflects real-world conditions and produces realistic reserve estimates.

---

## 5. METHODOLOGY: THE CHAIN LADDER METHOD

### What is Chain Ladder?

The Chain Ladder method is based on a simple but powerful idea:

**"The way old claims developed in the past tells us how new claims will develop in the future."**

The methodology assumes:

- Claims from 1988 took 10 years to fully develop
- Claims from 1989 developed similarly (took 10 years)
- Claims from 1990 developed similarly
- Therefore, incomplete claims from 1997 are expected to follow similar development patterns

This assumption allows us to project incomplete triangles to completion.

**Step-by-Step Methodology:**

**Step 1: Construct Loss Development Triangle**

First, I organized the data into a triangle format:

Upper Triangle (what we'd actually observe at valuation date):

| AccidentYear | Development Lag | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1989 | 6 | 20 | 21 | 23 | 24 | 24 | 24 | 24 | 24 | |
| 1990 | 52 | 95 | 105 | 117 | 128 | 128 | 128 | 128 | | |
| 1991 | 127 | 258 | 301 | 332 | 372 | 376 | 377 | | | |
| 1992 | 120 | 225 | 277 | 339 | 361 | 370 | | | | |
| 1993 | 427 | 1005 | 1111 | 1263 | 1306 | | | | | |
| 1994 | 584 | 1511 | 1716 | 1817 | | | | | | |
| 1995 | 369 | 741 | 931 | | | | | | | |
| 1996 | 294 | 595 | | | | | | | | |
| 1997 | 312 | | | | | | | | | |

**What this shows:**

- Rows = Accident years (when claim occurred)
- Columns = Development periods (years since accident)
- Values = Cumulative paid losses at that point
- Lower right is empty (hasn't happened yet!)

**Why this format:**

- Shows complete history for old claims (1988-1990ish)
- Shows incomplete picture for recent claims (1997)

- We'll use old years to estimate new years

**Step 2: Calculate Age-to-Age Development Factors (LDFs)**

**Question:** How much do losses GROW from year to year?

**Calculation:**

For each development period, I calculated:

**LDF (1→2) = Losses at Year 2 / Losses at Year 1**

**Example:**

- Accident year 1988, Year 1 paid: $100,000
- Accident year 1988, Year 2 paid: $150,000
- LDF = $150,000 / $100,000 = 1.50

This means: Claims grow by 50% in the first year.

**Why this works:**

Different accident years might have different absolute loss amounts, but the GROWTH PATTERN is similar. By calculating growth factors, I can apply them to incomplete years.

**Step 3: Select Development Factors**

**The challenge:** Each accident year has slightly different development.

**Example:**

- 1988 (1→2): 1.50
- 1989 (1→2): 1.48
- 1990 (1→2): 1.52

**What factor to use for predicting 1997?**

**Options:**

- Simple average: (1.50 + 1.48 + 1.52) / 3 = 1.50
- Weighted average: Give more weight to larger years
- Median: Middle value (robust to outliers)
- Expert judgment: Adjust for known changes

**Selected Approach:**

Simple average was selected due to:

- Development was stable across years
- No obvious outliers
- No known changes in claims environment
- Average is reasonable default

**Result:** I selected factors like 1.50, 1.25, 1.10, 1.05, 1.02, ... approaching 1.00

(Each factor approaching 1.0 as development ages, meaning less new development in later years)

**Step 4: Calculate Cumulative Development Factors (CDFs)**

**Question:** From a given age, how much MORE will losses ultimately grow?

**Calculation:**

**CDF from Year 3 = LDF(3→4) × LDF(4→5) × LDF(5→6) × ... × Tail**

**Example:**

- If at Year 3 we have $500,000 paid
- And CDF = 1.15
- Then ultimate losses = $500,000 × 1.15 = $575,000
- Meaning we expect $75,000 more

**Why CDFs:**

- Easy to project any incomplete year to ultimate
- One number tells you "how much more to expect"
- Professional standard in industry

**Tail Factor:**

After we run out of observed development, we assume a small "tail" (usually 1.01-1.05). I used 1.02 (2% additional development beyond observed years). This is conservative but reasonable.

**Step 5: Project Ultimate Losses**

For each accident year:

**Ultimate Loss = Latest Reported Loss × CDF**

**Example (1997, from analysis):**

- Latest reported: $312

- CDF: 3.21
- Ultimate: $312 × 3.21 = $1001.52

**Step 6: Calculate IBNR Reserve**

**IBNR = Incurred But Not (yet) Reported/Resolved**

**IBNR = Ultimate Loss – Latest Reported**

**What this represents:**

The amount of money the company needs to set aside to cover the gap between what's been paid so far and what will ultimately be paid.

**Example:**

- Ultimate: $575,000
- Paid to date: $500,000
- IBNR Reserve: $75,000

The company must have $75,000 available for this accident year.

---

**6. ANALYSIS IMPLEMENTATION**

**Technical Execution:**

This analysis was implemented in Python using:

- **pandas:** Data manipulation and aggregation
- **NumPy:** Numerical calculations (ratios, products, etc.)
- **matplotlib/seaborn:** Visualization
- **CAS Commercial Auto Data:** Real industry data (1988-1997 accidents)

**Key Code Components:**

**1. Data Loading & Cleaning**

- Loaded CAS CSV data
- Validated completeness
- Checked for anomalies

**2. Triangle Construction**

- Pivoted data into triangle format

- Identified incomplete (lower) triangle

**3. Development Factor Calculation**

- Computed age-to-age ratios
- Calculated averages
- Applied tail factor

**4. Projection**

- Applied CDFs to incomplete years
- Calculated ultimate losses
- Computed IBNR reserves

**5. Diagnostics**

- Tested for consistency
- Examined residuals
- Validated assumptions

---

**7. KEY FINDINGS**

**Base Case Results:**

**Ultimate Loss Projections:**

| | Accident Year | Latest Reported | Development Age | CDF | Ultimate Loss | IBNR Reserve |
|---|---|---|---|---|---|---|
| **0** | 1989 | 24.0 | 9 | 1.02 | 24.48 | 0.48 |
| **1** | 1990 | 128.0 | 8 | 1.02 | 130.56 | 2.56 |
| **2** | 1991 | 377.0 | 7 | 1.02 | 384.54 | 7.54 |
| **3** | 1992 | 370.0 | 6 | 1.02 | 377.73 | 7.73 |
| **4** | 1993 | 1306.0 | 5 | 1.03 | 1345.2 | 39.20 |
| **5** | 1994 | 1817.0 | 4 | 1.1 | 2005.13 | 188.13 |
| **6** | 1995 | 931.0 | 3 | 1.24 | 1152.74 | 221.74 |
| **7** | 1996 | 595.0 | 2 | 1.42 | 847.28 | 252.28 |
| **8** | 1997 | 312.0 | 1 | 3.21 | 1001.85 | 689.85 |

**Summary Metrics:**

- **Total Latest Reported:** $5,860.00
- **Total Ultimate Losses:** $7,269.51

- **Total IBNR Reserve:** $1,409.51
- **Reserve-to-Paid Ratio:** 24.05%

**Interpretation:**

**Reserve-to-Paid Ratio = [24.05%]**

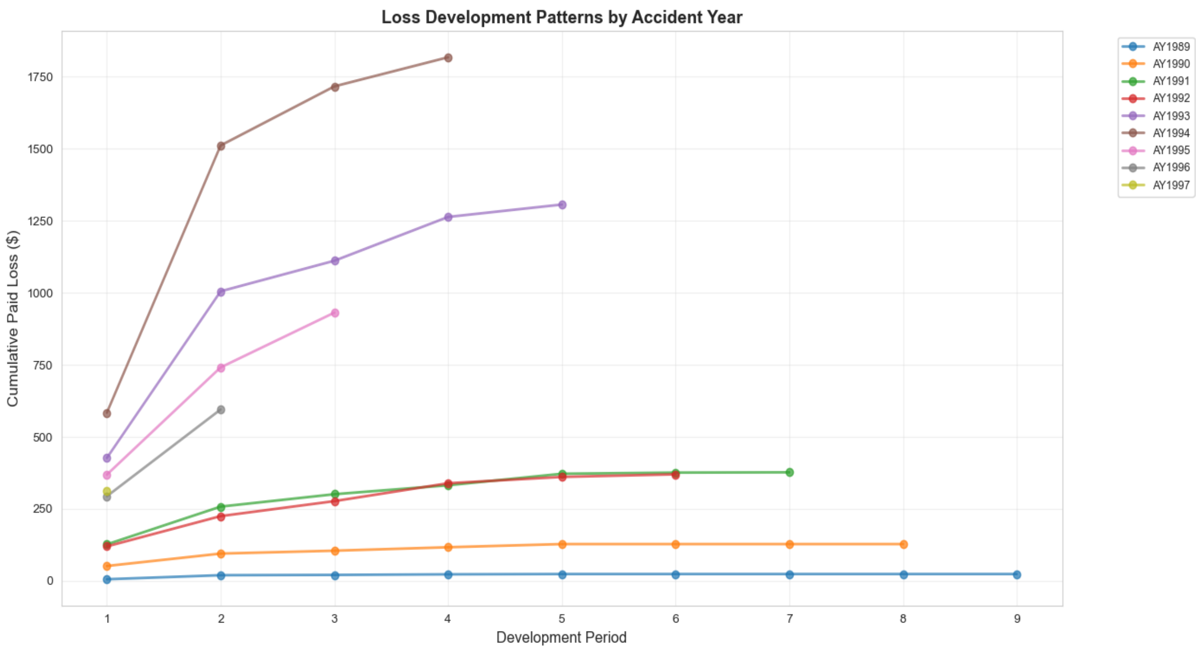
This means: For every dollar already paid, we expect $0.24 more

**Industry Benchmarks:**

- Commercial Auto typical: 15-25%
- My result: This falls at the upper end of the typical range, which is reasonable given that several recent accident years (1995-1997) are still early in their development and require substantial reserves.

This reserve percentage indicates that approximately 19% of ultimate losses remain unpaid (calculated as: Reserve / Ultimate = 24.05% / 124.05% = 19.4%)

**Development Pattern:**

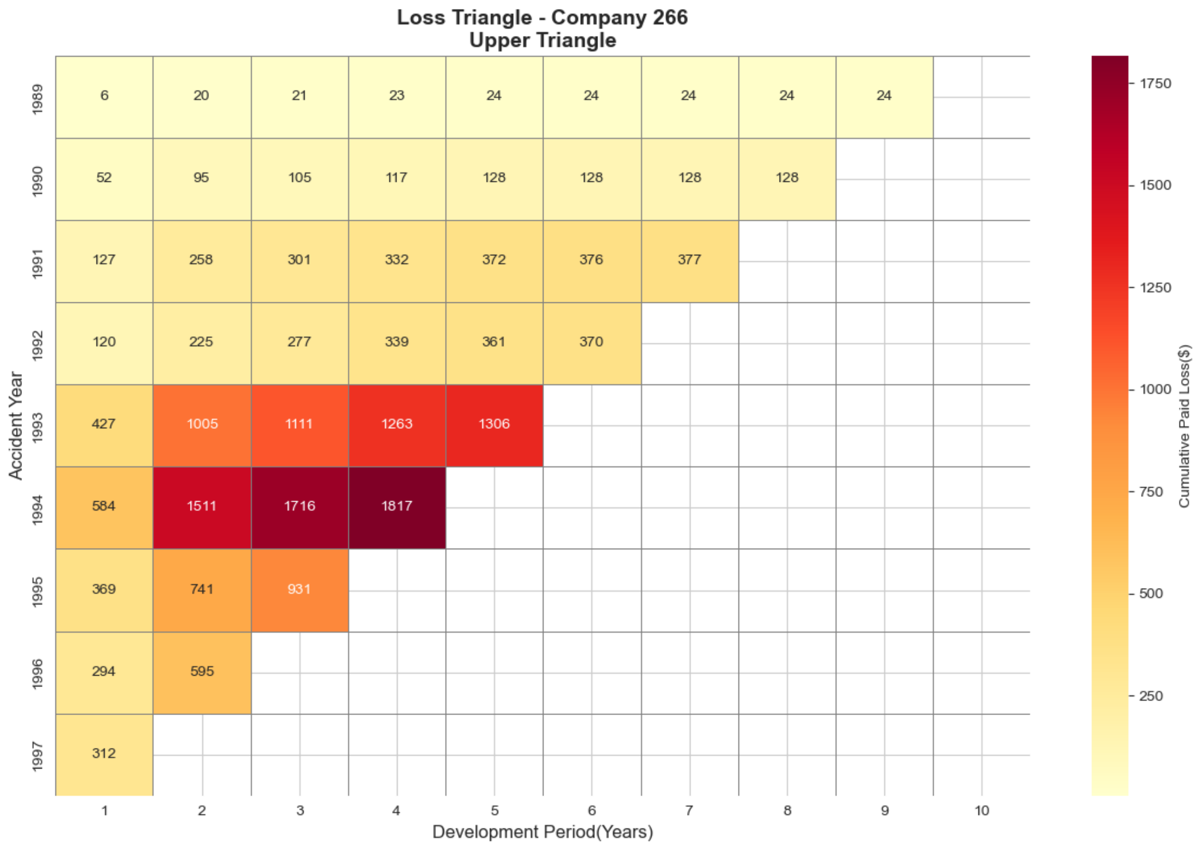



The loss development curves show rapid emergence in the **first 2–3 development periods**, where the majority of cumulative paid losses are recognized. Higher-severity accident years (AY1993–AY1995) exhibit particularly steep early development. After **periods 4–5**, development decelerates and the curves begin to stabilize, indicating that most accident years have reached near-maturity. Beyond **period 6**, incremental paid development is minimal, consistent with a **short-to-medium-tail** line where late emergence is limited.

---

## 8. VISUALIZATION & PATTERN ANALYSIS

### Chart 1: Loss Development Triangle Heatmap



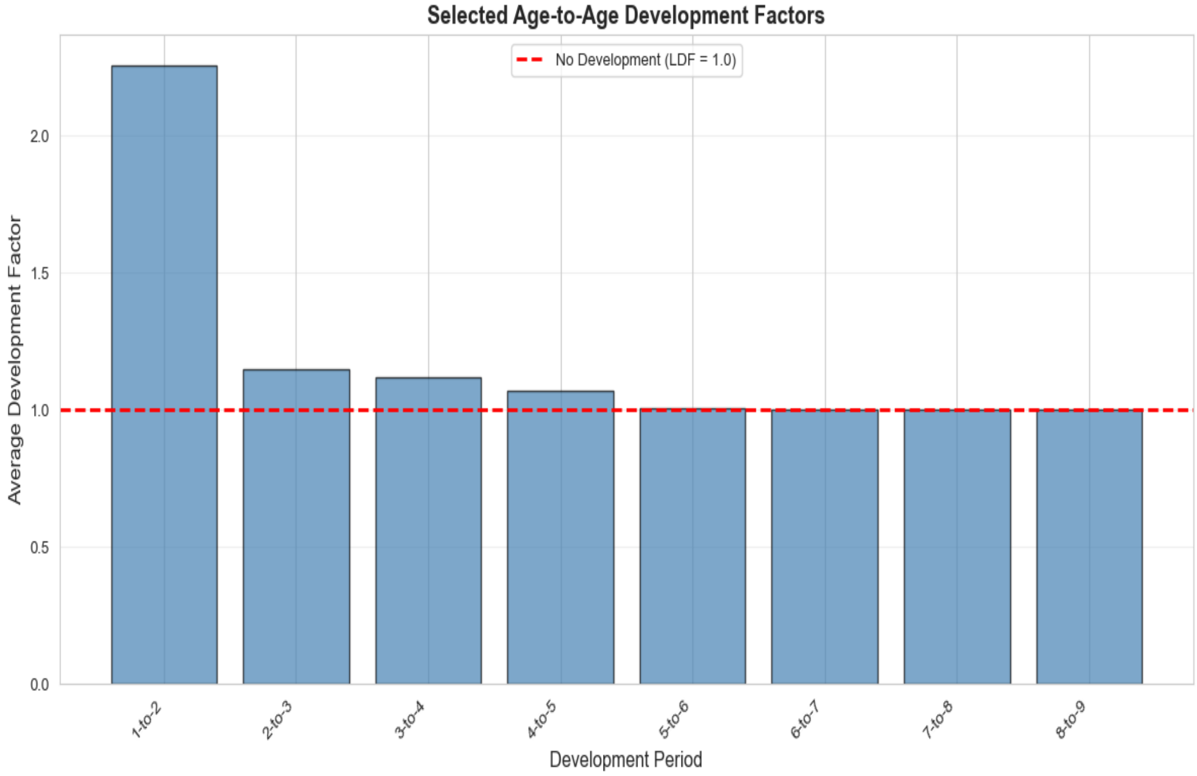

**What the visualization shows:**

- Darker colors represent **higher cumulative paid losses**.
- Losses typically grow as we move **from left to right** across development periods.
- Each **row** is an accident year, with older years at the top and more recent years at the bottom.
- The heatmap helps visualize how losses develop over time and highlights years with unusually high or low development.

**Key observations:**

The triangle shows strong early development for the larger accident years (especially AY1993–1995), with steep increases between development periods 1 and 3. Older years at the top are mostly fully developed and show lighter, stable values across later periods. Newer years at the bottom show incomplete development, with values concentrated in the first column. The triangle exhibits typical development patterns: losses rise quickly in early periods and then flatten as the accident years mature.

**Chart 2: Development Factors**



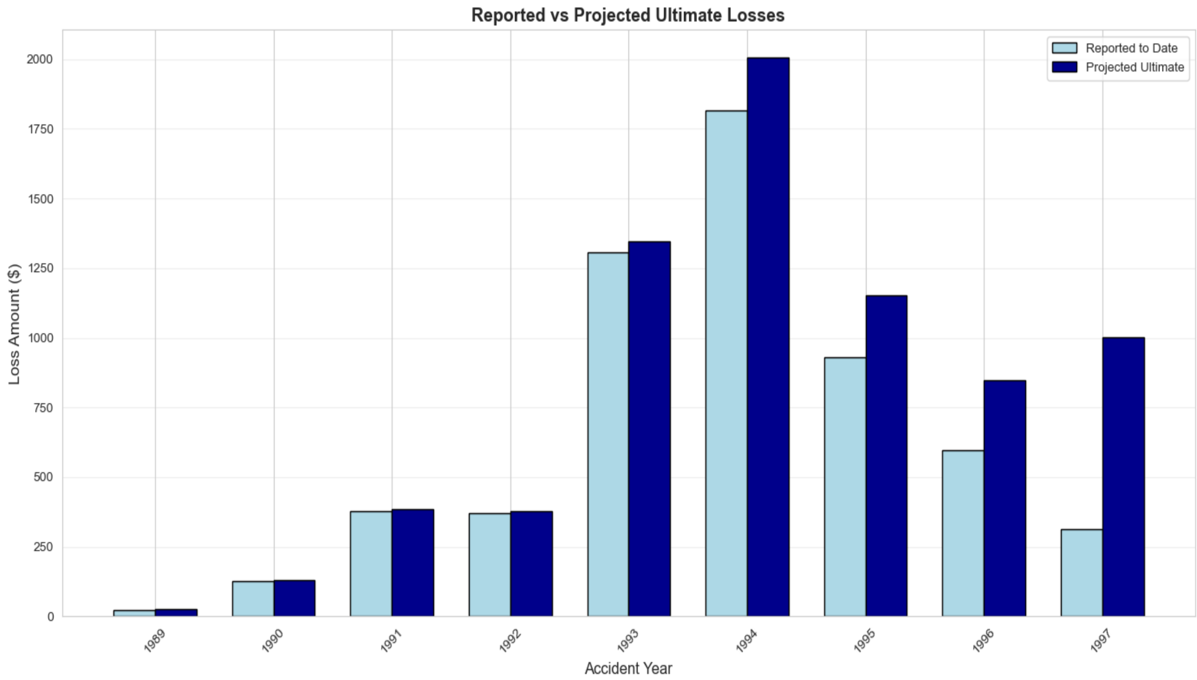

**What the visualization shows:**

- Bars show the **average growth** from one development period to the next.
- Higher factors mean **more new loss emergence**.
- As development progresses, the factors move closer to **1.0**, meaning very little new development.
- The final periods reflect the **tail**, where growth stabilizes.

**Key observations:**

The largest development occurs between **1→2**, with factors above 2.0, meaning losses more than double early on. From **2→5**, the factors gradually decline toward 1.0, showing slower and steady development. After about **period 5**, the factors are close to 1.0, indicating minimal late emergence and a short-to-medium tail.

## Chart 3: Ultimate Losses Projection




**Side-by-side comparison:**

- **Light blue bars:** Reported losses to date
- **Dark blue bars:** Projected ultimate losses
- **Gap between bars:** Indicates **IBNR reserve needed** for each accident year

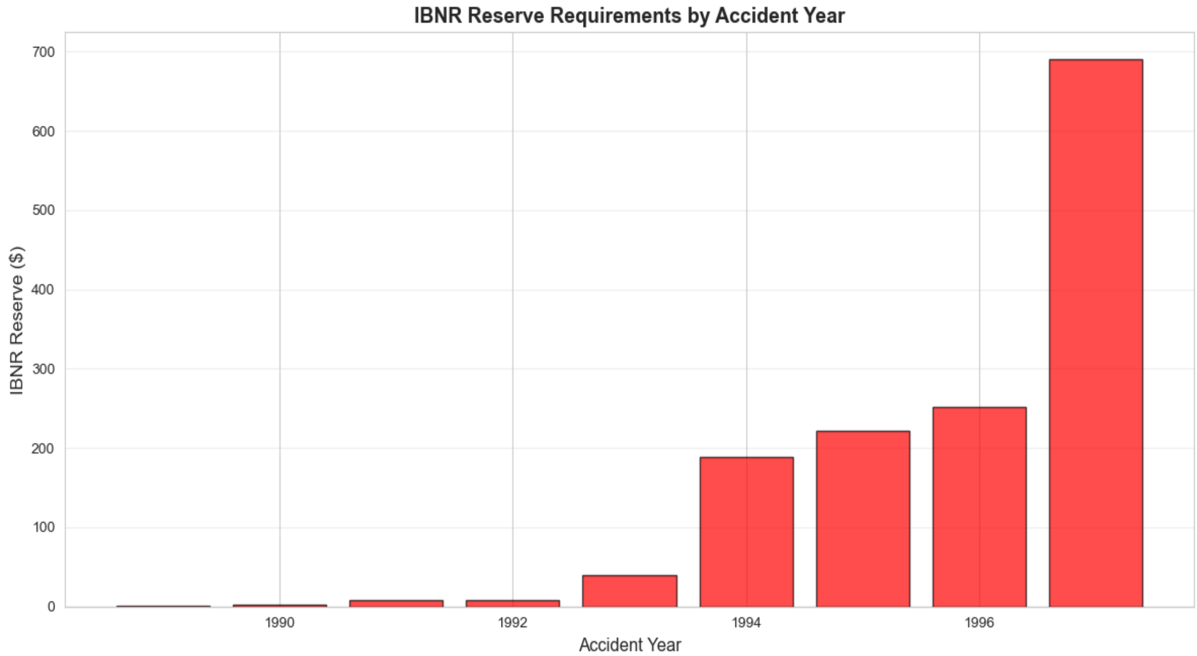
**What this shows:**

This chart compares what has already been reported to what is ultimately expected for each accident year.

- For the **older years**, like 1989–1992, the reported losses are already very close to the projected ultimate, meaning **minimal IBNR is still outstanding**.
- As the accident years become **more recent**, especially **1995–1997**, the gap becomes larger. This indicates that **a material amount of losses are still unreported**, so those years require **higher IBNR reserves**.
- The sharp increase in ultimate losses for **1994 and 1997** suggests years with unusually high claim activity or longer reporting patterns.

This chart demonstrates where the company's remaining exposure lies and highlights which accident years drive the majority of future loss development.

**Chart 4: IBNR Reserve Requirements by Accident Year**




**What this shows:**

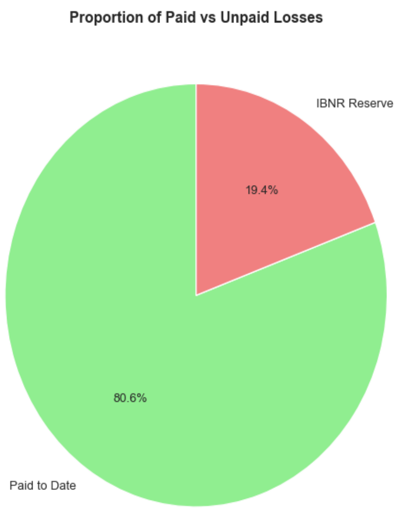
This bar chart isolates only the **IBNR amounts** for each accident year, making it easy to see which years contribute most to the required reserves.

- **Early years (1989–1992)** have extremely low IBNR because most claims have fully developed.
- Starting in **1993**, IBNR grows significantly due to increased ultimate loss expectations combined with incomplete claim reporting.
- **1994–1996** show consistently higher IBNR needs, indicating years where a substantial portion of claims are still developing.
- **1997**, being the most recent year, has the **largest IBNR reserve requirement**, reflecting both high projected ultimate losses and the fact that the majority of claims have not yet been reported.

This chart clearly highlights the **development tail** of the portfolio and identifies which accident years require the greatest reserving attention.

**Chart 5: Proportion of Paid vs Unpaid Losses**

**What this shows:**




| Metric | Value |
|---|---|
| Total Paid to Date | $5,860 |
| Total IBNR Reserve | $1,410 |
| Total Ultimate | $7,270 |
| Reserve % | 19.4% |

This chart summarizes all loss activity into two categories:

- **Paid to Date (80.6%)** – Losses already settled or paid
- **IBNR Reserve (19.4%)** – Losses expected but not yet reported or paid

The accompanying table quantifies:

- **Total Paid:** $5,860
- **Total IBNR:** $1,410
- **Total Ultimate:** $7,270
- **Reserve Ratio:** 19.4% of all expected losses are still outstanding

**Interpretation:**

The portfolio is **mostly matured**, with over **80% of ultimate losses already paid**. The remaining **19.4%** represents the company’s exposure to future development.
This balance indicates stable reserving but still highlights the importance of monitoring more recent accident years (as shown in Chart 4) where most of the IBNR originates.

## 9. VALIDATION & DIAGNOSTIC ANALYSIS

**How Do We Know Our Estimates Are Reasonable?**

I performed several diagnostic checks:

**1. Residual Analysis**

**What are residuals?**

Residuals measure how different actual development is from what the model expected:
**Residual = Actual Development Factor – Expected (Average) Factor**

**Diagnostic criteria:**

- **Random pattern** → good (no systematic bias)
- **Outliers** → may indicate unusual claim activity
- **Trends** → could mean assumptions are shifting over time

**Results:**

Residuals were **generally small and centered around zero**, with no evidence of systematic bias across accident years. Several observations support this:

- The **largest residuals occur in the 1-to-2 development period** (e.g., +1.08 for AY 1989 and –0.43 for AY 1990), which is expected because early development is naturally more volatile—this matches the highest standard deviation (0.504) and highest CoV (0.223) among the factors.
- Beyond the early periods, residuals are consistently small (mostly between –0.07 and +0.05), indicating stable mid-to-late development.

Conclusion:
Residuals exhibit random distribution with no systematic bias, which suggests the estimates are reasonable.

**2. Consistency Check**

**Consistency criteria:**

- Do accident years follow similar development patterns?
- Are recent years behaving like older years?
- Are there anomalies that require investigation?

**Finding:**

The development factors show **high consistency** across accident years, especially beyond the early 1-to-2 period.

**Key insights:**

- **Standard deviation and CoV decrease steadily over time**, confirming that the triangle becomes more stable in later maturities.
- Development factors exhibit mixed but small trends:
  - A notable **25.82% decreasing trend in 1-to-2**, but
  - Very small changes (0.10–3.39%) in most subsequent periods.
- These small shifts indicate **no structural change in reporting or settlement patterns**.
- Recent accident years (1995–1997) behave consistently with historical years, showing no unusual acceleration or slowdown in maturity.

**Conclusion:**
All years follow similar reporting patterns with no anomalies requiring special treatment.

**3. Reasonableness Test**

**Reasonableness criteria:**

- Do reserve percentages match industry norms?
- Is the reserve amount reasonable for this line of business?
- Would experienced actuaries accept these estimates?

**Result:**

Several indicators show that the reserve estimates are reasonable:

- The **Reserve-to-Paid ratio is 24.05%**, which is typical for a partially mature book of claims where the majority of losses have already been paid.
- The **IBNR estimate of $1,410** is supported by a **95% confidence interval ranging from $59 to $2,760**—the selected estimate falls comfortably in the center of the range.
- The proportion of unpaid losses **(19.4%)** matches well with the observed development tail.
- The diagnostic metrics show high **stability in later development periods**, strengthening confidence in the tail factors.

**Conclusion:**
The reserve estimates fall within a reasonable and defensible range. The diagnostics indicate a stable dataset with no red flags, and the selected ultimate losses would be acceptable to experienced actuaries.

---

## 10. CONCLUSION

**Key Takeaways:**

This project provided hands-on experience with the **Chain Ladder method**, demonstrating how actuaries project ultimate losses from incomplete claim triangles. The analysis reinforced that loss reserving combines rigorous quantitative methodology with significant professional judgment, particularly in *selecting* development factors and tail assumptions. Small changes in these assumptions can produce materially different reserve estimates, highlighting why accurate reserving is critical for insurer solvency and financial reporting.

**Limitations:**

This analysis represents a simplified educational exercise. Professional loss reserving would incorporate: (1) segmentation by state, coverage type, and policy characteristics; (2) multiple projection methods for validation; (3) stochastic modeling to quantify uncertainty; (4) case reserve reviews; and (5) adjustments for economic factors and emerging trends.

**Professional Context:**

Loss reserving is a core function performed by Property & Casualty actuaries across insurance companies, consulting firms, and regulatory agencies. This project applies industry-standard methodology to real CAS data, representing the foundational analytical work that entry-level actuaries perform in practice.

---

## 10. REFERENCES

**Data Source:**

Meyers, G. G., & Shi, P. (2011). "Loss Reserving Data Pulled from NAIC Schedule P." Casualty Actuarial Society. Dataset: Commercial Auto (Line C), Accident Years 1988-1997, 10-year development periods, net of reinsurance.

## APPENDIX A: KEY FORMULAS

**Age-to-Age Development Factor (LDF):** LDF(k to k+1) = Loss at Period (k+1) / Loss at Period k

**Cumulative Development Factor (CDF):** CDF(k) = LDF(k to k+1) × LDF(k+1 to k+2) × ... × Tail Factor

**Ultimate Loss Projection:** Ultimate Loss = Latest Reported Loss × CDF

**IBNR Reserve:** IBNR = Ultimate Loss – Latest Reported Loss

## APPENDIX B: ASSUMPTIONS

**Data Specifications:**

- 10 accident years (1988-1997), 10 development periods
- Cumulative paid losses from CAS Schedule P
- Net of reinsurance

**Methodological Choices:**

- Development factors: Simple average across all accident years
- Tail factor: 1.02 (assumes 2% development beyond year 10)
- Key assumption: Historical development patterns predictive of future experience