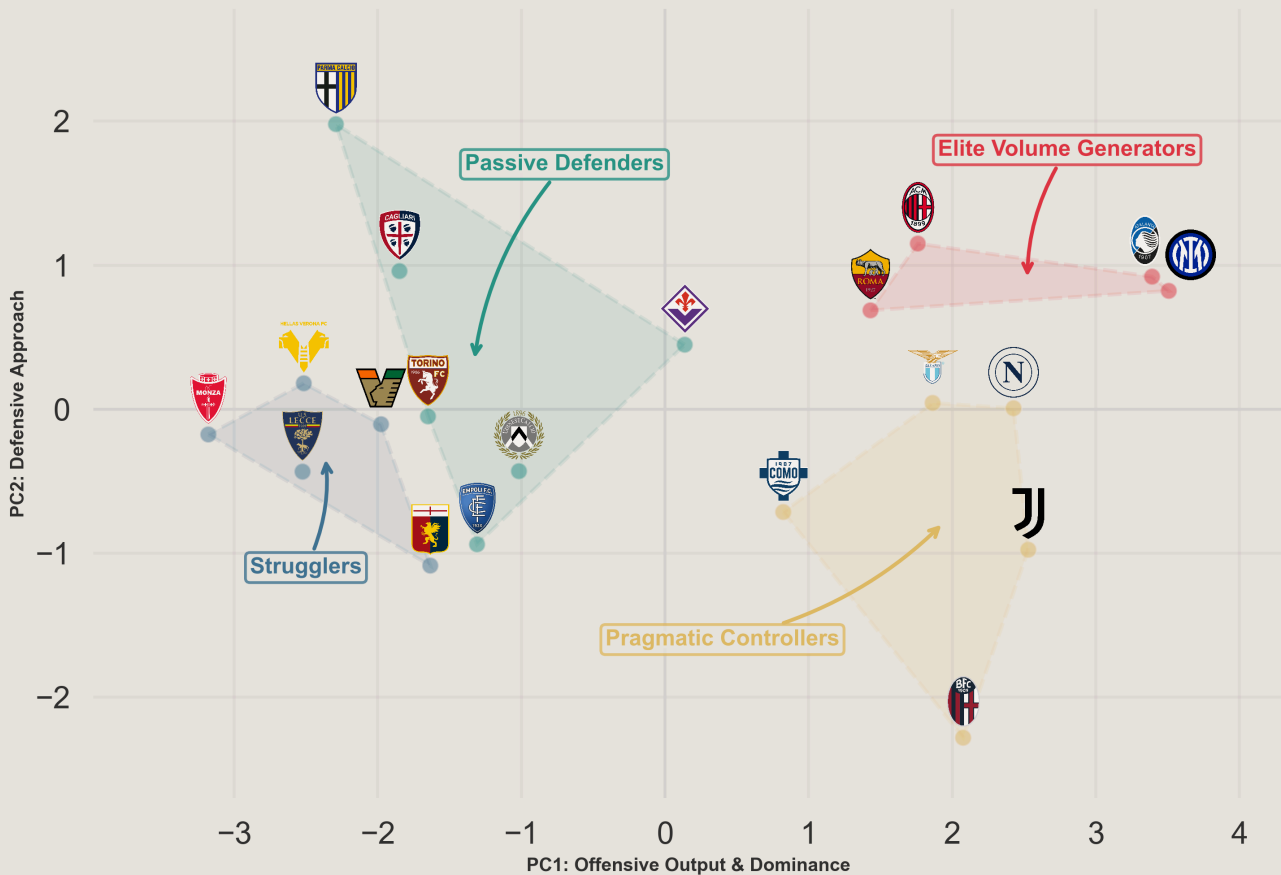


# THE PRESSURE GAME

Beyond Expected Goals:  
the hidden side of Serie A

## The Four Tactical Archetypes

PCA analysis groups teams by offensive output (xG & DC) and defensive approach (PPDA & xGA). Proximity reveals similar tactical identity.



# Research Questions

---

## THE CORE QUESTIONS DRIVING THIS ANALYSIS:

RQ1

### Tactical Identity

*"Do performance metrics reveal tactical clusters distinct from league standings?"*

RQ2

### Efficiency Analysis

*"How much does offensive/defensive efficiency impact final results compared to play volume?"*

RQ3

### Tactical Evolution

*"Can data-driven tactical adjustments reverse underperformance trends?"*

## WHY

---

Serie A is traditionally understood through a single lens: the league table. Yet modern football is shaped by multiple tactical philosophies, and results often reflect more than rankings alone.

This analysis examines whether league position might obscure the deeper reality that teams succeed through different strategic paths. By focusing on tactical identity, efficiency, and adaptability, it explores how data-driven evidence can refine, and at times challenge, conventional narratives about team performance and success.

# About Data

---

## Data Source

The analysis relies on advanced tactical metrics extracted from Understat, covering the 2024/25 Serie A season (RQ1 & RQ2) and early 2025/26 campaign (Milan case study).

## Primary Metrics

- **xG (Expected Goals):** Quality of chances created.
- **xGA (Expected Goals Against):** Quality of chances conceded.
- **DC (Deep Completions):** Completed passes within 20m of opponent's goal.
- **ODC (Opponent Deep Compl.):** Deep completions conceded.
- **PPDA:** Passes Allowed Per Defensive Action (pressing intensity).
- **xPTS:** Expected points based on xG and xGA.

## Data Quality & Cleaning

Raw export embeds differential values directly. Operations performed:

- **Regex Extraction:** Separated embedded differentials.
- **Type Conversion:** Converted string objects to floats.
- **Completeness:** Verified no missing values (20 teams).
- **Validation:** Cross-checked against published data.

# Methodology

## Collection & Preprocessing

Manual extraction from Understat (.xlsx) for 2024/25 and 25/26. Raw data cleaned to extract embedded differential values via Regex.

## Processing & Normalization

Utilized **Python** (Pandas, NumPy, Scikit-learn) for data processing and **Tableau** for preliminary exploration. Final visualizations generated via Matplotlib & Seaborn. All metrics normalized per 90 mins.

## Use of AI Tools

AI assisted with code generation (matplotlib logic) and debugging. Analytical decisions remain human-directed.

## Reproducibility

Source data publicly available via Understat; full analysis code, datasets, and transformations are documented and verifiable at: [🔗 View Repository](#)

## Analytical Techniques

- **Percentile Ranking:** Converts metrics to 0-100 scale.
- **PCA Analysis:** Reduces 6 tactical variables into 2 components.
- **K-Means Clustering:** Identifies 4 distinct tactical archetypes.
- **Z-Score Normalization:** Standardizes metrics.
- **Differential Analysis:** Calculates percentage change between seasons.

# Insights from the Data

---

## Four Tactical Identities

The league is not simply ranked by quality. Instead, PCA clustering reveals **4 distinct tactical archetypes**—from dominant controllers to passive low blocks. These clusters expose that winning is possible through multiple strategic paths, not just through superior dominance. Teams achieve results through different, sometimes opposite, tactical philosophies.

## Efficiency Over Volume

The gap between xPTS and Actual Points exposes a crucial truth: **creating chances matters less than converting them.** While xG establishes a baseline for team quality, actual outcomes depend on finishing efficiency and defensive discipline. Teams that systematically outperform their xG separate winners from mid-table teams through superior conversion and solidity.

## Tactical Shifts & Trends

Milan's transformation demonstrates that team identity is not fixed. By shifting from an aggressive, chaotic approach (24/25) to a compact, deeper defensive block (25/26), Milan restructured its profile. **Deliberate tactical adjustments produce measurable reversals**, proving that a data-driven redesign corrects underperformance trends when executed with absolute discipline and clarity.

PART 1

# TACTICAL IDENTITY

Mapping the League's DNA

# Performance Matrix: Quality vs Exposure

## 👑 The Dominators

**Inter & Atalanta.** The benchmark. They combine elite volume (**80+ xG**) with defensive solidity. Their balanced intensity allows them to dominate both boxes without exposing the backline to dangerous transitions.

## 🔴 High-Risk Entertainers

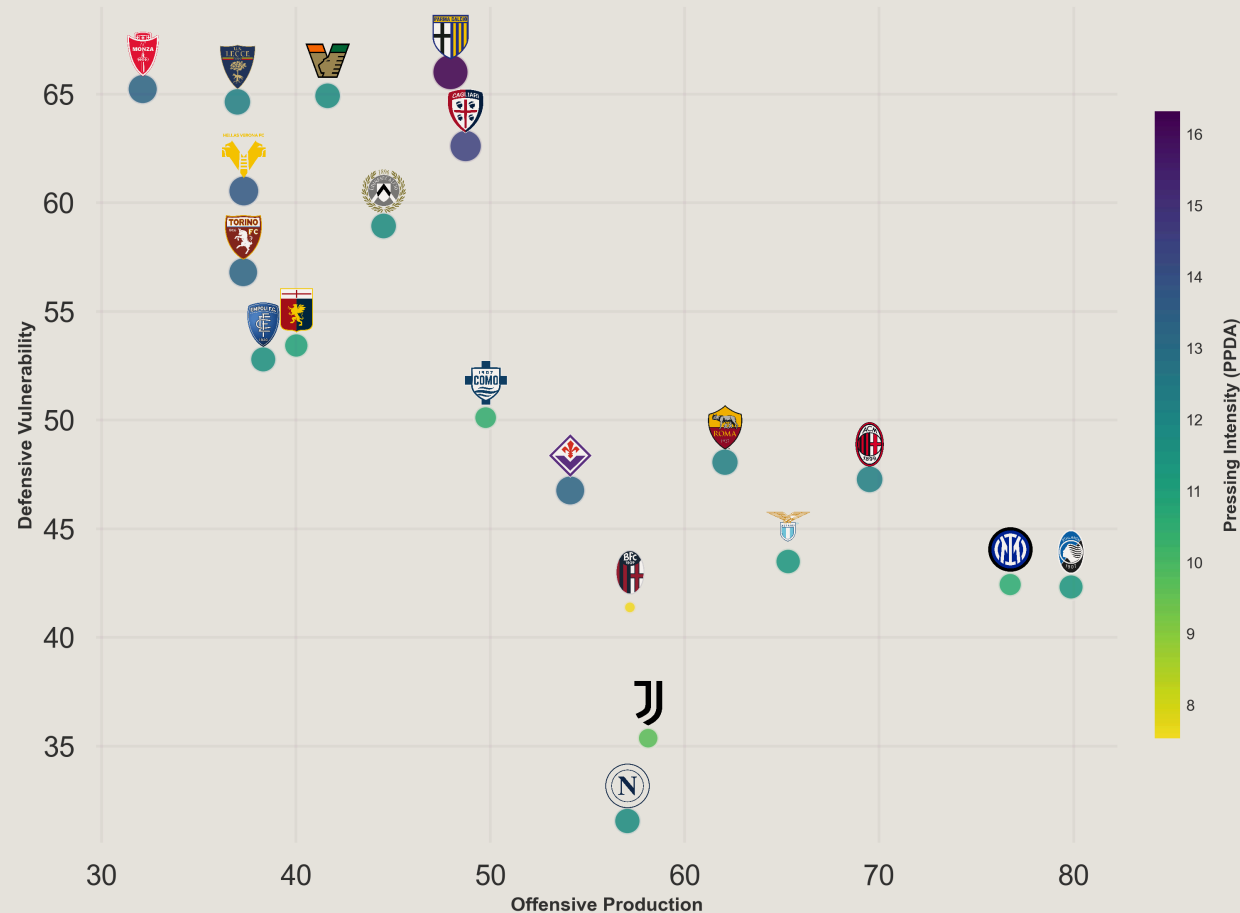
**Milan & Roma.** Elite creation (65+ xG) betrayed by structure. Their **Teal hue** suggests active intensity, yet it remains *chaotic*. This disorganization allows opponents to bypass the press easily, exposing the defense.

## 🛡️ The Iron Wall

**Juve & Napoli.** The defensive leaders (Bottom). Interestingly, they reach the same solidity via divergent paths: Juve suffocates play with extreme intensity, while Napoli relies on compact structural discipline.

## Performance Matrix

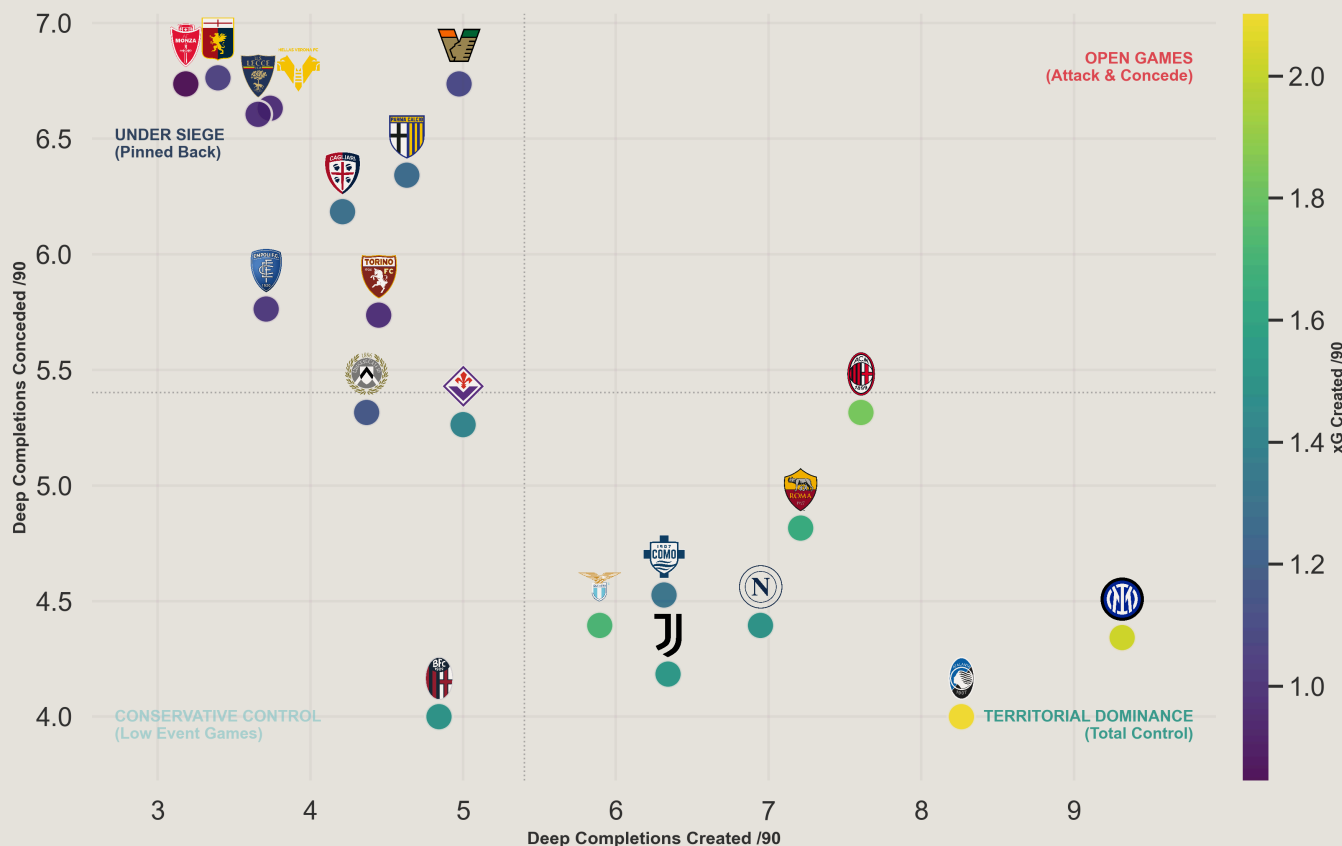
Offensive Quality (xG per season) vs. Defensive Risk (xGA per season). Color indicates Pressing Intensity (Lower value = Higher Aggression).



Data: Understat | Visualization: Lorenzo Zanotti

## Territorial Control Map

Deep Completions Created vs. Conceded. Color indicates Offensive Danger (xG: Yellow = High Threat).



Data: Understat | Visualization: Lorenzo Zanotti

# Territorial Dominance: Control vs Fragility

## ⚙️ The Elite Controllers

**Inter & Atalanta.** They dominate the bottom-right quadrant. Creating massive volume (>8 Deep Completions per match) while allowing minimal access to their own box. This is not just possession; it's **territorial supremacy**.

## ⚡ The Como Surprise

**Como's Ambition.** Unlike typical promoted teams that sit deep (Top-Left), Como plays positively (Center-Right). They create more danger than they concede, proving a proactive identity rarely seen in relegation fighters.

## 🛡️ Conservative Control

**Bologna.** Positioned in the "Low Event" zone (Bottom-Left). They concede very little due to possession defense, but struggle to penetrate deep. Safe, controlled, but often sterile.



# Tactical Clustering: The Four Identities

## Elite Dominators

**High Volume (>65 xG).** Groups "Complete Packages" (Inter/Atalanta) with fragile "Glass Cannons" (Milan/Roma). Same offensive ambition, opposite defensive discipline.

## Pragmatic Controllers

**Extreme Intensity (Low PPDA).** They transform moderate volume into elite results via suffocating pressure. Control is achieved through intensity, not possession.

## Passive Defenders

**Deep Block Strategy.** High PPDA indicates minimal pressing. They wait deep and absorb pressure. Survival depends entirely on clinical efficiency to punish opponents.

## The Strugglers

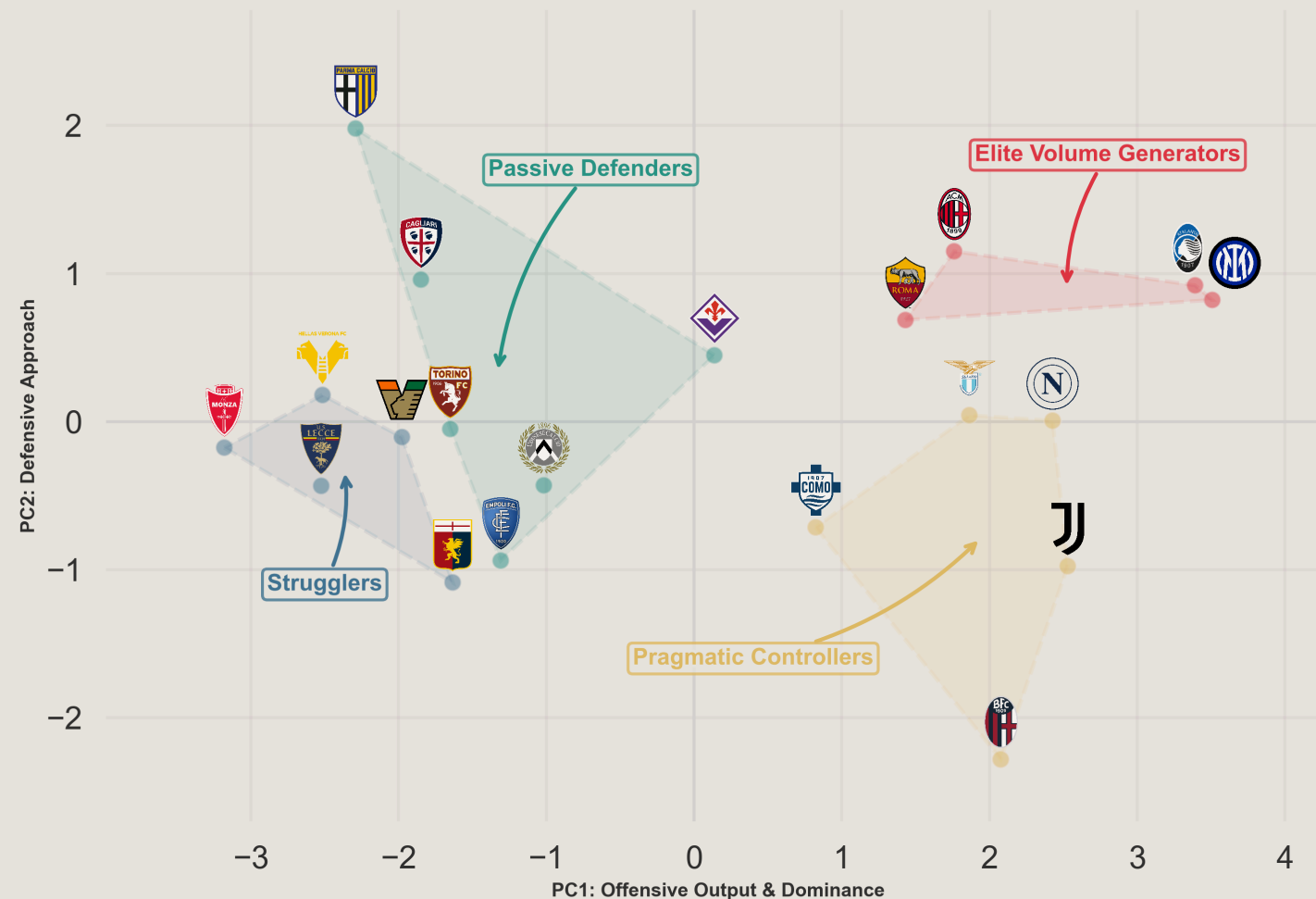
**The Danger Zone.** Insufficient output (<45 xG) plus structural weakness. A compounding spiral: unable to create danger, unable to prevent it.

### 💡 HIDDEN REALITY

"Do metrics reveal hidden identities? Yes. The PCA proves volume is a style, not a result. Milan and Inter share the Elite cluster, but structural stability separates efficient dominance from chaotic creation."

## The Four Tactical Archetypes

PCA analysis groups teams by offensive output (xG & DC) and defensive approach (PPDA & xGA). Proximity reveals similar tactical identity.



Data: Understat | Visualization: Lorenzo Zanotti

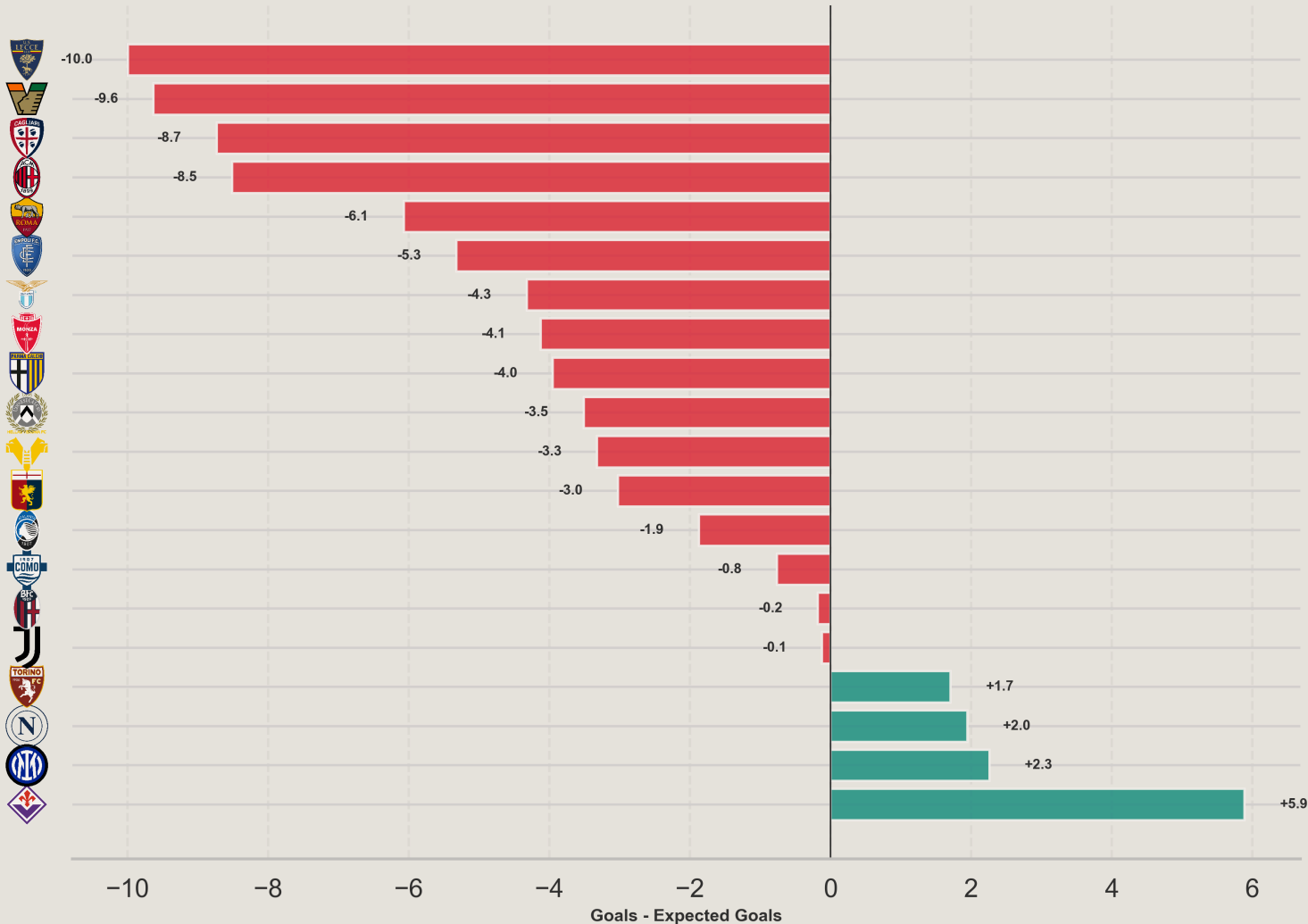
PART 2

# **EFFICIENCY ANALYSIS**

Volume is vanity, goals are sanity

## Clinicians vs Wasteful

Difference between Goals Scored and Expected Goals (xG). Positive = Overperformance. Negative = Underperformance.



Data: Understat | Visualization: Lorenzo Zanotti

## The xG Illusion: Conversion Reality

### 🎯 Clinical Precision

**Fiorentina (+5.9) & Inter (+2.3).**

Overperformance is a skill, not luck. These teams possess the **technical quality** to convert low-probability chances into goals consistently.

### 🗑️ Sterile Volume

**Milan (-8.5).** This is the critical failure. Milan creates at "Champion" level (Red Cluster) but finishes inefficiently. **Volume without precision** collapses the system.

### 📊 The Ranking Impact

A differential of -8.5 goals translates to **~6 lost league points**. This inefficiency metric explains why Milan sits outside the title race. **Elite volume requires elite finishing.**

# The Efficiency Verdict: Points vs Potential

## ↗ The Efficiency Multiplier

**Napoli (The Benchmark).** Situated well above the line, Napoli defies the algorithm. Despite creating *less* volume than Milan, they collected **significantly more points**. This proves that cynicism and defensive focus act as a "points multiplier."

## ↘ The "Wasted" Potential

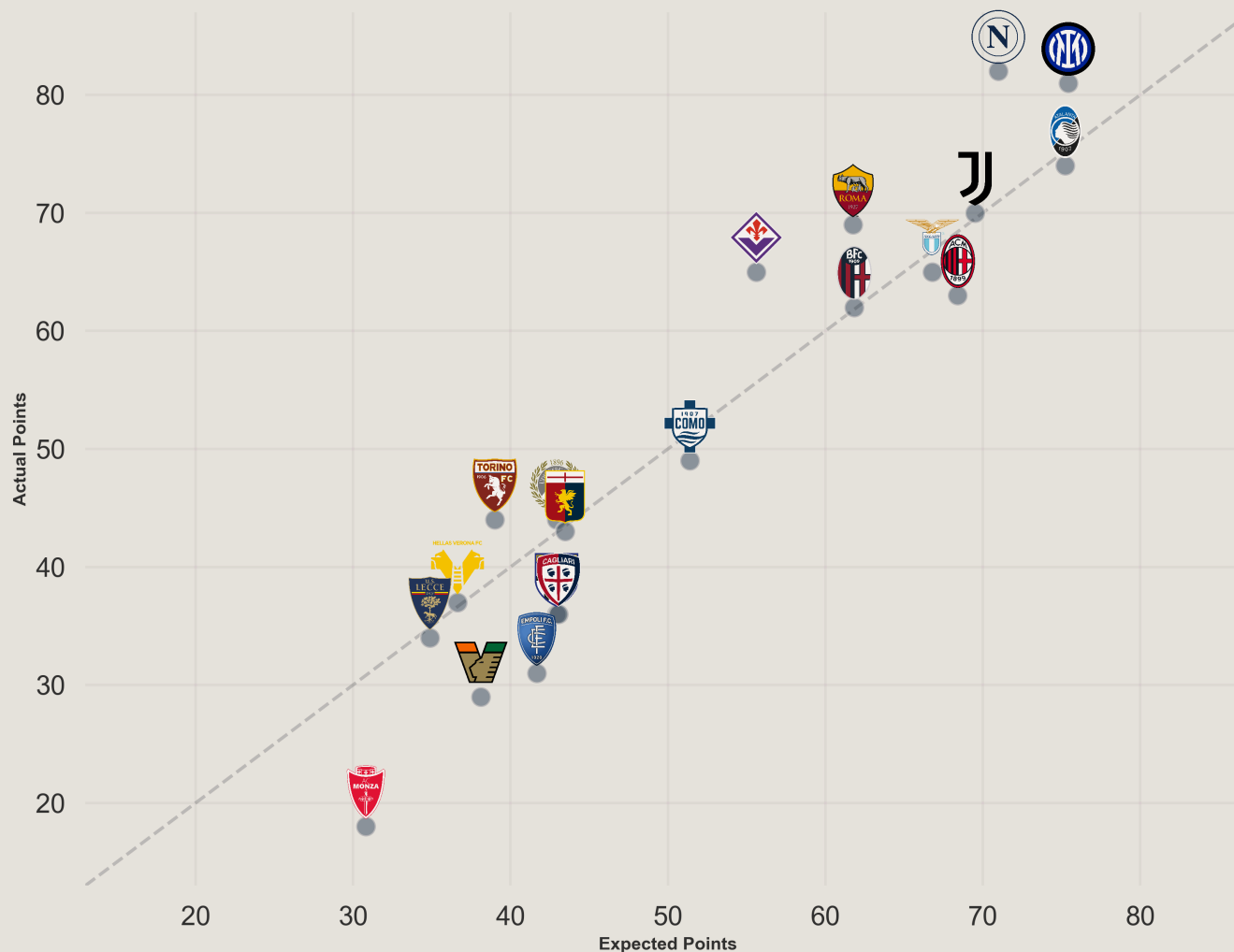
**Milan (The Failure).** The chart exposes a brutal reality. Milan generates "Title-Winning" metrics (High xPTS) but achieves "Mid-Table" results. Being so far below the diagonal confirms that their problem isn't tactics, but **execution and fragility**.

## ⚔ THE VERDICT

*"Napoli wins because they convert potential into reality. Milan loses because they leave points on the table. The distance from the diagonal is the true measure of a team's maturity."*

## Reality vs Expectation

Teams above the dashed line are overperforming their statistical projection. Efficiency acts as a point multiplier.



PART 3

# TACTICAL EVOLUTION

Case Study: AC Milan 24/25 → 25/26

# Context: Why AC Milan?

To answer the question *"Can data-driven adjustments reverse underperformance?"*, we need a team that underwent a **radical identity shift**.

Milan represents the perfect contradiction. Despite sitting in the **"Elite Cluster"** (RQ1), their 24/25 season failed due to extreme **inefficiency** (RQ2). This makes their tactical evolution the ideal case study.

*"Great data, poor reality. Milan's 24/25 season proves that volume without efficiency is meaningless. Their 25/26 evolution is the answer."*

## 01 THE STATISTICAL EXTREMES

Milan is the dataset's outlier. While ranking in **Cluster 1** for volume, they recorded the **most radical shift** in Pressing Intensity, explaining the disconnect between dominance and results.

## 02 ISOLATING VARIABLES

The **Efficiency Crisis (RQ2)** wasn't random bad luck. It correlates perfectly with their tactical metrics, allowing us to isolate the specific variables (High Line & Pressing) causing the broken system.

## 03 DEFINING "EVOLUTION"

Evolution isn't just "playing better". We track how **mathematically restructuring** the team's shape—from "Chaos" to "Solidity"—directly resolved the efficiency gap.

# The Trade-Off: Safety Over Chaos

## ⚡ The Strategic Shift

**Pressing Reset:** The **+35.4% increase in PPDA** confirms a deliberate tactical pivot. The team stopped chasing chaos and prioritized a compact, deeper defensive block.

## 🛡️ Defensive Stability

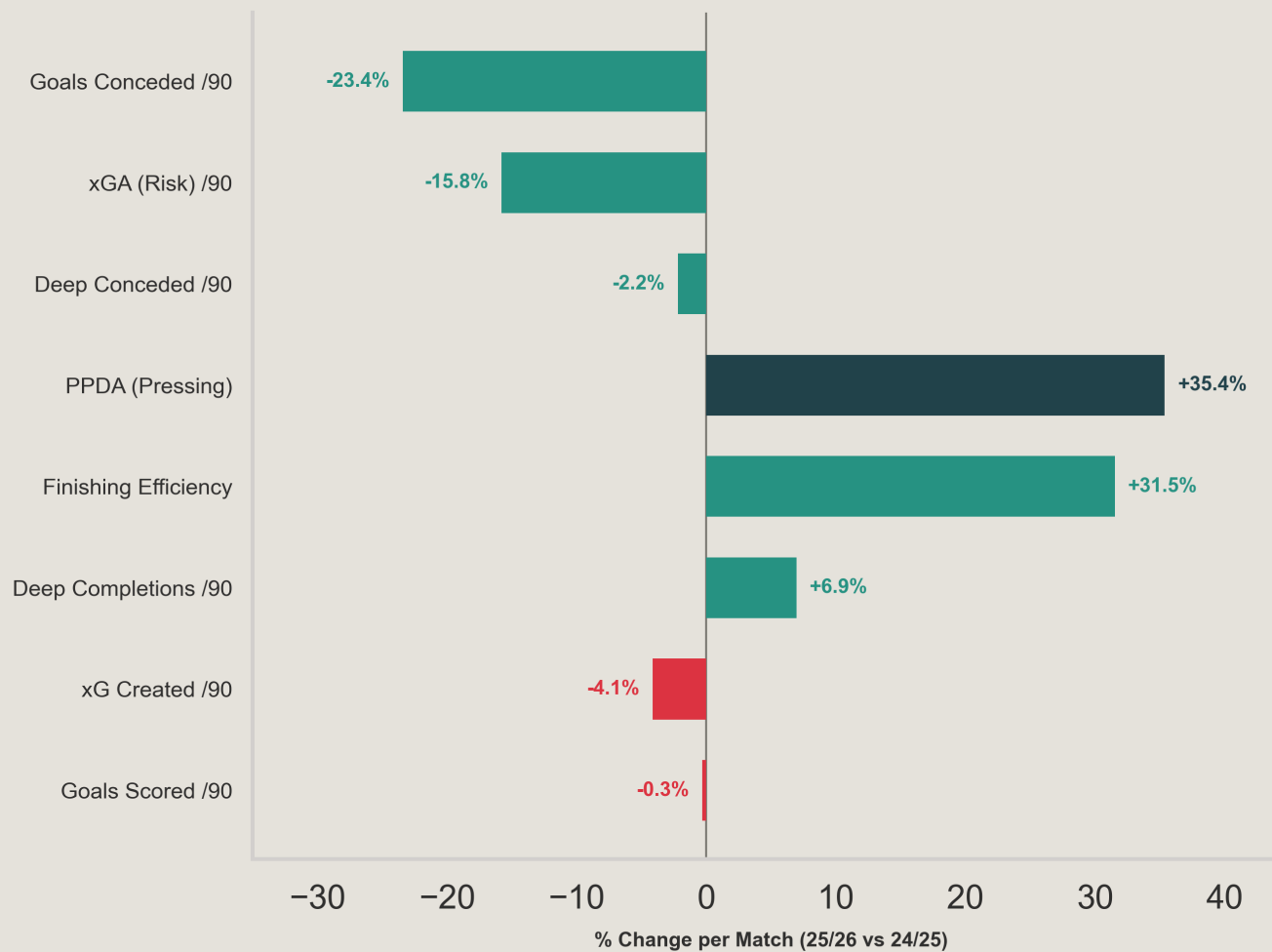
**Structural Gains:** It wasn't just luck. By lowering intensity, the team achieved a **23.4% reduction in Goals Conceded** and a **15.8% drop in xGA risk**.

## 🎯 Efficiency Surprise

**Quality over Quantity:** Despite creating slightly less volume (-4.1% xG), **Finishing Efficiency surged by +31.5%**. Fewer attacks, but with clearer minds.

## Milan's Trade-Off

Percentage change per 90 mins. What Milan sacrificed and gained.

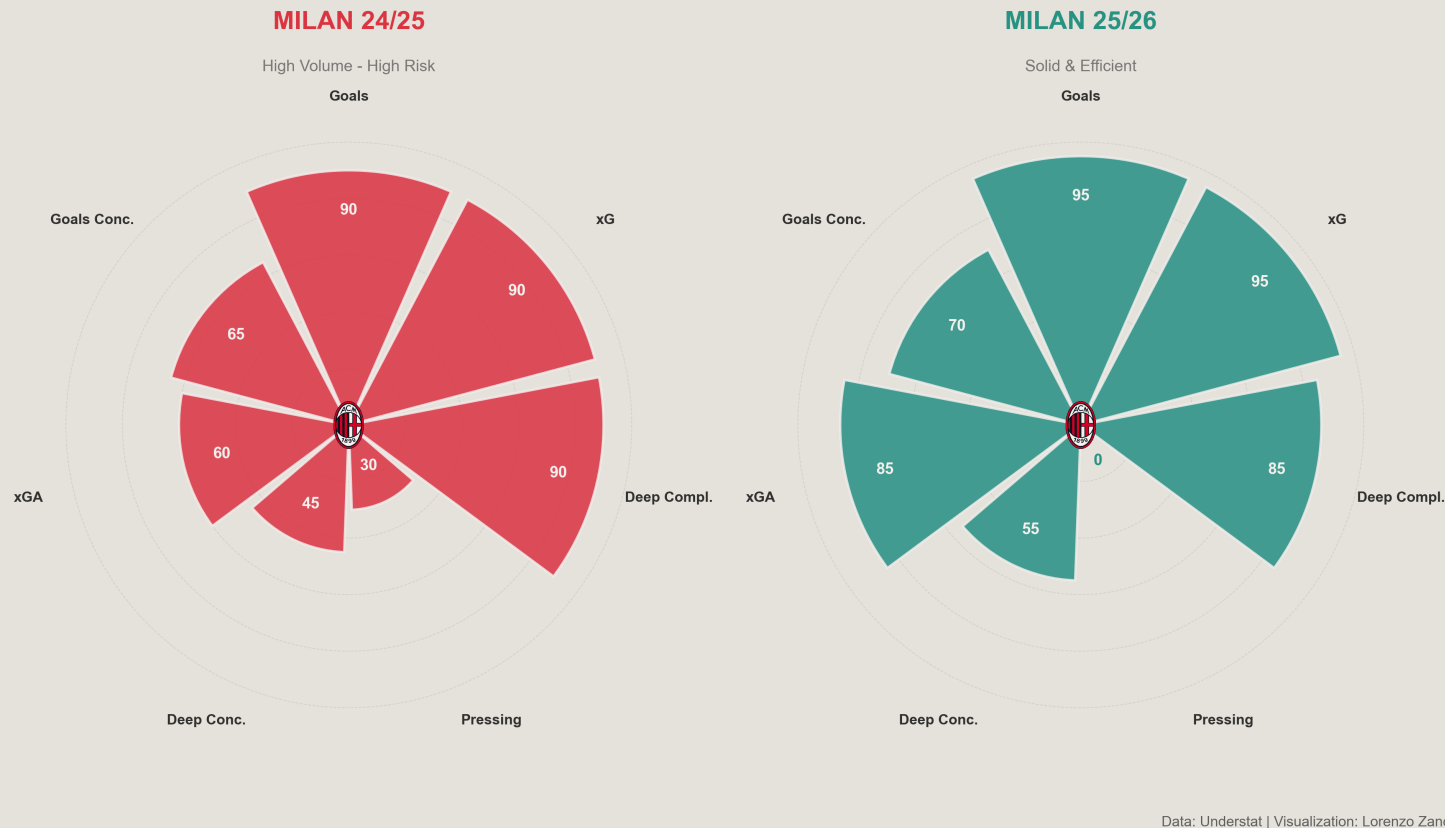


Data: Understat | Visualization: Lorenzo Zanotti

# The Transformation: Visualizing Balance

## The Structural Transformation

Percentile ranking vs Serie A average (0-100). How Milan restructured every dimension.



### ⚠️ 24/25: The "Broken" Profile

**Offensive Elite, Defensive Chaos.** The chart reveals a disconnected team: elite in Attack (90th percentile) but mediocre in Defense (60th). With **Goals Conceded** ranking only in the 65th percentile and chaotic pressing, the system had volume but no control.

**Verdict: Fun to watch, hard to defend.**

### 🛡️ 25/26: The Defensive Fortress

**Complete Evolution.** Attack remained elite, but defense transformed. **xGA solidity** hit the 85th percentile, **Deep Conceded** improved to the 55th, and **Pressing** dropped to the lowest rank. Chaos was replaced by compact structure.

**Verdict: Elite creation + elite solidity = Title-Ready.**



# | Conclusion: The Value of Adaptability

---

01

## THE DATA SHIFT

The metrics prove that volume is not the only path to victory. By accepting a **-4.1% drop in xG creation**, the team gained a massive **+31.5% in efficiency**. This confirms that quality shots and defensive confidence often outweigh pure offensive quantity in determining the final scoreline.

---

02

## TACTICAL PIVOT

The shift from "High Pressing" to "Mid-Block" was decisive. The data shows that solidity doesn't require parking the bus; it means **managing the spaces where the ball is won**. The team stopped chasing chaos and started controlling the transition moments.

---

03

## STRATEGIC LESSON

**Data-driven corrections work.** The case study demonstrates that identifying a "Broken Profile" early (High xG but High xGA) allows for targeted interventions that reverse underperformance trends within the same season context.

---

## 👁 FUTURE OUTLOOK: THE SUSTAINABILITY TEST

*"The tactical fix is proven, but can this intensity be maintained? The remaining challenge for Milan is transforming this newfound defensive stability into a consistent title run over 38 games, without losing their offensive spark."*

# | Key Findings & Takeaways



## Four Tactical Worlds

The league is not a simple ranking. Data proves the existence of **4 distinct archetypes**. High-pressing volume is just one strategic choice, not a universal law.



## Structure > Chaos

Pressing creates opportunities, but structure prevents disasters. The **"Glass Cannon"** profile (High Risk) statistically underperforms compared to balanced structures.



## The Efficiency Rule

Efficiency gaps explain **8-10 points** of variance. The Milan case proves that "Creating without Converting" is vanity. Efficiency is the true driver of the table.

### PROJECT CONCLUSION

**"While Pressing defines the style, Efficiency defines the winner."**

*This analysis reveals that high pressure is a powerful tool, but remains dangerous without control. The true masters of the game are not those who run the most, but those who successfully balance tactical risk with structural discipline and absolute clinical execution.*

# License

---

Slides are shared with the following license:

## **"THE PRESSURE GAME"**

by Lorenzo Zanotti,

is licensed under a

**Creative Commons Attribution–ShareAlike 4.0 International (CC BY-SA 4.0) .**