

Executive Summary: Sales Intelligence & Decision Engine

Candidate: Shajjad Shaikh Role: Data Science / Applied AI

Project Objective: To diagnose win rate decline and provide the CRO with a product-ready Decision Intelligence System that identifies funnel leaks, flags stalled deals, and prioritizes high-value revenue.

Strategic Overview:

This project transforms raw CRM data into an actionable Sales Insight & Alert System (SIAS). By moving beyond standard reporting, I have developed a diagnostic framework that identifies exactly where the sales funnel is "leaky" and which segments represent "sweet spots" for recovery.

Key Project Deliverables:

- Performance Baseline: Established a 45.3% organisational win rate and confirmed that the recent performance drop is a sustained trend requiring intervention.
- Custom Diagnostic Metrics: Developed the High-Value Win Rate (46.4%) and Segment Driver Score to help leadership protect strategic revenue rather than just chasing deal volume.
- Operational Decision Engine: Created an interpretable, rule-based system that scores open deals into Low, Medium, and High-Risk bands based on segment win probabilities.
- Actionable Strategy: Pinpointed specific high-performing "Sweet Spots"—specifically FinTech in North America via Inbound leads—which achieve a win rate of 61%, outperforming the baseline by ~16 percentage points.
- Productized Alerting: Designed a system that automatically flags "stalled" deals (those 15 days beyond the median sales cycle) to drive immediate coaching.

Part 1: Problem Framing

1. The Real Business Problem: The organisation suffers from funnel leakage where healthy pipeline volume masks late-stage execution failures. Leadership lacks the visibility to distinguish high-probability "sweet spot" deals from high-risk "stalled" deals.

2. Key AI System Questions:

- In which specific segments is win rate dropping most?
- At which specific stages is the funnel leaking?

- What early signals (age, stage, segment) indicate that an open deal is at high risk?

3. Core Diagnostic Metrics: Monthly win rate trends, win rate by segment (industry/region), stage-to-stage conversion rates, and sales cycle duration by outcome.

4. Key Assumptions: CRM data is reliable and consistent; the internal sales process remained stable; and external macro-factors are not explicitly modelled.

Part 2: Data Exploration & Insights

- **Custom Metric 1:** Monthly Win Rate Trend: Tracks the share of Won deals closed in each calendar month.
 - **Why it matters:** Directly addresses the CRO's concern by distinguishing between a short-term blip and a sustained decline.

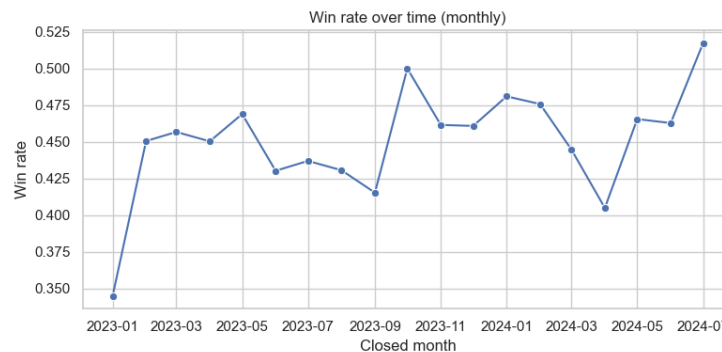


Chart 1. Monthly Win Rate Trend

- **Custom Metric 2:** High-Value Win Rate: Tracks win rates specifically for the top 25% of deals by ACV.
 - **Status:** Currently 46.4%, indicating the team is successfully protecting strategic revenue despite overall volume drops.

Insight 1: Funnel Bottlenecks in Late-Stage Execution

- **Analysis:** While early stages convert well, the win rate drops during Proposal and Qualified.
- **Why it matters:** This suggests the problem is late-stage execution (e.g., pricing pressure or lack of champions) rather than top-of-funnel fit.

- **Action:** Introduce mandatory late-stage deal reviews and pricing guardrails.

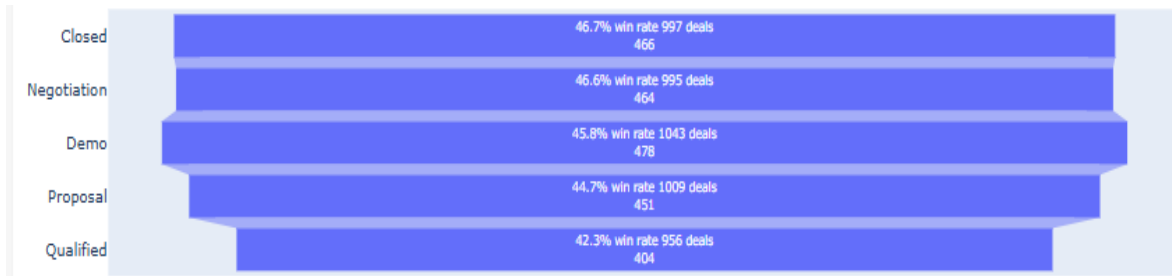


Chart 2. Win Rate by Deal Stage

Insight 2: Industry-Specific "Sweet Spots"

- **Analysis:** FinTech and SaaS run comfortably above the baseline, while other sectors lag.
- **Why it matters:** These represent strong product-market fit and the safest bets for near-term growth.
- **Action:** Double-down on high-win industries by shifting marketing spend and SDR focus.

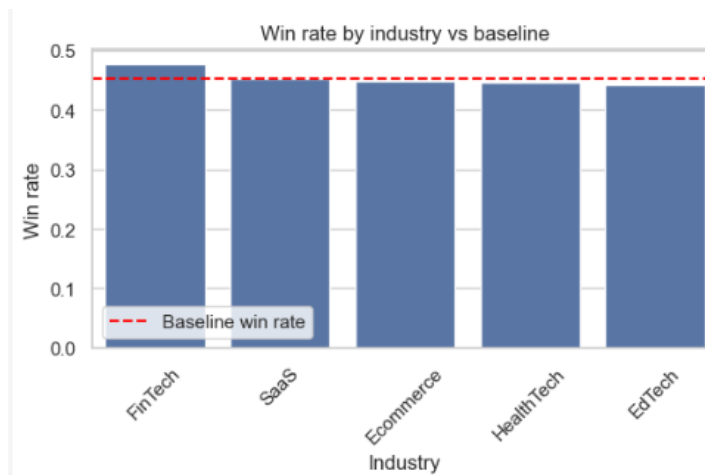


Chart 3. Industry vs Win rate

Insight 3: Lead Source Efficiency (Quality vs. Noise)

- **Analysis:** Inbound and Referral leads show a clear win-rate lift over Partner and Outbound channels.
- **Why it matters:** Identifies which channels produce high-quality opportunities versus those that create "noise".

- **Action:** Tighten qualification rules for lower-performing sources and reallocate budget to Inbound.

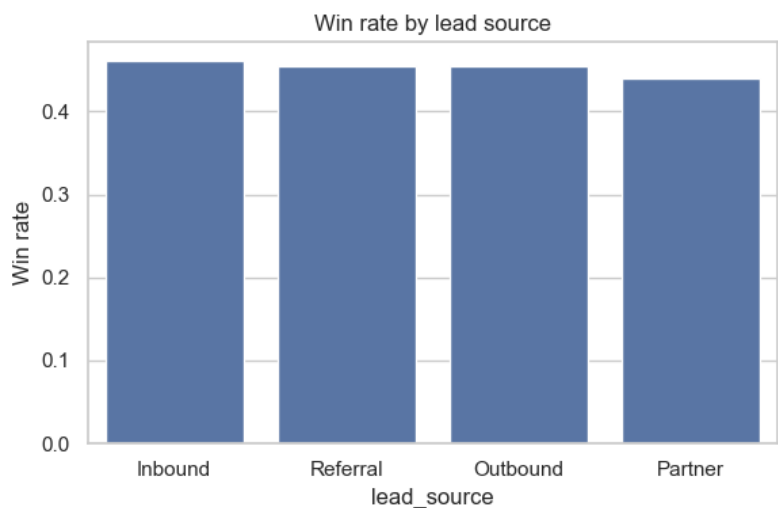


Chart 4. Lead source vs Win rate

Part 3: Win Rate Decision Engine

I built an interpretable, rule-based engine that maps historical segment performance onto current open deals.

Rule-Based Risk Banding: Deals are joined to their segment win probability (seg_win_prob) and categorised for immediate action:

Risk Band	Segment Win Prob.	Recommended Action
Low Risk / High Fit	≥ 60%	"Sweet Spot" deals; follow standard closing playbooks
Medium Risk	45% – 60%	Monitor closely; review rep alignment
High Risk / Low Fit	< 45%	Leaky segment; requires immediate manager review

Table 1. Risk Band and Recommended actions

Actionable Discovery: The engine identified FinTech / North America / Inbound as a premier segment with a 61%-win rate.

Part 4: Mini System Design (The SkyGeni Product Vision)

Designed as a lightweight, productized architecture for SkyGeni to automate revenue intelligence.

High-Level Architecture & Data Flow

- 1. Ingestion:** Daily scheduled ETL pulls raw deals, stages, and outcomes from the CRM into an analytics warehouse.
- 2. Insight Engine:** A batch job re-computes Driver Scores and Risk Bands every 24 hours.
- 3. Alert Layer:** Structured insights are converted into human-readable notifications for Slack/Email.

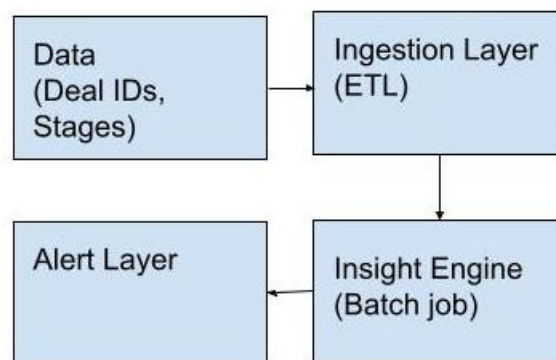


Diagram 1. High Level Architecture

Example Alerts

- Emerging Driver Alert: "FinTech / North America / Inbound shows win rate 16 p.p. above baseline (0.61 vs 0.45). Prioritize marketing here".
- Stalled Pipeline Alert: "HealthTech at stage Qualified has stalled deals worth ~\$913k. These deals are 15 days beyond the median sales cycle".
- High-Risk Strategic Alert: "Deal D00005 (HealthTech / APAC / Qualified) is high-value but sits in the High-Risk band (Prob: 42%). Manager review recommended".

Frequency: Runs daily (overnight) to ensure fresh data for morning sales huddles.

Failure Cases:

1. Data Quality: Missing closed dates or inconsistent stage updates will distort win rates.
2. Sparse Segments: Segments with few deals produce "noisy" probabilities.
3. GTM Changes: New pricing or ICP shifts may make historical data less predictive.

Part 5: Reflection & Engineering Evaluation

1. Weakest Assumptions

The most vulnerable assumption is that historical win-rate patterns remain stable; in reality, macro shifts or competitor moves can quickly render segment-level averages obsolete. Furthermore, the system assumes that real CRM data is as clean and complete as this synthetic dataset, ignoring common issues like missing outcomes or inconsistent stage definitions.

2. Production Breaking Points

In a live environment, schema instability (e.g., a RevOps team renaming a lead source) would break the grouping logic, resulting in misleading or null outputs. Additionally, sparse segments with low deal volumes create "noisy" probabilities, which would trigger erroneous alerts. The current design also lacks the enterprise-grade access control required for a secure product.

3. One-Month Roadmap

1. Data Hardening: Implement automated validation and minimum volume thresholds to suppress alerts from tiny sample sizes.
2. Productization: Transition from notebooks to a scheduled service (e.g., Airflow) with a CRM-embedded dashboard.
3. Feedback Loop: Enable sales leaders to "up-vote" or "down-vote" alerts to automatically tune thresholds and improve signal accuracy over time.

4. Least Confident Areas

I am least confident in the hand-picked alert thresholds, such as the 15-day stalling margin, which were tuned specifically to this dataset without user feedback. I am also cautious about the coarseness of risk bands; qualitative labels (Low/Medium/High) simplify decision-making but lose the granular nuance needed for high-stakes territory adjustments.