



Rakamin

The Promotion Paradox

Fixing Biased Talent Promotion Decisions Through Data-Driven HR Analytics
a project report by Syntax Society



Interactive Dashboard

Our team



**Keisya
Nisrina Aulia**
Data Scientist



**Febiansyah Annaufal
Ahnaf Fauzi**
Data Scientist



**Bayu
Maitra**
Project Manager



**Dian Ulhaq
Qurrata A'yun**
Data Analyst



**Mashlahatul
Husna**
Data Engineer

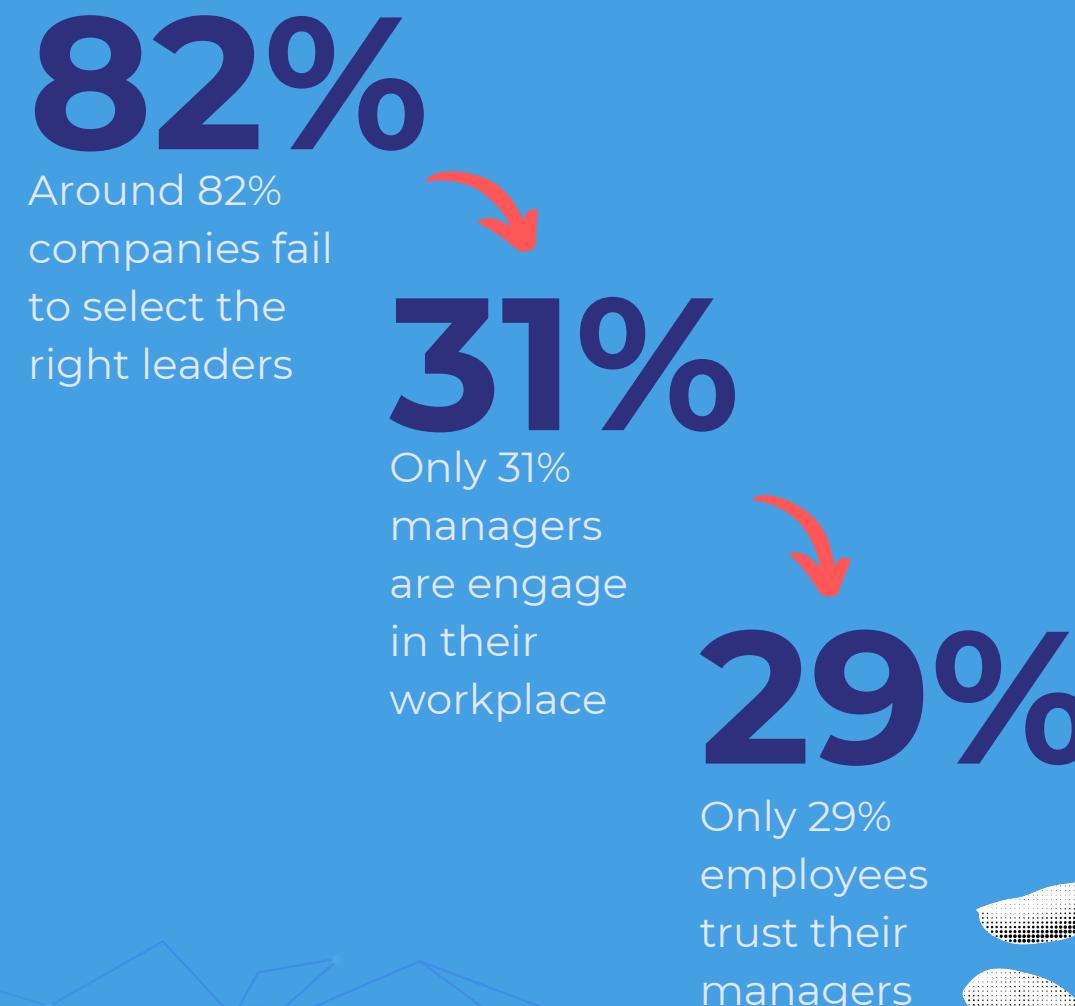
Table of Content

- | | | | |
|-----------|--------------------------|-----------|-------------------------|
| 03 | The Promotion Paradox | 16 | Machine learning & eval |
| 06 | Data anomalies & biasses | 26 | Github to dashboard |
| 11 | Cleaning & new features | 32 | Impact & action items |

Stage 1: Background.

The promotion paradox

Talent promotion is meant to fuel long-term organizational growth. When the system works, it strengthens efficiency, boosts retention, and builds trust across teams. Yet many organizations struggle to identify the right leaders. Chartered Management Institute (CMI) data shows low managerial engagement, weak trust, and widespread mismatches in leadership roles. This is not just a global issue. It also appears in Indonesia, including within Rakamin.



From bias to balance

The project aims to transform Rakamin's promotion system from a subjective, manual process into a fair, data-driven decision engine. By integrating machine learning with a real-time dashboard, HR can evaluate talent more accurately, reduce bias, and ensure leadership decisions truly support organizational growth.

Goal:

Develop a data-driven decision-support system that improves HR efficiency in monitoring talent performance and making fair, consistent promotion decisions.

Objectives:

- Build and validate ML models for promotion eligibility and employee characteristics.
- Integrate the models into a real-time insights interactive dashboard.
- Set up continuous monitoring to track system accuracy and impact.

Success Metrics

>95%

standardized and validated HR data.

>90%

model accuracy for promotion eligibility and employee characteristics.

24/7

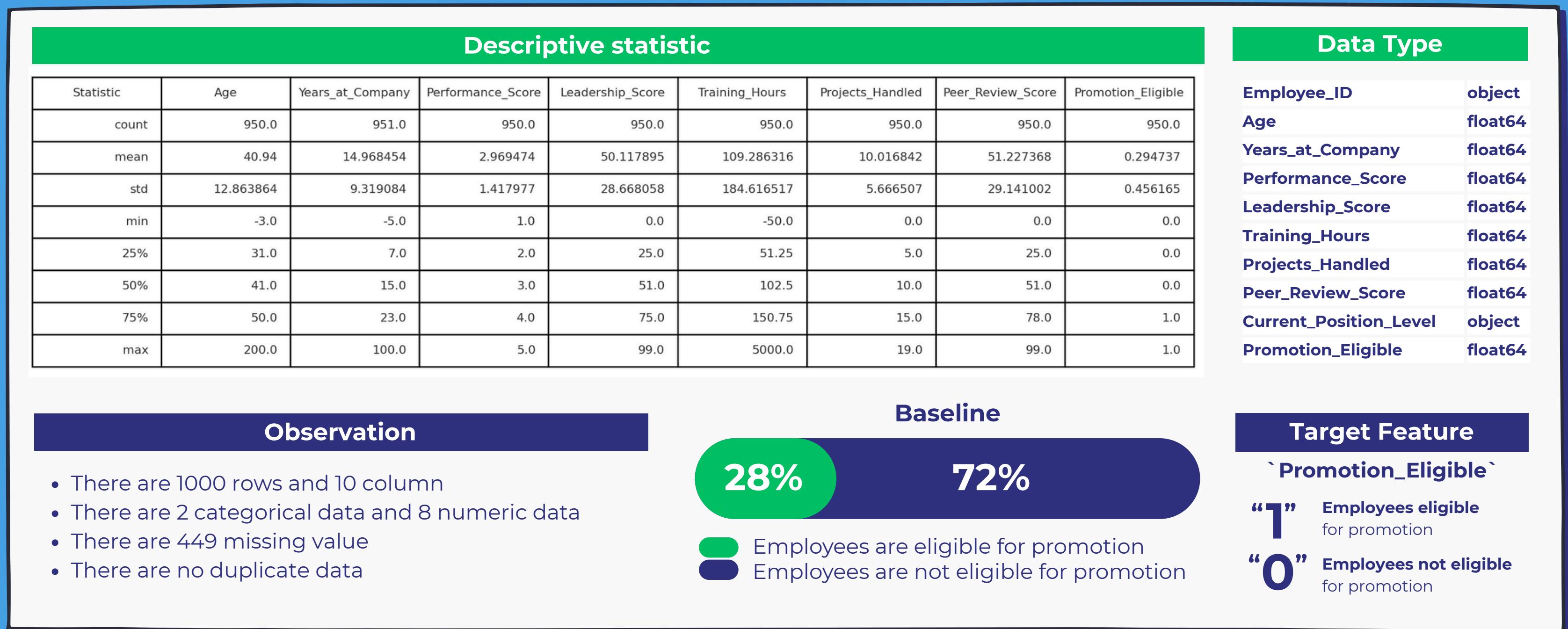
Real-time, always-on dashboard access.



Stage 2: Data analysis.

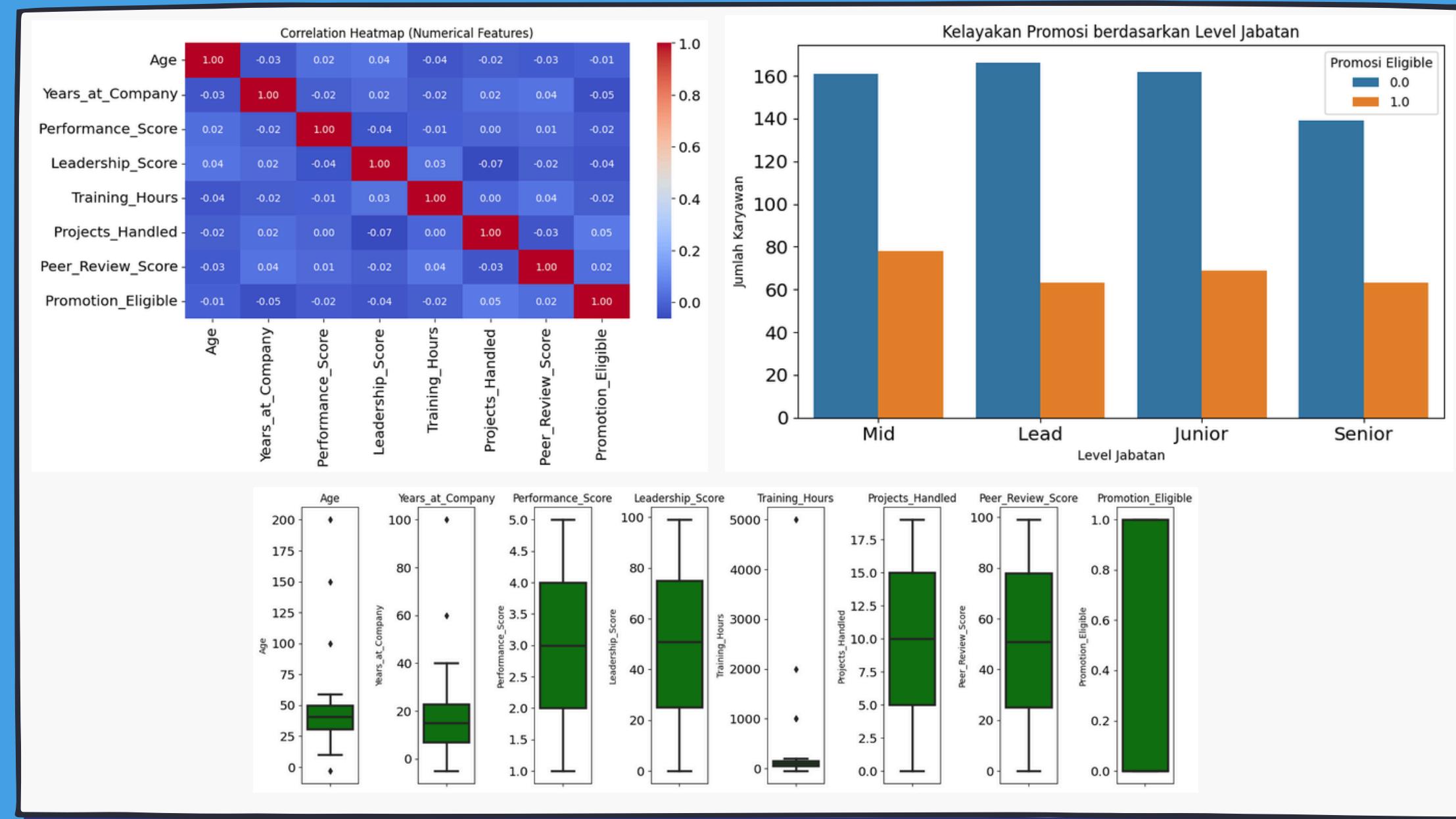
Dataset characteristic

Initial description of Rakamin's dataset: 1000 rows and 10 columns, 2 categorical and 8 numerical features, 449 missing values. We also have Promotion_Eligible as target feature that are imbalance.



Distribution & correlation

The data shows several **irregularities** in distribution, including outliers that disrupt normal patterns and indicate inconsistencies in data entry.



- About 1.3% of observations are outliers, contributing to skewed distributions and data quality issues.
- Promotion eligibility is low across all job levels, indicating system-level barriers rather than individual-level differences.
- Projects Handled and Peer Review Score show the highest correlation with promotion, but the associations are weak and not statistically meaningful.
- The weak correlations suggest that **current promotion criteria may not effectively capture true performance** or potential.



Data anomalies

The dataset contains several critical anomalies that affect data quality and reliability. These issues range from impossible values, such as negative ages, tenure exceeding age, or extreme training hours, to inconsistencies in leadership scores and job levels. Such anomalies indicate manual entry errors and undermine the validity of any HR decision-making process unless properly cleaned and standardised.

Employee_Id	Training_Hours	
EMP0345	-50	
EMP0850	999	
EMP0379	1000	
EMP0394	2000	
EMP0926	5000	
Employee_Id	Leadership_Score	Current_Position_Level
EMP0016	1.0	Lead
EMP0153	1.0	Lead
EMP0366	0.0	Lead
EMP0488	0.0	Lead
EMP0608	1.0	Lead
Employee_Id	Age	Years_at_Company
EMP0049	-3	20
EMP0036	22	24
EMP0061	23	29
EMP0095	23	29
EMP0117	23	27
EMP0137	24	26
EMP0164	24	28
EMP0097	26	27

Implausible numbers of Training_Hours

Negative value in Age, also a lot of Years at Company that is bigger than Age.

Talents with extremely low Leadership_Score becomes Lead.

Biased target feature

The data shows that promotion decisions rely too heavily on Projects Handled and Peer Review Score, even though both have weak correlations with actual promotion outcomes. These inconsistencies, where strong performers are labeled ineligible and weaker ones are promoted, indicate that the **Promotion_Eligible target itself is unreliable** and does not reflect real employee capability.

Employee_ID	Projects_Handled	Peer_Review_Score	Promotion_Eligible
EMP0050	19	98	ya
EMP0324	19	94	tidak

Employee_ID	Performance_Score	Leadership_Score	Peer_Review_Score	Promotion_Eligible
EMP0939	1	12	4	ya
EMP0939	5	96	90	tidak

At least 418 row has bias related to target feature

Two employee has similar performance, but one is eligible and the other is not.

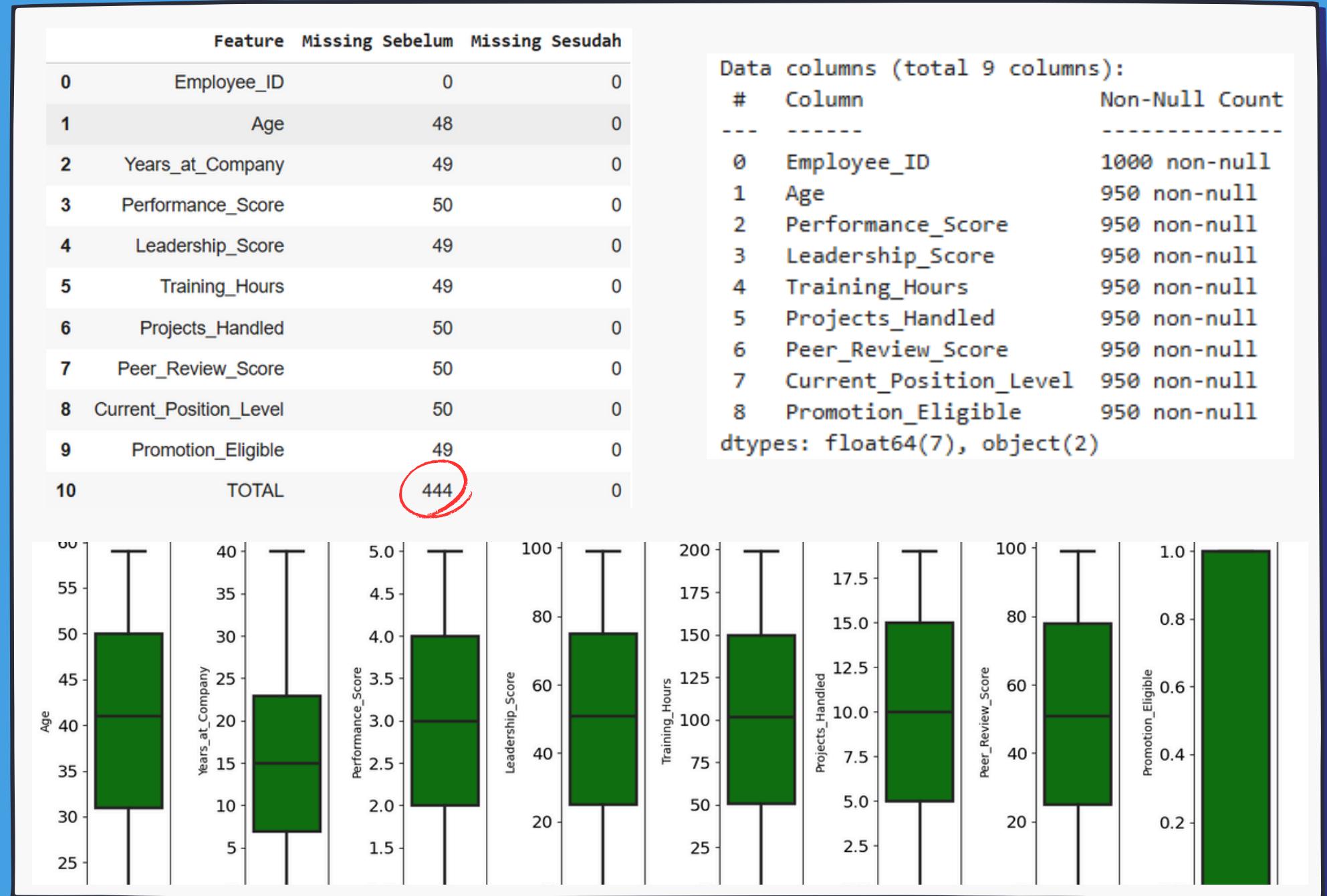
The less qualified employee is the one who is eligible for promotion.

Stage 3: Data preprocessing.

Data cleaning

To ensure the dataset was reliable for modeling, we cleaned missing values and removed invalid outliers. Numeric features were imputed using the median and categorical features using the mode, resulting in a fully complete dataset.

We also removed extreme outliers, particularly in Age and Training Hours, to restore a more realistic distribution and prevent distortion in the machine learning results.



Feature selection

We selected features based on data reliability and their relevance to employee performance. Features with implausible values, weak conceptual links, or no predictive contribution were removed. We also identified that the **Promotion_Eligible target is biased and inconsistent, meaning it might cannot be used as a trustworthy ground truth** for modeling. However, we still use it for exploratory purposes.

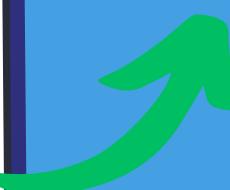


DROPPED

- Employee_Id: Has no effect
- Age : Not reflecting employee quality
- Current_Position_Level: Does not reflect quality
- Years_at_Company: 295 (30%) of data are implausible
- Promotion_Eligible: 418 (40%) Biased and inconsistent

USED

- Performance_Score: Employee performance indicators
- Training_Score: Employee knowledge indicators
- Peer_Review_Score: Indicators of teamwork and skills.
- Project_Handled: Indicators of work experience.
- Leadership_Score: Leadership quality indicators.
- Promotion_Eligible: target feature.



Although their correlations are very low, these features are retained because they may still provide useful information once processed through feature engineering.

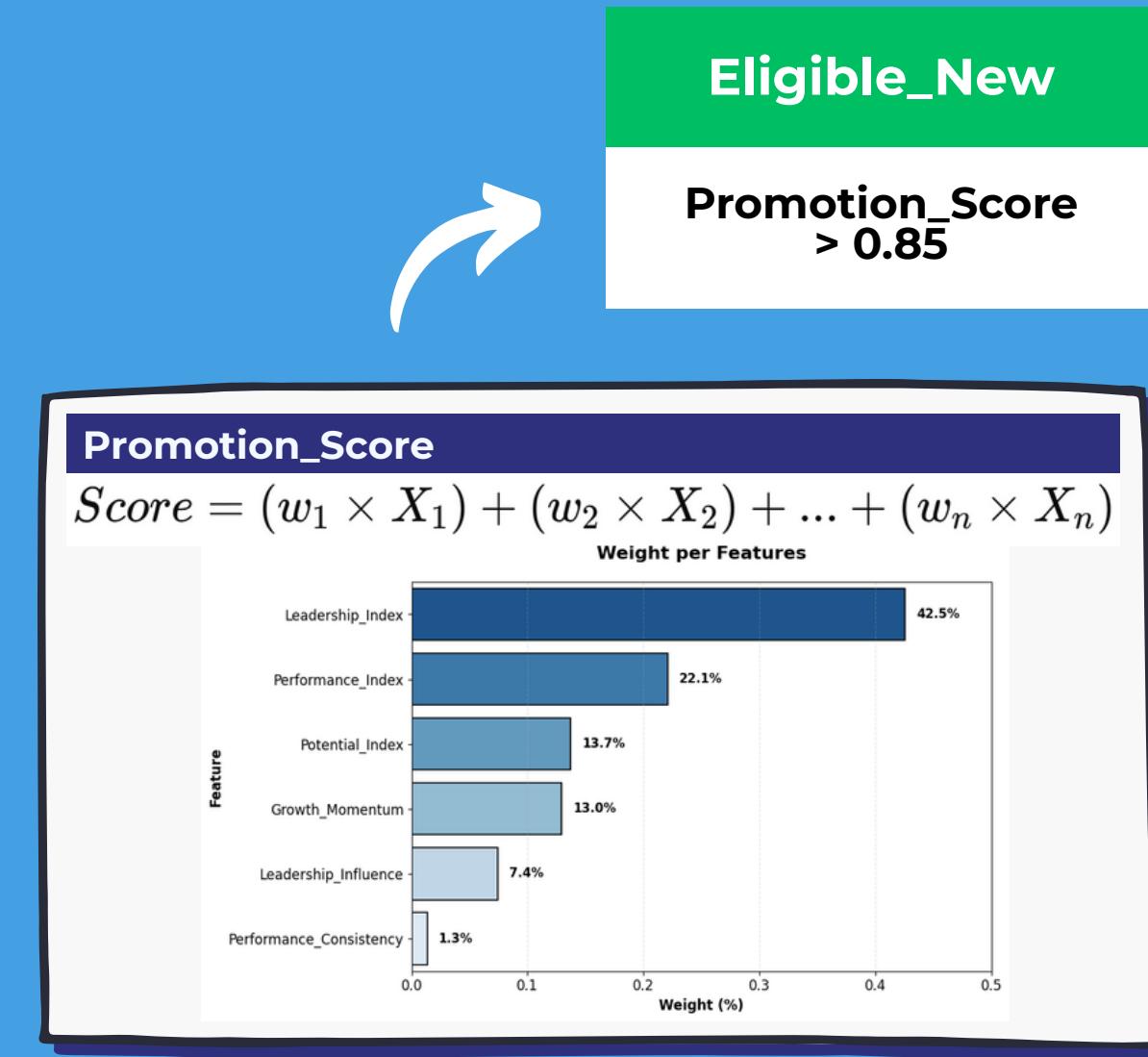
Feature engineering

We engineered six new composite features from the original data to better capture employee potential and performance. These features provide more reliable signals for our clustering and promotion logic than raw metrics alone. Details on how we use these to calculate **Promotion_Score** and **Eligible_New** will be covered in the modeling stage.

6 New features

We engineered 6 new features to capture deeper signals of performance, leadership, growth, and consistency, providing more reliable inputs for evaluating promotion readiness than the original raw features.

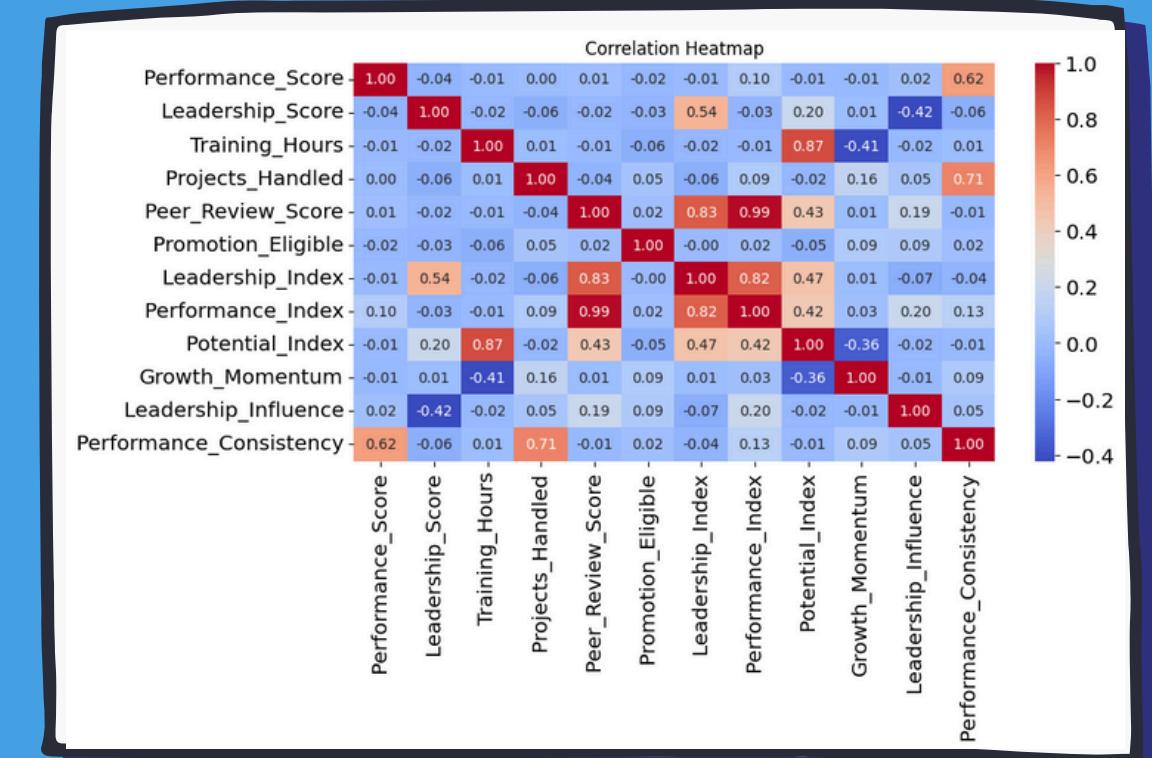
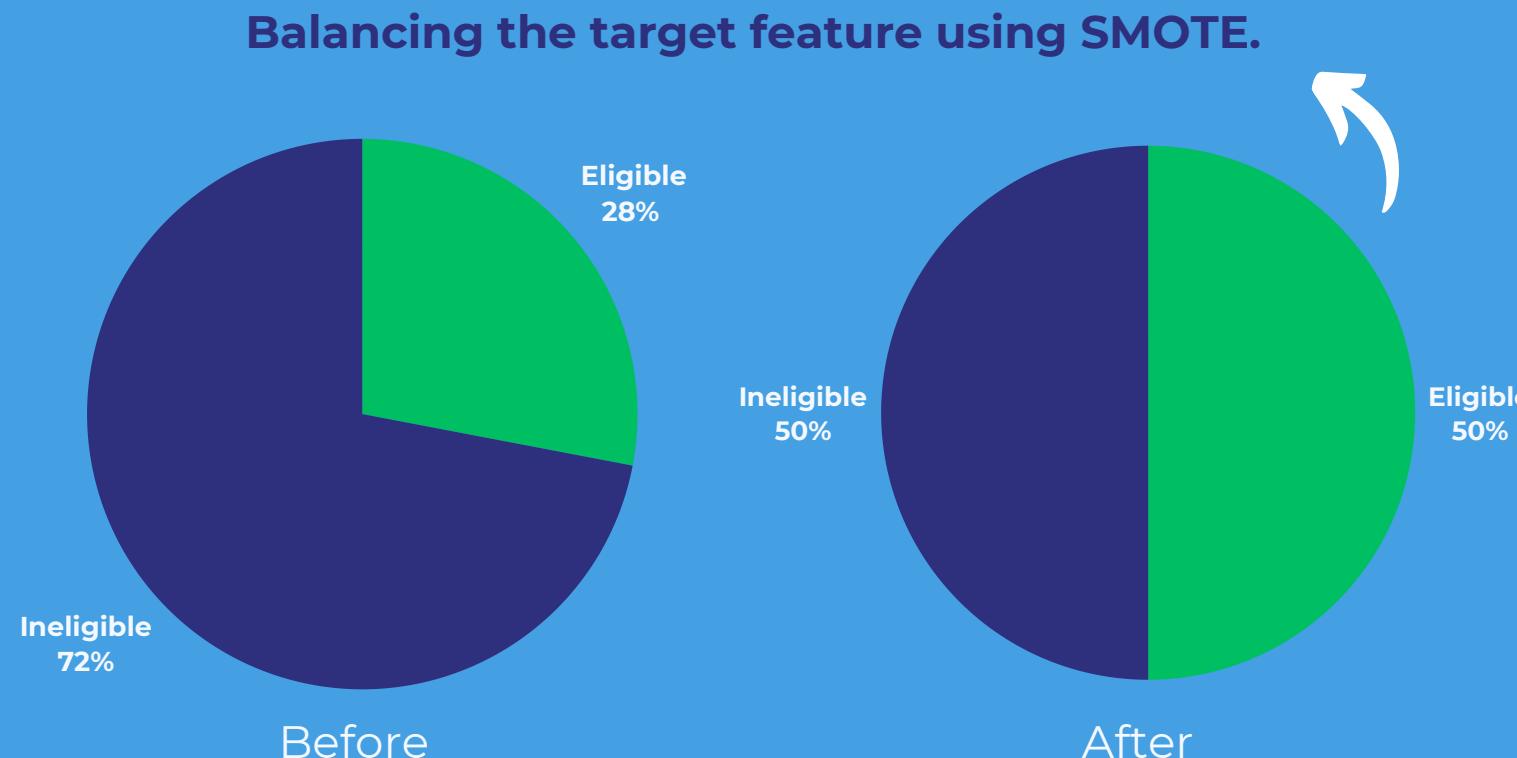
- **Leadership_Index:** Combines leadership and peer review signals to reflect influence and team perception.
- **Performance_Index:** Aggregates performance, experience, and training into a single quality indicator.
- **Potential_Index:** Measures motivation and perceived capability using behavioral and leadership factors.
- **Growth_Momentum:** Captures how quickly an employee converts training into tangible output.
- **Leadership_Influence:** Indicates how colleagues view and trust the employee in collaborative settings.
- **Performance_Consistency:** Highlights sustained performance and reliability over time.



Data standarization

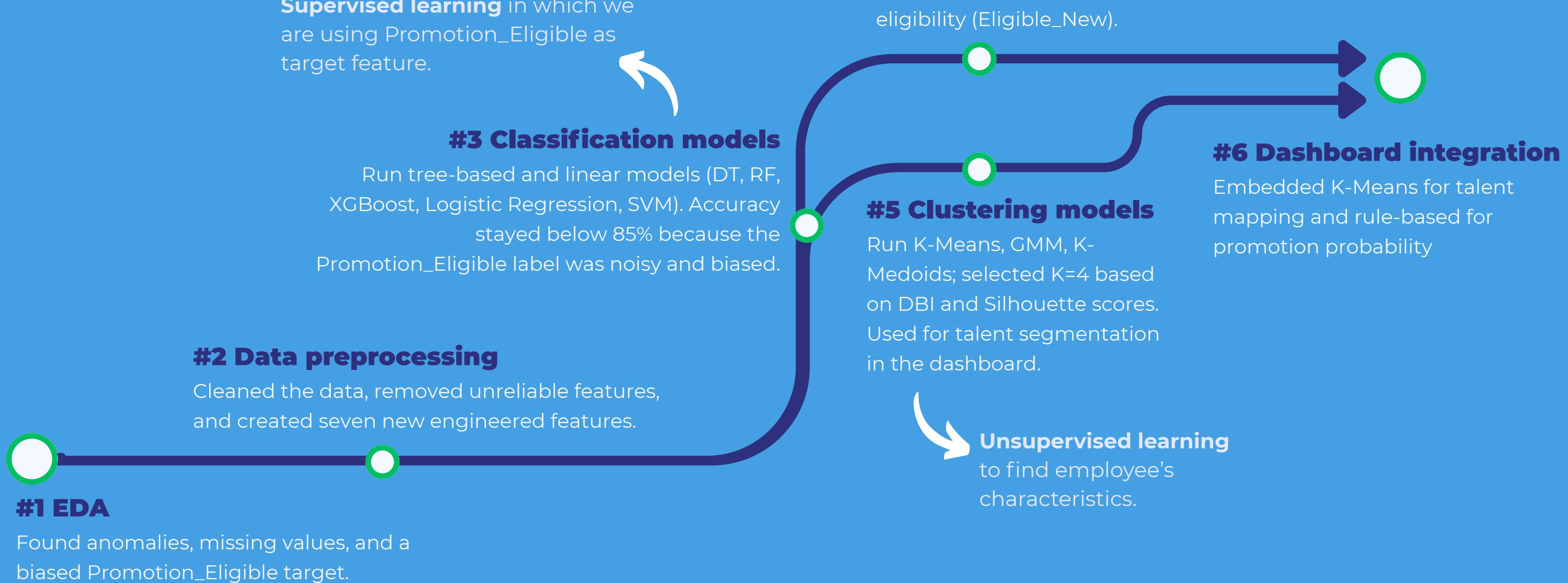
Performance_Score	Leadership_Score	Training_Hours	Projects_Handled	Peer_Review_Score	Promotion_Eligible	Leadership_Index	Performance_Index
-0.7010183561	-0.220183522	-1.03597658	1.627891733	0.3490112651	0	0.1723933596	0.4992035113
1.465699131	-0.1843396928	-0.6100273461	0.1782658027	-1.591787396	0	-1.443854664	-1.437583173
1.465699131	-1.116279251	-0.07759080319	-0.7277504037	1.231192475	0	0.4202841608	1.24591645
0.7434599685	0.03072328213	-1.497421584	-0.7277504037	0.03142602971	0	0.04349014299	-0.002494244397
0.7434599685	0.4966930612	-0.7697583089	1.084282009	0.2784367684	0	0.5095248492	0.4758687319
-1.423257518	0.926819011	0.7565597806	1.627891733	-0.6743189379	1	-0.05566617748	-0.574196338
0.7434599685	0.8909751819	-0.04209503367	-0.908953645	1.407628717	0	1.679569431	1.339255567
0.02122080619	0.6042245486	0.7920555502	-0.3653439211	-1.133053167	0	-0.6208572041	-1.169233211
-0.7010183561	0.8909751819	1.484223056	1.627891733	-0.5331699444	0	0.04349014299	-0.3758507137

Scaling all the numerical features using Standard Scaler.



Stage 4: Modeling + Eval.

Development process



ML modeling and evaluation

We explored supervised, unsupervised, and rule-based modeling to support fair promotion decisions. Model performance was assessed using ROC-AUC, F1, recall, and precision, with Silhouette Score and DBI used to validate clustering quality.

Tree-based & linear model (exploratory only)

Model: Decision Tree, Random Forest, XGBoost, Logistic Regression, SVM. **Evaluation metrics:** F1 Score, ROC-AUC.

Clustering model

Model: K-Means, GMM, K-Medoids

Evaluation metrics: Silhouette Score, DBI

Rule-based model

Model: Built from engineered features and weighted coefficients. Replaces the biased Promotion_Eligible target with an objective Promotion_Score.

Threshold: Promotion_Score > 0.85



NOTES: to validate our assumption regarding potential bias in **Promotion_Eligible**, we used this target feature within all classification model.



Treebased & linear model

We tested tree-based models, logistic regression, and SVM on the promotion eligibility task. The results were disappointing. All models were still unstable and didn't give us the confidence to deploy. This taught us that no algorithm can fix a broken ground truth, which is why we pivoted to rule-based logic instead.

Classification Model	F1 Score	ROC-AUC	Outcome
Tree-based (DT, RF, XGB)	17–27%	48–59%	Weak patterns; struggled even after tuning
Logistic Regression	35–40%	52–55%	Weak patterns; struggled even after tuning
SVM	41–60% (unstable)	Peaked at 59%	Best performer, but not strong enough for reliable prediction

- **Tree-based models:** weak patterns, even after tuning.
- **Logistic Regression:** low signal, struggled to learn any meaningful pattern.
- **SVM:** best result but still unstable and insufficient for deployment.
- **Overall:** most metrics fell below 50%, confirming that the target label is unreliable.

Best tree-based: Random Forest

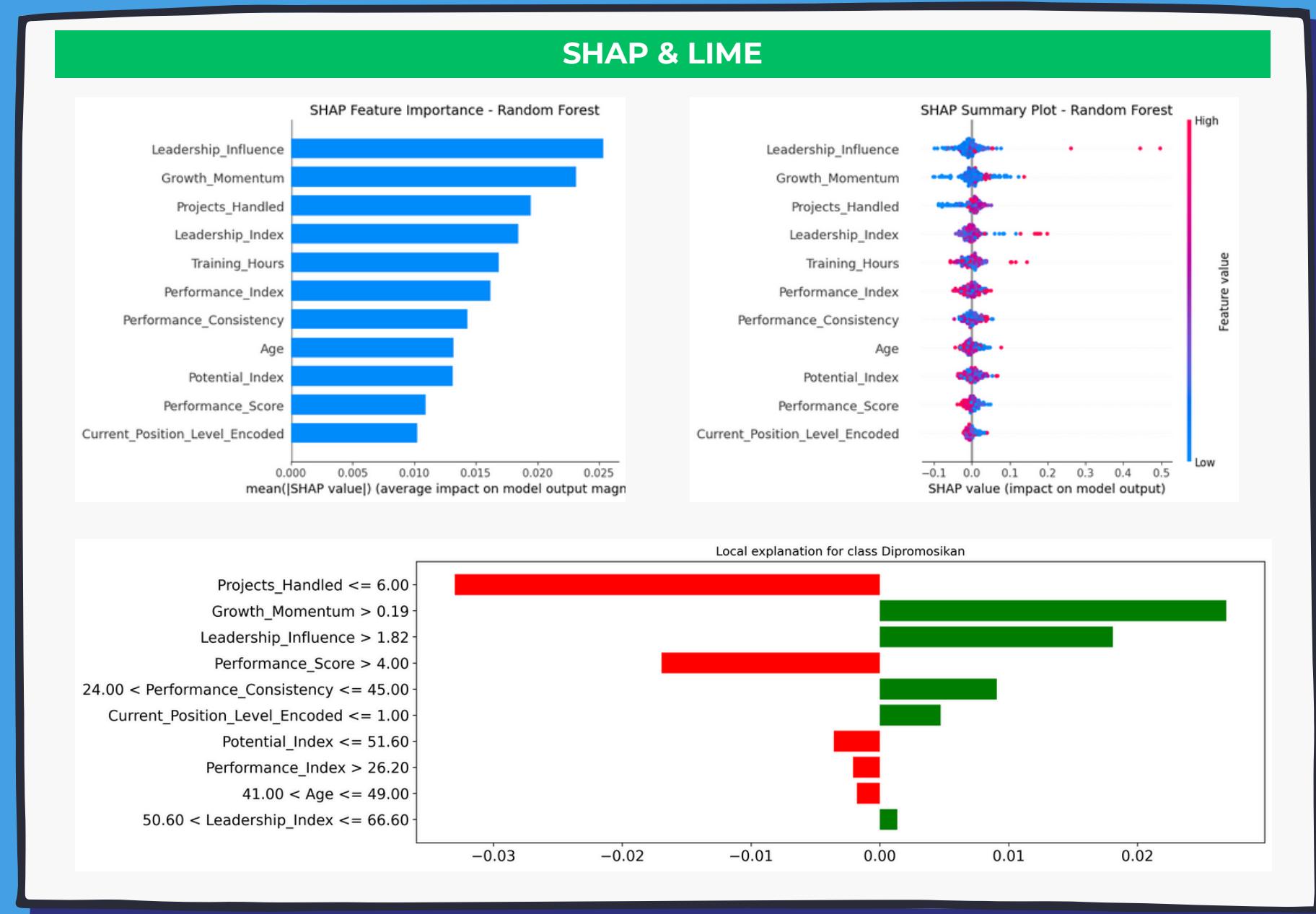
Train	Pred 0	Pred 1	Test	Pred 0	Pred 1
Actual 0	412	156	Actual 0	85	58
Actual 1	38	183	Actual 1	30	25

Best linear: SVM

Train	Pred 0	Pred 1	Test	Pred 0	Pred 1
Actual 0	227	270	Actual 0	87	127
Actual 1	51	142	Actual 1	40	43

Feature importance

So did we just discard those experiments? No. The tree-based and linear models taught us which features are truly predictive. This knowledge became the foundation for our clustering and rule-based approach, ensuring our logic is grounded in real feature importance.

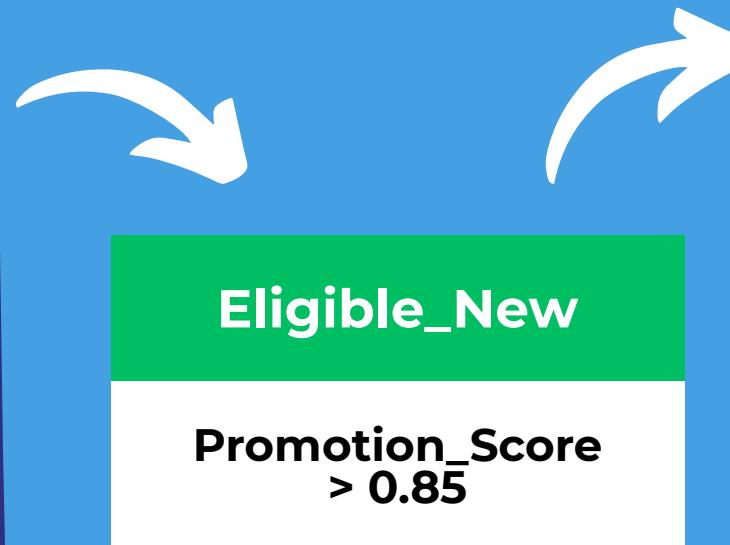
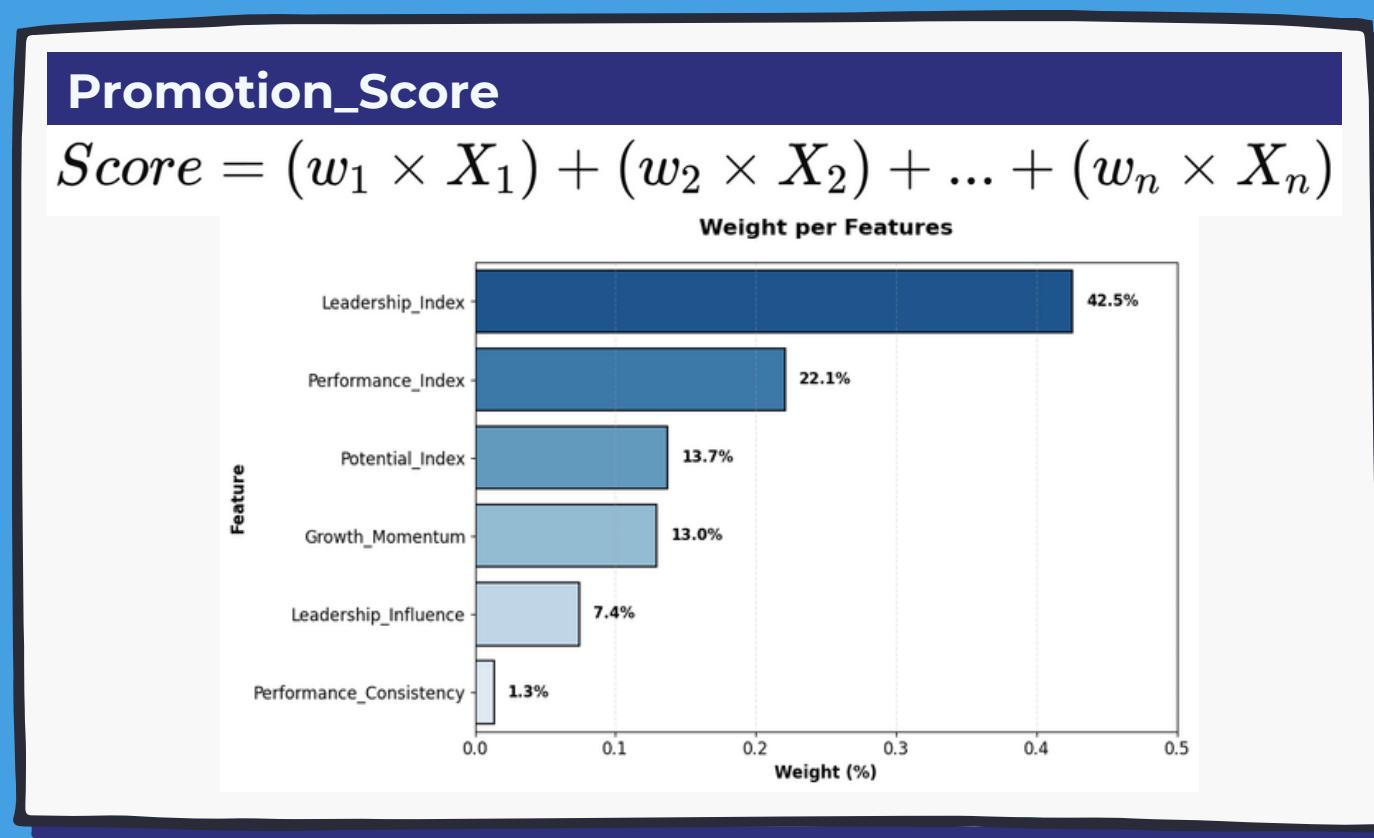


The **SHAP** results show that leadership and growth-related features, especially Leadership Influence, Growth Momentum, and Projects Handled are the strongest global drivers of the promotion predictions.

The **LIME** shows that for the selected employee, strong growth and leadership indicators support the promotion prediction, while low project count and inconsistent performance reduce it.

Rule-based model

We built a transparent rule-based system instead of a black-box model. It combines our features with weights, Leadership Influence (42.3%), Performance Index (22.1%), and others into a single Promotion Score. Scores above 0.85 indicate promotion readiness. Every recommendation is explainable and auditable by HR.

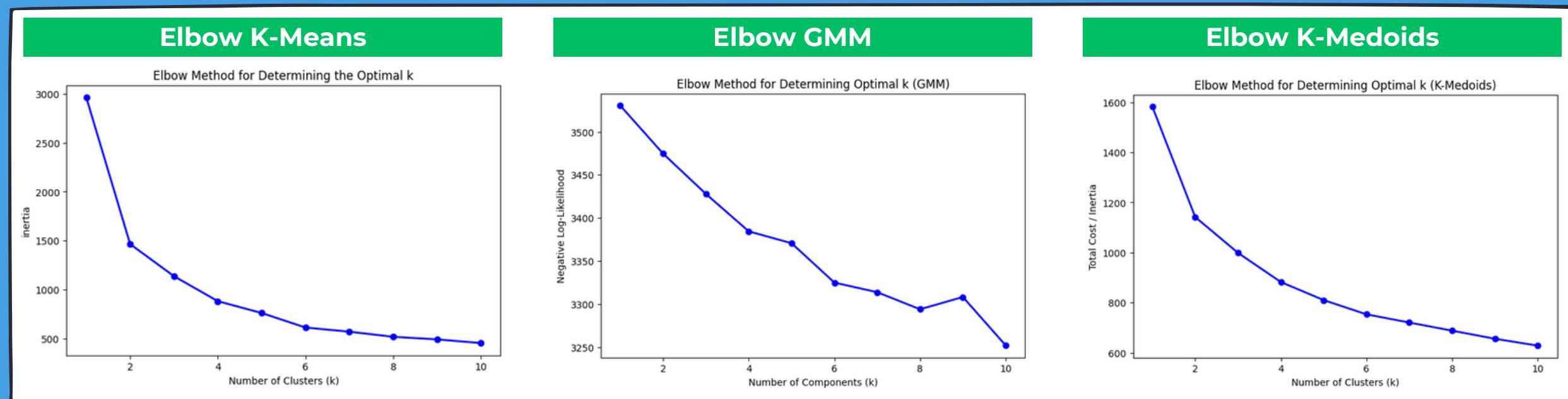


	JUNIOR	LEAD	MID	SENIOR
1	42	32	42	32
0	196	207	254	182

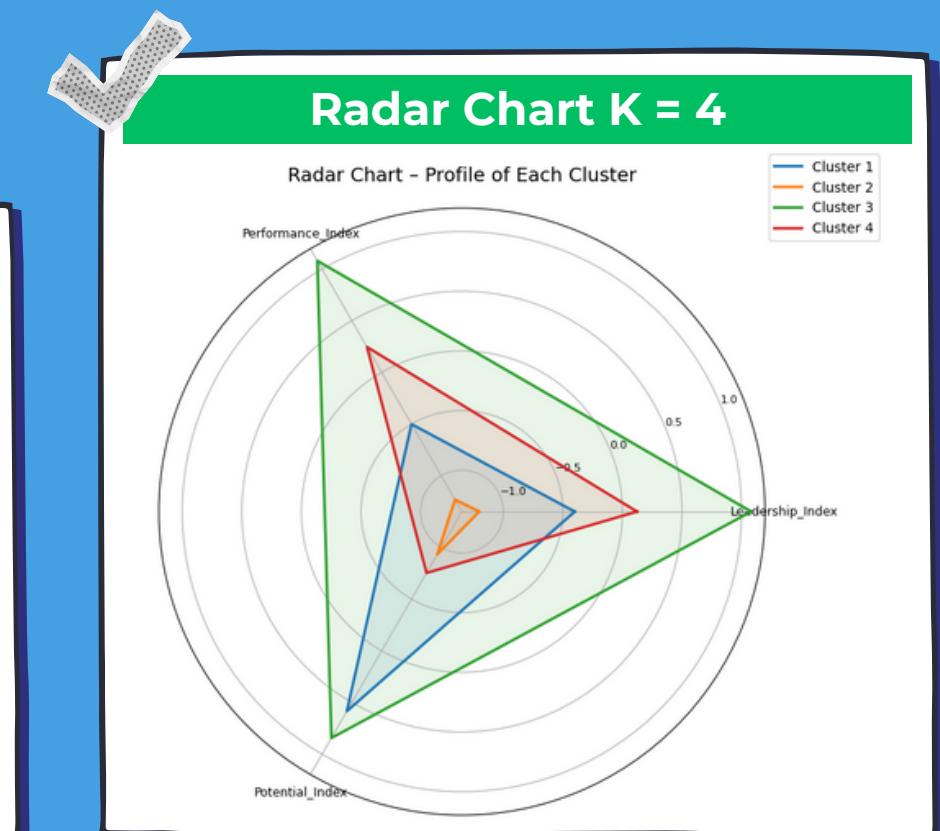
Our rule-based threshold is backed by data. Industry standards show 6-8% is healthy, 8-12% is good, and tech companies average 14% promotion rates. Our system is tuned to match these realistic benchmarks, not arbitrary thresholds.

Clustering model

Criteria	K-Means	GMM	K-Medoids
Best K	K = 4	K = 5	K = 5
Silhouette Score	0.320	0.247	0.293
DBI Score	1.00	1.173	1.071
Key Observation	Clearest and most compact clusters.	Lowest performance, overlapping clusters.	Fairly good and stable but not as good as K-Means.

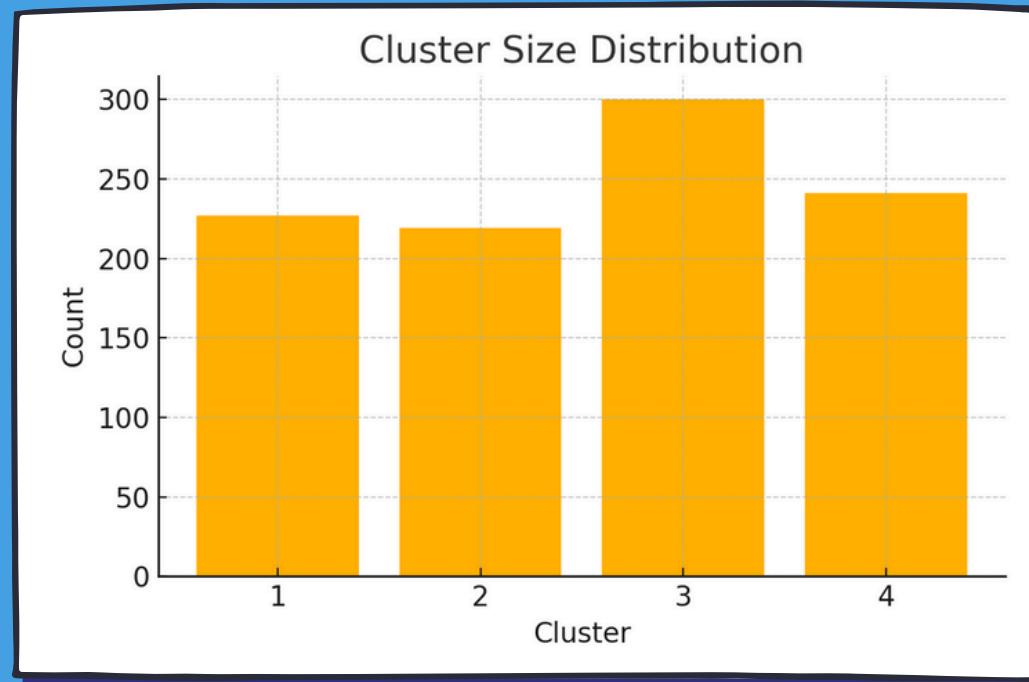


We tested three clustering algorithms: K-Means, GMM, and K-Medoids. K-Means with K = 4 won. It has the highest Silhouette Score (0.320) and best DBI (1.00), with clear, well-separated, and compact clusters. The elbow plots confirm K = 4 is optimal. This interpretability and cluster quality make K-Means perfect for segmenting employee talent profiles.



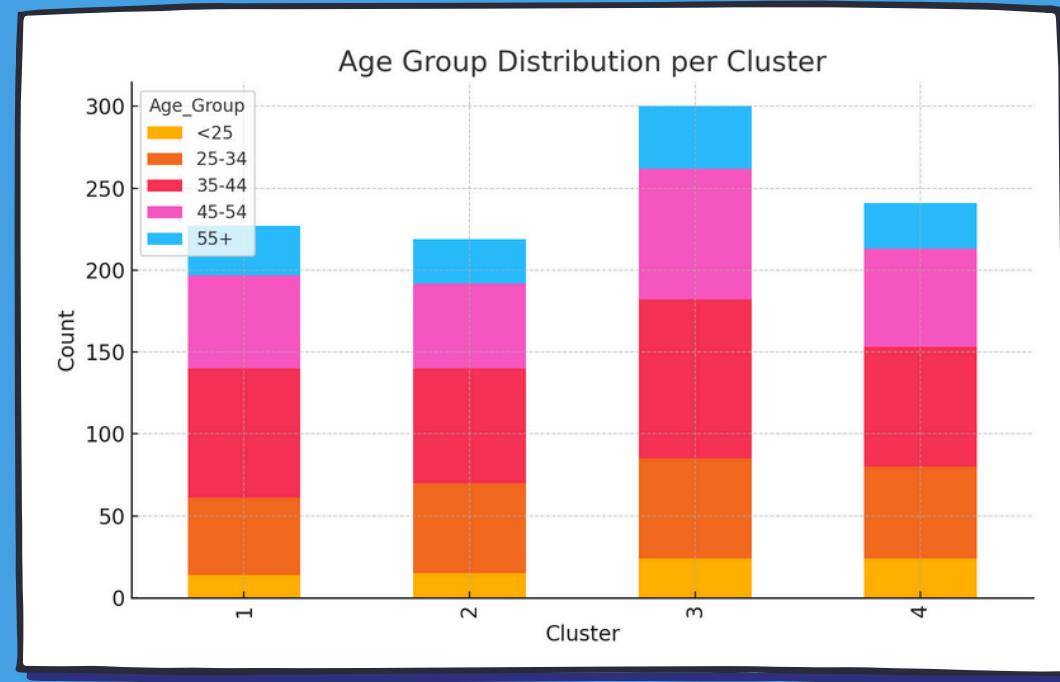
Clustering bias analysis

We validated that K-Means didn't introduce bias. Cluster sizes are balanced, age is evenly distributed, and job levels are mixed, meaning the model segments based on actual performance and capability, not demographics.



Cluster Size Distribution

All four clusters are well-balanced in size. This is important because it means K-Means didn't favor one group over another, reducing bias and ensuring fair representation across all segments.



Age Group Distribution Across Clusters

Age is evenly distributed across clusters—we see early, mid, and senior-career employees in each segment. This tells us the clustering is driven by actual performance and capability, not by age bias.



Job Level Distribution per Cluster

Job levels are also balanced across clusters. While some roles appear slightly more in certain segments, no cluster is dominated by a single job tier. The model captured behavioral differences and capability signals rather than just replicating the organizational hierarchy.

Cluster interpretation

K-Means with K=4



Cluster 1: Under Developed With Potential

Technical: Moderate scores in Performance (-0.50) and Leadership (-0.39), but relatively higher scores in Potential (0.58). Stable and fairly consistent, but has yet to demonstrate strong execution skills.

Business: Moderate performance and leadership, but high potential.

Cluster 2: At-Risk & Underpowered

Technical: Lowest values across all three indexes, which are performance (-1.23), leadership (-1.20), and potential (-0.93). Shows stagnation and lack of momentum.

Business: Low across all metrics. They need immediate intervention, such as reskilling, reassignment, or transition planning.



Cluster 3: All Around Top Performer

Technical: Achieves the highest scores in performance (1.08), leadership (1.08), and potential (0.84); highly adaptive and shows strong momentum.

Business: The best candidates for career acceleration and leadership pipeline. A strategic source of potential for the organization's future.

Cluster 4: Consistent Performer or Leader

Technical: Demonstrates consistent execution with average performance (0.24) and leadership (0.12), but low potential (-0.75). This means they are reliable in their day-to-day work but do not yet show strong signs of long-term growth.

Business: Stable execution and reliable. They're the backbone of operations. Keep them engaged and valued.

Final models to be utilized

K-Means model

K-Means ($k = 4$) segments employees into four talent groups, providing HR clear, interpretable categories for decision-making. Clustering turns scattered employee data into structured talent profiles with distinct strengths and risks.



Rule-based

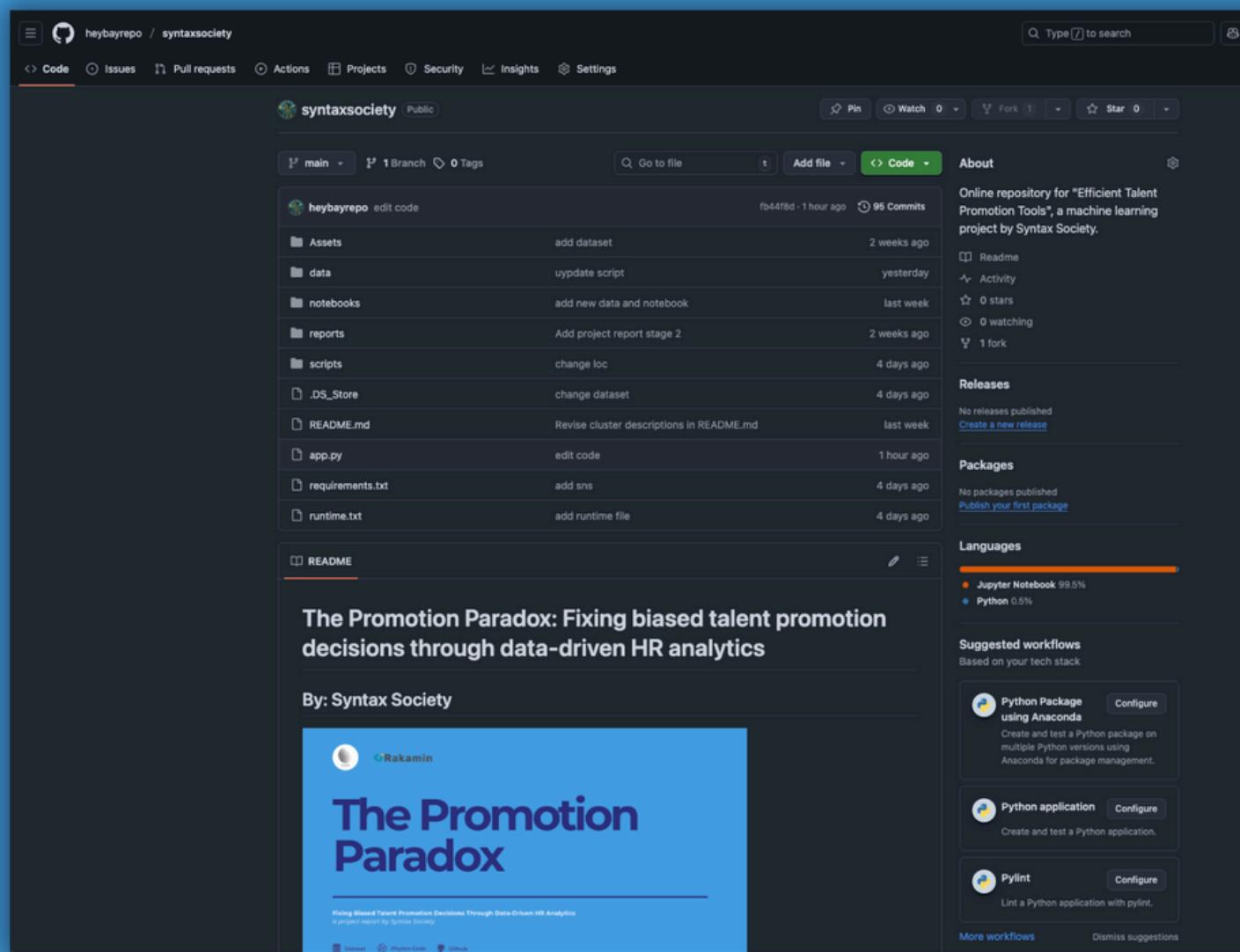
The rule-based method calculates a promotion score using weighted features: Leadership Influence (0.425), Performance Index (0.221), Performance Consistency (0.137), Growth Momentum (0.130), Leadership Index (0.074), and Potential Index (0.013). This transparent, non-black-box approach ensures promotion decisions are explainable and consistent.



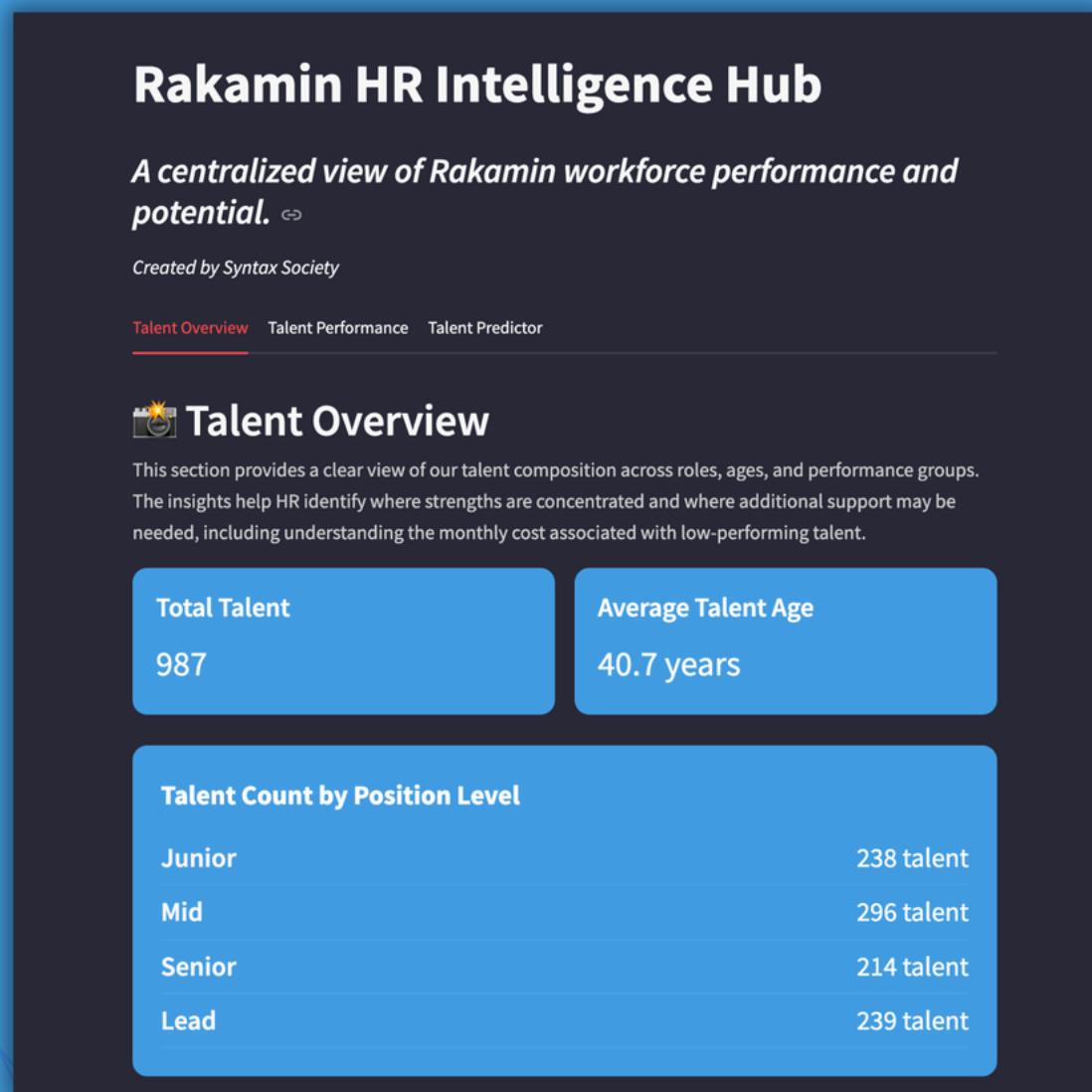
Stage 5: Deployment.

From Github to Streamlit

We built the dashboard on Streamlit with GitHub integration. Updates deploy automatically with each code push. The key design philosophy: keep it simple for non-technical users. Clean interface, clear metrics, intuitive workflows. And simplicity enables scalability: easier to maintain, modify, and expand as your needs grow. Check our prototype at <https://rakaminhrdashboard.streamlit.app/>



A screenshot of a GitHub repository page for 'syntaxsociety'. The repository name is 'syntaxsociety' and it is public. The 'Code' tab is selected. The commit history shows several recent changes, including 'edit code', 'add dataset', 'update script', and 'change loc'. The 'About' section describes the project as an online repository for 'Efficient Talent Promotion Tools', a machine learning project by Syntax Society. The 'Talent Overview' section contains a snippet of text: 'The Promotion Paradox: Fixing biased talent promotion decisions through data-driven HR analytics'.



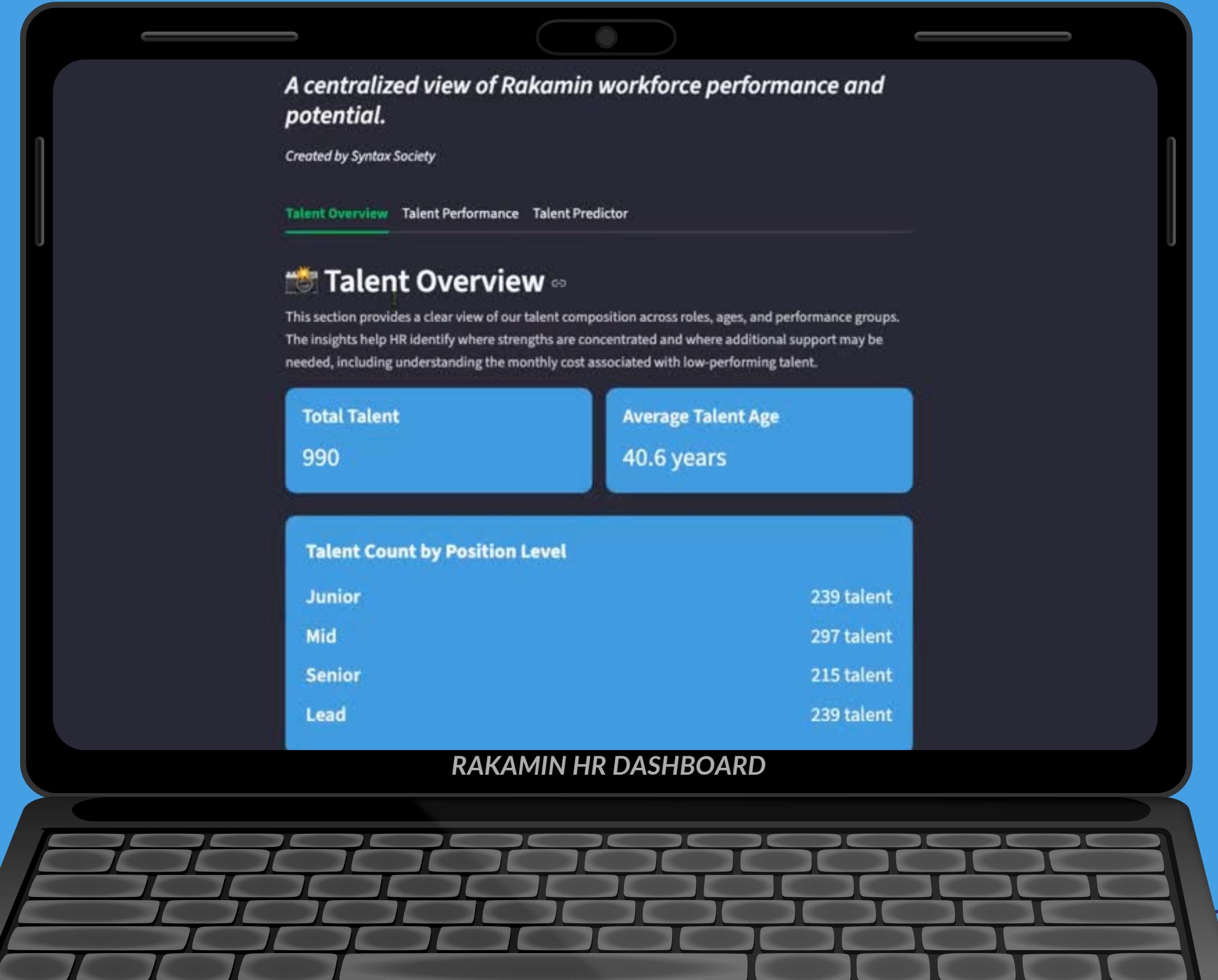
The dashboard has a dark theme with a title 'Rakamin HR Intelligence Hub' and a subtitle 'A centralized view of Rakamin workforce performance and potential.' It is created by Syntax Society. The 'Talent Overview' section includes a camera icon, the title 'Talent Overview', and a brief description: 'This section provides a clear view of our talent composition across roles, ages, and performance groups. The insights help HR identify where strengths are concentrated and where additional support may be needed, including understanding the monthly cost associated with low-performing talent.' Below this are two cards: 'Total Talent' (987) and 'Average Talent Age' (40.7 years). A larger card titled 'Talent Count by Position Level' lists the following data:

Position Level	Talent Count
Junior	238 talent
Mid	296 talent
Senior	214 talent
Lead	239 talent



Source: [Streamlit](#)

App features



Tab 1: Talent Overview. Shows a snapshot of the workforce: total talent, position-level distribution, and potential attrition indicators.

Tab 2: Talent Performance. Displays top and bottom performers across key indexes (performance, leadership, potential) with optional filters for deeper exploration.

Tab 3: Talent Predictor. ML-powered clustering that reveals each employee's group, traits, and risks. Users can input data manually or upload records via CSV.

Flexibility & scalability

HR can choose to upload their own dataset

Select employee ID to see specific employee. We're not using name because it is not in the dataset.

Use the provided CSV template to ensure uniformity

Data upload success notification

Browse uploaded data instantly on any device

Promotion eligibility calculator can predict employee's readiness, strength and aspects that need to be improved.

Radar chart is shown for quick assessment of employee's quality.

The screenshot shows a Streamlit application interface. On the left, a 'Select Talent Input Method' page offers three options: 'Select employee ID', 'Predict employee cluster and characteristics', and 'Upload employee data in bulk using CSV'. The 'Upload employee data in bulk using CSV' option is selected. Below this are fields for 'Upload CSV' and 'Download template CSV', followed by a file upload area where 'bm_test.csv' is selected. A green success message states '✓ Uploaded 3 rows successfully. Go to \'Select employee ID\' to view/inspect.' Below this is a table titled 'Preview of your processed and clustered data' with three rows of data. On the right, a 'Promotion Prediction' page displays results for an employee with a score of 54.56, eligible for promotion. It includes sections for 'Why This Result?', 'Top Strengths', 'Can Be Improved', and a 'Talent Radar Chart'. The radar chart has axes for Peer Review Score, Leadership Score, Performance Score, Growth Momentum, Projects Handled, and Training Hours. A 'Succession Potential Indicator' section suggests providing advanced leadership exposure, starting formal succession mentoring, and assessing readiness for expanded scope. To the right of the radar chart are three gold stars and a thumbs-up icon.

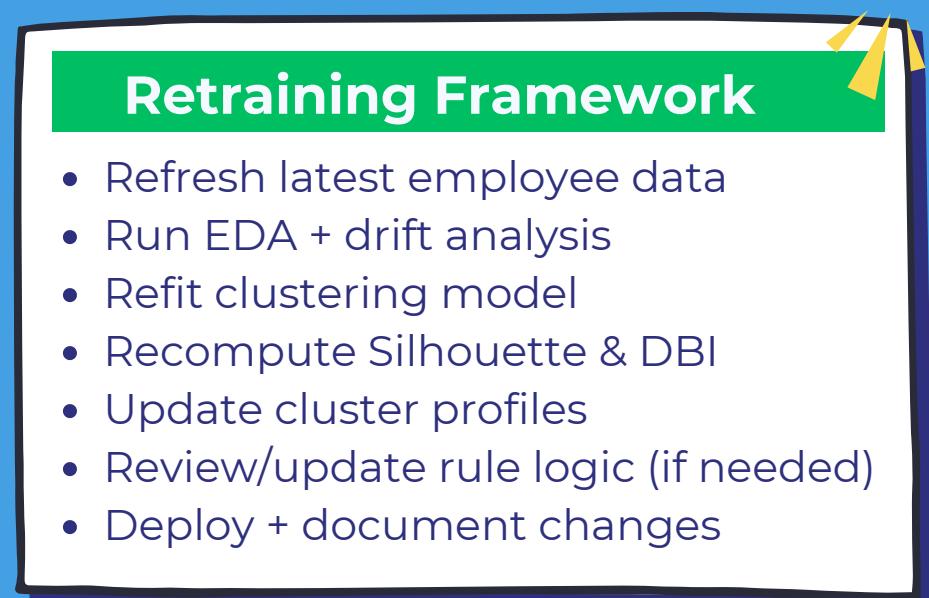
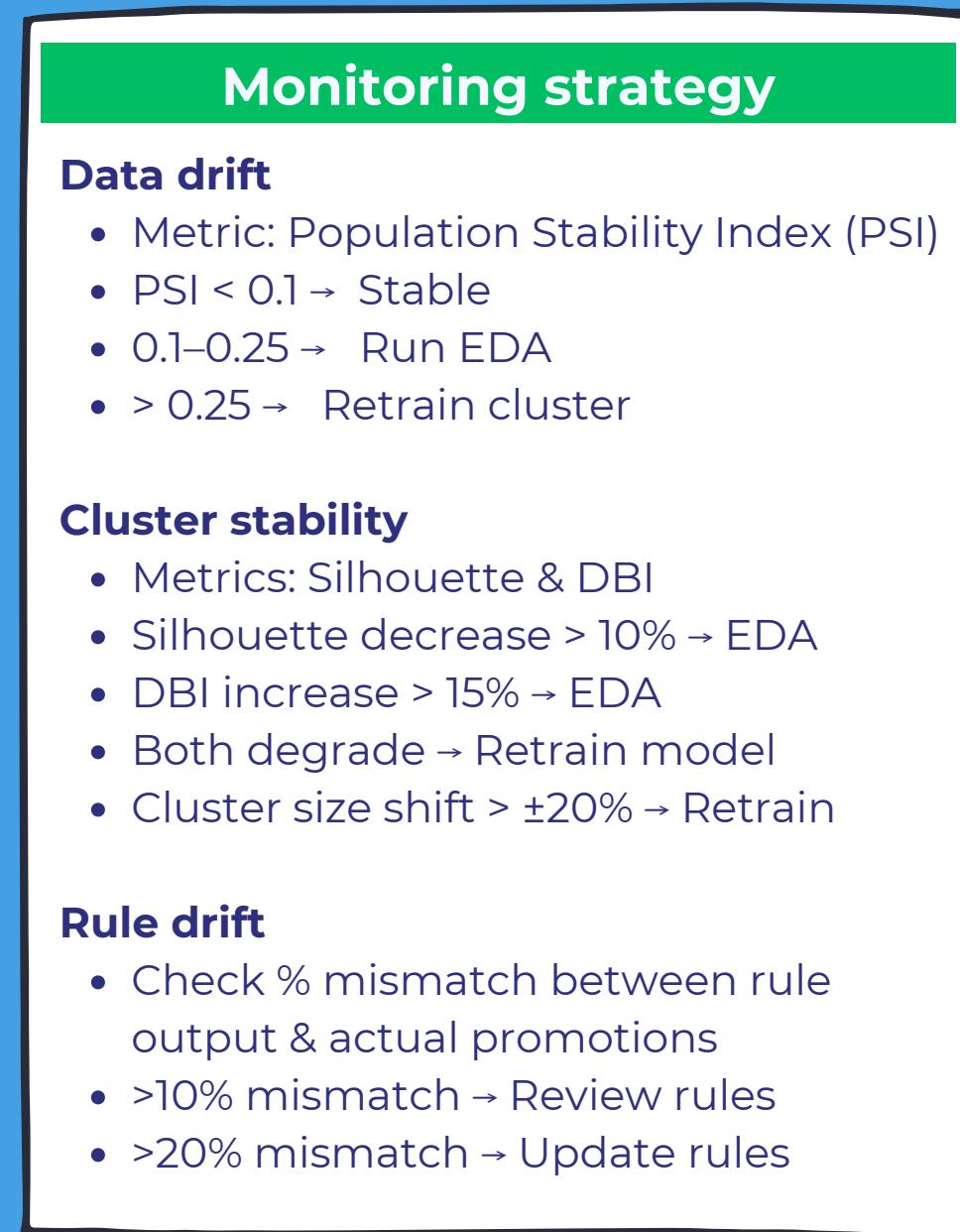
#	z_Hours_scaled	Performance_Consistency	Growth_Momentum	Cluster	Characteristics	D
0	-0.508	-4.2146	-1.7133	3	All Around Top Performer	T
1	-0.889	-1.1239	-5.0627	3	All Around Top Performer	T
2	1.397	7.0244	0.5861	2	At-Risk and Underpowered	T

Model usage flow



Monitoring & retraining

We continuously monitor three things: shifts in employee data patterns (data drift), the health of the cluster structure, and how well the system's recommendations align with actual promotion decisions. When any metric crosses its threshold, it triggers a deeper analysis.



Stage 6: Impact.

Business impact

Talent management isn't just about fairness. It's about money. Every time you promote someone who isn't ready, you're paying a senior-level salary for below-par performance. Every time you pass over someone capable, you risk losing them.

⚠️ High Risk Talent				
These employees fall into lower performance ranges and may need targeted development, mentoring, or closer support.				
Select ranking category:		Filter by Position Level:		
Low Performing		All Levels		
Employee_ID	Current_Position_Level	Performance_Index	Performance_Score	Cluster
EMP0010	Mid	1	1	1
EMP0178	Senior	1.3	1	1
EMP0974	Junior	1.8	1	1
EMP0910	Mid	2.1	1	1
EMP0810	Lead	2.6	3	1
EMP0241	Lead	2.7	1	1
EMP0939	Junior	2.7	1	1
EMP0251	Lead	2.8	2	1
EMP0656	Senior	2.8	4	1
EMP0829	Mid	3.1	3	1

AVG salary based on Deals!

Level	AVG Salary
Junior	Rp6.500.000
Middle	Rp11.500.000
Senior	Rp22.500.000
Lead	Rp27.500.000

Most of the workforce is in the Average Talent category, followed by High Performing Talent, and a fairly large proportion is Low Performing Talent at 219 employee which is the main cause of the estimated loss.

Save
↑ 3.5 billions!

↑ 22%

Saving by not paying incompetent employees

↑ 42%

Improved accuracy in previously biased determinations of promotion eligibility.

Business recommendation

Action items



#1 Overhaul eligibility determination system.

Update it using reliable and reasonable metric calculations to minimize potential bias with online-based data collection.



#2 Conduct a rigorous performance review.

Specifically for employees who fall into Cluster 2 (At-Risk & Underpowered) and make a decision between skill development or, at the worst, employment termination.



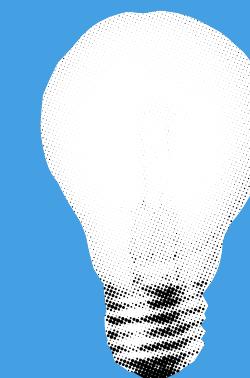
#3 Leverage the result of the Talent Clustering.

Identify and map existing skills gaps. Focus development interventions on Clusters 1 and 2 to address the identified gaps. For Clusters 3 and 4, provide targeted training for retention.

Strategic roadmap

Next-Phase Opportunities

Expand predictive models to additional HR domains like attrition risk, promotion readiness



Scalability & Integration Roadmap

Develop APIs for payroll and attendance systems to enable real-time data flow.

Future Capability Development

Create targeted upskilling programs aligned with model insights and skill gaps.



Rakamin

Thank You!

The Promotion Paradox:

Fixing biased talent promotion decisions through data-driven HR analytics

A project proposal by Syntax Society | Rakamin - Data Science batch 59

