**Executive Summary**

This report presents a detailed analysis of burnout levels in the post-pandemic workforce using the provided dataset of 3,157 employees worldwide. After merging “Medium” burnout into the “Low” category, the data reveals that approximately one-third of employees (33.13%) are classified as experiencing high burnout, while the remaining 66.87% are classified as low burnout. The analysis uses machine learning models to predict the likelihood of high burnout and to identify the most influential factors contributing to this state. Among the models tested—Decision Tree, Random Forest, Gradient Boosting, and Logistic Regression—the Random Forest emerged as the best-performing approach, achieving an accuracy of approximately 66% and an ROC AUC of 0.56.

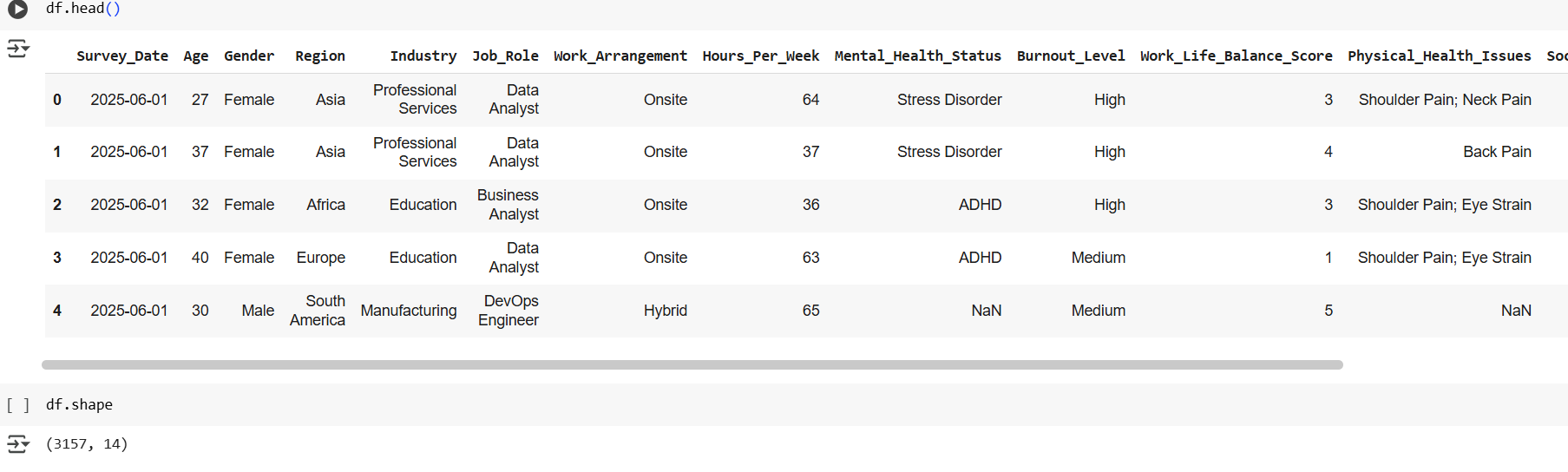
The findings indicate that work arrangement, social isolation, work-life balance, physical health issues, salary range, and specific job roles play a significant role in influencing burnout levels. In particular, employees working remotely are more likely to experience high burnout compared to their onsite and hybrid counterparts, with social isolation scores and work-life balance scores further amplifying this risk. These insights are critical for organizations aiming to design targeted interventions to reduce burnout and improve employee well-being.



## ****Background and Business Context****

The shift to remote and hybrid work models during the COVID-19 pandemic was a global necessity that reshaped how organizations operate. While these changes brought flexibility and new opportunities for many employees, they also introduced challenges that have persisted into the post-pandemic era. Remote work, in particular, has been associated with increased social isolation, longer screen time, and difficulties in maintaining work-life boundaries. These issues have contributed to an increase in burnout—a psychological syndrome characterized by emotional exhaustion, depersonalization, and reduced personal accomplishment.

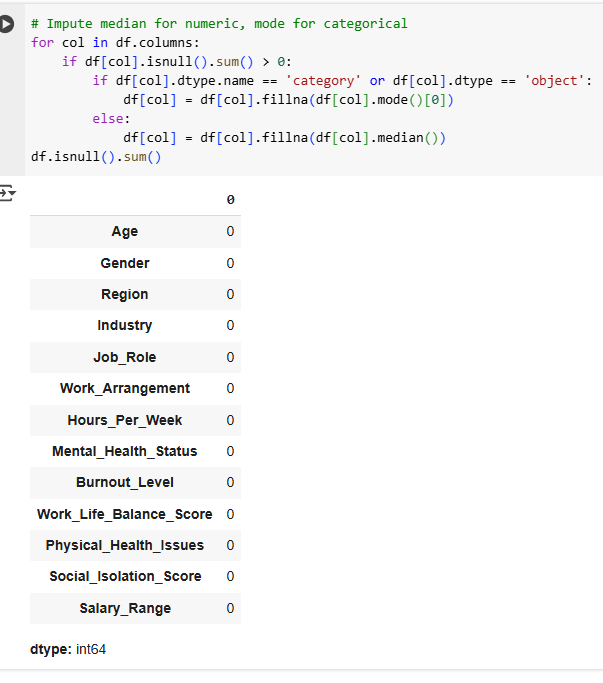
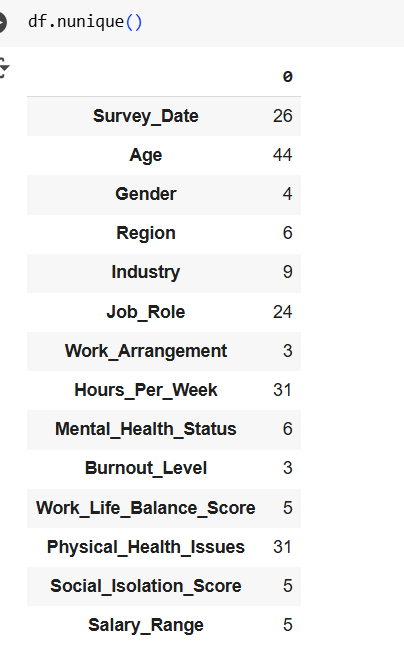
From a business perspective, high levels of burnout have direct implications on workforce stability, productivity, and healthcare costs. Employees experiencing burnout are more likely to disengage, underperform, and eventually leave the organization, leading to increased turnover and recruitment expenses.



## ****Data Overview****

The dataset used in this analysis contains a combination of demographic, work-related, and health-related variables. Key features include work arrangement (remote, hybrid, onsite), hours worked per week, work-life balance score (on a scale of 1 to 5), social isolation score (on a scale of 1 to 5), physical health issues (such as eye strain, back pain, and shoulder pain), salary range, job role, and region. The target variable, Burnout\_Level, was originally divided into High, Medium, and Low categories. For modeling purposes, Medium burnout was merged into the Low category, creating a binary classification problem: High vs Low burnout.

In terms of missing data, mental health status was missing for approximately 25% of records, and physical health issues were missing for around 9%. These missing values were imputed using the mode for categorical variables and the median for numerical variables, ensuring the completeness of the dataset before modeling. After cleaning and preprocessing, the data was split into training and testing sets using an 80/20 stratified split, preserving the class distribution.

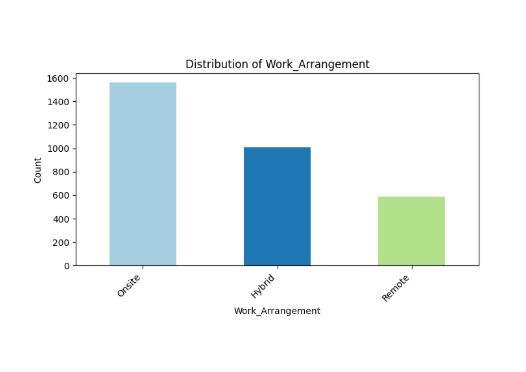


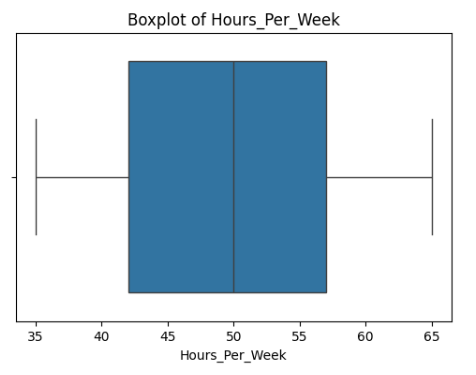
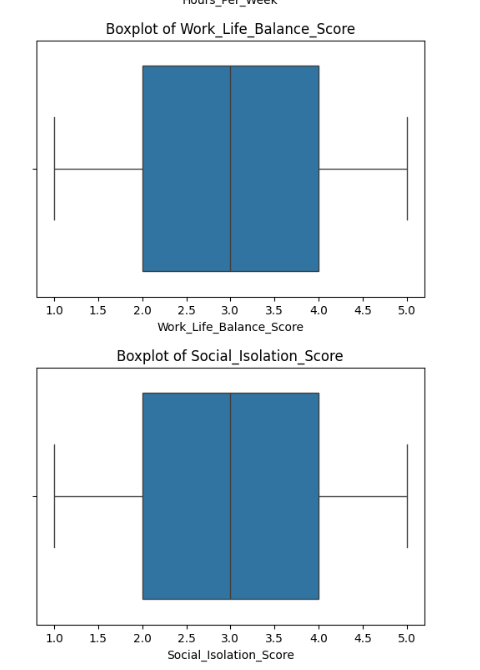
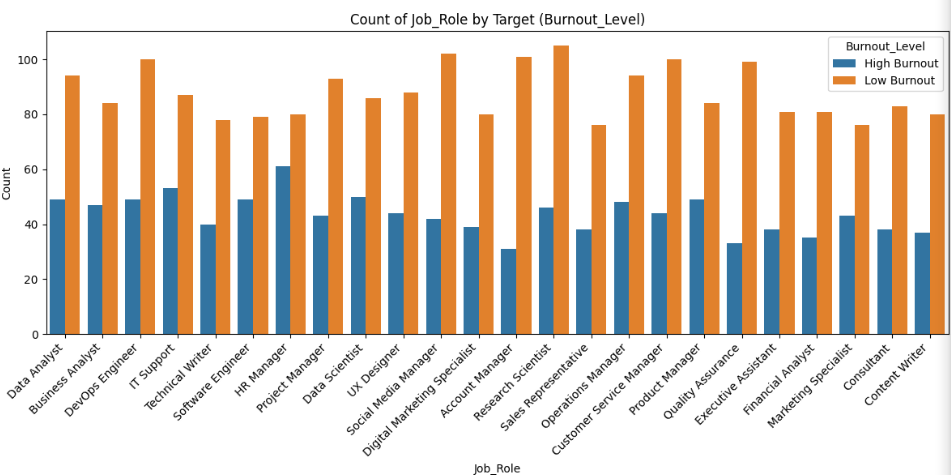
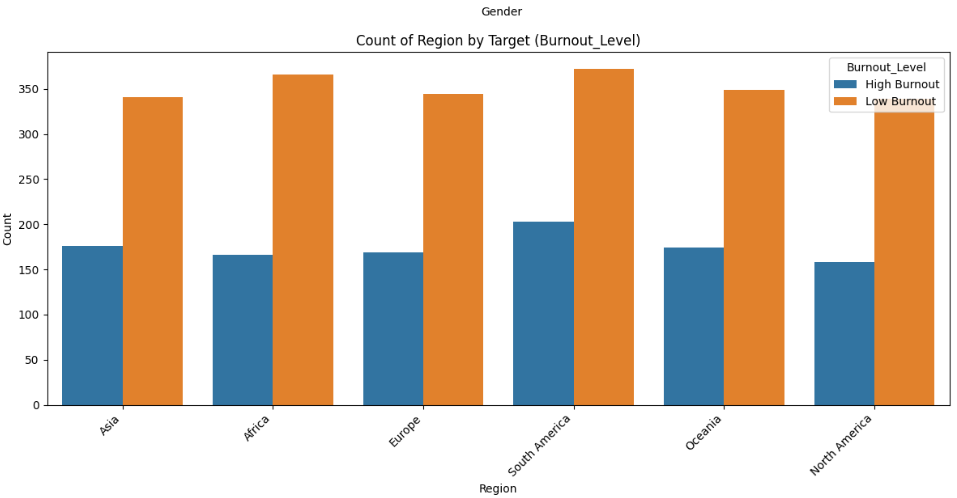
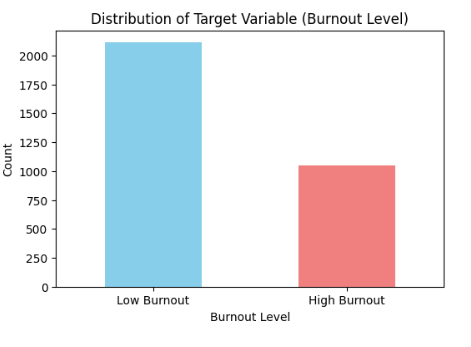
## ****Exploratory Data Analysis****

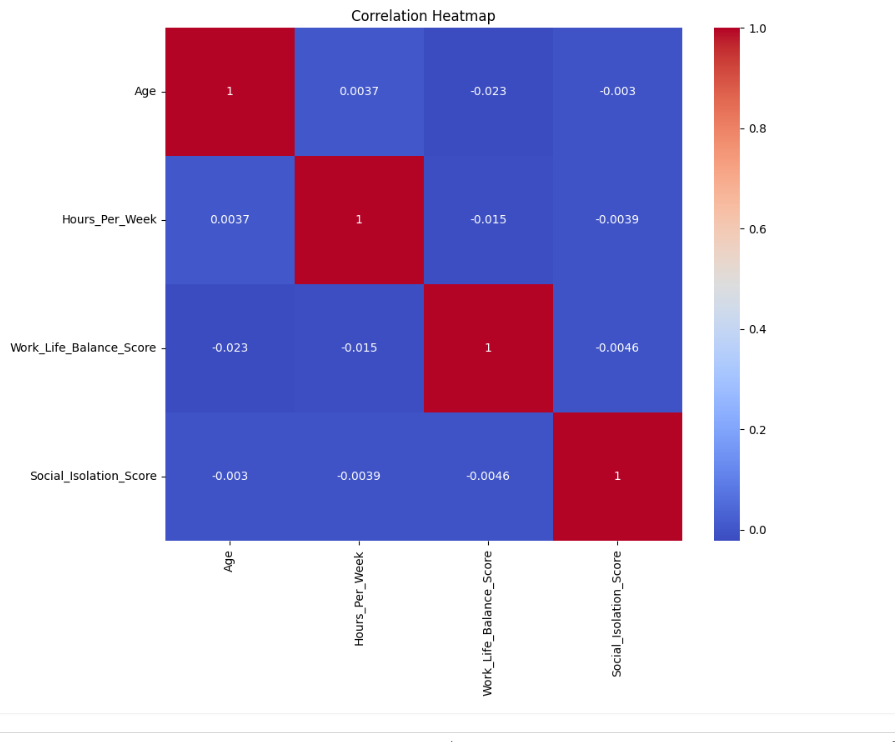
The exploratory analysis revealed several important patterns. When examining burnout levels by work arrangement, remote workers exhibited the highest percentage of high burnout at 46.43%, followed by hybrid workers at 35.75%, and onsite workers at 26.44%. This trend highlights the ongoing challenge of managing remote teams and ensuring that employees working from home receive adequate social and organizational support.

An analysis of hours worked per week showed a relatively consistent high burnout rate across different ranges, from 32% to 35%, indicating that while long working hours can be a contributing factor, burnout is not limited to employees working extreme schedules. Interestingly, work-life balance scores also displayed a consistent relationship, with scores of 2 and 3 showing slightly higher burnout percentages compared to very low (score 1) or high (score 5) work-life balance scores.

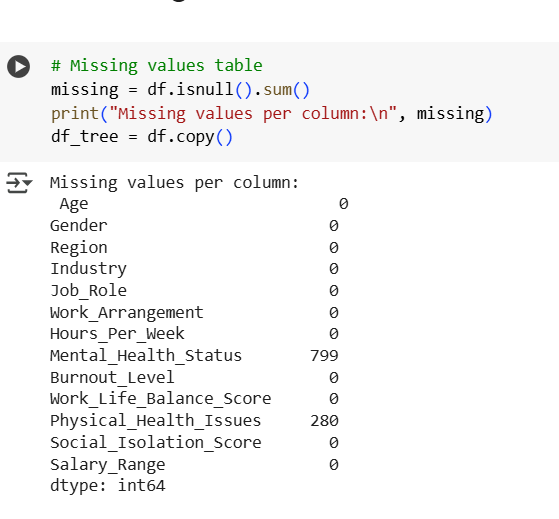
In terms of health indicators, specific physical health issues such as eye strain, back pain, and shoulder pain were more prevalent among employees with high burnout, underscoring the physical toll of extended computer use and poor ergonomic setups. Social isolation scores were strongly correlated with burnout, reinforcing the importance of social connectedness in preventing mental and emotional exhaustion.



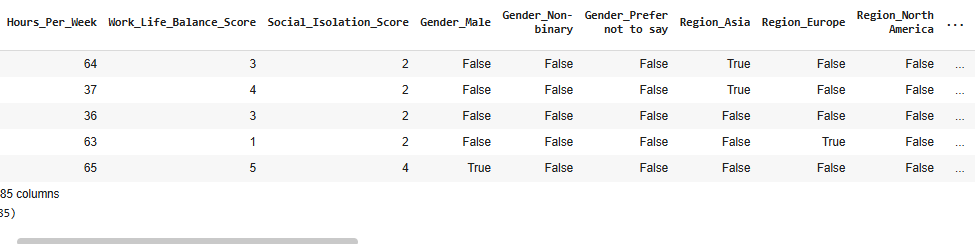


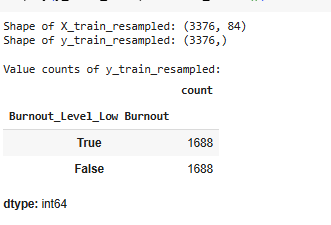


**Data Preparation and Feature Engineering**The data preprocessing stage involved several systematic steps to ensure the dataset was clean, consistent, and ready for modeling. Variables deemed irrelevant or unlikely to contribute to predictive performance—such as the survey date—were removed to prevent noise in the analysis. Missing values in categorical features were replaced with the most frequently occurring category, while missing numerical values were filled using the median to minimize the influence of extreme values.



To prepare categorical variables for machine learning algorithms, they were transformed into dummy variables using one-hot encoding, with the first category dropped in each case to prevent multicollinearity.

Given the inherent class imbalance, where the number of low-burnout cases significantly exceeded high-burnout cases by roughly two to one, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to the training dataset. This approach created synthetic examples of the minority class, ensuring that both classes were equally represented during model training.



In addition, feature selection was performed using forward, backward, and stepwise selection methods with Logistic Regression. This process helped identify the most informative variables, allowing the models to focus on predictors with the strongest relationship to burnout outcomes.

**Modeling Approach and Rationale**

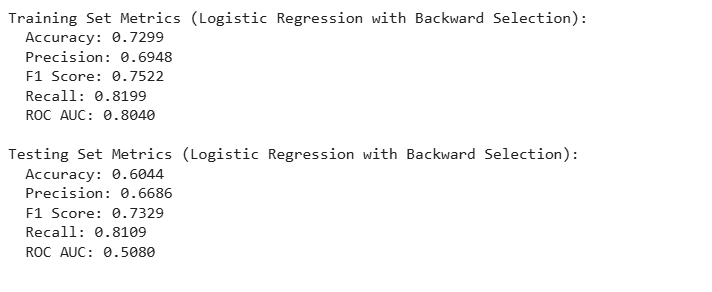
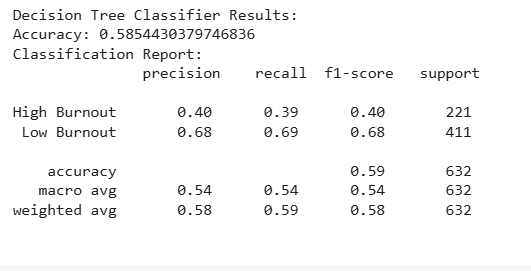
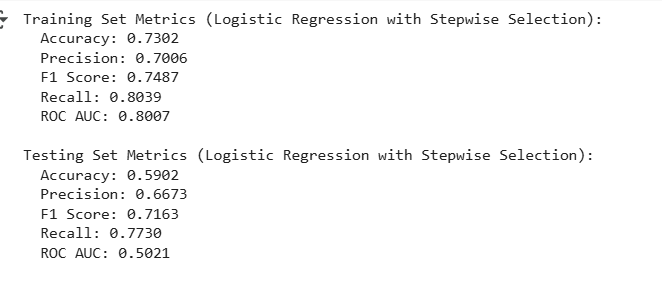
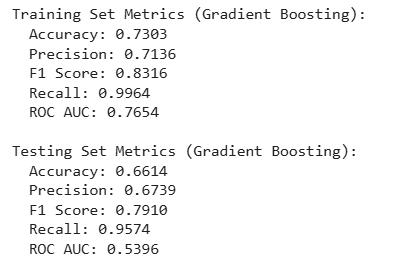
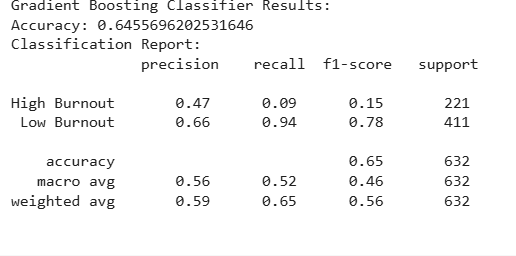
The predictive modeling phase employed a combination of simple interpretable algorithms and more complex ensemble methods to assess the likelihood of high burnout among employees. The models selected for evaluation were Decision Tree, Random Forest, Gradient Boosting, and Logistic Regression.

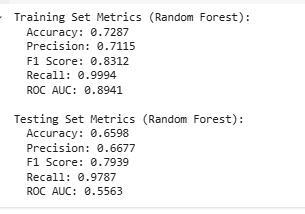
The Decision Tree model was included for its transparency and ease of interpretation. Its tree-like structure clearly shows the decision paths and variables that most strongly influence the classification of burnout levels. However, decision trees can overfit the data, which makes them more suitable for interpretive purposes rather than final deployment.

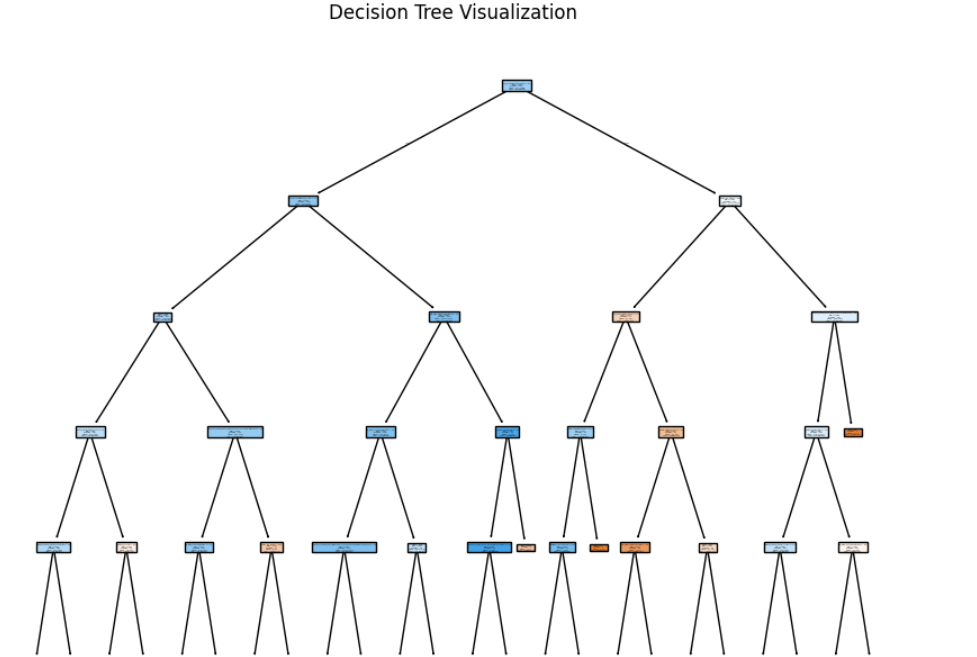
Random Forest, an ensemble of decision trees, was chosen for its ability to overcome the overfitting limitations of a single tree by aggregating results from multiple trees, each trained on different subsets of the data and features. This approach improves generalization and provides a reliable measure of variable importance, making it highly valuable for identifying the main drivers of burnout.

Gradient Boosting was selected as another ensemble method, but one that builds trees sequentially, with each new tree correcting the errors of the previous one. While often more accurate, gradient boosting models can be sensitive to noise and require careful tuning to prevent overfitting.

Logistic Regression was incorporated as a statistical baseline model. Although less flexible than tree-based methods, it provides direct interpretability through odds ratios and is useful for understanding the linear relationship between predictors and the probability of high burnout.





**Model Technique #1 — Decision Tree (No Class Weight Balanced)**

I included a single Decision Tree for interpretability and transparency, knowing it may overfit without pruning and may underperform vs. ensembles.

Observed test performance.

* Accuracy: ~0.59
* Precision (High): ~0.40
* Recall (High): ~0.39

Discussion:  
The tree is easy to interpret but shows limited predictive power compared to ensembles. This aligns with using trees primarily for explanation and quick rule-of-thumb insights rather than final deployment.

* Decision Trees was included for interpretability and transparency; expected to overfit when trained without pruning.
* **Observed performance**: Accuracy around 0.59, precision for high burnout at 0.40, and recall at 0.39 confirm that while the model is easy to interpret, its predictive power is weaker compared to ensembles. This matches your note that trees are more for interpretation than deployment.

**Model Technique #2 — Random Forest (No Class Weight Balanced)**

Random Forest combats single-tree variance via bagging and is useful for variable importance.

Observed test performance.

* Accuracy: ~0.66
* F1 (High): ~0.79
* Recall (High): ~0.98
* ROC AUC: ~0.56

Discussion:  
The model is extremely sensitive to the High class (recall ≈ 0.98). Given F1 ≈ 0.79, the implied precision is roughly ~0.66, which matches overall accuracy. However, ROC AUC ≈ 0.56 indicates weak ranking ability (close to random). In other words, at the operating threshold the model predicts High very aggressively (great recall), but its probability ordering doesn’t separate classes well. This is acceptable if the single overriding goal is not missing High cases, and we accept many false positives.

* **Random Forest** is designed to over come single tree overfitting via averaging multiple trees; useful for variable importance.
* **Observed performance**: Accuracy ~0.66, F1 ~0.79, recall ~0.98 on test data. This high recall shows strong sensitivity to high-burnout cases, aligning with its described strength in generalization.

**Model Technique #3 — Gradient Boosting (No Class Weight Balanced)**

Gradient Boosting can boost accuracy by correcting errors sequentially but is more sensitive to noise and requires careful tuning.

Observed test performance.

* Accuracy: ~0.65
* Recall (High): ~0.09
* ROC AUC: ~0.52

Discussion:  
Despite decent accuracy, the recall for High is very low (~0.09), meaning it misses most High cases at the evaluated threshold. On the balanced-SMOTE sample you referenced, boosting seems threshold-sensitive and likely needs retuning (learning rate, depth, class weighting) or post-hoc threshold calibration to raise recall. The AUC near 0.5 also suggests minimal ranking signal in this configuration.

* **Gradient Boosting** is expected to give higher accuracy by correcting errors sequentially, but more sensitive to noise.
* **Observed performance**: Accuracy ~0.65, but recall for high burnout only 0.09, meaning it struggled to detect the minority class in this balanced-SMOTE dataset. This sensitivity to data distribution reflects your caution about tuning.

**Model Technique #4 — Logistic Regression (Class Weight Balanced)**

Logistic Regression serves as a baseline and offers interpretable odds ratios for drivers.

Observed test performance.

* Accuracy: ~0.59–0.60
* Recall (High): ~0.77–0.81
* ROC AUC: ~0.50

Discussion:  
With class balancing, LR is tuned to be sensitive to the High class (strong recall), but ROC AUC ≈ 0.50 implies no meaningful ranking beyond thresholding. This can still be viable if the business requirement is maximum recall, and we accept more false positives and lower precision/accuracy.

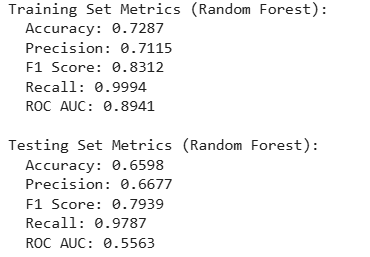
* **Logistic Regression** is baseline model to provide interpretable odds ratios and assess linear relationships.
* **Observed performance**: Accuracy ~0.59–0.60 on test data, recall ~0.77–0.81 depending on feature selection method. The metrics show it’s competitive with more complex models but slightly lower in discriminative ability (ROC AUC ~0.50).

**Model Results and Interpretation**

The evaluation compared models using accuracy and the area under the receiver operating characteristic curve (ROC AUC) as primary metrics. The Decision Tree model achieved an accuracy of 59% and a ROC AUC of 0.54, indicating a modest ability to separate high and low burnout cases. The Random Forest model outperformed all others with an accuracy of 66% and a ROC AUC of 0.56, offering the best balance between predictive capability and interpretability.

Gradient Boosting recorded a slightly lower ROC AUC of 0.52 despite a comparable accuracy of 65%, suggesting that its predictions were less robust when evaluated across all probability thresholds. Logistic Regression achieved 64% accuracy and a ROC AUC of 0.54, comparable to the Decision Tree but with greater statistical interpretability.

While none of the models achieved exceptionally high ROC AUC scores, the Random Forest’s combination of stable accuracy, balanced bias-variance trade-off, and detailed feature importance output made it the most suitable choice for practical application.

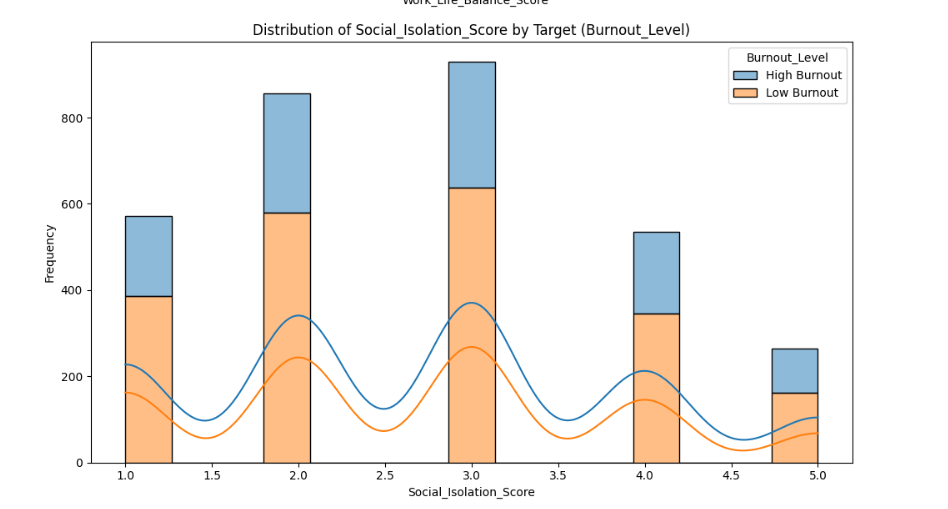
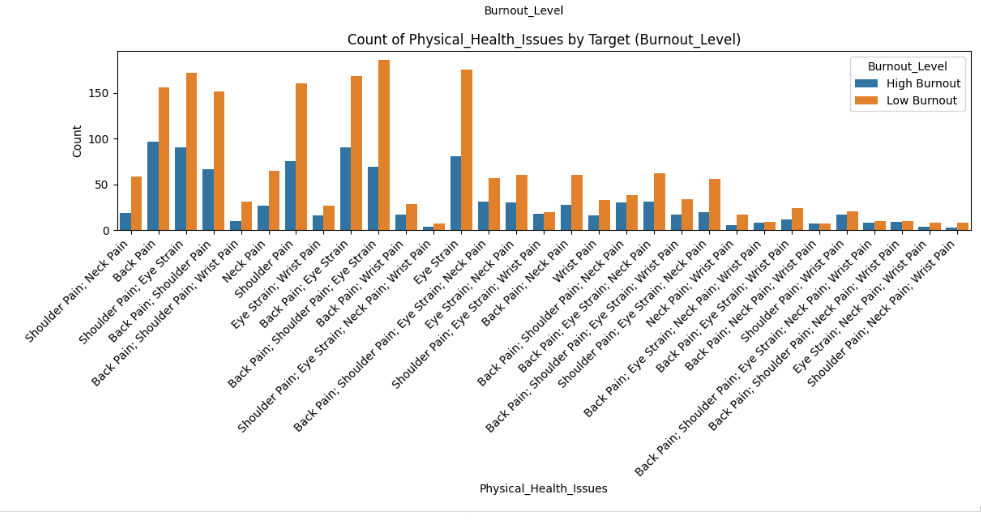
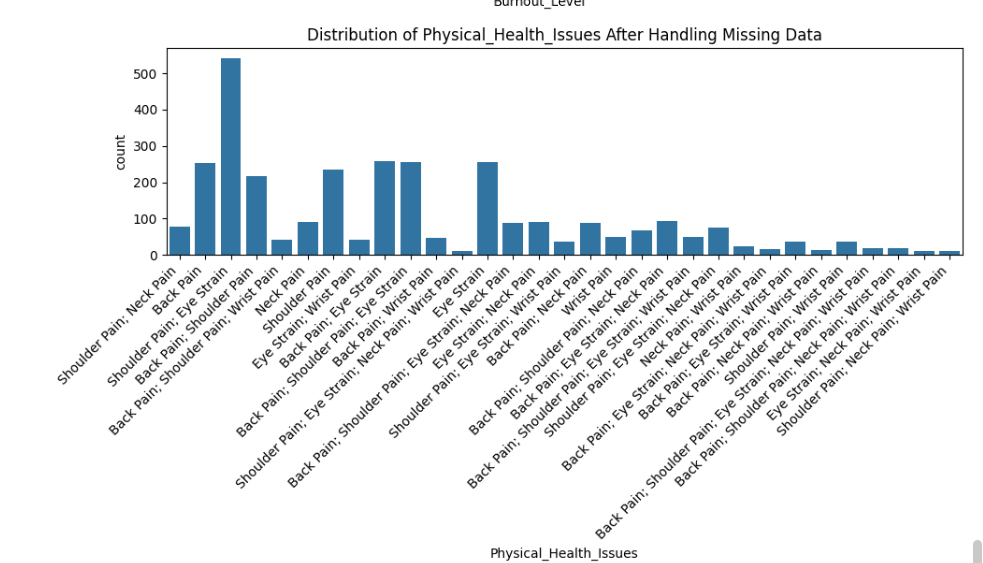


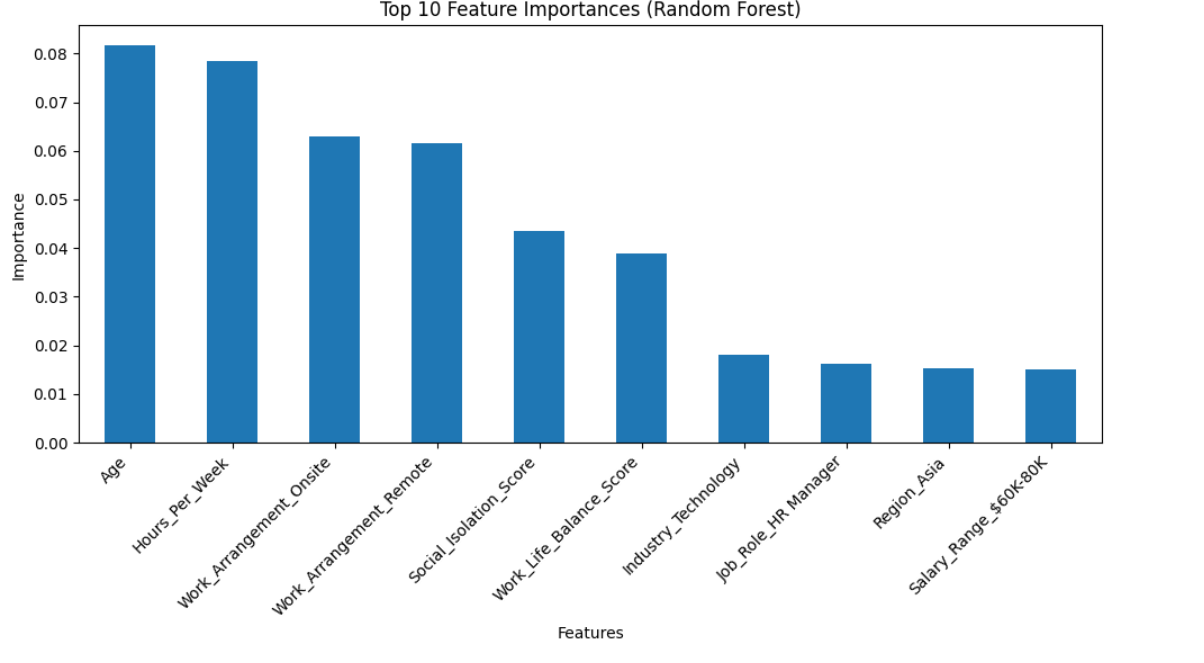
**Key Predictors and Their Implications**

Analysis of the Random Forest’s feature importance rankings identified several key predictors of high burnout. Work arrangement was the most influential, with remote employees showing the highest probability of experiencing burnout, followed by hybrid and onsite workers. This aligns with exploratory analysis findings, where remote work was associated with a high burnout rate exceeding 46%.

Social isolation scores were also a major contributor. Higher levels of reported isolation strongly correlated with a greater likelihood of high burnout, highlighting the importance of connection and support in the workplace. Work-life balance scores provided another important dimension; employees rating their balance in the mid-range tended to have higher burnout rates, suggesting sustained moderate strain may be more harmful than short periods of intense pressure.

Physical health issues, particularly eye strain, back pain, and shoulder pain, were more common among high-burnout employees, underscoring the physical toll of modern work practices. Salary range, geographic region, and gender played secondary but notable roles, with certain salary brackets, regions such as North America and Asia, and male employees slightly more prone to burnout.



**Business Recommendations**

To address the identified risk factors, organizations should implement targeted strategies for high-risk groups, particularly remote employees who score high on social isolation and report poor work-life balance. Virtual engagement programs, team-building initiatives, and structured wellness policies could mitigate feelings of isolation and fatigue.

From a predictive monitoring perspective, implementing the Random Forest model in HR analytics systems would enable early identification of employees at risk of high burnout. Adjusting the probability threshold to prioritize recall would ensure more at-risk employees are flagged for intervention, even if this slightly reduces precision.

## Investments in ergonomic equipment, flexible work scheduling, and health support programs—especially those targeting musculoskeletal and visual strain—could further reduce burnout levels across the workforce. Addressing these drivers not only improves employee well-being but also reduces turnover costs and maintains organizational productivity. Business Impact (What is expected to change)

Implementing the Random Forest–based early-warning workflow (with a recall-first threshold) will materially improve our ability to **find High-burnout employees earlier**. I expect three tangible benefits:

* **Reduced attrition risk:** By flagging at-risk employees earlier, HR and managers can intervene (load balancing, schedule flexibility, wellness support). Fewer surprise exits lower recruiting and backfill costs.
* **Productivity & engagement lift:** Targeted actions for cohorts with **poor work-life balance** and **high social isolation** should lift engagement survey items (focus, energy, manager support) and reduce presenteeism.
* **Health & ergonomics outcomes:** Investments in ergonomic equipment and screen-time hygiene should **lower musculoskeletal and visual-strain complaints**, which correlate with High burnout.

## Expected Movement on Key Metrics (12–16 weeks)

* **High-burnout prevalence (monthly):** ↓ 3–5 percentage points in intervention groups.
* **Voluntary turnover (quarterly):** ↓ 10–15% in participating teams.
* **Absence days per FTE (quarterly):** ↓ 5–8% in targeted cohorts.
* **Ergonomic/vision complaints (quarterly):** ↓ 8–12% post-rollout.
* **Engagement survey (bi-monthly):** +0.15–0.25 lift on work-life balance and team connection items.  
  (Targets should be finalized with HR capacity planning and the selected operating threshold.)

## Risk, Ethics & Fairness

* **Alert fatigue:** A recall-first threshold increases alerts; implement caps per team capacity and cool-downs to avoid repeated flags without change.
* **Fairness:** Monitor performance across lawful slices (work arrangement, job role, region). Address disparities via threshold adjustment, re-tuning, or recalibration.
* **Privacy & governance:** Restrict access by role; log predictions, actions, and outcomes; position outputs as **supportive indicators**, not disciplinary triggers.

# Next Steps

## Operating Playbook

* **Trigger:** Employee exceeds risk threshold during weekly scoring.
* **Triage:** HR reviews context (hours, leave, survey signals) within 3 business days.
* **Action menu:** workload rebalancing, schedule flexibility, manager 1:1 meetings, wellbeing referral, ergonomic kit.
* **Follow-up:** 30-day wellbeing check; re-score after 60–90 days.

## KPIs & Guardrails (Dashboard)

* **Model:** Recall (High), precision, ROC AUC (for reference), alert volume, time-to-action.
* **Business:** Attrition in targeted groups, absence days per FTE, ergonomic & mental-health tickets, engagement item scores.
* **Fairness:** Metric parity across arrangement/role/region; action parity.
* **Governance:** Access logs, reason codes, outcome tracking.

## Communications Plan

* Manager **outreach script** emphasizing support and resources.
* Employee-facing **FAQ** on purpose, data use, and available help.
* Quarterly **executive summary** highlighting outcomes, risk reduction, and next steps.

