ISYS3447|ISYS3448: INTRODUCTION TO BUSINESS ANALYTICS

ASSIGNMENT 2 - GROUP PRESENTATION

VISUALIZATION AND PREDICTIVE ANALYSIS

EMPLOYEE SURVEY

Diep Yen Nhung - S3974867
Ly Tuong Vy - S4129695
Nguyen Thuy Hien - S4051203
Luc My Ha - S4119009
Nguyen Thanh Thuy Nhi - S4099021

"We declare that in submitting all work for this assessment we have read, understood and agree to the content and expectations of the Assessment declaration."

AGENDA

01
Introduction

02
Problem Statement

03
Methodology & Data Pre-processing

04
Finding & Discussion

05
Recommendation

06
Conclusion

I. INTRODUCTION

BUSINESS PROBLEM

Organisations experience rising turnover & workforce costs, but a lack of data-driven insights

Struggle to design sustainable workloads, target interventions, and prevent avoidable attrition.

IMPACT

OVERTIME

With 36–42-hour workweeks linked to 54% of departures and 23% of employees at risk of leaving (Tiwari S. 2025; Aswale and Mukul 2020)

=> Overtime emerges as a driver of employee turnover risk and workforce instability.

JOB SATISFACTION

- Job dissatisfaction increases turnover intention and productivity loss
- Dissatisfaction reflects failures in workload design, work environment, and balance, but key determinants remain unclear.

=> Requiring predictive models for targeted interventions

OBJECTIVES

OVERTIME

Identify key factors driving overtime risk and quantify employees' likelihood of working overtime to support targeted workforce interventions.

JOB SATISFACTION

Determine the key drivers of job satisfaction, and identify where improvements in individual attributes workload, and work place fail to translate into higher satisfaction.

II. METHODOLOGY

GGPLOT2 PACKAGE	To create four visualisations and explore relationships between JobSatisfaction and other candidate predictors
GLM() FUNCTION	To model the Have_OT binary logistic regression and examine HaveOT determinant predictors
POLR() FUNCTION	To model the JobSatisfaction ordinal logistic regression and examine JobSatisfaction determinant predictors
LIKELIHOOD RATIO TESTS (LRTS)	To assess each predictor's marginal contribution and test whether a more complex model fits the data significantly better than a simpler, nested model
CONFUSION MATRIX	To evaluate the HaveOT binary logistic regression model and JobSatisfaction ordinal logistic regression performance

III. DATA PREPROCESSING

VARIABLE RESTRUCTURE

The dataset was reclassified into numerical, nominal, and ordinal variables to preserve information content, enhance interpretability, and improve model accuracy

(Islam and Dhanekula, 2024)

ERROR DETECTION

No duplicate entries or orthography mistakes were detected

MISSING VALUES

JobLevel (13.02%) exhibited the highest missingness, followed by **EduLevel (8.63%)**, while other variables showed minimal missingness **(0.10–0.53%)**. The missingness is most consistent with MCAR or weak MAR, with no evidence of MNAR

Variable Type	Method	Rationale
Nominal categorical	kNN	Superior accuracy under low missingness (Memon et al. 2023; Paczkowski 2021:127–157)
Numeric	MICE – PMM	Avoids distributional distortion; realistic values (Nehler and Schultze 2025)
Ordinal	MICE – POLR	Preserves order; prevents mean inflation (Rigol et al. 2025; Emmanuel et al. 2021)
Binary	Logistic / Mode	Appropriate for binary structure (Nwakuya and Biu 2022)

Table 1: Imputation Strategy

III. DATA PREPROCESSING

OUTLIERS

Only **22 outliers (0.73%)** were identified via the IQR. As variability in Sleep_Duration and Physical_Activity is expected across populations, no outliers were removed

(Predo et al. 2012; Nicola et al. 2019)

SKEWNESS

The Shapiro-Wilk test was conducted → All variables returned **p-values <0.05** due to the test's sensitivity in large samples, rejecting the null hypothesis of normality despite minor deviation

(Ahad et al. 2011; Kelvin 2026)

DISTRIBUTION

Q-Q plots and box plots confirming only mild deviations from symmetry and minimal impact on modelling

(Adam et al 2014; Mohit 2025)

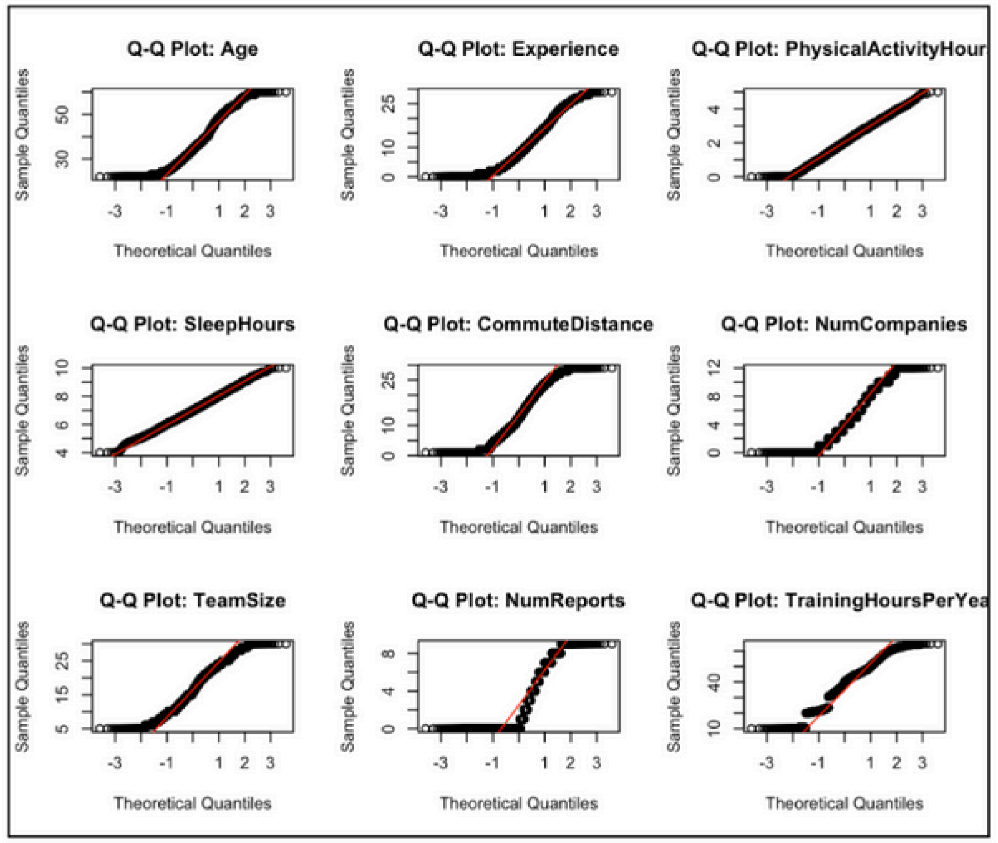


Figure 1: Q-Q Plots (Before modification)

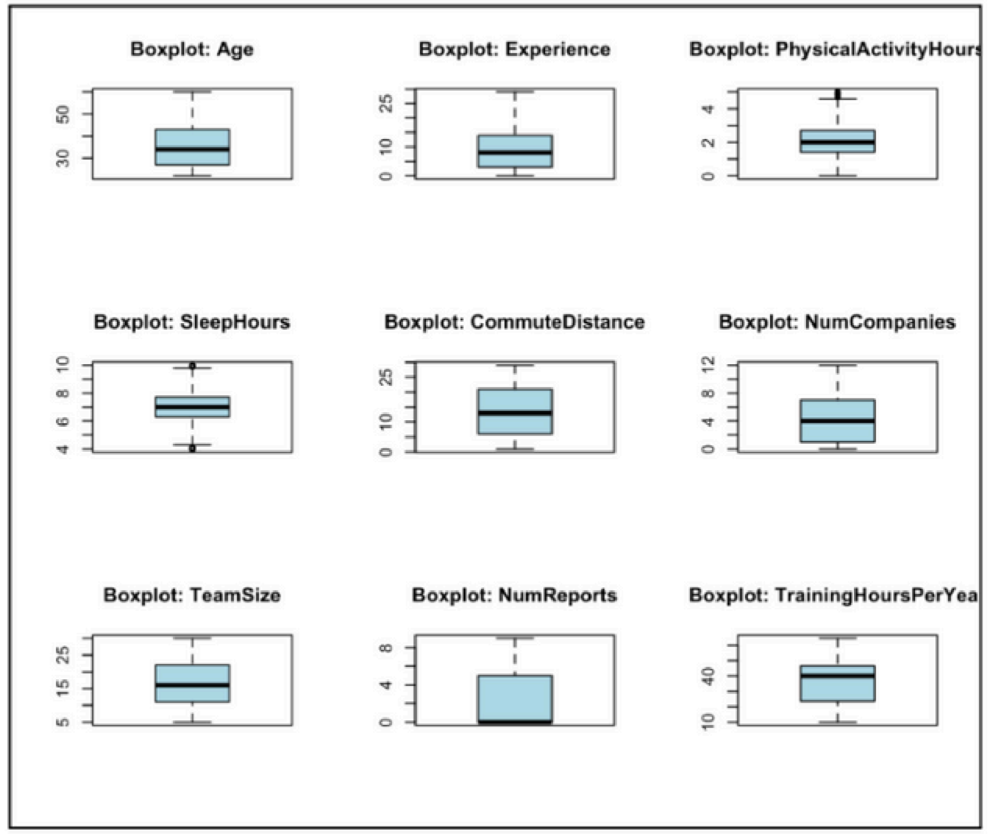


Figure 2: Boxplots (Before modification)

III. DATA PREPROCESSING

SKEWNESS FIXING

Square-root transformations were applied for mildly right-skewed variables like **Experience (0.73)** and **Num_Companies (0.53)**, and **zero-inflated Num_Reports (0.91)**, where log transformations are inappropriate

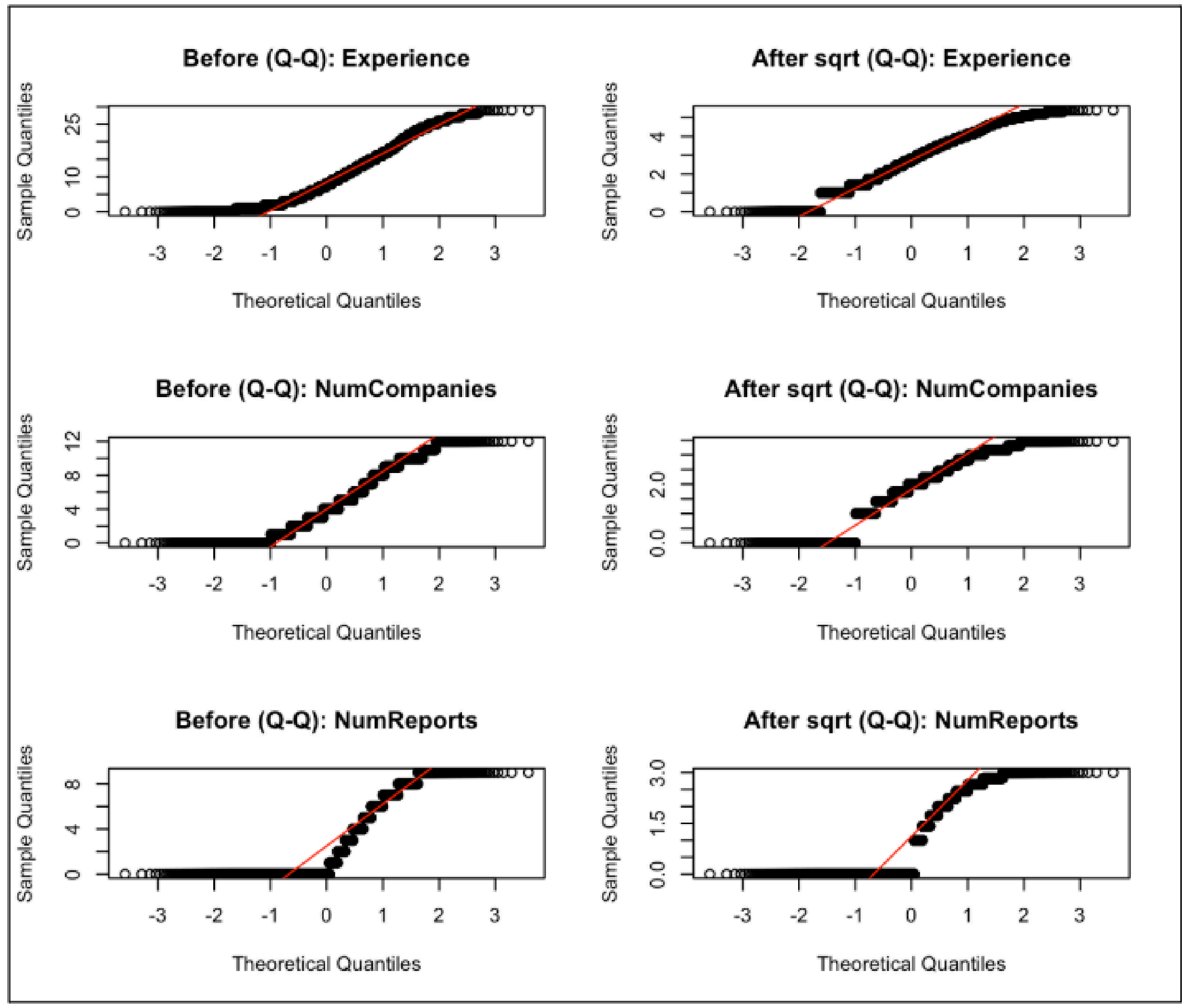


Figure 3: Q-Q Plots before and after transformation

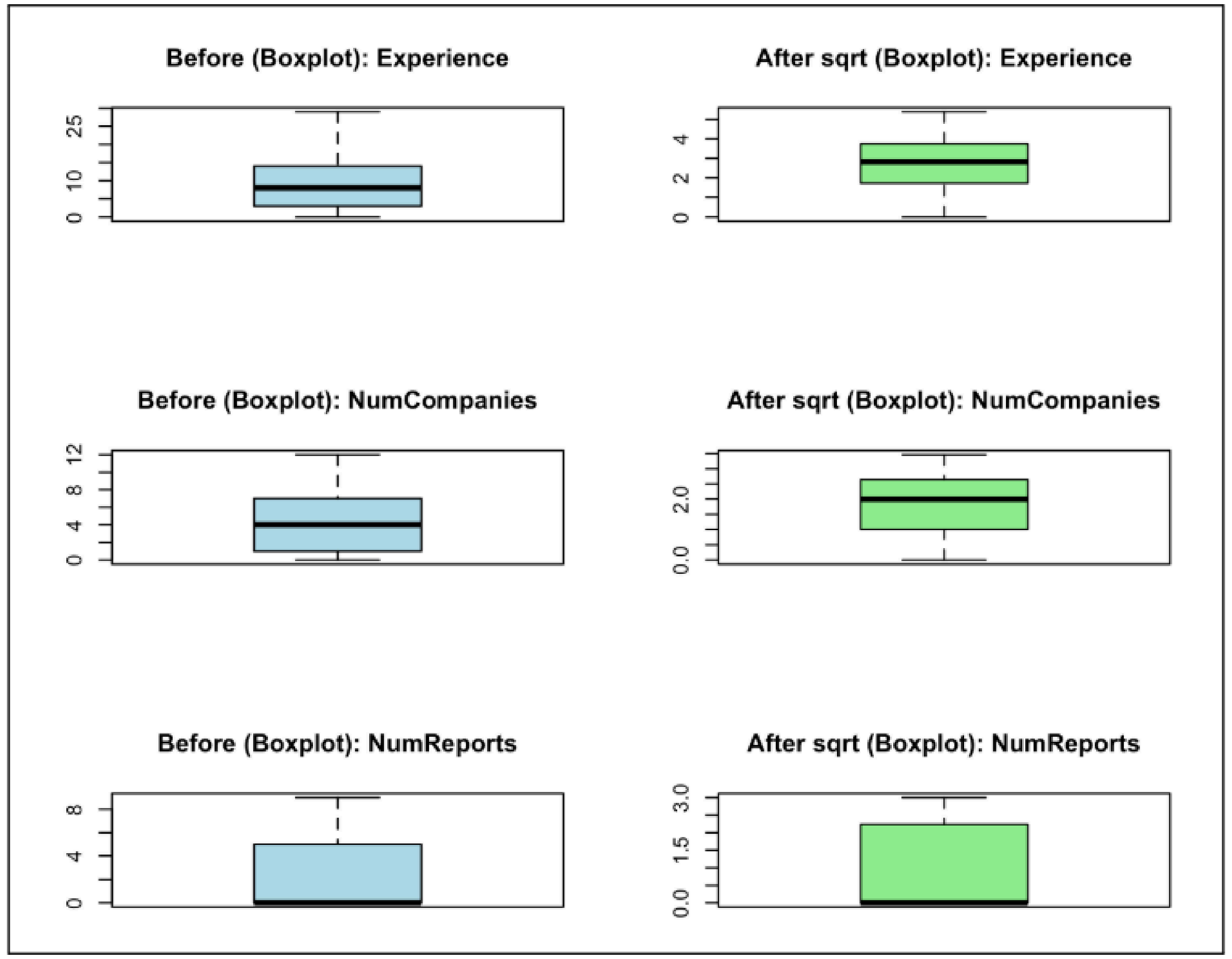


Figure 4: Boxplots before and after transformation

RESULTS

Skewness was substantially reduced, with **Experience (-0.15)** and **Num_Companies (-0.40)** approaching symmetry, and **Num_Reports (0.40)** showing a meaningful reduction despite remaining moderately skewed

IV. DESCRIPTIVE ANALYSIS

We choose **Job Satisfaction** as **Dependent Variable** as it has meaningful correlation relationship with important job-related variables like WorkLifeBalance, WorkEnv, v.v

	Variable	Test_Used	X² / F Value	p-value	Significant
1	Age	ANOVA	111	739	No
2	Experience	ANOVA	468	494	No
3	PhysicalActivityHours	ANOVA	256	613	No
4	SleepHours	ANOVA	**101.253**	**0**	**Yes**
5	CommuteDistance	ANOVA	6	939	No
6	NumCompanies	ANOVA	33	856	No
7	TeamSize	ANOVA	30	863	No
8	NumReports	ANOVA	1.086	297	No
9	TrainingHoursPerYear	ANOVA	84	772	No
10	Gender	Chi-square	5.028	755	No
11	MartialStatus	Chi-square	14.15	291	No
12	Dept	Chi-square	36.568	129	No
13	Emptype	Chi-square	**16.173**	**40**	**Yes**
14	CommuteMode	Chi-square	14.851	536	No
15	HaveOT	Chi-square	**95.715**	0	**Yes**
16	JobLevel	Chi-square	**95.715**	0	**Yes**
17	WorkLifeBalance	Chi-square	**271.097**	0	**Yes**
18	WorkEnv	Chi-square	**298.626**	0	**Yes**
19	Workload	Chi-square	**266.461**	0	**Yes**
20	Stress	Chi-square	**196.244**	0	**Yes**
21	EduLevel	Chi-square	**221.344**	0	**Yes**

Table 2: ANNOVA and Chi-squared test results

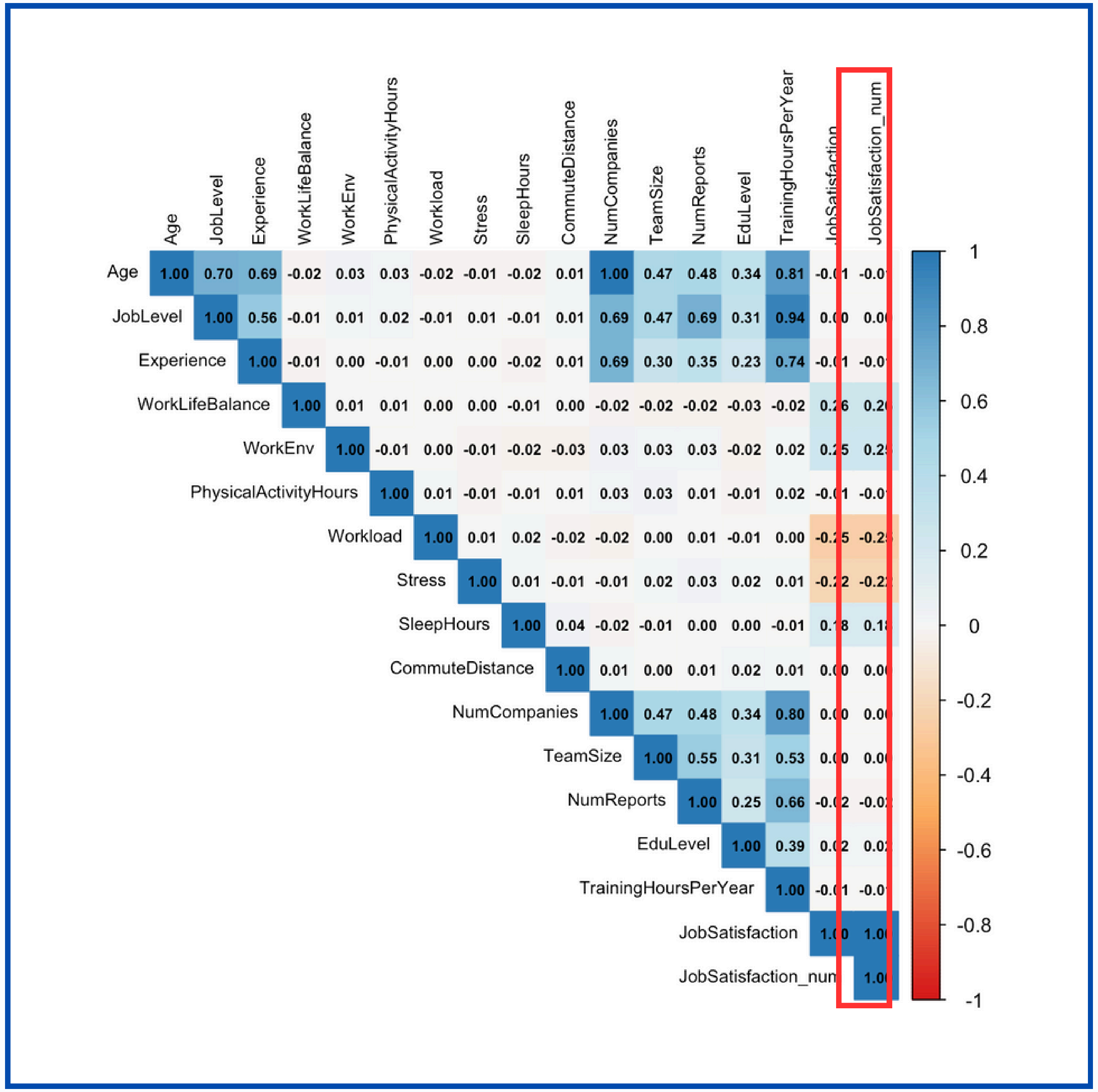


Figure 5: Correlation Matrix

Scenario 1 | Work Environment works as the hygiene factor at Level 3, with little or no changes to job satisfaction if increasing higher

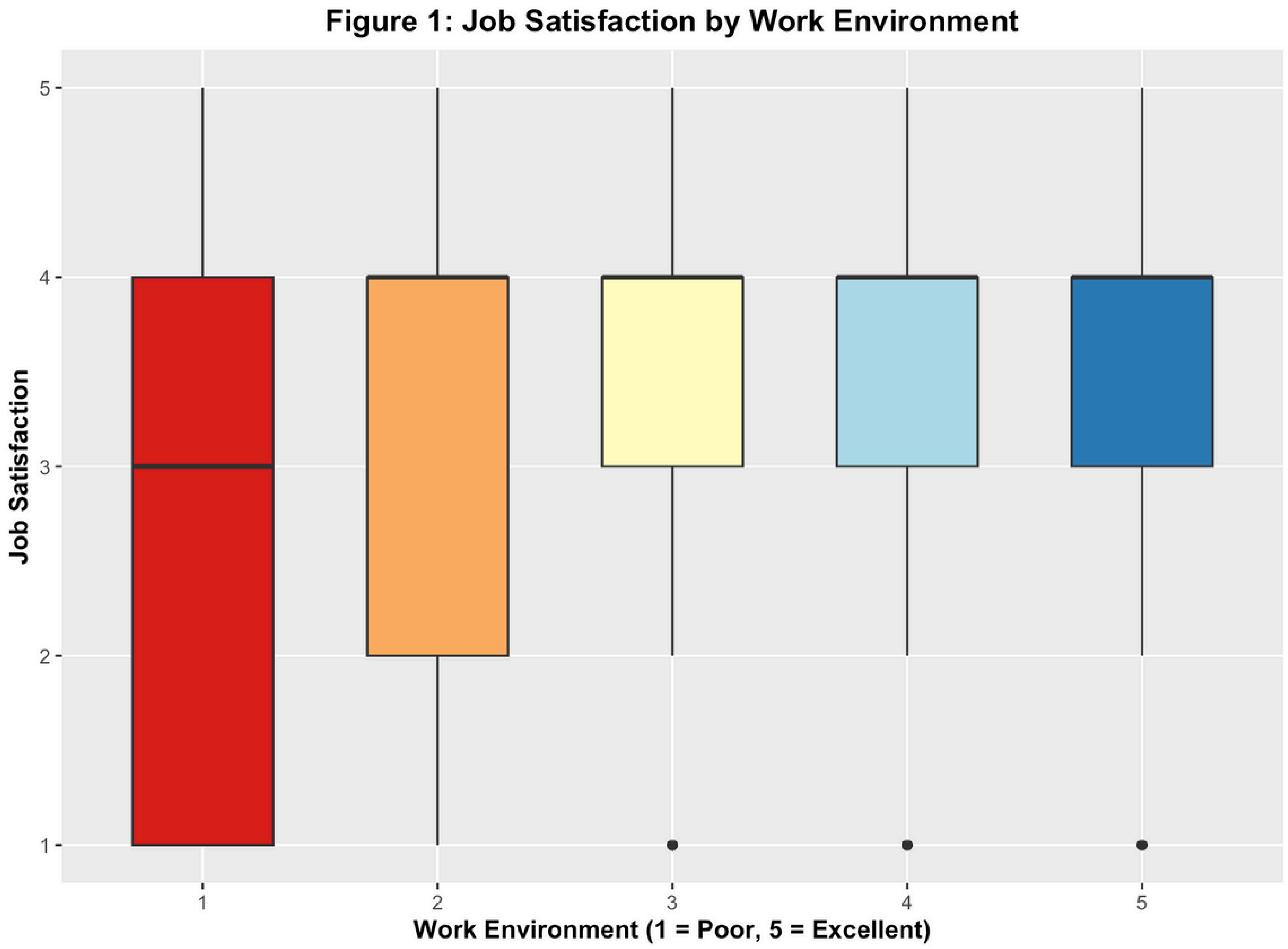


Figure 6: JobSatisfaction vs WorkEnvironment

STATISTICAL SIGNIFICANCE

Work environment significantly influences job satisfaction **(χ^2 = 298.626, p < 0.001)**

KEY FINDINGS

- Poor environments (ranks 1-2) leads to substantially lower satisfaction
- Most significant improvement at middle work environment level (rank 3)
- **HOWEVER,** beyond rank 3 witnesses minimal to little change, which means that other factors drive satisfaction

This aligns with **Herzberg's two-factor theory**: work environment as hygiene factor preventing dissatisfaction with diminishing returns beyond baseline (Lundberg et al. 2009; Tyll et al. 2025)

ACTIVITY PATTERN

HR should **prioritize improving substandard environments** to acceptable levels, and focus on other factors when acceptable levels is reached (Imran et al. 2024)

Scenario 2 | Overtime negatively impact job satisfaction across all employment types

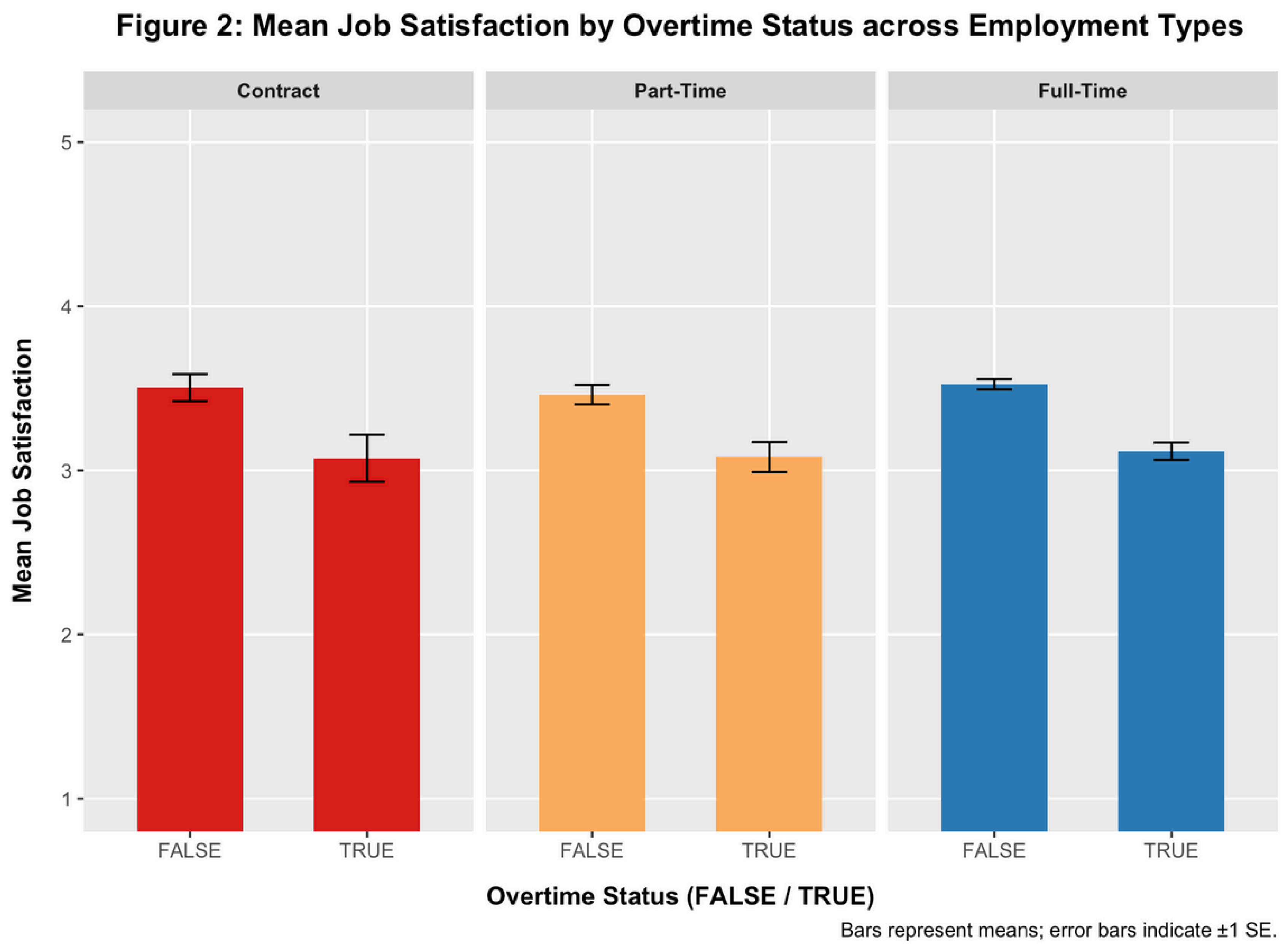


Figure 7: JobSatisfaction vs HaveOT across EmpType

STATISTICAL SIGNIFICANCE

- Overtime consistently associated with lower job satisfaction across all employment types
- HaveOT: **$\chi^2 = 95.715$, $p < 0.001$**
- EmpType: **$\chi^2 = 16.173$, $p = 0.040$**

KEY FINDINGS

- Employees with overtime experience a uniform decline of **0.4 Job Satisfaction points** across all employment types
- Employment type does not buffer overtime's negative impact (Mammadzada 2025)
- Overtime undermines WLB and increases burnout (Barck-Holst et al. 2020)
- Stress is a primary pathway linking overtime to reduced satisfaction (Gupta and Singh 2025)

ACTIVITY PATTERN

Better overtime allocation across all employee types is suggested to ensure better job satisfaction.

Scenario 3

Work-life balance positively impacts job satisfaction across all levels, but shows limited influence on moderate dissatisfaction (levels 2-3)

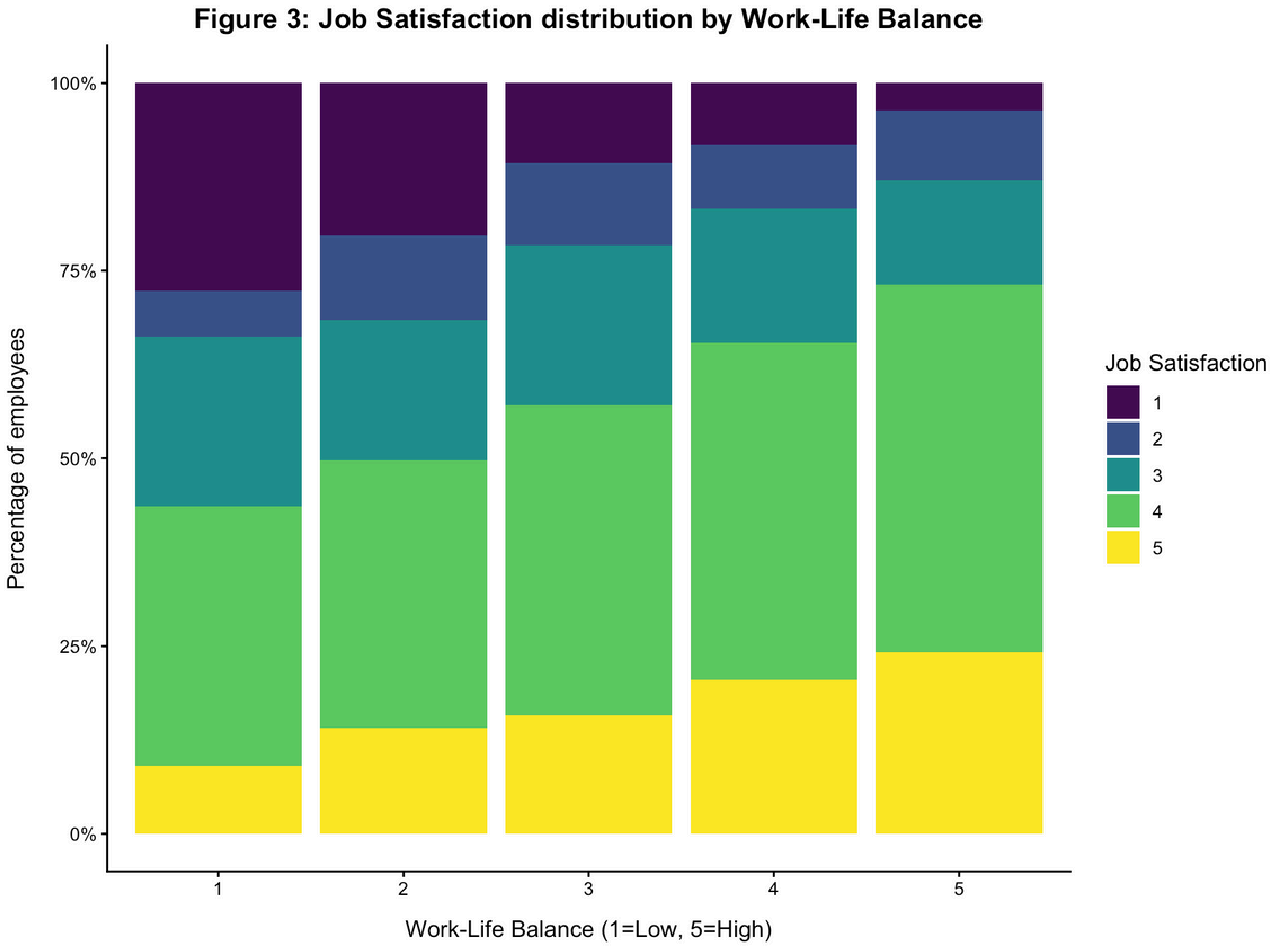


Figure 8: JobSatisfaction vs WorkLifeBalance

STATISTICAL SIGNIFICANCE

Chi-square test **(χ^2 = 271.097, $p < 0.001$)**

→ Systematically influenced by WLB.

KEY FINDINGS

Job Satisfaction Response to WLB

- Level 1: **declines** notably as WLB improves
- Levels 2-3: **persists** regardless of WLB level
- Levels 4-5: **increases** significantly with better WLB

→ **Moderate Job Satisfaction persists even with good WLB**

ACTIVITY PATTERN

- Moderate Job satisfaction persists without a supportive, non-toxic work environment requiring holistic approaches addressing both WLB and workplace culture (Muhamad and Ajmal 2021).
- Work-life balance is insufficient without supportive work environment

Scenario 4

Workload negatively impacts job satisfaction, but effect intensifies sharply under high stress

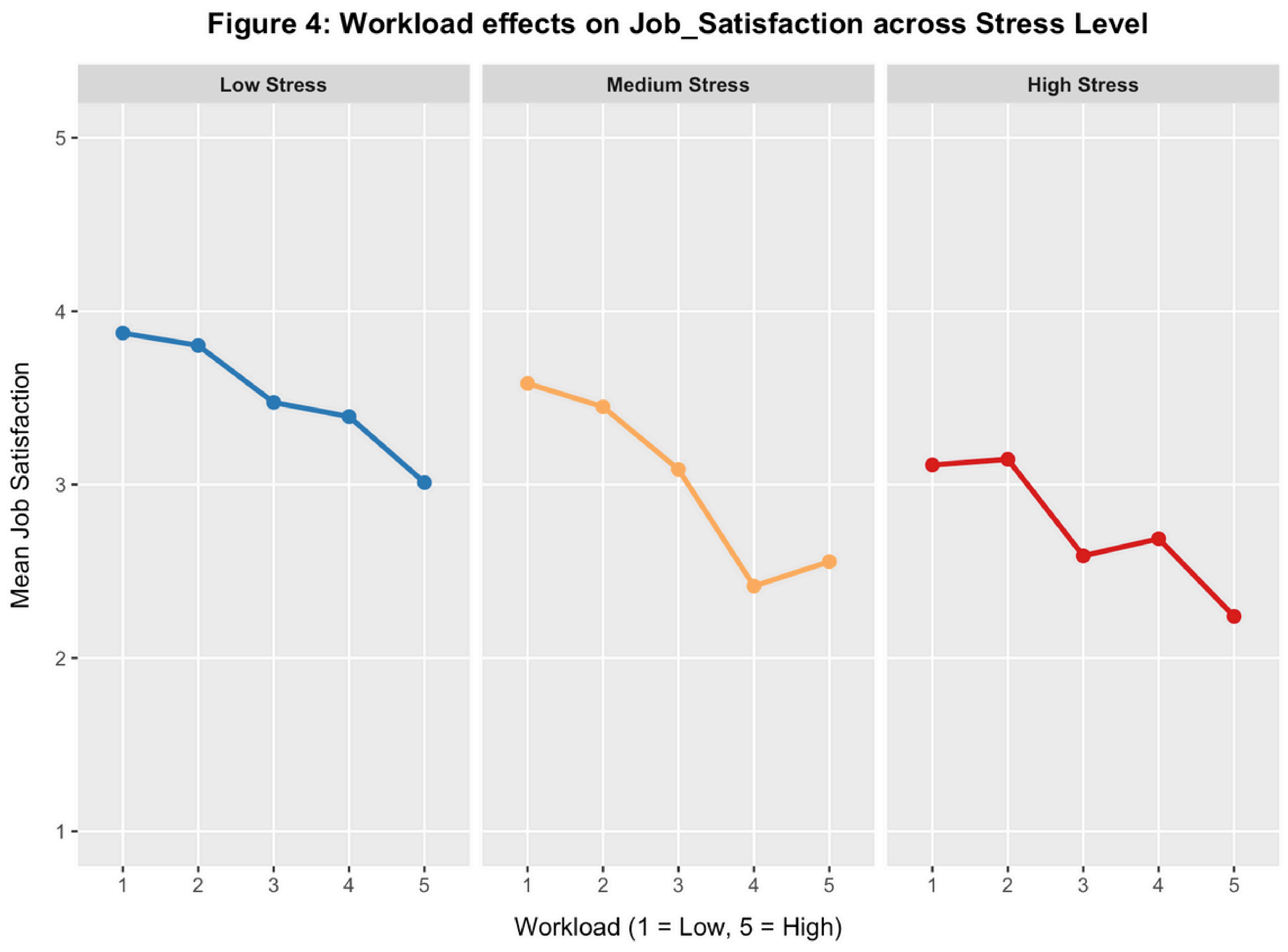


Figure 9: JobSatisfaction vs Stress

STATISTICAL SIGNIFICANCE

Chi-square tests for both Workload and Stress are highly significant

- **Workload**: χ^2=266.461 ; p-value < 0.001
- **Stress**: χ^2=196.244; p-value < 0.001

KEY FINDINGS

Workload alone has minimal impact on satisfaction, but when coupled with high stress, **satisfaction will decline sharply**

→ **Stress is the key amplifier of dissatisfaction under high workload** (Kusuma et al. 2021; Kittisak et al. 2021).

ACTIVITY PATTERN

Integrating stress management and workload optimization strategies is suggested, as high stress amplifies the negative impact of heavy workload on job satisfaction

V. PREDICTIVE ANALYSIS

MODEL DEVELOPMENT - HAVEOT

A **full binary logistic regression model** was estimated using all candidate predictors to identify significant predictors to HaveOT

Objective: to identify the key drivers of employees overtime

Appendix 1

Predictor	Coefficient
Job_Satisfaction	$\beta = -0.404$, $p < 2e-16$
Dept(IT)	$\beta = 3.156$, $p < 2e-16$
Num_Reports	$\beta = 0.066$, $p = 0.0202$

Table 3: Significant predictors to Have_OT

Log odds and odd ratios are used since log odds are the correct inferential scale, and odd ratios enable a clearer multiplicative interpretation

Predictor	OR	Interpretation
Job_Satisfaction	0.73	Each one-point increase in Job_Satisfaction reduces the odds of overtime by approximately 33%, independent of department
Dept(IT)	23.5	Employees in IT have over 23 times higher odds than the Customer_Service employees (reference group) → IT's job design features drive overtime
Other Dept	Near unity with confidence intervals spanning 1	No statistically distinguishable difference from Customer_Service
Num_Reports	0.0656	Each additional direct report increases the odds of having OVT by about 6–7%, holding all other variables constant

Table 4: Significant predictors to HaveOT

JOBSATISFACTION

Statistical result
- $\beta = -0.404$, $p < 2e-16$
- OR = 0.73

JobSatisfaction retains a statistically significant independent negative association with HaveOT, though at a smaller effect

DEPT (IT)

Statistical result
- $\beta = 3.156$, $p < 2e-16$
- OR = 23.5

Dept (IT) exhibits an exceptionally large and highly significant effect

NUMREPORTS

Statistical result
- $\beta = 0.066$, $p = 0.0202$
- OR = 0.0656

NumReports shows a positive and statistically significant effect, but modest compared to JobSatisfaction

OTHER VARIABLES

Gender (Male) ($p = 0.471$)
Age ($p = 0.828$)
Marital Status (p-values > 0.75)
Remaining variable (p-value > 10)

Do not exhibit statistically significant associations with HaveOT once Dept was controlled for

HaveOT is predominantly a structural phenomenon embedded in departmental job design

MODEL DEVELOPMENT - HAVEOT

A **likelihood ratio test (LRT)** was applied to assess each predictor's marginal contribution

	LRT
Dept	898.48, $p < 2e{-}16$
JobSatisfaction	74.06, $p < 2e{-}16$
Num_Reports	5.42, $p = 0.0199$

Table 5: Likelihood Ratio Test

DEPT

ΔDeviance = 898.48, p < 2e-16

AIC increases from 2714.6 to 3599.0

Dept is the dominant structural driver of overtime

Removing Dept would cause a substantial loss of explanatory power

JOBSATISFACTION

ΔDeviance = 74.06, p < 2e-16

AIC increases to 2786.6

Removing JobSatisfaction_num results in a loss of explanatory power, though the effect is smaller than Dept

NUMREPORTS

ΔDeviance = 5.42, p = 0.0199

NumReports has a relatively small influence on overtime due to its comparatively small effect

OTHER VARIABLES

JobLevel (ΔDeviance = 7.32, p = 0.062)
Workload (ΔDeviance = 8.81, p = 0.066)
Remaining variables (all p-values > 0.10)

Their removal yields minimal changes in deviance and AIC

Including them would add complexity without substantive explanatory benefit

The final model will retain only Dept, JobSatisfaction, and NumReports for **parsimony** and **structural coherence**

BUSINESS MODEL - HAVEOT

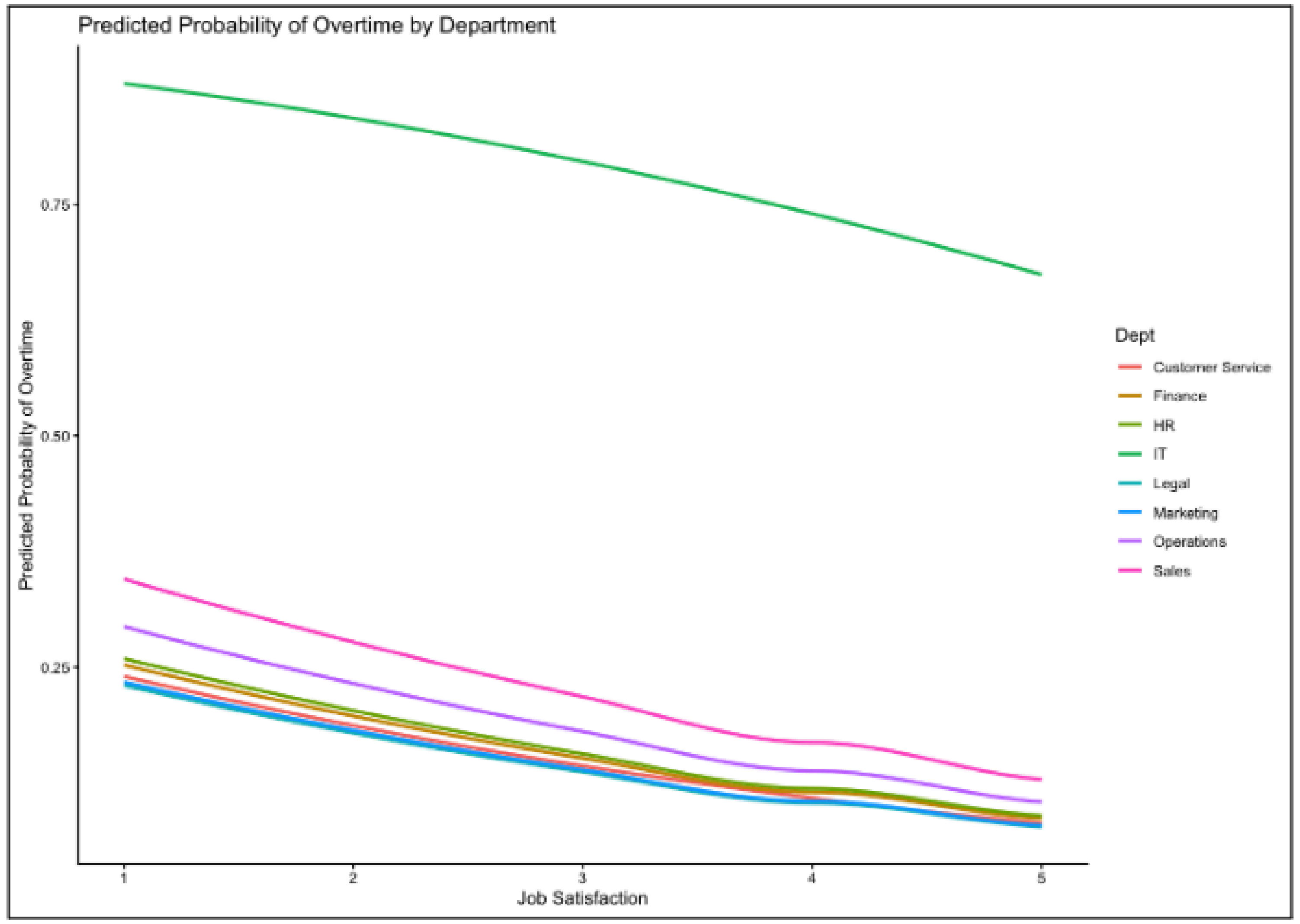


Figure 10: HaveOT final model

* Note: Through independent variables interaction testing, no interaction among variables found

```
>                          lrt_summary
                         Interaction..Dept.JobSatisfaction_num. Interaction..Dept.NumReports. Interaction..JobSatisfaction_num.NumReports.      Chisq df   p_value
Dept:JobSatisfaction_num                   Dept:JobSatisfaction_num              Dept:NumReports              JobSatisfaction_num:NumReports 10.3582357  7 0.1691651
Dept:NumReports                            Dept:JobSatisfaction_num              Dept:NumReports              JobSatisfaction_num:NumReports  6.2243647  7 0.5138095
JobSatisfaction_num:NumReports             Dept:JobSatisfaction_num              Dept:NumReports              JobSatisfaction_num:NumReports  0.2626751  1 0.6082883
> |
```

Figure 11: HaveOT interaction checking

	Full Model	Final Model
Null Deviance	3749.2	3749.2
Residual Deviance	2583.9	2664.1
AIC	2689.9	2684.1
Number of variables	>20	3

Figure 12: Model comparison

FULL MODEL

- Achieves a lower residual deviance (2583.9), but this improvement adds complexity
- Suffers from estimation instability, indicating multicollinearity and reduced interpretability

FINAL MODEL

- Performs better, with a lower AIC (2684.1) than the full model (2689.9), indicating more model efficiency

FINAL BUSINESS MODEL

Logit(P(HaveOT=1)) = -0.857 + 0.068(Dept_Finance) + 0.103(Dept_HR) + 3.156(Dept_IT) -0.052(Dept_Legal) - 0.036(Dept_Marketing) 0.274(Dept_Operations) +0.514(Dept_Sales) - 0.318(Job Satisfaction) + 0.009(NumReports)

MODEL EVALUATION

CONFUSION MATRIX

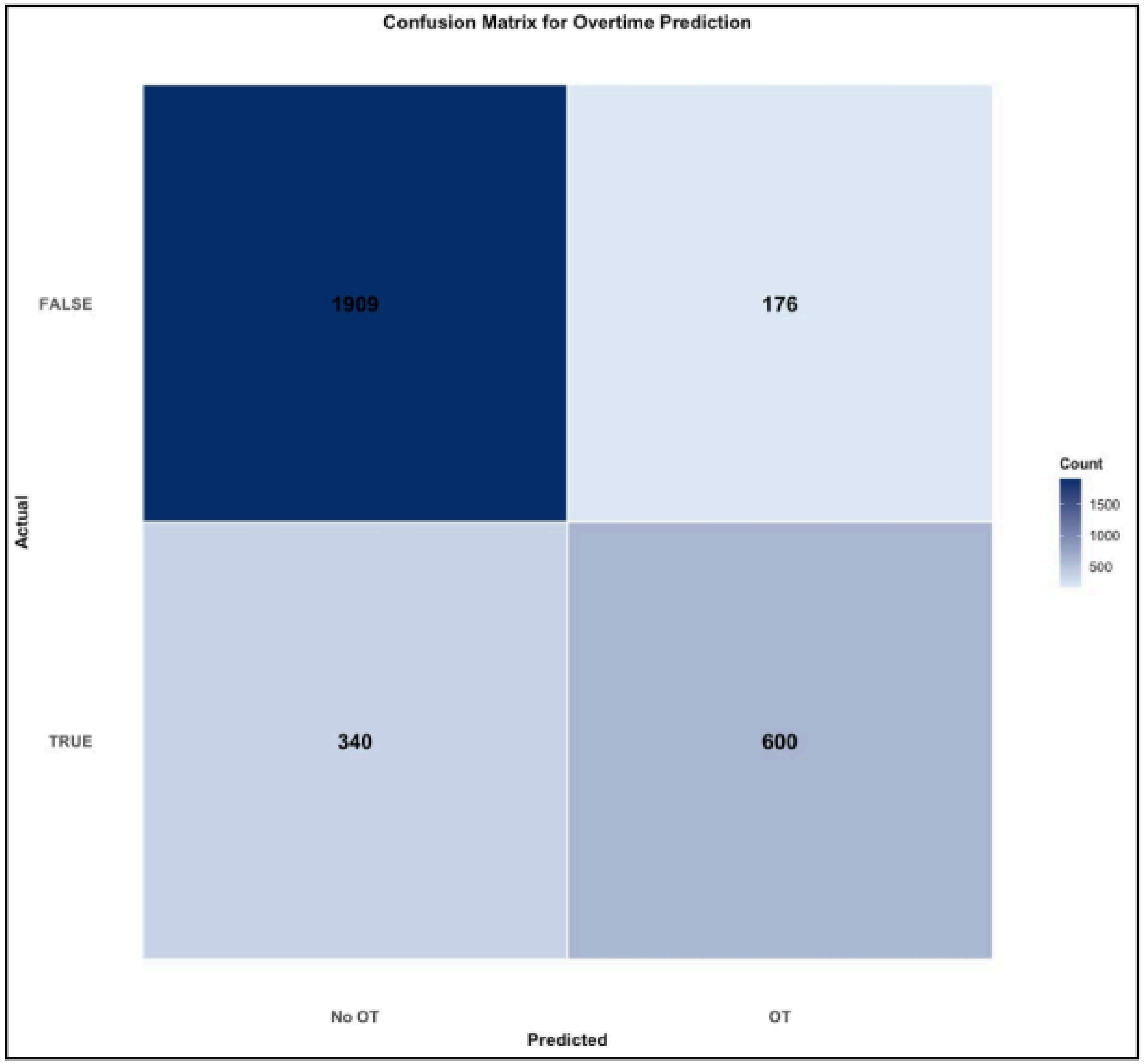


Figure 13: Confusion matrix for HaveOT prediction

KEY METRICS ASSESSED

- Overall Accuracy: **82.94%**
- Specificity: **91.56%**
- Recall: **63.83%**
- Precision: **77.32%**

INSIGHT

- High accuracy **(82.94%)** specificity **(91.5%)** → reliable in ruling out overtime and prevents unnecessary managerial intervention
- Moderate recall **(63.83%)** → over ⅓ of true overtime cases missed which limits early risk detection
- High precision **(77.32%)** → reflects a conservative classification strategy that prioritises reducing false positives and yields high-confidence predictions

1. MODEL DEVELOPMENT

MODEL OVERVIEW & VARIABLE SELECTION

Objective

- Identify key drivers of Job Satisfaction using ordinal logistic regression

Key Results (Full Model)

- **Strong positive predictors:**
  WorkLifeBalance **(β = 1.245, t > 15)**
  WorkEnvironment **(β = 1.312, t > 15)**

- **Strong negative predictors:**
  Workload **(β = -1.289)**
  Stress (β = -1.336)
  Overtime (HaveOT) **(β = -0.804, t = -8.67)**

Key Control Effects

- SleepHours: positive effect **(β = 0.457, t = 12.74)**
- Age, Education: significant negative effects

2. MODEL REFINEMENT

LRT-BASED SELECTION & INTERACTION TESTING

LRT-Based Variable Selection

- Predictors retained **(Pr(>Chi) ≤ 0.05)**
- **Selected Independent Variables:**
  WorkLifeBalance, WorkEnv, Workload, Stress, HaveOT
- **Controls Variables:**
  Age, MaritalStatus, SleepHours, NumCompanies, EduLevel

(Appendix 1)

→ Improved parsimony & interpretability

Model	Predictors Included	Residual Deviance	AIC
Full Model	All variables	**7,780.87**	**7,886.87**
Final Model (LRT-refined)	WorkLifeBalance, WorkEnv, Workload, Stress, HaveOT, Control Variables	**7,807.01**	**7,865.01**

Table 6: Model comparison

Increase Job Satisfaction (Positive Drivers)	Decrease Job Satisfaction (Negative Drivers)
Work Environment (β = 1.313, t = 16.51) → Strongest positive impact	Stress (β = -1.326, t = -8.71) → Most damaging factor
Work–Life Balance (β = 1.244, t = 16.01) → Critical for retention	Workload (β = -1.265, t = -15.82) → Major dissatisfier
Sleep Hours (β = 0.461, t = 12.90) → Wellbeing significantly matters	Overtime (HaveOT) (β = -0.714, t = -9.55) → Burnout indicator

Table 7: Key Drivers of Job Satisfaction

2.MODEL REFINEMENT

Business Model

Job Satisfaction Score = 1.244(WorkLifeBalance)+1.313(WorkEnvironment) -1.265(Workload) -1.326(Stress) -0.714(Overtime)-0.078(Age) +0.461(SleepHours)+0.235(Number of Companies) -0.384(Education Level – Linear) -0.593(Education Level – Quadratic)

** Note:Through independent variables interaction testing, no interaction among variables found*

Interaction Testing

- All pairwise interactions tested using LRT
- No interaction reduced AIC **(Final AIC = 7,865.164)**

→ Effects are stable and additive, not conditional

```
                  Interaction Chisq df p_value AIC_base AIC_int Delta_AIC                                              Status
1      WorkLifeBalance:WorkEnv    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
2     WorkLifeBalance:Workload    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
3       WorkLifeBalance:Stress    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
4       WorkLifeBalance:HaveOT    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
5   WorkLifeBalance:SleepHours    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
6             WorkEnv:Workload    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
7               WorkEnv:Stress    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
8               WorkEnv:HaveOT    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
9           WorkEnv:SleepHours    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
10             Workload:Stress    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
11             Workload:HaveOT    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
12         Workload:SleepHours    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
13               Stress:HaveOT    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
14           Stress:SleepHours    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
15           HaveOT:SleepHours    NA NA      NA 7865.164      NA        NA Fit_error: arguments imply differing number of rows: 1, 0
>                 keep_interactions
[1] Interaction Chisq       df          p_value     AIC_base    AIC_int     Delta_AIC   Status      Keep
<0 rows> (or 0-length row.names)
```

Figure 14: JobSatisfaction Interaction Checking

Threshold	Value	Boundary Interpretation
τ_1	-0.531	Between Job Satisfaction 1 and ≥ 2
τ_2	0.245	Between Job Satisfaction ≤ 2 and ≥ 3
τ_3	1.368	Between Job Satisfaction ≤ 3 and ≥ 4
τ_4	3.77	Between Job Satisfaction ≤ 4 and 5

Table 8: Estimated Threshold Values

3.MODEL EVALUATION

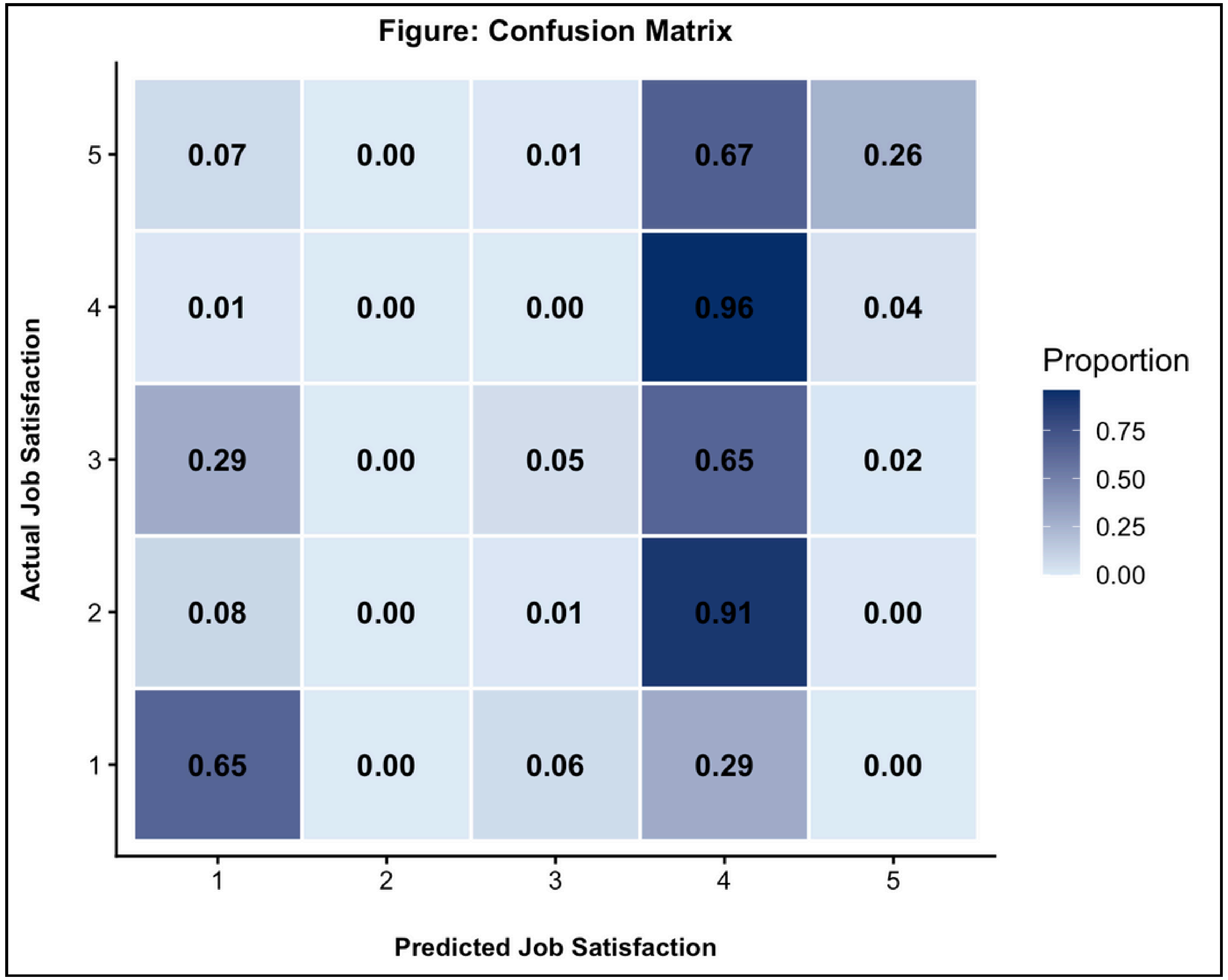


Figure 15: Confusion Matrix

KEY METRICS ACCESSED

- Overall Accuracy: **53.85%**
- Statistically Significant: **$p < 2.2e{-16}$**

(better than baseline)

Strengths	Limitations
• Strong prediction for Level 4 (Sensitivity ≈ 96%)	• Weak discrimination for mid-range levels (2–3)
• High accuracy for Level 5 (High satisfaction)	• Level 2 sensitivity: 0.00
• Effective at identifying conditions leading to high job satisfaction	• Level 3 sensitivity: ~0.025
• Suitable for managerial decision-making focused on positive outcomes	• Substantial overlap between adjacent satisfaction categories

4.MODEL INTERPRETATION - JOBSATISFACTION

Descriptive Analytics Insight

When work environment reached the acceptable level (Level 3), Job Satisfaction is affected by another factors

Descriptive Analytics - Scenario 1

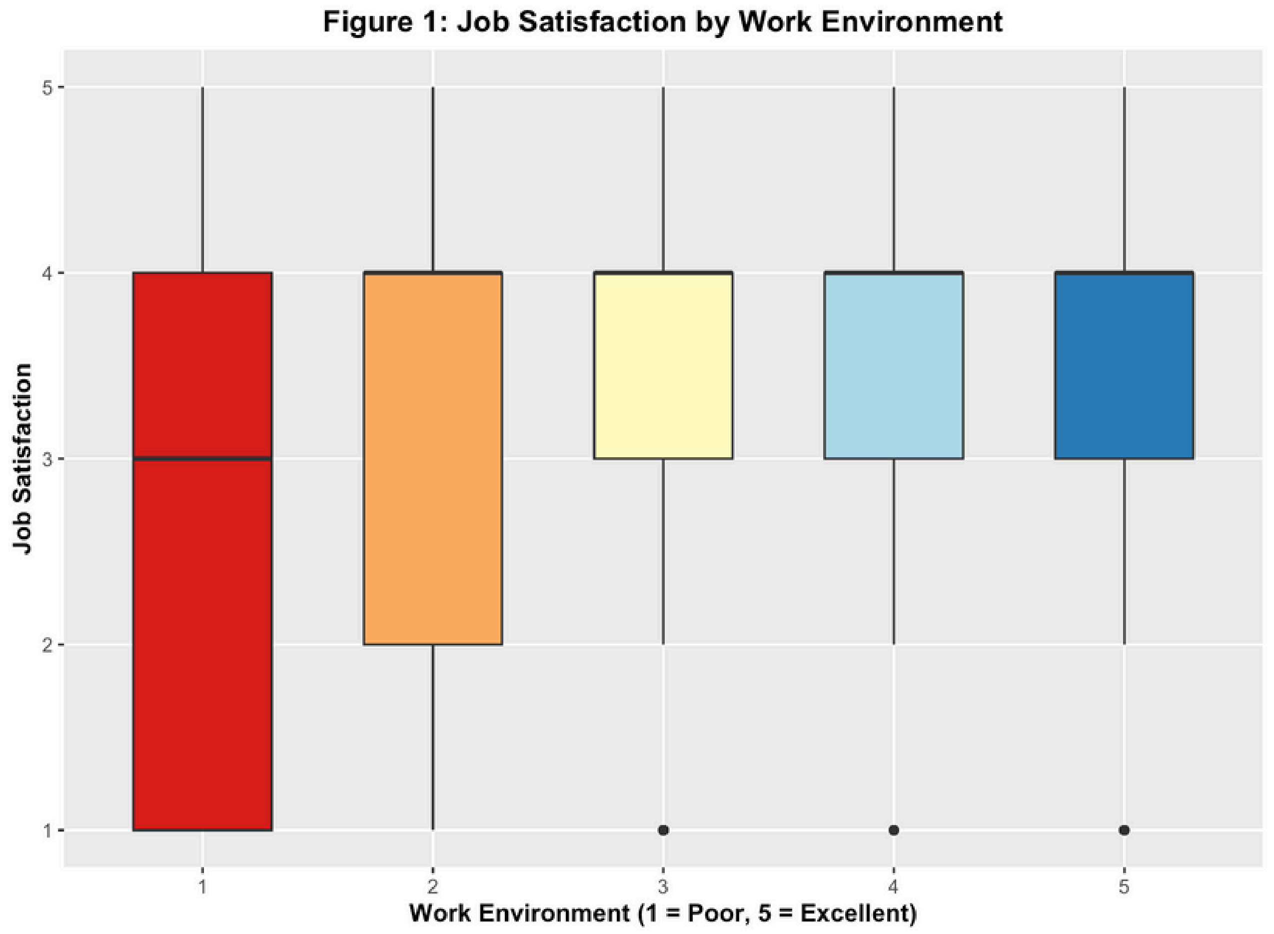


Figure 16: JobSatisfaction vs WorkEnvironment

Predictive Analytics Testing

Question: When the work environment reached acceptable leve (Level 3), what factors influence Job Satisfaction the most?

Testing methodology: Set the baseline scenario (WE = 3) and controls other non-job-related variables, testing through change one unit of selected job-related variables.

Predictive Analytics

Attributes	J0 Baseline	J1 Reduce OT	J2 Reduce Workload	J3 Reduce Stress	J4 Increase SleepHours	J5 Increase WLB
WorkLifeBalance	3	3	3	3	3	4
WorkEnvironment	3	3	3	3	3	3
Workload	3	3	2	3	3	3
Stress	3	3	3	2	3	3
HaveOT	TRUE (1)	FALSE (0)	TRUE (1)	TRUE (1)	TRUE (1)	TRUE (1)
SleepHours	7	7	7	7	8	7
Controls (Age, MaritalStatus, EduLevel, NumCompanies)	Fixed at median/mode	Fixed at median/mode	Fixed at median/mode	Fixed at median/mode	Fixed at median/mode	Fixed at median/mode

Table 9: Estimated Threshold Values

SCENARIO 1: REDUCE OT

Each one-unit reduction in OT increases the probability of higher JS by 45.58%.

Attribute	Baseline	J1 Reduce OT
WorkLifeBalance	3	3
WorkEnv	3	3
Workload	3	3
Stress	3	3
HaveOT	TRUE (1)	FALSE (0)
SleepHours	7	7
Controls (Age, MaritalStatus, EduLevel, NumCompanies)	Fixed at median/mode	

Table 10: Scenarios 10

Outcome Level	Baseline	J1 Reduce OT
JS=1	26.87%	15.19%
JS=2	17.49%	12.80%
JS=1	26.64%	26.42%
JS=4	25.43%	38.52%
JS=5	3.57%	7.06%
P (JS>= 4)	29.00%	45.58%
Conclusion	**JS= 1**	**JS= 4**

Table 11: Predicted probabilities

SCENARIO 2: REDUCE WORKLOAD

Each one-unit reduction in Workload increases the probability of higher JS by 41.98%.

Attribute	Baseline	J2 Reduce Workload
WorkLifeBalance	3	3
WorkEnv	3	3
Workload	3	2
Stress	3	3
HaveOT	TRUE (1)	TRUE (1)
SleepHours	7	7
Controls (Age, MaritalStatus, EduLevel, NumCompanies)	Fixed at median/mode	

Table 12: Scenarios 2

Outcome Level	Baseline	J2 Reduce Workload
JS=1	26.87%	17.18%
JS=2	17.49%	13.86%
JS=1	26.64%	26.98%
JS=4	25.43%	35.82%
JS=5	3.57%	6.16%
P (JS>= 4)	29.00%	41.98%
Conclusion	**JS= 1**	**JS= 4**

Table 13: Predicted probabilities

SCENARIO 3: REDUCE STRESS

Each one-unit reduction in Stress increases the probability of higher JS by 42.23%.

Attribute	Baseline	J3 Reduce Stress
WorkLifeBalance	3	3
WorkEnv	3	3
Workload	3	3
Stress	3	2
HaveOT	TRUE (1)	TRUE (1)
SleepHours	7	7
Controls (Age, MaritalStatus, EduLevel, NumCompanies)	Fixed at median/mode	

Table 14: Scenarios 3

Outcome Level	Baseline	J3 Reduce Stress
JS=1	26.87%	17.04%
JS=2	17.49%	13.79%
JS=1	26.64%	26.95%
JS=4	25.43%	36.01%
JS=5	3.57%	6.22%
P (JS>= 4)	29.00%	42.23%
Conclusion	**JS= 1**	**JS= 4**

Table 15: Predicted probabilities

SCENARIO 4: INCREASE SLEEPHOURS

Each one-unit increase in sleep hours increases the probability of higher JS by 39.38%.

Attribute	Baseline	J4 Increase SleepHours
WorkLifeBalance	3	3
WorkEnv	3	3
Workload	3	3
Stress	3	3
HaveOT	TRUE (1)	TRUE (1)
SleepHours	7	8
Controls (Age, MaritalStatus, EduLevel, NumCompanies)	Fixed at median/mode	

Table 16: Scenarios 1

Outcome Level	Baseline	J4 Increase SleepHours
JS=1	26.87%	18.76%
JS=2	17.49%	14.62%
JS=1	26.64%	27.23%
JS=4	25.43%	33.82%
JS=5	3.57%	5.57%
P (JS>= 4)	29.00%	39.38%
Conclusion	**JS= 1**	**JS= 4**

Table 17: Predicted probabilities

SCENARIO 5: INCREASE WLB

Each one-unit increase in worklife balance increases the probability of higher JS by 36.70%.

Attribute	Baseline	J5 Increase WLB
WorkLifeBalance	3	4
WorkEnv	3	3
Workload	3	3
Stress	3	3
HaveOT	TRUE (1)	TRUE (1)
SleepHours	7	7
Controls (Age, MaritalStatus, EduLevel, NumCompanies)	Fixed at median/mode	

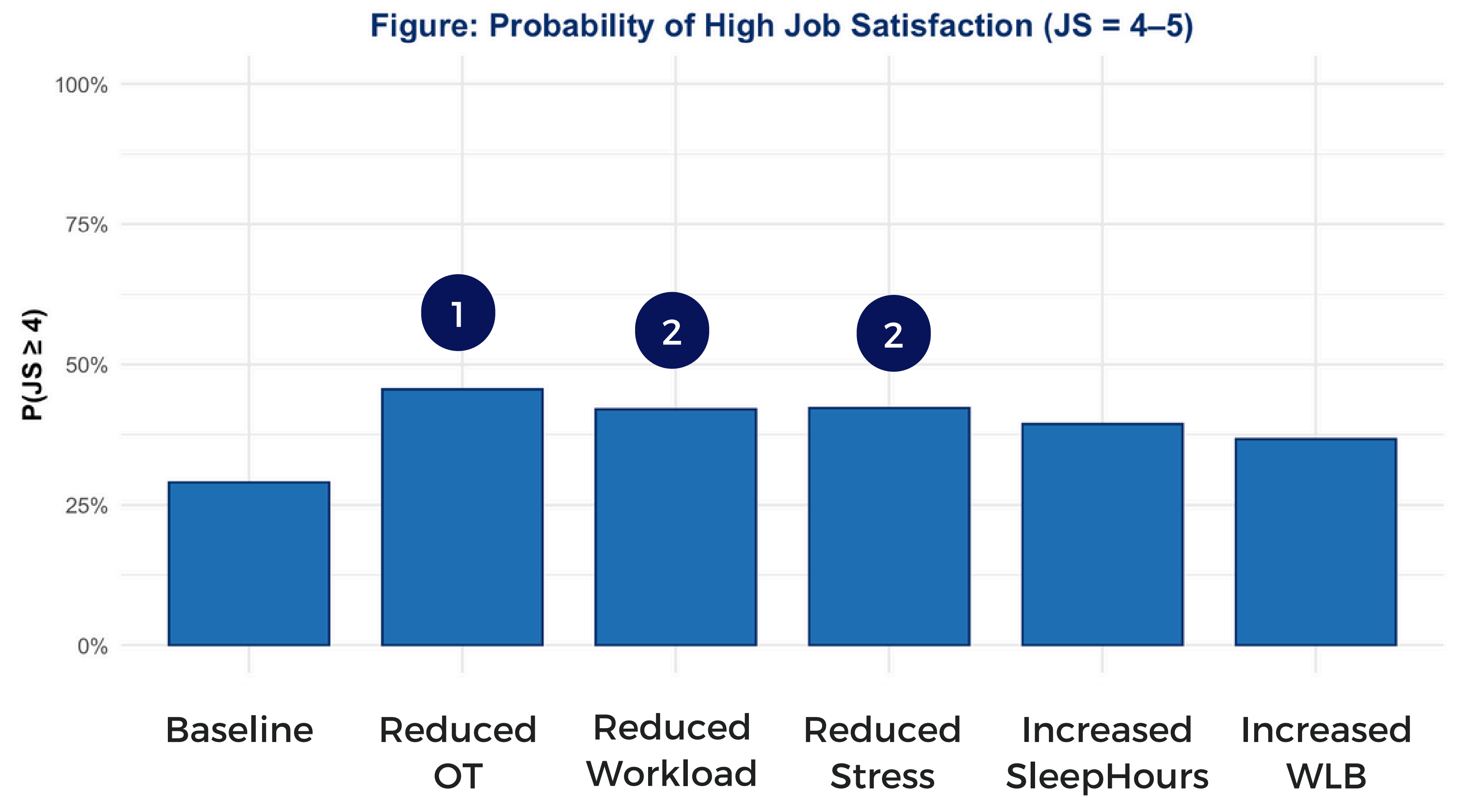
Table 18: Scenarios 5

Outcome Level	Baseline	J5 Increase WLB
JS=1	26.87%	20.57%
JS=2	17.49%	15.41%
JS=1	26.64%	27.33%
JS=4	25.43%	31.70%
JS=5	3.57%	5.00%
P (JS>= 4)	29.00%	36.70%
Conclusion	**JS= 1**	**JS= 4**

Table 19: Predicted probabilities

MODEL INTERPRETATION - JOBSATISFACTION

Predictive Analytics



Findings for testing:

When the **work environment reach the acceptable levels**, **to increase the job satisfaction,** HR should **focus on reduced working overtime the most**, and other job related-factor following this order

Reduced OT › Reduced Stress › Reduced Workload › Increased SleepHours › Increased WLB

Through this predictive pattern, we, then apply a suitable framework for HR to have better human resources management.

VI. RECOMMENDATIONS & ACTIONABLE PLAN

01 RECOMMENDATION 1: SIMON'S 4 LEVERS OF CONTROL FRAMEWORK

Simons' Levers of Control framework is applied to guide actionable managerial interventions that sustain job satisfaction despite high demand, by aligning values, boundaries, performance measurement, and learning mechanisms.

02 RECOMMENDATION 2: PPP FRAMEWORK

The model shows that job demands reduce job satisfaction while job resources improve it, supporting a structured **PPP approach** to managing workload sustainably.

RECOMMENDATION 1

INTEGRATE SIMON'S FOUR LEVERS FRAMEWORK INTO THE WORKFORCE MANAGEMENT

	Simons' Four Levers Framework
Goal	Align job demands and control mechanisms to manage high-demand work sustainably, reduce excessive overtime and burnout risk, and improve long-term job satisfaction.
Methodology	• **Stage 1: Belief System –** Set clear expectations on workload, overtime, and role responsibilities. • **Stage 2: Boundary Systems –** Enforce non-negotiable limits on working hours, on-call periods, and recovery time. • **Stage 3: Diagnostic System –** Monitor effort–reward balance and apply wellbeing programs after demand controls. • **Stage 4: Interactive Control System –** Use workload and overtime reviews as learning and feedback mechanisms**.**
How it helps	• Prevents normalization of chronic overwork • Reduces burnout and unmanaged health risks • Translates policy into operational practice • Sustains motivation, psychological safety, and job satisfaction

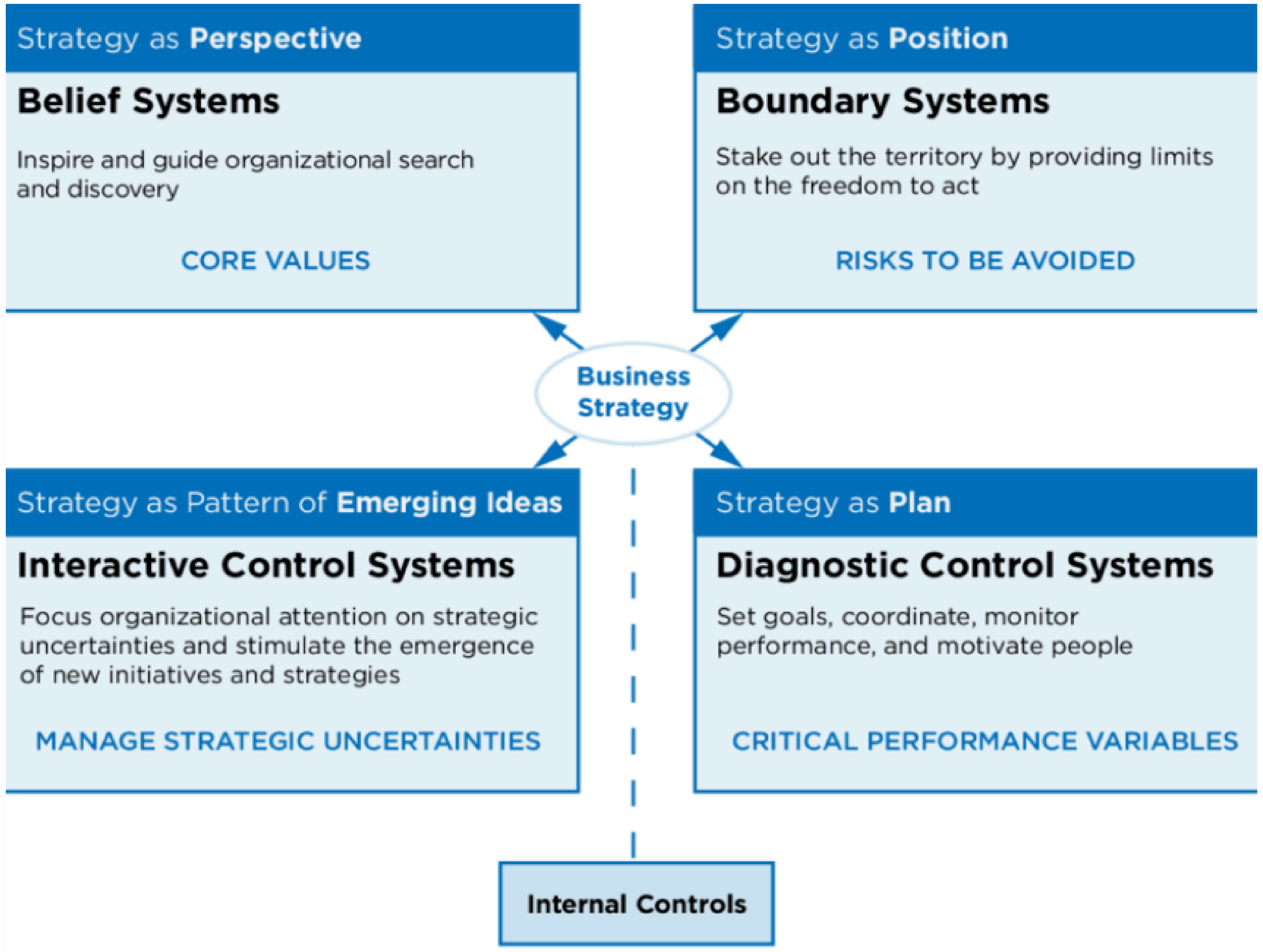


Figure 16: Simon's four levers framework

ACTIONABLE PLAN 1

Stage	Objectives	Strategies	Outcome Benefits
Stage 1: Belief System (3 months)	Align meaning and identity with High-Demand Work, such as ITs	Define clear limits on workload, overtime thresholds, and responsibility scopes for each role. Prevent normalization of excessive demands, targeting high-impact factors like overtime identified in the model.	Increases intrinsic motivation and perceived meaningfulness of work, reduces resentment associated with unavoidable overtime.
Stage 2: Boundary Systems (3 months)	To protect employee wellbeing by establishing clear non-negotiable limits	Set formal caps on continuous on-call periods and extended working hours, define clear escalation rules when workload exceeds team capacity, and mandate minimum recovery periods	Limits the normalisation of chronic overwork and reduces unmanaged health and burnout risks
Stage 3: Diagonstic System (3 months)	Ensuring fair exchange between effort and reward.	Implement wellbeing programs for sleep quality, energy recovery, and psychological safety - effective only after prior layers address demands.	Maximized impact from resources like sleep and work-life balance, leading to higher job satisfaction when demands are controlled, aligning with model findings on limited standalone effects.
Stage 4: Interactive Control System (3 months)	To prevent overtime from becoming invisible or permanent by using it as a strategic learning signal	Conduct regular post-incident and workload review sessions involving IT staff and leadership, identify recurring causes of overtime	Gradually reduces avoidable overtime while sustaining job satisfaction in the interim and encourages psychological safety and upward feedback from employees.

Figure 16: Action plan for Recommendation 1

RECOMMENDATION 2

INTEGRATE PPP FRAMEWORK INTO THE WORKFORCE MANAGEMENT

	PPP FRAMEWORK
Goal	Align job demands and job resources in line with the JD-R framework, targeting to reduce demand on stress, overtime, workload to sustainably improve operations and ultimately enhance job satisfaction.
Methodology	• **Policy Layer (M1-2)**: Set workload/overtime limits, prevent demand creep • **Process Layer (M3-4)**: Redesign work allocation, embed realistic capacity planning • **People Layer (M5-6)**: Add wellbeing programs after structural fixes in place
How it helps	• Establishes sustainable workload foundation • Transforms policy into operational reality • Maximizes wellbeing ROI

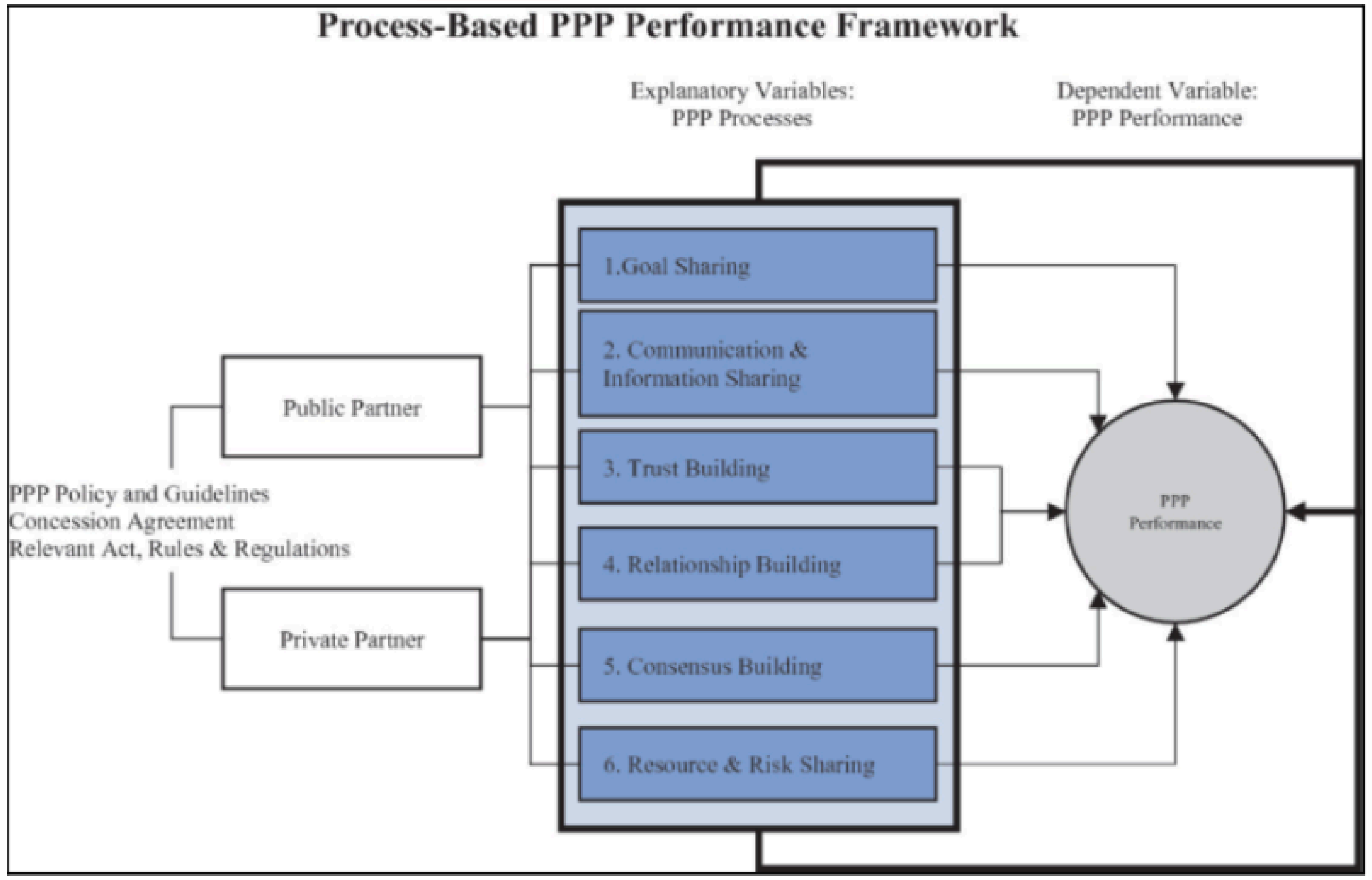


Figure 17: PPP framewrok

ACTIONABLE PLAN 2

Stage	Objectives	Strategies	Outcome Benefits
Stage 1: Policy Layer (Month 1-2)	Establish structural controls on job demands.	Define clear limits on workload, overtime thresholds, and responsibility scopes for each role. Prevent normalization of excessive demands, targeting high-impact factors like overtime identified in the model.	Direct reduction in key job demands, creating a foundation for sustainable workload management and preventing burnout.
Stage 2: Process Layer (Month 3-4)	Translate policies into daily operations.	Redesign work allocation, project planning, and performance evaluation cycles based on realistic capacity, ensuring consistent reflection of workload and time limits in management practices.	Actual reductions in employee workload and stress, turning policy commitments into tangible operational improvements.
Stage 3: People Layer (Month 5-6)	Enhance employee welfare and mental health as performance boosters.	Implement wellbeing programs for sleep quality, energy recovery, and psychological safety - effective only after prior layers address demands.	Maximized impact from resources like sleep and work-life balance, leading to higher job satisfaction when demands are controlled, aligning with model findings on limited standalone effects.

Figure 18: Action plan for Recommendation 2

VI. CONCLUSION

KEY INSIGHTS → DATA INSIGHTS → CONCLUSION

KEY INSIGHTS

OVERTIME

- Overtime is driven primarily by job structure rather than individual characteristics.

JOBSATISFACTION

- Job satisfaction functions as a buffer, reducing overtime risk within structural constraints

DATA INSIGHTS

DESCRIPTIVE ANALYSIS

- Shows systematic differences in JobSatisfaction across WorkEnvironment, WorkLifeBalance, Workload, Stress, Sleep Hours, and Overtime.

PREDICTIVE MODELS

- **Dept(IT)** → strongest predictor of HaveOT, with JobSatisfaction remaining a significant negative predictor after controlling for Dept
- JobSatisfaction, WorkEnvironment, and WorkLifeBalance increase satisfaction, while Workload, Stress, and HaveOT reduce it; SleepHours provide a modest uplift

CONCLUSION

RECCOMENDATIONS

- Improving employee outcomes requires structural interventions, not marginal environmental upgrades alone
- Management should prioritise overtime control, workload rebalancing, and recovery support, especially in overtime-intensive departments.

Thank You

We appreciate your time and attention.

REFERENCES

- Adam L, Follett L and Hofmann H (2016) 'Variations of Q–Q Plots: The Power of Our Eyes!', *The American Statistician*, 70(2):202–214, doi:10.1080/00031305.2015.1077728.
- Ahad N, Yin TS, Othman AR, and Yaacob CR (2011) 'Sensitivity of Normality Tests to non-normal Data', *Sains Malaysiana,* 40(6): 637–641.
- Alfred KM and Arnut P (2014) 'Shapiro–Francia test compared to other normality test using expectedp-value', *Journal of Statistical Computation and Simulation,* 85(15):3002–3016, doi:10.1080/00949655.2014.947986.
- Aruldoss A, Kowalski KB, Travis ML and Parayitam S (2021) 'The Relationship between Work–life Balance and Job satisfaction: Moderating Role of Training and Development and Work Environment', *Journal of Advances in Management Research*, 19(2):240–271, doi:10.1108/jamr-01-2021-0002.
- Aswale N and Mukul K (2019) 'Role of Data Analytics in Human Resource Management for Prediction of Attrition Using Job Satisfaction', *Data Management, Analytics and Innovation*, 1:57–67, doi:10.1007/978-981-32-9949-8_5.
- Austin P and Buuren S (2025) 'Imputation of incomplete ordinal and nominal data by6 predictive mean matching', *Statistical Methods in Medical Research*, 1-20, doi:10.1177/09622802251362642
- Barck-Holst P, Nilsonne Å, Åkerstedt T and Hellgren C (2020) 'Reduced Working Hours and work-life Balance', *Nordic Social Work Research,* 12(4):1–14, doi:10.1080/2156857x.2020.1839784.
- Borgogni L, Russo S, Petitta L, Vecchione M (2010) 'Predicting Job Satisfaction and Job Performance in a Privatized Organization', *International Public Management Journal*, 13(3),275-296, doi:10.1080/10967494.2010.504114
- Buchanan W (1974) 'Nominal and Ordinal Bivariate Statistics: The Practitioner's View', *American Journal of Political Science*, 18(3):625-646, doi:10.2307/2110637
- Coomber B and Louise B (2007) 'Impact of job satisfaction components on intent to leave and turnover for hospital-based nurses: A review of the research literature', *International Journal of Nursing Studies*, 44(2):297–314, doi:10.1016/j.ijnurstu.2006.02.004.
- Das B and Resnick SI (2008) 'QQ Plots, Random Sets and Data from a Heavy Tailed Distribution', *Stochastic Models*, 24(1):103–132, doi:10.1080/15326340701828308.
- Eisenberger R, Shanock LR and Wen X (2020) 'Perceived Organizational Support: Why Caring About Employees Counts', *Annual Review of Organizational Psychology and Organizational Behavior*, 7(1):101–124, doi:10.1146/annurev-orgpsych-012119-044917.

REFERENCES

- Emmanuel T, Maupong T, Mpoeleng D, Semong T, Mphago B and Tabona O (2021) 'A Survey on Missing Data in Machine Learning', *Journal of Big Data*, 8(1):1–37, doi:10.1186/s40537-021-00516-9.
- Evangelia and Demerouti (2025) 'Job demands-resources and conservation of resources theories: How do they help to explain employee well-being and future job design?', *Journal of Business Research*, 192, doi:10.1016/j.jbusres.2025.115296.
- Gabr MI, Helmy YM and Elzanfaly DS (2023) 'Effect of Missing Data Types and Imputation Methods on Supervised Classifiers: an Evaluation Study', *Big Data and Cognitive Computing*, 7(1):55, doi:10.3390/bdcc7010055.
- Gniazdowski (2022) 'On the Analysis of Correlation Between Nominal Data and Numerical Data', *Warsaw School of Computer*, 27(16):57-82, doi:10.26348/znwwsi.27.57
- Gowrishankar KL, N Kanagaraj and Krishnan SG (2017) 'AGE DIVERSITY OF THE WORKFORCE AND EMPLOYEES' PERFORMANCE -A DESCRIPTIVE STUDY', *International Journal of Marketing & Financial Managemen*t, 5(10):1–11, doi:hal-04850319
- Gupta NP and Singh A (2025) 'Examining Human Resource Management Practices and Policies for Enhancing Employee Work-Life Balance', *Deleted Journal,* 3(04):1453–1459, doi:10.47392/irjaem.2025.0236.
- Idika D, Owan V and Agama V (2023) 'The Application of the Nominal Scale of Measurement in Research Data Analysis', *Prestige Journal of Education,* 6(1):190-198 https://philpapers.org/archive/IDITAO.pdf
- Imran UD, Ghazwan MF and Firmansyah F (2024) 'The Effect of Recognition and Appreciation on Employee Motivation and Performance', *Economics and Digital Business Review*, 6(1):1–18, doi:10.37531/ecotal.v6i1.1890.
- Islam M and Dhanekula A (2024) 'PREDICTIVE ANALYTICS AND DATA-DRIVEN ALGORITHMS FOR IMPROVING EFFICIENCY IN FULL-STACK WEB SYSTEMS', *Interdisciplinary Research in Education, Technology, and Equity*, 5(2):226-260, doi:10.63125/q75tbj05
- Islam T and Rizwan M (2021) 'Comparison of correlation measures for nominal data', *Commucations in Statisticss - Simulation and Computation,* 51(3):698-714, doi:10.1080/03610918.2020.1869984
- Janse RJ, Hoekstra T, Jager KJ, Zoccali C, Tripepi G, Dekker FW and van Diepen M (2021) 'Conducting correlation analysis: Important limitations and pitfalls', *Clinical Kidney Journa*l, 14(11):2332–2337, doi:10.1093/ckj/sfab085.
- Jeyanthi P, Choudhury T, Polay D, Singh T (2022) 'Analytics Techniques: Descriptive Analytics, Predictive Analytics, and Prescriptive Analytics', *EAI/Springer Innovations in Communication and Computing*, 1-14, doi:10.1007/978-3-030-82763-2

REFERENCES

- Judge T, Zhang S, Glerum D (2020) 'JOB SATISFACTION', Essentials of Job Attitudes and Other Workplace Psychological Constructs, 207-241 https://www.taylorfrancis.com/chapters/edit/10.4324/9780429325755-11/job-satisfaction-timothy-judge-shuxia-carrie-zhang-david-glerum
- Kevin LB (2016) 'Sensitivity and specificity of normality tests and consequences on reference interval accuracy at small sample size: a computer-simulation study', *Veterinary Clinical Pathology*, 45(4):648–656, doi:10.1111/vcp.12390.
- Kittisak J, Panarat P, Pattanant P and Siridech K (2021) 'The Impact of the Workload on the Job Satisfaction: Does the Job Stress Matter?', *Academy of Strategic Management Journal*, 20(5), doi:10205074/11388464/s200_siridech.
- Ko YJ and Choi JN (2018) 'Overtime Work as the Antecedent of Employee satisfaction, Firm productivity, and Innovation', *Journal of Organizational Behavior,* 40(3):282–295, doi:10.1002/job.2328.
- Kumari R and Kataria A (2021) 'The mediating role of work-life balance on the relationship between job resources and job satisfaction: perspectives from Indian nursing professionals', *International Journal of Organizational Analysis*, 31(5), doi:10.1108/ijoa-04-2021-2722.
- Kusuma CK, Ambar L and Dorce ML (2021) 'ANALYSIS OF THE IMPACT OF WORKLOAD AND WORK STRESS ON JOB SATISFACTION WITH MOTIVATION AS INTERVENING VARIABLE', *Jurnal Aplikasi Manajemen*, 19(4):791–803, doi:10.21776/ub.jam.2021.019.04.08.
- Liu X, Qin N and Wei X (2025) 'The Complex Relationship between Sleep Quality and Job Satisfaction: a Machine Learning-Based Bayesian Rule Set Algorithm', *Behavioral Sciences,* 15(3):276, doi:10.3390/bs15030276.
- Lundberg C, Gudmundson A and Andersson TD (2009) 'Herzberg's Two-Factor Theory of Work Motivation Tested Empirically on Seasonal Workers in Hospitality and Tourism', *Tourism Management*, 30(6):890–899, doi:10.1016/j.tourman.2008.12.003.
- Mammadzada M (2025) 'Overtime Work and satisfaction: Review and Directions for Future Research', *The Journal of Economic Research & Business Administration,* 1(151):158–171, doi:10.26577/be2025151112.
- Memon SMZ, Wamala R and Kabano IH (2023) 'A comparison of imputation methods for categorical data', *Informatics in Medicine Unlocked,* 42:101–382, doi:10.1016/j.imu.2023.101382.
- Michelucci U (2025) 'Skewness, Kurtosis, and Modality', 95–108, doi:10.1007/978-3-031-78147-6_9.
- Mohammadinejad A, Yazdanirad S, Dehdashti A, Kakaei P, Keshizadeh F and Farokhzad M (2025) 'Mediating Effects of Job Stress on the Relationship between Sleep Quality and Job Satisfaction among Construction Workers', *Iranian Journal Of Health Sciences*, 13(1):43–56, doi:10.32598/ijhs.13.1.1067.1.

REFERENCES

- Mohit C (2025) *Statistical Analysis and Visualization,* Google Books website, accessed 20 December 2025. https://books.google.com.vn/books?hl=en&lr=&id=P5dIEQAAQBAJ&oi=fnd&pg=PT8&dq=boxplots+visualize+distributional+spread+and+skewness.
- Nicola LB, Rowley S, Robson A, Akram U and Myachykov A (2020) ‘Sleep duration, sleep variability, and impairments of visual attention’, *Quarterly Journal of Experimental* Psychology, 73(6):174702181989577, doi:10.1177/1747021819895771.
- Okta IF and Perdana MA (2024) ‘The Effect of Overtime and Workload Towards Employee Productivity with Burnout as Mediating Variable’, *Jurnal Manajemen dan Ekonomi Kreatif,* 2(3):282–296, doi:10.59024/jumek.v2i3.398.
- Paczkowski WR (2021) 'Business Analytics, 1st Edition', *Springer International Publishing, Cham*, doi:10.1007/978-3-030-87023-2.
- Pavithrakannan R, Fenn NB, Raman S, Kalyanaraman V, Murugananthan VK and Janarthanan J (2021) ‘Imputation Analysis of Central Tendencies for Classification’, *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, 1–7, doi:10.1109/iemtronics52119.2021.9422507.
- Pedro CH, Andersen LB, Bull FC, Guthold R, Haskell W and Ekelund U (2012) ‘Global physical activity levels: surveillance progress, pitfalls, and prospects’, *The Lancet*, 380(9838):247–257, doi:10.1016/s0140-6736(12)60646-1.
- Rusbult CE and Farrell D (1983) ‘A longitudinal test of the investment model: The impact on job satisfaction, job commitment, and turnover of variations in rewards, costs, alternatives, and investments’, *Journal of Applied Psychology*, 68(3):429–438, doi:10.1037/0021-9010.68.3.429
- Sabuj HH, Nigar Sultana N, Richie Gomes P, Lameesa A and Alam MdA (2022) ‘Interpretable Garment Workers’ Productivity Prediction in Bangladesh Using Machine Learning Algorithms and Explainable AI’, *2022 25th International Conference on Computer and Information Technology (ICCIT),* doi:10.1109/iccit57492.2022.10054863.
- Sarker A, Shamim SM, Rahman M and Zama SZ (2018) ‘View of Employeeas Performance Analysis and Prediction Using K-means Clustering’, *Global Journal of Computer Science and Technology: C Software & Data Engineering*, 18(1).
- Sayadat N (2021) ‘Process-Based Performance Framework for Public-Private Partnership (PPP) Policy Implementation in Bangladesh’, *Policy Response, Local Service Delivery, and Governance in Bangladesh, Nepal, and Sri Lanka*, 207–232, doi:10.1007/978-3-030-66018-5_9.

REFERENCES

- Szumilas M (2010) 'Explaining Odds Ratios', *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 19(3):227, https://pmc.ncbi.nlm.nih.gov/articles/PMC2938757/.
- Naidu G, Zuva T and Sibanda EM (2023) 'A Review of Evaluation Metrics in Machine Learning Algorithms', *Lecture notes in networks and systems,* 724:15–25, doi:10.1007/978-3-031-35314-7_2.
- Obi JC (2023) 'A comparative study of several classification metrics and their performances on data', *World Journal of Advanced Engineering Technology and Sciences,* 8(1):308–314, doi:10.30574/wjaets.2023.8.1.0054.
- Yang S, Chen L and Bi X (2023) 'Overtime work, job autonomy, and employees' subjective well-being: Evidence from China', *Frontiers in Public Health,* 11, doi:10.3389/fpubh.2023.1077177.
- Lindsay RM (2015) ' MANAGEMENT CONTROL SYSTEM DESIGN FOR MEETING THE CHALLENGE', *Developing Robust Strategy for Uncertain Times*: *Expanding our Concept of Management Control to Deal with Dynamic Markets* , https://www.researchgate.net/publication/337570386_Developing_Robust_Strategy_for_Uncertain_Times_Expanding_our_Concept_of_Management_Control_to_Deal_with_Dynamic_Markets_PART_II_MANAGEMENT_CONTROL_SYSTEM_DESIGN_FOR_MEETING_THE_CHALLENGE.

- .Szumilas M (2010) 'Explaining Odds Ratios', *Journal of the Canadian Academy of Child and Adolescent Psychiatry,* 19(3):227, https://pmc.ncbi.nlm.nih.gov/articles/PMC2938757/.
- Naidu G, Zuva T and Sibanda EM (2023) 'A Review of Evaluation Metrics in Machine Learning Algorithms', *Lecture notes in networks and systems*, 724:15–25, doi:10.1007/978-3-031-35314-7_2.
- Obi JC (2023) 'A comparative study of several classification metrics and their performances on data', *World Journal of Advanced Engineering Technology and Sciences*, 8(1):308–314, doi:10.30574/wjaets.2023.8.1.0054.
- Yang S, Chen L and Bi X (2023) 'Overtime work, job autonomy, and employees' subjective well-being: Evidence from China', *Frontiers in Public Health*, 11, doi:10.3389/fpubh.2023.1077177.
- Lindsay RM (2015) ' MANAGEMENT CONTROL SYSTEM DESIGN FOR MEETING THE CHALLENGE', *Developing Robust Strategy for Uncertain Times*: *Expanding our Concept of Management Control to Deal with Dynamic Markets* , https://www.researchgate.net/publication/337570386_Developing_Robust_Strategy_for_Uncertain_Times_Expanding_our_Concept_of_Management_Control_to_Deal_with_Dynamic_Markets_PART_II_MANAGEMENT_CONTROL_SYSTEM_DESIGN_FOR_MEETING_THE_CHALLENGE.

REFERENCES

- Silan MA (2025) ‘When can we treat Likert type data as interval?’, Hal.science, doi:hal.science/hal-04927866.
- Sur P, Chen Y and Candès EJ (2019) ‘The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled Chi-square’, P*robability Theory and Relat*ed Fields, 175(1-2):487–558, doi:10.1007/s00440-018-00896-9.
- Timothy GG, Qi Feng Lin, Boudreau J and Nelson R (2008) ‘Regression Estimation Following the Square-Root Transformation of the Response’, *Forest science,* 54(6):597–606, doi:10.1093/forestscience/54.6.597.
- Tiwari S (2025) *Average Working Hours 2025*: *Global Stats, Trends & Insights*, Mera Monitor website, accessed 20 December 2025.https://meramonitor.com/working-hours/
- True Geometry (2025) *What is the key difference in scale compression between the log and square root transformations*, True Geometry website, accessed 20 December 2025. https://blog.truegeometry.com/api/exploreHTML/47456faff9bde4d43df08d51577199d8.exploreHTML.
- Tyll L, Srivastava M and Bruder Badie N (2025) ‘The Industry Specific Effect of Ethical Organisational Culture on Job Response and Job Commitment: A Case of Czech Firms’, *Journal of East-West Business*, 1–33, doi:10.1080/10669868.2025.2463669.
- Vancouver JB, Li X, Weinhardt JM, Steel P and Purl JD (2016) ‘Using a Computational Model to Understand Possible Sources of Skews in Distributions of Job Performance’, *Personnel Psychology*, 69(4):931–974, doi:10.1111/peps.12141.
- Vinutha HP, Poornima B and Sagar BM (2018) ‘Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset’, *Advances in Intelligent Systems and Computing,* 701:511–518, doi:10.1007/978-981-10-7563-6_53.
- Wolniak R (2023) ‘THE CONCEPT OF DESCRIPTIVE ANALYTICS’, *Scientific Papers of Silesian University of Technology Organization and Management,* 172:699-715, doi:10.29119/1641-3466.2023.172.42
- Wongkamthong C and Akande O (2021) ‘A Comparative Study of Imputation Methods for Multivariate Ordinal Data’, *Journal of Survey Statistics and Methodology*, 11(1):189–212, doi:10.1093/jssam/smab028.
- Yücel I (2012) ‘Examining the Relationships among Job Satisfaction, Organizational Commitment, and Turnover Intention: An Empirical Study’, *International Journal of Business and Management,* 7(20), doi:10.5539/ijbm.v7n20p44.
- Zhu Y (2012) ‘A Review of Job Satisfaction’, *Canadian Center of Science and Education*, 9(1):293-298, doi:10.5539/ass.v9nlp293
- Zou KH, Tuncali K and Silverman SG (2003) ‘Correlation and Simple Linear Regression’, *Radiology,* 227(3):617–628, doi:10.1148/radiol.2273011499.

APPENDIX

APPENDIX 1

Predictor	Pr(>Chi)	LRT-Based Decision	Analytical Justification
Work-Life Balance	< 0.001	**Retained**	Removing this variable leads to a substantial deterioration in model fit, indicating a strong contribution to predicting job satisfaction.
Work Environment	< 0.001	**Retained**	Demonstrates a significant likelihood loss when excluded, confirming its critical role as a core job-related predictor.
Workload	< 0.001	**Retained**	Contributes significantly to overall model fit, highlighting workload as a key determinant of job satisfaction outcomes.
Stress	< 0.001	**Retained**	Exclusion results in a large and statistically significant reduction in model fit, indicating strong predictive relevance.
Overtime Status (HaveOT)	< 0.001	**Retained**	Overtime status provides meaningful explanatory power and significantly improves model likelihood.
Age	0.017	**Retained**	Age contributes significantly to model fit and captures individual heterogeneity affecting job satisfaction.
Marital Status	0.027	**Retained**	Demonstrates a statistically significant contribution to model fit, warranting retention as a control variable.
Sleep Hours	< 0.001	**Retained**	One of the strongest individual-level predictors, with exclusion causing a substantial deterioration in model fit.
Number of Companies	0.026	**Retained**	Improves model fit and captures career mobility effects relevant to job satisfaction.
Education Level	< 0.001	**Retained**	Exhibits a strong likelihood contribution, suggesting a non-linear relationship with job satisfaction.
Employment Type	0.466	**Removed**	Exclusion does not significantly affect model fit, indicating limited predictive contribution once other variables are controlled.
Experience	0.725	**Removed**	Provides no additional improvement in model fit beyond age and career mobility variables.
Job Level	0.322	**Removed**	Does not significantly enhance model likelihood after accounting for job-related and demographic factors.
Department	0.537	**Removed**	Departmental differences do not materially improve predictive performance in the presence of core predictors.
Training Hours per Year	0.784	**Removed**	Fails to contribute to model fit and offers limited predictive value for job satisfaction.
Commute Distance	0.920	**Removed**	Shows negligible impact on model likelihood and predictive accuracy.
Team Size	0.057	**Excluded (Borderline)**	Marginal contribution to model fit; excluded to improve parsimony and interpretability.
Number of Reports	0.281	**Removed**	Does not significantly improve model fit once job demand variables are included.
Commute Mode	0.226	**Removed**	Provides limited explanatory power and does not enhance predictive performance.
Gender	0.504	**Removed**	No significant contribution to model fit after controlling for job conditions.
Physical Activity Hours	0.930	**Removed**	Exclusion has no meaningful effect on model likelihood.

Table : Predictor Retention Results and Analytical Justification from the Likelihood Ratio Test (LRT)

APPENDIX

APPENDIX 2

```
Call:
glm(formula = HaveOT ~ Gender + Age + MaritalStatus + JobLevel +
    Experience + Dept + EmpType + WorkLifeBalance + WorkEnv +
    PhysicalActivityHours + Workload + Stress + SleepHours +
    CommuteMode + CommuteDistance + NumCompanies + TeamSize +
    NumReports + EduLevel + TrainingHoursPerYear + JobSatisfaction_num,
    family = binomial(link = "logit"), data = df2)

Coefficients:
                             Estimate Std. Error z value Pr(>|z|)
(Intercept)                 -0.947142   1.596153  -0.593   0.5529
GenderMale                   0.080502   0.111761   0.720   0.4713
GenderOther                 -0.013173   0.207268  -0.064   0.9493
Age                         -0.013165   0.060685  -0.217   0.8283
MaritalStatusMarried        -0.062219   0.219011  -0.284   0.7763
MaritalStatusSingle          0.077838   0.244442   0.318   0.7502
MaritalStatusWidowed       -12.942449 299.244782  -0.043   0.9655
JobLevel.L                  -0.092104   0.552701  -0.167   0.8677
JobLevel.Q                  -0.312147   0.278313  -1.122   0.2620
JobLevel.C                  -0.211413   0.145525  -1.453   0.1463
Experience                   0.003398   0.015483   0.219   0.8263
DeptFinance                  0.031901   0.273588   0.117   0.9072
DeptHR                       0.028561   0.345860   0.083   0.9342
DeptIT                       3.156019   0.266349  11.849   <2e-16 ***
DeptLegal                   -0.087130   0.309214  -0.282   0.7781
DeptMarketing               -0.051943   0.300850  -0.173   0.8629
DeptOperations               0.240502   0.279534   0.860   0.3896
DeptSales                    0.489015   0.291134   1.680   0.0930 .
EmpTypeFull-Time             0.027740   0.168336   0.165   0.8691
EmpTypePart-Time             0.029406   0.190166   0.155   0.8771
WorkLifeBalance.L            0.067840   0.114934   0.590   0.5550
WorkLifeBalance.Q           -0.106982   0.111769  -0.957   0.3385
WorkLifeBalance.C           -0.064316   0.112671  -0.571   0.5681
WorkLifeBalance^4            0.036573   0.114492   0.319   0.7494
WorkEnv.L                    0.266563   0.117747   2.264   0.0236 *
WorkEnv.Q                   -0.028062   0.110837  -0.253   0.8001
WorkEnv.C                    0.195447   0.114733   1.703   0.0885 .
WorkEnv^4                    0.111545   0.111157   1.003   0.3156
PhysicalActivityHours        0.014106   0.051592   0.273   0.7845
Workload.L                  -0.175236   0.119020  -1.472   0.1409
Workload.Q                  -0.138892   0.112999  -1.229   0.2190
Workload.C                   0.191405   0.110447   1.733   0.0831 .
Workload^4                  -0.159574   0.111063  -1.437   0.1508
Stress.L                    -0.064319   0.213840  -0.301   0.7636
Stress.Q                    -0.048528   0.191594  -0.253   0.8000
Stress.C                    -0.098757   0.176485  -0.560   0.5758
Stress^4                    -0.088843   0.150545  -0.590   0.5551
SleepHours                   0.054779   0.051844   1.057   0.2907
CommuteModeCar               0.265413   0.182632   1.453   0.1462
CommuteModeMotorbike         0.297440   0.216671   1.373   0.1698
CommuteModePublic Transport  0.299988   0.187247   1.602   0.1091
CommuteModeWalk              0.325890   0.210514   1.548   0.1216
CommuteDistance             -0.005826   0.007945  -0.733   0.4634
NumCompanies                -0.001998   0.180563  -0.011   0.9912
TeamSize                    -0.009358   0.019325  -0.484   0.6282
NumReports                   0.065568   0.028221   2.323   0.0202 *
EduLevel.L                  -0.017717   0.234402  -0.076   0.9397
EduLevel.Q                  -0.388486   0.169849  -2.287   0.0222 *
TrainingHoursPerYear         0.008145   0.022974   0.355   0.7229
JobSatisfaction_num         -0.403597   0.047721  -8.457   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3749.2  on 3024  degrees of freedom
Residual deviance: 2614.6  on 2975  degrees of freedom
AIC: 2714.6

Number of Fisher Scoring iterations: 12
```