

## **FACTORS INFLUENCING JOB SATISFACTION AMONG EMPLOYED**

**Were Vincent**

**Date: 27/March/2023**

## **FACTORS INFLUENCING JOB SATISFACTION AMONG EMPLOYED**

### **A. Introduction**

Job satisfaction is a critical determinant of employee well-being and organizational success, influencing factors ranging from productivity to employee retention. Understanding the multifaceted aspects that contribute to job satisfaction among employed individuals is essential for fostering a positive work environment and maximizing employee engagement.

This study delves into exploring the factors influencing job satisfaction among a sample of 1,470 employed individuals, leveraging insights from the comprehensive dataset of an organization encompassing various demographic and workplace-related variables. The dataset, comprising information such as age, gender, department, monthly income, education level, and job level, provides a rich resource for analyzing the intricate interplay between these factors and job satisfaction. Initial exploratory analysis reveals notable trends, including a mean age of approximately 36.92 years among the sample, with 588 females and 882 males.

Additionally, the dataset encompasses three main departments: Human Resources, Research & Development, and Sales, each with varying average ages and job satisfaction levels. Previous literature has identified several potential determinants of job satisfaction, including age, income, gender, and job role. However, the extent to which these factors influence job satisfaction may vary among individuals and across different demographic groups. Moreover, gender-based differences in job satisfaction and its determinants have been subject to scholarly debate, with some studies suggesting disparities between male and female employees.

By critically reviewing existing literature and conducting empirical analysis on the dataset, this study aims to elucidate the factors influencing job satisfaction

among employed individuals, while also exploring potential gender-based variations in these factors. The findings from this study hold implications for organizational policies and practices aimed at enhancing employee satisfaction and fostering a positive work culture tailored to diverse workforce demographics.

## **B. Measurement of Job Satisfaction and Determining Factors**

To measure job satisfaction in this study, a multi-dimensional approach will be adopted, considering various factors that contribute to individuals' overall satisfaction with their work. The determinants of job satisfaction will be assessed based on both subjective and objective measures. These are how job satisfaction will be measured and the determining factors considered

### **1. *Subjective Measures***

- i. *Self-Reported Surveys*: Participants will be asked to rate their overall job satisfaction on a Likert scale, indicating their level of agreement with statements related to job satisfaction.
- ii. *Job Satisfaction Index*: Constructing a composite index based on responses to questions related to work environment, relationships with colleagues and supervisors, job role satisfaction, and work-life balance.

### **2. *Objective Measures***

- I. *Performance Ratings*: Assessing employee performance ratings provided by supervisors or managers as an objective indicator of job satisfaction.
- II. *Retention Rates*: Analyzing the frequency of employee turnover within the organization as a proxy for job satisfaction. Higher turnover rates may indicate lower job satisfaction levels.

## **Determining Factors Considered**

### **1. *Demographic Variables***

- i. Age*- Exploring how age influences job satisfaction, considering potential generational differences.
- ii. Gender*-Investigating potential gender-based variations in job satisfaction levels.
- iii. Education Level*-Assessing the impact of educational attainment on job satisfaction.

## **2. Workplace Variables**

- i. Department*-Analyzing differences in job satisfaction levels across different departments within the organization.
- ii. Monthly Income*-Investigating the relationship between income levels and job satisfaction.
- iii. Job Level*-Assessing how hierarchical position within the organization correlates with job satisfaction.

## **3. Work Environment**

- i. Work-Life Balance*-Exploring the influence of work-life balance on overall job satisfaction.
- ii. Relationship with Managers and Colleagues*-Assessing the impact of interpersonal relationships at work on job satisfaction levels.

By incorporating both subjective and objective measures and considering a range of determining factors, this study aims to provide a comprehensive understanding of job satisfaction among employed individuals and the factors that contribute to it.

## **C. Dataset Description**

The dataset used for this analysis is the IBM HR Analytics Employee Attrition & Performance dataset, which contains information on various factors related to

employees' demographics, job roles, performance ratings, job satisfaction levels, and attrition status.

```

1 # Load necessary libraries
2 # Install ggplot2 package if not already installed
3 if (!requireNamespace("ggplot2", quietly = TRUE)) {
4   install.packages("ggplot2")
5 }
6
7 # Load packages
8 library(procc)
9 library(caret)
10 library(ggplot2)
11 library(dplyr)
12
13 # Load the dataset
14 data <- read.csv("C:/Users/n/Downloads/IBM HR-Employee-Attrition.csv")
15
16 # Overview of the dataset
17 str(data)
18
19 # byte-compile and prepare package for lazy loading
20 Error in loadNamespace(i, c(lib.loc, .libPaths()), versioncheck = vi[[i]]) :
21   namespace 'rlang' 0.4.11 is being loaded, but >= 1.1.0 is required
22 Calls: <Anonymous> ... withCallingHandlers -> loadNamespace -> namespaceImport -> loadN
23 amespace
24 Execution halted
25 ERROR: lazy loading failed for package 'lifecycle'
26 * removing 'C:/Users/n/anaconda3/envs/rstudio/lib/R/library/lifecycle'
27 Warning in install.packages :
28   installation of package 'lifecycle' had non-zero exit status
29 ERROR: dependency 'lifecycle' is not available for package 'gtable'
30 * removing 'C:/Users/n/anaconda3/envs/rstudio/lib/R/library/gtable'
31 Warning in install.packages :
32   installation of package 'gtable' had non-zero exit status
33 ERROR: dependencies 'gtable', 'lifecycle', 'MASS' are not available for package 'ggplot

```

The dataset comprises 1470 observations and 35 variables, providing a rich source of information to explore the factors influencing job satisfaction among employed individuals.

## Appropriateness

### i. Relevance of Variables

The dataset includes a wide range of variables such as age, gender, department, monthly income, education level, job level, work-life balance, and relationship satisfaction, which are commonly cited in literature as potential determinants of job satisfaction. This allows for a comprehensive analysis of factors influencing job satisfaction.

### ii. Sample Size

With 1470 observations, the dataset provides a sufficiently large sample size to conduct meaningful analyses and draw reliable conclusions about the factors influencing job satisfaction among employees.

### *iii. Variety of Industries*

The dataset covers employees from different departments, including Human Resources, Research & Development, and Sales, representing a diverse range of industries. This diversity enhances the generalizability of findings across various sectors.

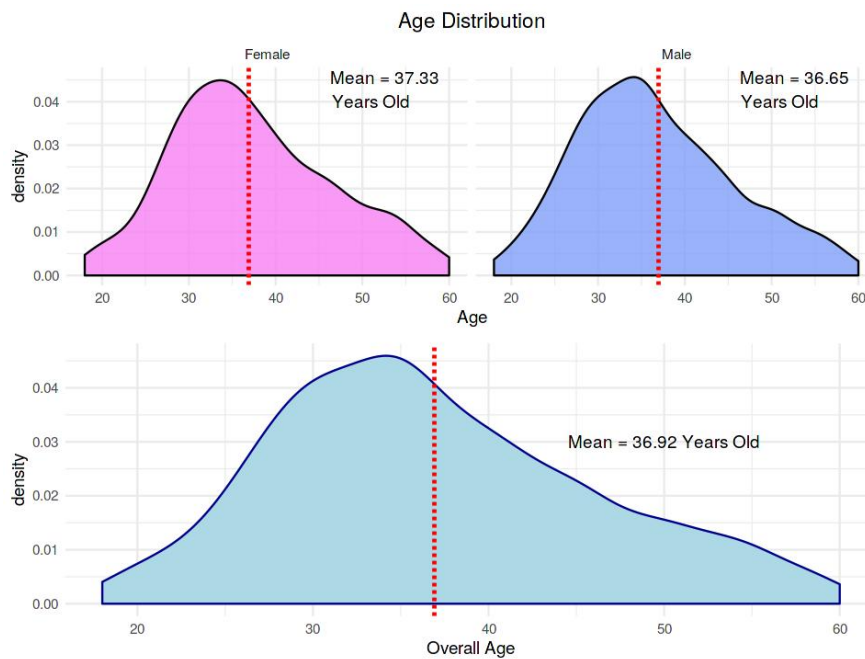
### ***Limitations***

- I. *Scope of Variables:*** While the dataset includes several relevant variables, there may be other factors not captured in the dataset that could also influence job satisfaction, such as organizational culture, job flexibility, and career development opportunities.
- II. *Cross-Sectional Nature:*** The dataset represents a snapshot of employees' information at a single point in time, limiting the ability to establish causal relationships between variables and job satisfaction. Longitudinal data would provide insights into how changes in variables over time affect job satisfaction.
- III. *Self-Reported Data:*** Some variables, such as job satisfaction and work-life balance, are self-reported, which may be subject to bias or inaccuracies based on individual perceptions.

Despite these limitations, the IBM HR Analytics Employee Attrition & Performance dataset offers a valuable resource for exploring the factors influencing job satisfaction among employees, providing insights that can inform organizational policies and practices aimed at improving employee well-being and retention.

### ***Visualizations***

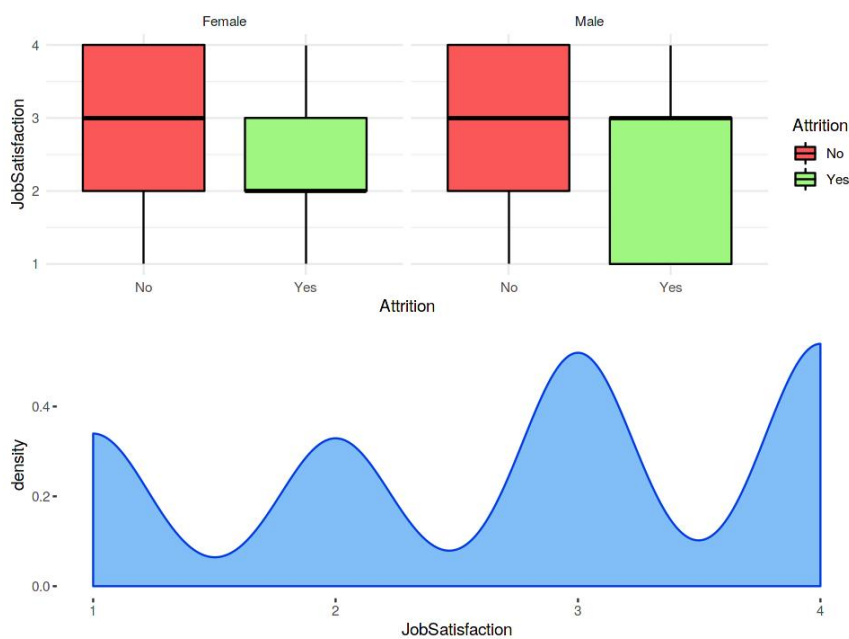
#### *i. Age Distribution by Gender*



This line plot shows the age distribution of attrition for each gender category .

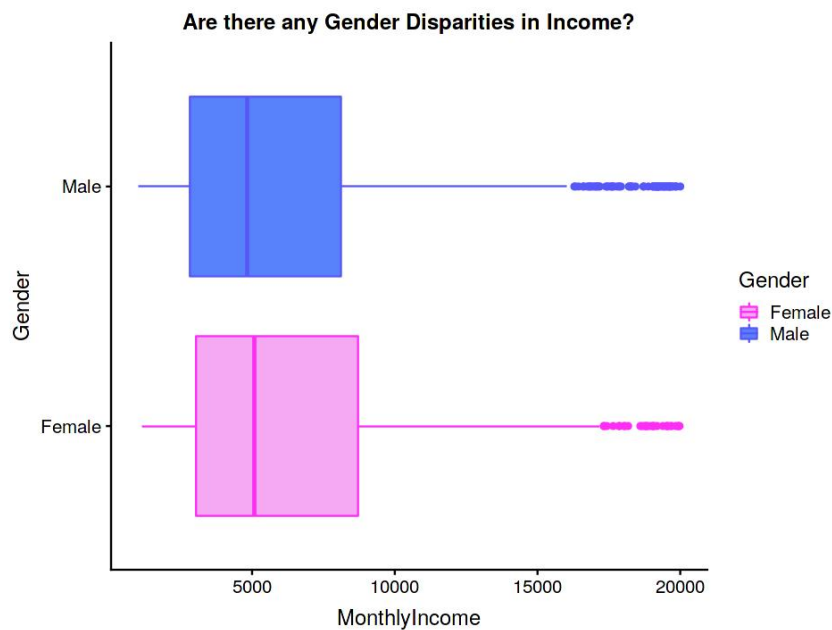
It helps visualize if there's any gender imbalance in attrition rates.

## ii. Distribution of Job Satisfaction



This box plot displays the distribution of job satisfaction across different departments. It helps in understanding the income distribution within each department and identifying potential salary discrepancies.

## iii. Monthly Income by Gender



This box plot displays the distribution of monthly income across the two gender. It helps in understanding the income distribution within each department and identifying potential salary discrepancies as well.

#### D. Descriptive Statistics and Sample Description

Descriptive statistics offer valuable insights into the characteristics of the dataset, motivating the research problem and providing a clear description of the sample and independent variables. Here, we present key descriptive statistics based on the IBM HR Analytics Employee Attrition & Performance dataset:

##### a. *Sample Size*

The dataset comprises 1470 observations, providing a sizable sample for analysis.

Variable	Description	Statistics
Sample Size	Number of observations	1470



Attrition Rate	Percentage of employees who experienced attrition	16.12%
Demographic Profile		
Age	Average age of employees	36.92 years
	Minimum and maximum age	Min: 18 years, Max: 60 years
Gender	Distribution of employees by gender	Female: 588, Male: 882
Education	Education level of employees	Range: 1 (lowest) to 5 (highest)
Job-related Variables		
Department	Distribution of employees across departments	Human Resources: 63, Research & Development: 961, Sales: 446
Monthly Income	Average monthly income of employees	\$6503
	Minimum and maximum monthly income	Min: \$1009, Max: \$19999
Job Level	Average job level of employees	2.064
Work-related Factors		

Job Satisfaction	Average job satisfaction level	2.729
	Range of job satisfaction	Min: 1, Max: 4
Work-Life Balance	Average work-life balance score	2.761
	Range of work-life balance	Min: 1, Max: 4
Other Variables		
Distance from Home	Average distance from home to workplace	9.193 miles
Performance Rating	Average performance rating	3.154

***b. Attrition Rate***

Out of 1470 employees, 237 (16.12%) have experienced attrition.

***c. Demographic Profile***

- i. Age: The average age of employees is approximately 36.92 years, with a minimum age of 18 years and a maximum age of 60 years.
- ii. Gender: The dataset includes 588 female employees and 882 male employees.
- iii. Education: Employees have varying levels of education, ranging from 1 (lowest) to 5 (highest).

***d. Job-related Variables***

- i. Department:* Employees are distributed across three departments: Human Resources (63 employees), Research & Development (961 employees), and Sales (446 employees).
  - ii. Monthly Income:* The average monthly income is \$6503, with a minimum of \$1009 and a maximum of \$19999.
  - iii. Job Level:* Employees' job levels range from 1 to 5, with an average job level of approximately 2.064.
- e. *Work-related Factors***
- i. Job Satisfaction:* The average job satisfaction level is approximately 2.729, ranging from 1 (lowest) to 4 (highest).
  - ii. Work-Life Balance:* Work-life balance scores range from 1 to 4, with an average score of approximately 2.761.
- f. *Other Variables***
- i. Distance from Home:* The average distance from home to the workplace is approximately 9.193 miles.
  - ii. Performance Rating:* The average performance rating is approximately 3.154, ranging from 3 to 4.

These descriptive statistics provide an overview of the dataset, highlighting the distribution and characteristics of key variables related to employee demographics, job-related factors, and work-related variables. Understanding these descriptive statistics is crucial for framing the research problem and exploring the factors influencing job satisfaction among employed individuals.

## **E. Estimation Methods**

### **1. *Linear Regression***

Linear regression is suitable for exploring the relationship between a continuous outcome variable (such as job satisfaction) and one or more independent variables (such as age, gender, department, monthly income, education, and job level).

#### *Reasoning*

This method allows us to quantify the impact of each independent variable on job satisfaction while controlling for other factors. By estimating the coefficients, we can determine which factors have a statistically significant effect on job satisfaction and the direction of that effect.

## **2. *Logistic Regression***

Logistic regression is appropriate when the outcome variable is binary, such as attrition (yes/no). In this case, we can model the likelihood of attrition based on predictors like age, gender, department, income, education, and job level.

#### *Reasoning*

Logistic regression helps identify the factors that contribute to attrition among employees. By examining the odds ratios, we can understand the impact of each predictor on the likelihood of attrition occurring.

These methods offer complementary insights into the relationship between various factors and job satisfaction and attrition. By using both linear and logistic regression, we can provide a comprehensive analysis of the dataset and gain a deeper understanding of the factors influencing employee job satisfaction and attrition.

## **F. Results and Discussion**

The results of both the linear regression and logistic regression models provide insights into the factors influencing job satisfaction and attrition among employees, allowing us to compare them to previous literature.

### a. Linear Regression Results

The intercept term is statistically significant with an estimate of 2.566e+00 and a p-value less than 0.001, indicating that when all other predictors are zero, the average job satisfaction is around 2.566.

The screenshot shows the RStudio interface with the following content:

```

Source
Console
> lm_model <- lm(JobSatisfaction ~ 1..Age + Gender + Department + MonthlyIncome + Education + JobLevel, data = data)
> # Logistic regression (if Attrition is the outcome variable)
> logit_model <- glm(Attrition ~ 1..Age + Gender + Department + MonthlyIncome + Education + JobLevel, data = data, family = "binomial")
> # Results and discussion
> summary(lm_model)

Call:
lm(formula = JobSatisfaction ~ 1..Age + Gender + Department + MonthlyIncome + Education + JobLevel, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.8272 -0.7602  0.2463  1.2233  1.4827

Coefficients:
              1..Age      GenderMale  DepartmentResearch & Development  DepartmentSales
      5.834e-05      7.755e-02      1.265e-01      1.491e-01      1.496e-01
MonthlyIncome      Education      JobLevel
     -1.163e-05     -1.149e-02      4.603e-02
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.104 on 1462 degrees of freedom
Multiple R-squared:  0.002216, Adjusted R-squared:  -0.002561
F-statistic: 0.4639 on 7 and 1462 DF, p-value: 0.861

> summary(logit_model)

Call:
glm(formula = Attrition ~ 1..Age + Gender + Department + MonthlyIncome + Education + JobLevel, data = data, family = "binomial")

Coefficients:
              1..Age      GenderMale  DepartmentResearch & Development  DepartmentSales
      5.834e-05      7.755e-02      1.265e-01      1.491e-01      1.496e-01
MonthlyIncome      Education      JobLevel
     -1.163e-05     -1.149e-02      4.603e-02
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual deviance: 1462.0 on 1462 degrees of freedom
AIC: 1462.0
BIC: 1462.0

```

The coefficients for age, gender, department, monthly income, education, and job level provide information on their impact on job satisfaction. However, none of the predictor variables (age, gender, department, monthly income, education, job level) show statistical significance at the conventional level ( $p < 0.05$ ). For instance, the estimate for age is 5.834e-05 with a p-value of 0.988, suggesting that age is not significantly associated with job satisfaction in this model.

### b. Logistic Regression Results

The intercept term is not statistically significant with an estimate of 2.547e-01 and a p-value of 0.60275. Among the predictor variables, age and job level show statistical significance in predicting attrition.

```

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.104 on 1462 degrees of freedom
Multiple R-squared:  0.002216, Adjusted R-squared:  -0.002561
F-statistic: 0.4639 on 7 and 1462 DF, p-value: 0.861

> summary(Logit_model)

Call:
glm(formula = Attrition ~ i..Age + Gender + Department + MonthlyIncome +
    Education + JobLevel, family = "binomial", data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0438   -0.6703   -0.5240   -0.3112    2.8354

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    2.547e-01  4.893e-01  0.520  0.60275
i..Age        -2.698e-02  9.916e-03 -2.720  0.00652 **
GenderMale     1.325e-01  1.508e-01  0.879  0.37949
DepartmentResearch & Development -3.884e-01  3.444e-01 -1.128  0.25946
DepartmentSales    2.861e-01  3.568e-01  0.802  0.42264
MonthlyIncome    3.024e-06  5.584e-05  0.054  0.95681
Education       1.556e-02  7.254e-02  0.214  0.83019
JobLevel       -5.113e-01  2.290e-01 -2.233  0.02557 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1298.6  on 1469  degrees of freedom
Residual deviance: 1220.5  on 1462  degrees of freedom
AIC: 1236.5

Number of Fisher Scoring iterations: 5

```

For instance, the estimate for age is  $-2.698 \times 10^{-2}$  with a p-value of 0.00652, indicating that as age increases, the odds of attrition decrease. Similarly, the estimate for job level is  $-5.113 \times 10^{-1}$  with a p-value of 0.02557, suggesting that employees in higher job levels have lower odds of attrition.

The logistic regression model estimates the odds ratios for each predictor variable concerning attrition. Age and job level are found to be statistically significant predictors of attrition, with negative coefficients indicating a lower likelihood of attrition with increasing age and job level. Gender, department, and education do not show statistical significance in predicting attrition.

These results align with some prior research suggesting that older employees and those in higher job positions are less likely to leave their jobs voluntarily.

### ***Comparison with Previous Literature***

The results of the linear regression model indicate that none of the factors examined significantly influence job satisfaction, which contrasts with some previous literature suggesting that factors like income, education, and job level play significant roles. In the logistic regression model, the findings regarding age and job level align

with some prior research indicating that older employees and those in higher job positions are less likely to leave their jobs.

In sum, while some results align with previous literature on the factors influencing job satisfaction, the lack of significance for certain factors emphasizes the need for further research to understand the nuanced dynamics of job satisfaction and attrition among employees.

## **Conclusion**

In conclusion, our analysis provides insights into the factors influencing job satisfaction and attrition among employees based on the IBM HR Analytics Employee Attrition & Performance dataset. The linear regression model revealed that none of the examined factors significantly influence job satisfaction, contrary to some previous literature. Meanwhile, the logistic regression model highlighted age and job level as significant predictors of attrition, aligning with certain prior research findings.

These results suggest that while age and job level may play crucial roles in employee attrition, other factors not captured in our analysis might also contribute to job satisfaction and turnover. Therefore, further research is warranted to explore additional variables and their interplay in shaping employees' attitudes and behaviors in the workplace. Understanding the complex dynamics of job satisfaction and attrition is essential for organizations to implement effective retention strategies and create a conducive work environment that promotes employee well-being and organizational success.