



---

## **CCDS223 – Group Project Report**

### **Employee Attrition Prediction**

November 22, 2025

**Group 2:**

Abdulallah Abdullah Nasser 2340466

Abdullah Atiah Alzahrani 2340088

Mohannad Abdullah Alamri 2340047

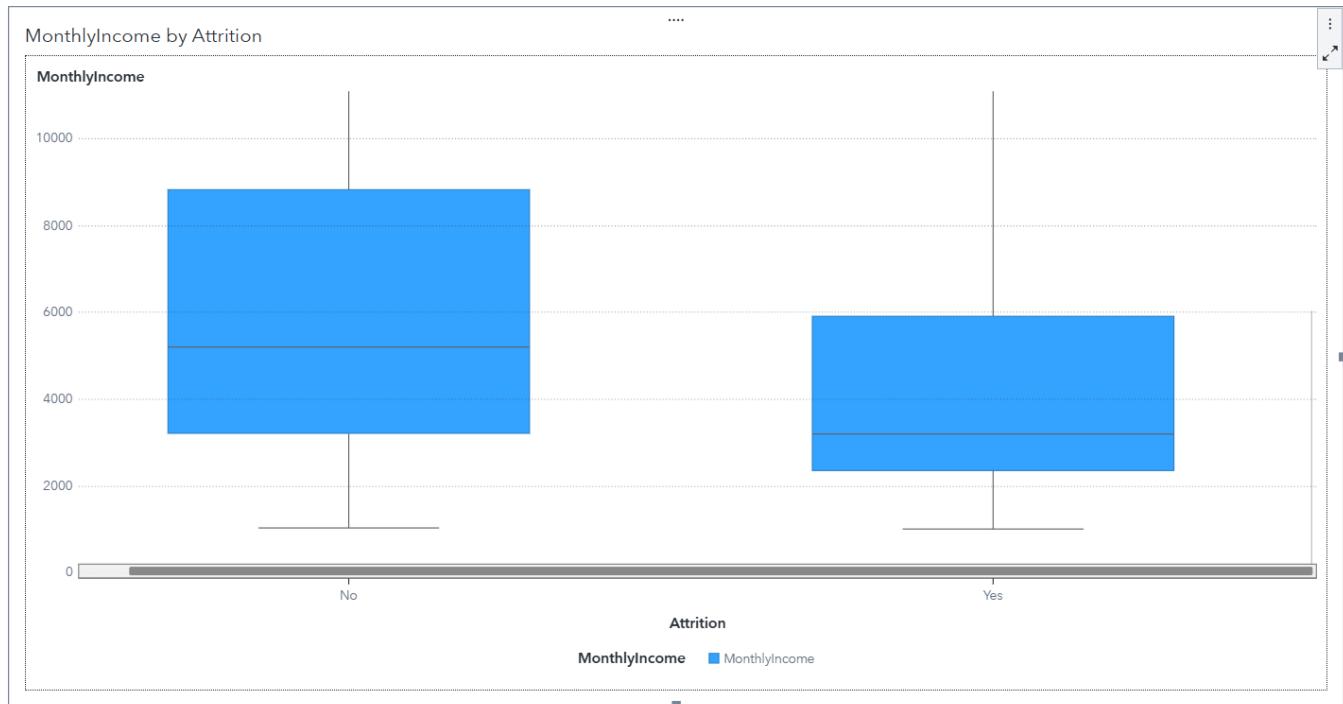
Tariq Zaid Alqhtani 2340458

**Instructor:** Abdullah Alghoson

## **Introduction:**

**The third phase, Exploratory Data Analysis (EDA), aims to deepen our understanding of the dataset that was cleaned and narrowed down in the previous phase to 20 variables for our project. Statistical summaries and visual representations, including correlation matrices, will be used to extract key patterns and identify the factors that most influence an employee's decision to leave (attrition). This valuable information will undoubtedly assist us and insights will guide the feature selection and modeling decisions in the next phase where we will build predictive models.**

## Box Plot:

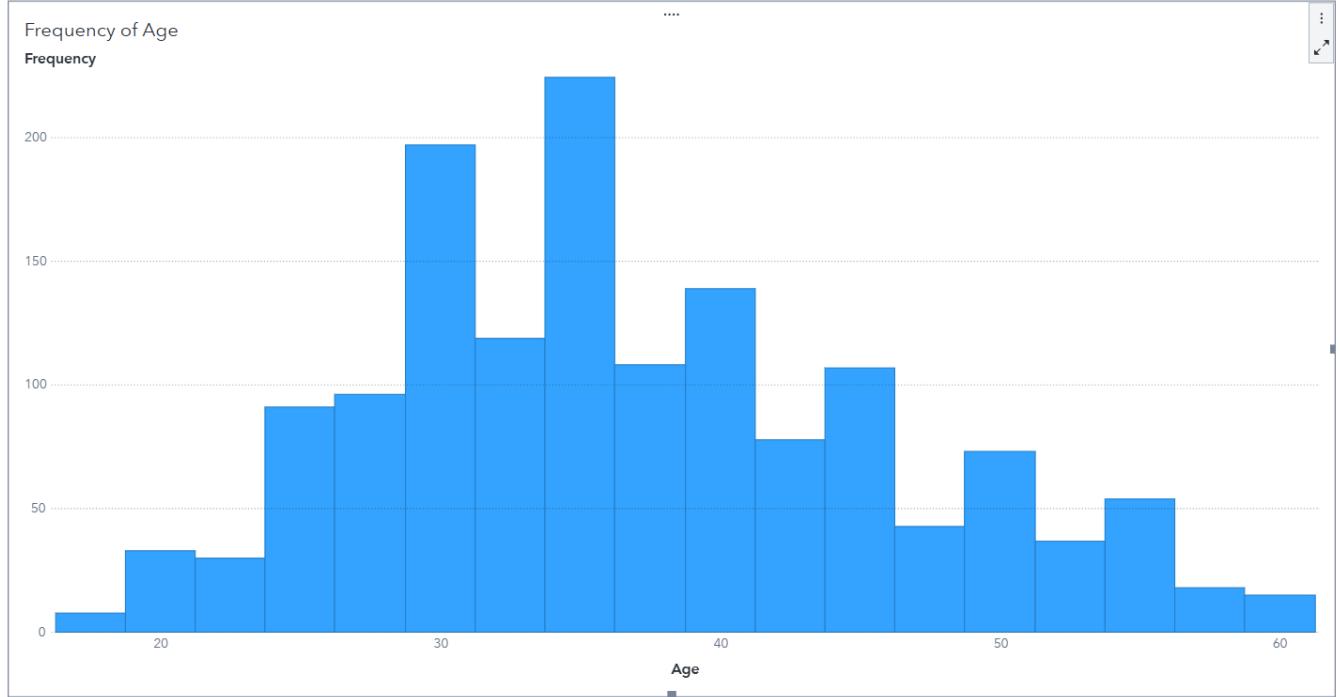


**Measure:** Monthly Income

**Category:** Attrition

**Insight:** The employees that have lower salaries are more likely to leave the company.

## Histogram:

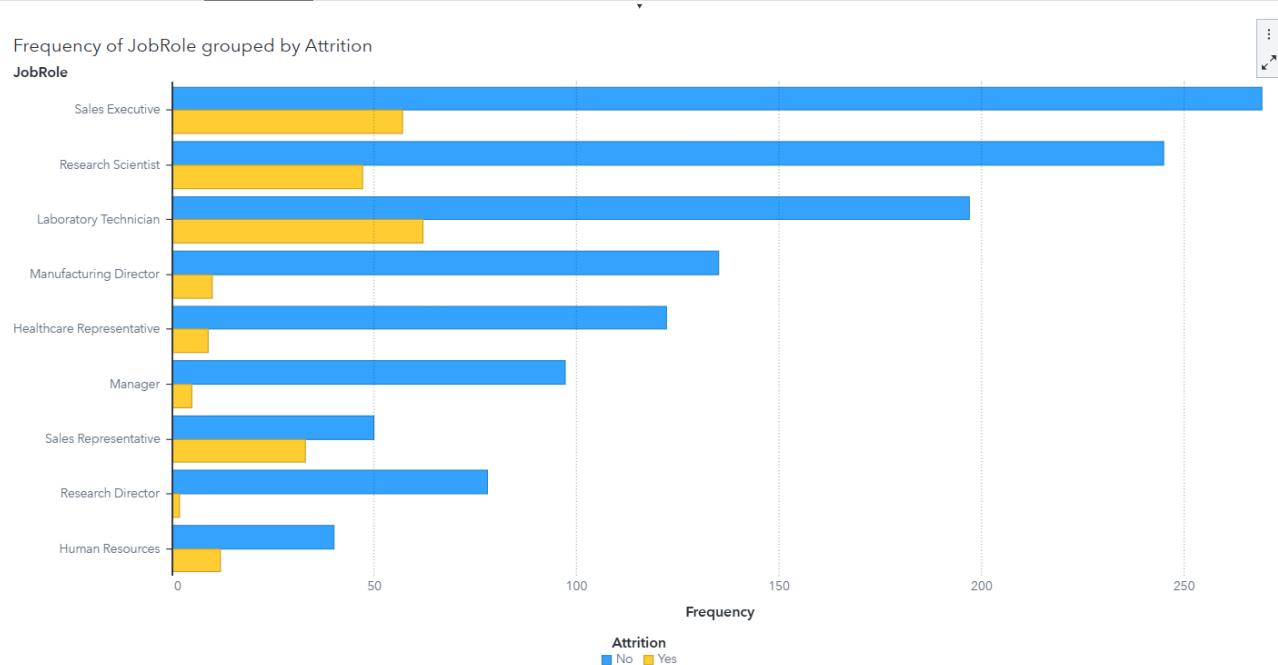


**Measure:** Age

**Insight:**

- Employees at the very early career stage <22 and nearing retirement >55 appear only in small numbers.
- Most employees are concentrated in the age group between 25 and 40 years old approximately.

## Bar chart:



**Category:** job role

**Measure:** frequency

**Group:** Attrition

**Insight:** Other jobs, like Sales Representative and Laboratory Technician, have a big proportion of their workers leaving compared to how many people they have. So, even if they don't have most people leaving overall, a large part of their team is leaving. That's high relative leaving. Some jobs, like Sales Executive and Research Scientist, have a lot of people leaving, which means many employees in those roles say "Yes" when asked if they are leaving. This is called high overall leaving. But some jobs, like Manager and Research Director, don't see many people leaving at all. Their numbers of employees leaving are low.

## Scatter Plot:



Measure X is MonthlyIncome

Measure Y is Age

**Group:** Attrition

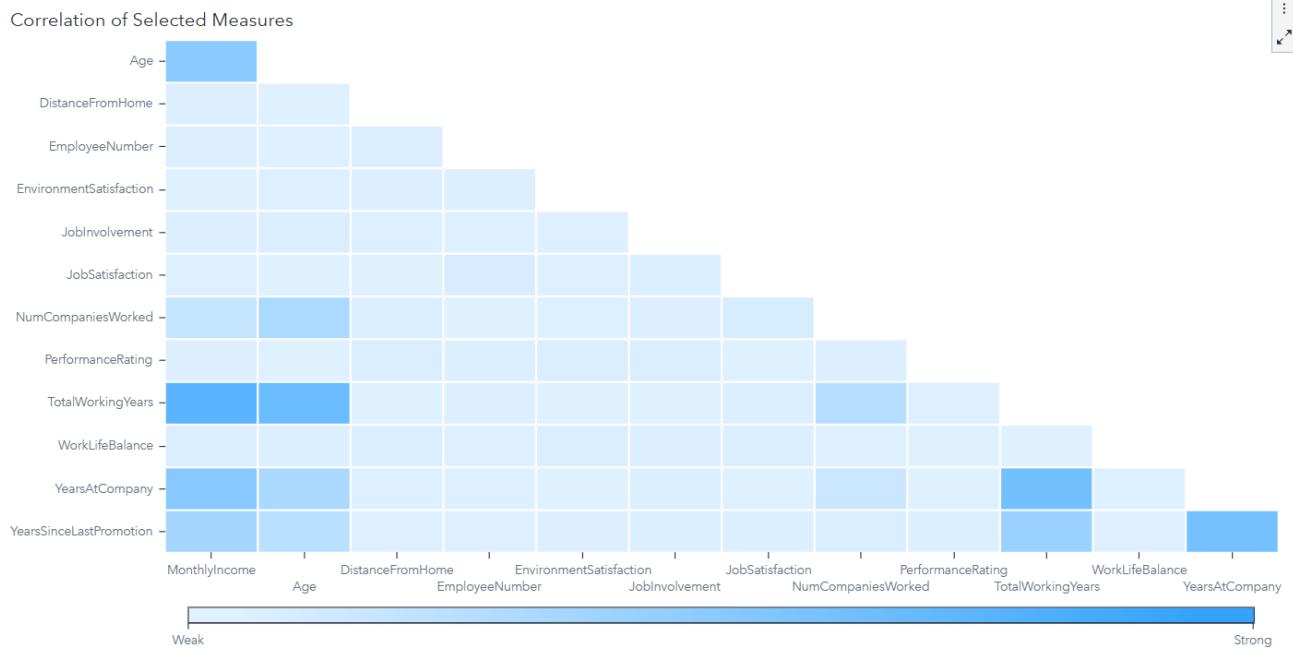
**Insight:**

Most people in the data are between 25 and 40 years old and earn less than 6,000 each month.

Generally, as people get older, they tend to earn more money because they have more experience.

So, older workers usually make higher salaries.

## Correlation Matrix:



**Measures:** all numeric variables

**Insight:** MonthlyIncome strongly correlates with TotalWorkingYears of

employment and monthly income, as well as between age and total years of employment.

Moderate correlation: There is a moderate positive correlation between years of employment at the company and years since last promotion.

Weak correlation: Most job satisfaction and workplace satisfaction variables show weak or negligible correlation with baseline measures such as age and monthly income.

## **Conclusion:**

EDA helps us understand and identify data that may influence the target variable Attrition. We looked closely at data about employees to understand why some might leave their jobs. We found that employees who earn lower money income each month are more likely to leave. Also, people work as Sales Representatives and Laboratory Technicians tend to leave more often than others. Additionally, we saw that if someone has worked for many years and he is older, they usually earn more money and have more work experience.

