**Prediction and Data Analysis on Employee Attrition**

Logo, company name

Description automatically generated

Visual Analytics & Business Intelligence (BAN 5573)

Supervised by:

Prof. Hamidreza Ahady Dolatsara

Group Members

Aashiya Aryal, Rosy Shrestha, Vinayak Jeyabalakrishnan, Yashwant Raut

April 20th, 2022

Project Deliverable 2

Table of Contents

[Introduction 3](#_Toc98622356)

[Research Questions 3](#_Toc98622357)

[1. Descriptive Analysis 3](#_Toc98622358)

[2. Diagnostic Analysis 3](#_Toc98622359)

[3. Predictive Analysis 3](#_Toc98622360)

[4. Prescriptive Analysis 4](#_Toc98622361)

[Significance of the Project 4](#_Toc98622362)

[Literature Review 4](#_Toc98622363)

[Methodology and Analysis 5](#_Toc98622364)

[1. Logistic Regression, 5](#_Toc98622365)

[2. Random Forest Regression, and 5](#_Toc98622366)

[3. Decision Tree Algorithm 5](#_Toc98622367)

[Descriptive Analysis 6](#_Toc98622368)

[1. Key Performance Indicators 6](#_Toc98622369)

[2. Attrition by Department 6](#_Toc98622370)

[3. Attrition by Job Role 7](#_Toc98622371)

[4. Attrition by Department 7](#_Toc98622372)

[5. Attrition by Income 8](#_Toc98622373)

[6. Attrition by Marital Status 8](#_Toc98622374)

[7. Tableau Dashboard 9](#_Toc98622375)

[9](#_Toc98622376)

[References 10](#_Toc98622377)

[Bibliography 10](#_Toc98622378)

# **Abstract**

Human Resource analytics is the method of collecting and analyzing the available data to improve an organization’s performance. It is usually referred to as talent analytics, allocation analytics people analytics, performance analytics, or workforce analytics. The data analysis is done by taking data that is routinely collected by the human resource which correlates with the organization’s objectives. After the data is compared, organized, and further analyzed, it will provide meaningful insights into the organization. From these different analyses can be done such as the time taken to hire an employee, the changes in employee turnover, the amount of investment to be made to get productive output from employees, the likelihood of employees leaving the organization within the year, and more. HR analytics is done to gain problem-solving insights that the organization promises. For HR analysis data must be collected, monitored, and measured against other data. This analysis will help the organization identify trends and patterns and apply insights to the decision of the organization.

Machine learning could be of great help to determine the factors for employee attrition. Using machine learning to detect the rate of attrition and factors affecting the employee’s performance can help to reduce the attrition rate. In this study, we applied four machine learning algorithms such as Random Forest, Logistic Regression, and The Decision Tree in this dataset. After getting the results, a performance evaluation and comparison are done through visualization. The main objective of this project is to predict and diagnose the rate of attrition rate and the factors affecting it, using machine-learning algorithms to find out the most effective respect to the accuracy, precision, and confusion matrix. WRITE ABOUT Visualization TOO. In this project, we have done our analysis and prediction by some specific tools like “Python”, and “Tableau”. To present the results, we have used visualizations, predictions, and a business plan is provided.

# **Introduction**

Employees are the most valuable assets of an organization. Any organization is as successful as its employees, and to stand strong in a long term, they need dedicated employees who work above and beyond to achieve a common organizational goal. Motivated employees make a significant difference in an organization, and it is the duty of the organization to provide a positive environment for their employees to make them stay.

The practice of gathering and analyzing Human Resource (HR) data in order to improve an organization's workforce performance is known as HR analytics. With the help of HR analytics, the human resource department in an organization can determine the rate of employee attrition, the variables that cause employee attrition, which employees are likely to leave, and which are likely to stay. Further, the organization can work on the factors that are causing the attrition and maintain an efficient work environment.

# **Objective**

The objective is to determine the rate of employee attrition, the variables that cause employee attrition, which employees are likely to leave, and which are likely to stay. We will achieve this by fitting a model that can predict the discrete class of new input using machine learning classification algorithms and visualization. Also, we can use this dataset for discovering and visualizing key factors affecting the rate of employee attrition. Moreover, by providing for the prediction can result in an efficient model which can be used by different organizations to decrease the rate of attrition and improve the factors that affected the rate of attrition in the company.

## Descriptive Analysis

The primary goal of this project is to conduct an in-depth analysis of the employee attrition rate using the company’s data and the employee engagement survey. Through the data we will analyze the following questions:

* What is the situation of the attrition rate?
* What are the various factors which affect the employee attrition rate?
* Why has such an attrition rate occurred in the company?

## Diagnostic Analysis

The project will have a diagnostic analysis to understand why the employee attrition rate is high and find a reason why some of the departments have higher resignations. The project will help to understand

1. Which factor have a higher impact on employee attrition rate?
2. The overall correlation between job satisfaction and attrition rate.

## Predictive Analysis

The project will use the dataset and employee profiles to predict the employees who may resign in the coming year.

## Prescriptive Analysis

Lastly, the project will seek prescriptive analysis and determine if there are any chances of top performers employees might leave the company.

# **Data Dictionary**

# The dataset is obtained from Kaggle collected from random people in a specific period. The data set contains 5 rows and 35 columns. Link <https://www.kaggle.com/code/iambideniz/hr-analytics-prediction-why-do-people-resign/data>

# **Significance of the Project**

The project will assist in determining which variables contributed the most to employee turnover and in developing a model that can forecast whether a certain employee would quit the

organization. The purpose is to develop or improve various retention methods for certain personnel and have an idea of resource planning. Overall, the use of this approach will enable any business management to make better decisions.

# **Literature Review**

**(Abdulquadri Ade, 2015)** expressed the issue of employee turnover has raised the burden in the construction industries which affects productivity. It was mentioned that only a few resources have been used in recruiting, training, and developing the employees, due to which they leave for other organizations. Their research was aimed at determining the impact of employee turnover rate in small and medium construction firms in Nigeria.

**(N & Paramashivaiah, 2015)** articulated that in a BPO Sector, employees leave the organization either entirely or change their industry. Since all the research is focused on the issues and antecedents of attrition the consequences were not much focused on, so the study focuses on a holistic view rather than a specific portion.

**(Shine, 2015)** in the research said that the employees working in the IT sector in various cities of India (Mumbai, Pune, Bangalore, Delhi, and Indore) to analyze the case of attrition. Samples were collected with 21 factors/items which resulted in 5 factors. Findings of the study suggest organizational culture, working conditions, career growth opportunities, work pressure, and mutual trust. The conclusion that was provided was employee friendly organizational culture with positive working conditions, low pressure, and higher opportunities for career growth.

**(Farooq, 2015)** investigated several factors that are affecting the high attrition rate in the hotel industry. A study conducted by an industry chamber located in New Delhi reveals that the attrition rate in the hospitality industry in India is set to double to nearly 50 percent by 2010, up from the earlier 25 percent growing at an alarming rate of 10 percent per annum. The result showed that the lack in salary, the stress in job, compensation, better opportunities, and stake issues were the major reasons behind the problem of attrition in the company.

**(Dorance, 2014)** mentioned that many organizations have concerned about employee turnover. It is highly destructive to both the organization as well as the employees. The research was conducted for the retail industry in Bangalore. Despite the incentives, motivational techniques, and old practices of HRM there is still a high attrition rate, so the study showcases ways to reduce the intentions of employees leaving the organization. This is performed by distributing questionnaires to retail outlets in Bangalore. The study concluded that turnover intention has an influence on attrition factors such as QWL, career growth, working hours, personal/family reasons, relation with internal co-workers, welfare, working condition, and salary.

**(Manjula, 2013)** articulated that the objective of the study is to understand the growth of the ITES sector in India. The study in the BPO industry is to understand and gauge the attrition rate, and its intensity and make a causal analysis, to design a strategy to stabilize the sector by suggesting mitigating the attrition.

**(Mishra & Mishra, 2013)** articulated that the study carried out is for the shipping industry of Kutch, Gujarat. The purpose of the study is to analyze the turnover and the commitment to identify several domains of organization, human resource practices, and other employee characteristics and environmental factors, which may have a positive or negative impact on employees ‘intention to stay with an organization.

**(Gupta, 2013)** stated that in recent years the turnover is high in BPO sectors either to absenteeism or employees absconding without any prior notice. Turnover rates for permanent Agents/Executives were 15.6% in 2009 and 35% in 2012. Department of Human Resources which also tracks attrition of temporary employees measured the turnover rate for temporary employees to be 77% in 2012. Therefore, the study is focused on recruitment and retention challenges that the IT/BPO industry currently faces and examines ways to reduce high turnover rates among first-year Employees in the leading Domestic Call Center based in Indore.

**(Negi, 2013)** stated that the study explains that the attrition rate is inevitable but manageable because it’s hard to provide a permanent and promising workforce. This research shows the cause of attrition from various points of view and what are the effects for both the employer and employee. It also gives us a view of the positive side of attrition and the role of leadership skills in controlling attrition.

# **Methodology and Analysis**

This project will use three methods of machine learning.

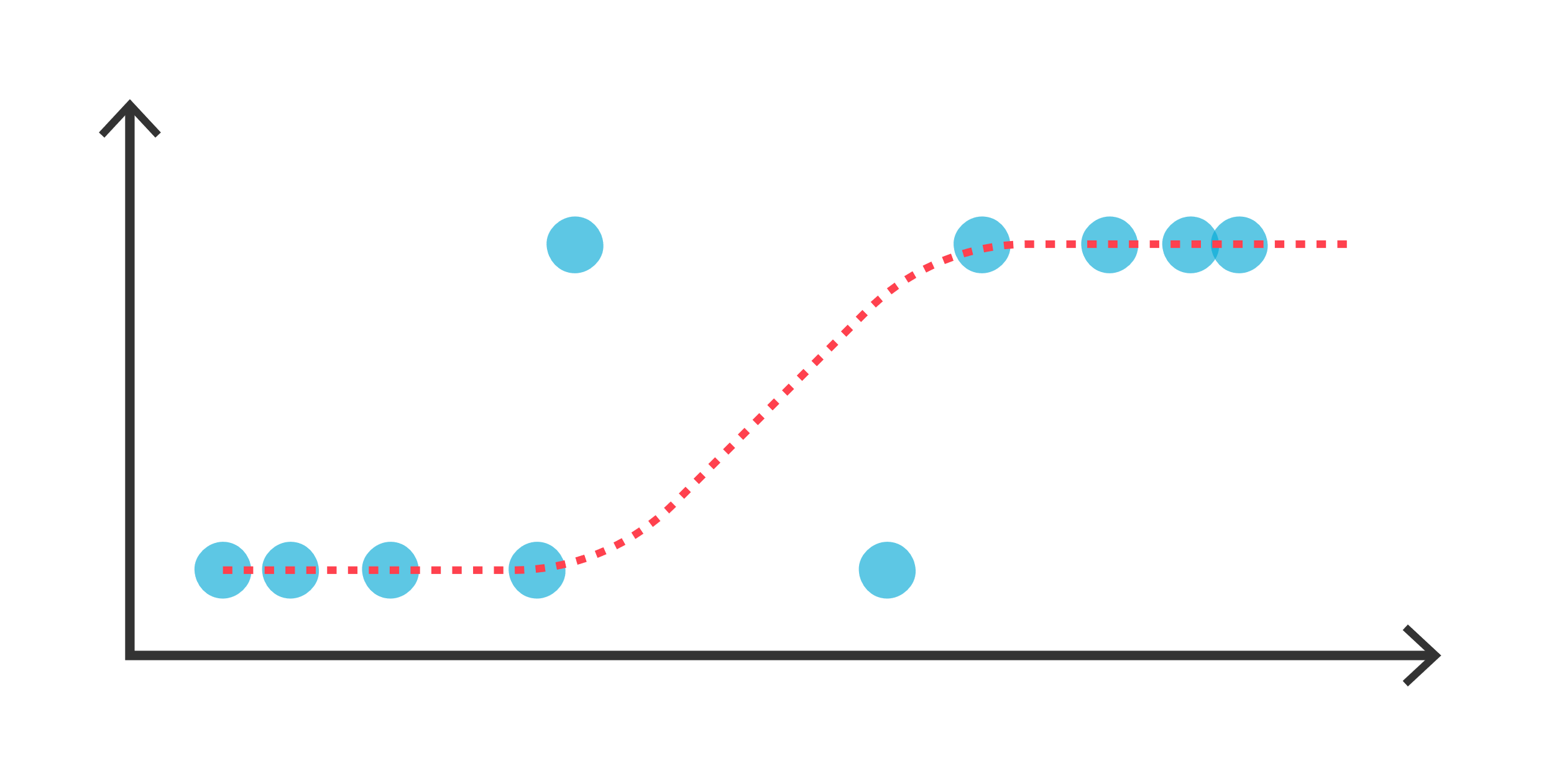
# Logistic Regression,

# Random Forest Regression, and

# Decision Tree Algorithm

These three methods are further explained below:

**Logistic regression** is a very powerful modeling tool and is a generalization of linear regression. Logistic Regression is used to assess the likelihood of a disease or health condition as a function of a risk factor (and covariates). Both simple and multiple logistic regression, assess the association between the independent variable(s) (Xi) -- sometimes called exposure or predictor variables — and a dichotomous dependent variable (Y) -- sometimes called the outcome or response variable. It is used primarily for predicting binary or multiclass dependent variables.



**Random decision forests** are an ensemble method for classification, regression, and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of overfitting to their training set.

Chart, diagram

Description automatically generated

**The Decision Tree algorithm** belongs to the family of supervised learning algorithms. Unlike other supervised learning algorithms, the decision tree algorithm can be used for solving regression and classification problems too. The goal of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable by learning simple decision rules inferred from prior data (training data).

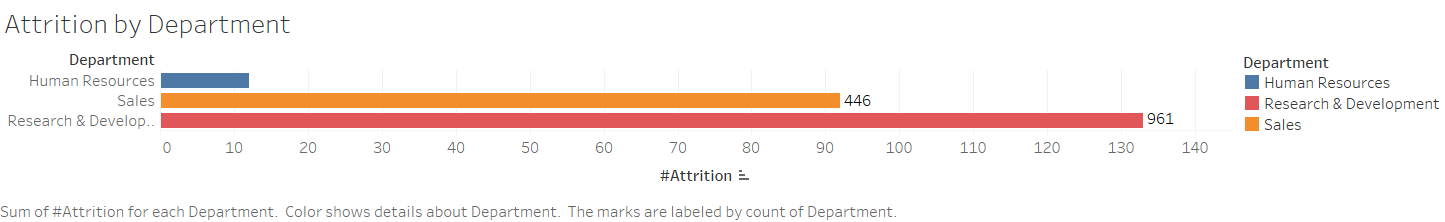


# Text, table Description automatically generated**Descriptive Analysis**

## 1. Key Performance Indicators

**Observations:** Out of 1470 employees, 237 employees responded for change of their job. The average age of these employees is 37 and their average monthly income is $4919. Also, the average total working years of these employees is 11 years and their average years at the company is 7 years.

## 2. Attrition by Department



**Observations:** The graph illustrates that Research & Development department has the higher attrition rate, followed by Sales department, then Human Resources.

## 3. Attrition by Job Role

Chart

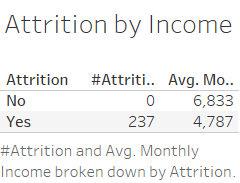
Description automatically generated

**Observations:** As seen in the chart above, Laboratory Technicians have the highest attrition rate. It is followed by Sales Executive, Research Scientist, Sales Representative, Human Resources, Manufacturing Directive, Health care Representative, and Manager respectively. The Research Director role has the least attrition rate.

## Attrition by Department4. Attrition by Department

**Observations:** Only three major departments are included in this report, with the Sales Department having the highest attrition rates, followed by the Human Resource Department.

## 5. Attrition by Income



**Observations:** The attrition rate is very high at extremely low-income levels—average $4787 per month, as shown in the table above. The odds of an employee quitting the company is zero when the monthly salary is fair – average $6,833.

## 6. Attrition by Marital Status

Chart, pie chart, bubble chart

Description automatically generated

**Observations:** Looking at the graph, Single employees have the highest attrition rate compared to married and divorced employees. However, the difference is not very significant, thus, we can conclude that marital status is not an important factor in employee attrition.

## **7. Tableau Dashboard**Dashboard 1

# References

<https://www.kaggle.com/paramarthasengupta/hr-analytics-prediction-why-do-people-resign/data>

<https://www.business.com/articles/how-employees-make-or-break-business-success-and-how-you-can-lead-the-way/>

<https://tex.stackexchange.com/questions/503883/illustrating-the-random-forest-algorithm-in-tikz>

<https://www.tibco.com/reference-center/what-is-logistic-regression>

# Bibliography

Abdulquadri Ade, A. M. (2015). *“Impact of Employee Turnover in Small and Medium Construction Firms: A Literature Review”, .* International Journal of Engineering Research & Technology (IJERT).

Dorance, B. (2014). A Study on Attrition – Turnover Intentions in retail Industry. *International Journal of Business and Administration Research Review*, 55.

Farooq, A. (2015). “Employee Attrition Analysis of Hotel Industry. *International Journal of Advancement in Engineering Technology, Management & Applied Science*, 34.

Gupta, V. (2013). An Analysis of Attrition: Retention Strategy forIT/BPO Industry. *International Journal of Advance Research in Computer Science and Management Studies*, 434.

Manjula, N. (2013). Employee Attrition in ITES sector - A casual Aanalysis. *ASM’s International E-Journal of Ongoing Research in Management and IT*, 70.

Mishra, S., & Mishra, D. (2013). Review of Literature On Factors Influencing Attrition And Retention. *International Journal of Organizational Behaviour &Management Perspective © Pezzottaite Journals*, 435.

N, B., & Paramashivaiah, P. (2015). The Real Problem Behind the Growth of Business Process Outsourcing (BPO) Companies. *Journal of Social Welfare & Management*.

Negi, G. (2013). Employee Attrition: Inevitable yet manageable. *International Monthly Refereed Journal of Research In Management & Technology*, 50.

Rosy. (2015). abc. *dkkd*.

Shine, D. (2015). Attrition in IT. *International Journal Of Core Engineering & Management (IJCEM)*, 236.