
Data Science Project

IBM HR Analytics Employee Attrition

FY 2022-2023

ANNEXURE-I
DECLARATION

I hereby declare that this Project Report titled IBM HR Analytics Employee Attrition submitted by me is a bonafide work undertaken by me and it is not submitted to any other University or Institution for the award of any degree/ diploma/certificate or published any time before.

Name of the student: Harshal Ramesh Wadke

Signature of the Student

ANNEXURE-II
CERTIFICATION

This certificate is given to Mr. Harshal Ramesh Wadke for successful submission of the project on the topic IBM HR Analytics Employee Attrition under the department of Data Analytics for the academic year 2023

Name of the Supervisor

Signature of the Supervisor

Abstract

The IBM HR Analytics Employee Attrition project was a comprehensive data analytics project aimed at identifying and predicting the factors that contribute to employee attrition in the company. The project used a large dataset of employee information, including attrition, job satisfaction, environmental satisfaction, performance, age etc. to analyze and identify patterns and correlations that could help understand why employees leave the company.

The project involved various steps, including data cleaning, exploratory data analysis, and feature engineering. The team used programming language Python and its library, such as Pandas, NumPy, Matplotlib, and Seaborn to understand the factor effecting the employee attrition.

The results of the project showed that the most significant factors that contribute to employee attrition were job satisfaction, age, job role and Monthly Income.

The project's findings provided insights into the factors that contribute to employee turnover and helped to identify areas where they could improve their employee retention strategies. By addressing these factors, the company could reduce its attrition rate, improve employee satisfaction, and ultimately improve its bottom line. The project also helped to understand the importance of employee engagement and retention and highlighted the need for ongoing data analysis to monitor and improve HR strategies. The project's success demonstrated the value of using data analytics techniques to analyze employee data and improve HR strategies.

The project's approach can be applied to other companies to help them understand their employee turnover patterns and improve their HR strategies. Additionally, the project highlighted the importance of data privacy and security, with IBM taking steps to ensure that employee data was handled securely and in compliance with regulations.

The project's success also demonstrated the importance of collaboration between HR professionals and data analysts to identify and address employee retention issues. The project's findings helped to create a more data-driven approach to HR and provided insights that were used to improve employee retention strategies across the company.

Overall, the IBM HR Analytics Employee Attrition project serves as an example of how companies can leverage data analytics to improve HR strategies and employee engagement. By analyzing employee data and identifying patterns, companies can gain insights into the factors that contribute to employee attrition and develop effective retention strategies that improve employee satisfaction and ultimately, their bottom line.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who contributed to the successful completion of the IBM HR Analytics Employee Attrition project.

We would also like to thank the Data Science and HR teams who collaborated with us on this project and provided valuable input and feedback. Their expertise and knowledge of the company's HR policies and practices were instrumental in helping us understand the context and interpret the results of the analysis.

We are grateful to the data engineering team, who ensured that the data was properly cleaned, formatted, and prepared for analysis. Their efforts enabled us to work with a high-quality dataset and achieve accurate results.

We would like to extend our thanks to the IT team, who provided technical support and ensured that the necessary infrastructure and tools were in place for the project. Their efforts were critical in enabling us to conduct the analysis efficiently and effectively.

We would also like to acknowledge the contribution of the Data Privacy and Security teams, who ensured that the employee data was handled securely and in compliance with regulations. Their efforts were critical in maintaining the integrity and confidentiality of the data throughout the project.

Finally, we would like to thank our team members for their hard work, dedication, and collaboration throughout the project. Their insights, expertise, and teamwork were critical in delivering the project successfully.

In conclusion, we would like to express our gratitude to all those who contributed to the IBM HR Analytics Employee Attrition project. Their efforts and support were instrumental in enabling us to achieve our goals.

Table of Contents

Sr.no	Chapter	Pg.no
1	List of Figure	7
2	Introduction	8
3	Scope	9
4	Literature Review	10
5	Methodology	12
6	Historical Retrospect	14
7	Data Analysis	16
	7.1. Introduction	16
	7.2. Data Acquisition	16
	7.3. Data Specification	16
	7.4. Univariate Analysis	17
	1. Attrition	17
	2. Age	17
	3. Gender	18
	4. Marital Status	18
	5. Department	18
	6. Performance Rating	19
	7. Job Satisfaction	19
	8. Business Travel	19
	9. Distance from Home	19
	10. Monthly Income	20
	11. Education Field	21
	12. Performance and Over Time	21
	13. Year with Current Manager	21
8	Summary	23
9	Conclusion and Recommendation	25
10	Bibliography	27
11	Appendices	28

List of Figures

Sr.no.	Figures	Pg.no
1	Count of Attrition	17
2	Relation of Age with Attrition	18
3	Relation of Gender with Attrition	18
4	Relation of Marital Status with Attrition	18
5	Relation of Department with Attrition	18
6	Relation of Performance Rating with Attrition	19
7	Relation of Job Satisfaction with Attrition	19
8	Relation of Business Travel with Attrition	20
9	Relation of Distance from Home with Attrition	20
10	Relation of Monthly Income with Attrition (Box Plot)	20
11	Relation of Monthly Income with Attrition (Count Plot)	21
12	Relation of Education Field with Attrition	21
13	Year at company and Year with last promotion	21
14	Relation of Over time with Attrition	21
15	Relation of Year with Current Manager with Attrition	22

Introduction

Analytics can be defined as a “logical progression and set of statistical tools”. In a simple way Analytics is a science of analysis. HR analytics is defined as “demonstrating the direct impact of people data on important business outcomes”. “HR Analytics is the systematic identification and quantification of the people drivers of the business outcomes”. HR Analytics is the need of the company so that they can spend the money on the right employees rather than spending on wrong ones. With the help HR Analytics, HR Managers can take numerous decisions on investment on employees to get the excellent outcomes that benefits the stakeholders and customer. HR Analytics deals with the Human Resource Management processes and can be used by HR Managers, Project Managers and Line Managers. Employee Turnover is the most important factor that causes the huge loss to the company. There are many reasons for which the employees leave the company, such as; salary dissatisfaction, stagnant career growth, etc. The loss is not only in terms of the money but also the company sometimes loses the skilled employees who are the most valuable assets to the company. If the company can predict the employee attrition in near future, they can also work on retention beforehand and avoid the loss of valuable employee. The prediction of attrition and retention is the part of the HR Analytics. Employee Turnover is the most important factor that causes the huge loss to the company. There are many reasons for which the employees leave the company, such as; salary dissatisfaction, stagnant career growth, etc. The loss is not only in terms of the money but also the company sometimes loses the skilled employees who are the most valuable assets to the company. If the company can predict the employee attrition in near future, they can also work on retention beforehand and avoid the loss of valuable employee. The prediction of attrition and retention is the part of the HR Analytics.

Using Predictive Analytics, we can predict the employee attrition. Python programming language and its libraries such as Pandas, NumPy, Matplotlib and Seaborn are used for analysis. “Predictive analytics is the practice of extracting information from existing data sets to determine patterns and predict future outcomes and trends”. Employee Attrition Prediction helps HR Managers to predict how many employees will resign from the company during a particular period. Thus, Managers can work on deciding on the valuable employees and try to retain them.

Scope

The scope of the IBM HR Analytics Employee Attrition project was to investigate the factors contributing to employee attrition in the company and develop effective retention strategies to reduce the attrition rate. The project aimed to use a large dataset of employee information, including job satisfaction, and performance metrics, to identify patterns and correlations that could help understand why employees leave the company.

The project's primary objective was to identify the factors contributing to employee attrition in the company and develop strategies to address them. The project involved analyzing employee data to identify patterns and trends that could help understand why employees were leaving the company. The data analysis included looking at demographic factors, such as age, gender, and tenure, as well as job-related factors, such as job satisfaction and work-life balance.

The research project aimed to address several research questions, including what factors contribute to employee attrition, how to improve job satisfaction and work-life balance, and how to develop effective retention strategies to reduce the attrition rate. The project also aimed to assess the performance of different retention strategies and identify the most effective strategies for the task.

The project's findings provided insights into the factors that contribute to employee turnover and helped to identify areas where they could improve their employee retention strategies. The project showed that factors such as job satisfaction, work-life balance, and performance evaluation were significant contributors to employee attrition. Based on these findings, HR could develop retention strategies that address these factors, such as improving work-life balance programs or increasing job satisfaction through better management practices.

The scope of the research project was limited to the organization and may not be generalizable to other industries or organizations. However, the approach and methodology used in the project could be applied to other companies to help them understand their employee turnover patterns and develop effective retention strategies. The project highlighted the importance of employee retention and the need for organizations to invest in strategies to retain their best employees.

Literature Review

Employee attrition has been a longstanding challenge for organizations of all sizes and industries. It is a complex issue that affects the bottom line of organizations, as well as the morale and productivity of remaining employees. There is a wealth of research on employee turnover and retention strategies that can inform the analysis of the IBM HR Analytics Employee Attrition.

One of the main theories that inform the study of employee attrition is the job satisfaction theory. This theory suggests that employees who are satisfied with their job are more likely to stay in their current position. Conversely, employees who are dissatisfied with their job are more likely to leave the organization. This theory has been supported by numerous studies, which have found that job satisfaction is a key factor in employee retention.

Several studies have found that job satisfaction is positively related to employee commitment to the organization, employee performance, and employee retention (Hom, Tsui, Wu, Lee & Zhang, 2009; Petrucci, 2016; Tett & Meyer, 1993). Job satisfaction is influenced by various factors such as pay, recognition, job autonomy, leadership, job security, and work-life balance (Robbins & Coulter, 2018). Employee satisfaction is an important factor in employee retention as it is more cost-effective to retain current employees than to hire and train new ones.

Another theory that has been used to explain employee turnover is the organizational commitment theory. This theory suggests that employees who feel a sense of commitment to the organization are more likely to stay in their current position. Conversely, employees who do not feel a sense of commitment to the organization are more likely to leave. Studies have found that organizational commitment is positively associated with job satisfaction and negatively associated with employee turnover (Hom et al., 2009; Petrucci, 2016; Tett & Meyer, 1993).

In addition to these theories, there are several factors that have been identified as contributing to employee attrition. One such factor is employee compensation. Studies have found that employees who feel that they are not fairly compensated for their work are more likely to leave the organization (Hom et al., 2009; Robbins & Coulter, 2018). Other factors that have been identified include lack of career advancement opportunities, poor management, and job stress.

Several retention strategies have been identified as effective in reducing employee attrition. One such strategy is employee training and development. Studies have found that employees who receive training and development opportunities are more likely to stay in their current position (Hom et al., 2009; Petrucci, 2016). Another effective retention strategy is employee recognition and rewards. Studies have found that employees who feel that their work is recognized and appreciated are more likely to stay with the organization (Robbins & Coulter, 2018).

Organizational culture has also been identified as a factor that affects employee retention. A positive organizational culture is one where employees feel valued, respected, and supported. Studies have found that organizations with a positive culture have higher employee retention rates (Petrucci, 2016). In contrast, organizations with a negative culture, where employees feel unsupported and undervalued, have higher rates of employee attrition.

Conclusion:

Employee attrition is a significant challenge for many organizations, and several factors contribute to employees leaving their jobs. To address this issue, many organizations have implemented retention strategies, such as training and development programs, flexible work arrangements, and employee recognition programs. However, the effectiveness of these strategies depends on various factors, such as organizational culture, employee characteristics, and job characteristics. The literature review provides an overview of previous research on employee turnover and retention strategies to inform the analysis of the IBM HR Analytics Employee Attrition project.

Methodology

Introduction:

This section provides the theoretical and technical walkthrough of the research method used for building an analytical application using Python for understand the factor effecting employee attrition and how to recognize the valuable employee and retain them, thereby saving the company HRM budget on hiring new employee. This chapter describes the set of methods used for carrying out the research. It also covers the data collection and data analysis methods.

Research Philosophy:

The purpose of this research is to predict the employee attrition and improve retention of valuable employee, thereby saving the HRM cost. For the research study, both the qualitative and quantitative methodologies are used. Qualitative methodology is used for gathering the employee attributes and demographics for predicting attrition as well as finding the retention factors for retaining a valuable employee, whereas we used the quantitative methodology for weighting the factors influencing the attrition.

Research Approach:

For completing our research, we followed the inductive approach. Inductive approach is the one in which there is a systematic study on observation and previous research for proposing the theories and fining the patterns with the use of various assumptions and hypotheses. The assumptions and the approach for research changes with the progress in the research.

Research Strategy:

For successfully completing the research we took the help of survey which was carried among the HR Professionals to collect the factors and attributes of employee for predicting employee attrition and improving employee retention.

Data Collection:

The first step is to collect relevant data on employee from the organization's HR department. We collect the data from GitHub and this data include information on employee demographics, job characteristics, work history, job satisfaction, job involvement, and monthly income.

Data Cleaning and Preprocessing:

The second step is to clean and preprocess the data to remove any missing or inconsistent values and ensure

that the data is in a usable format. This involves tasks such as null values treatment, data validation, data transformation, and data normalization.

Data Analysis:

The third step is to analyze the data using descriptive statistics, such as mean, median, and mode, to identify trends and patterns in the data. This involves creating charts and graphs to visualize the data and to identify significant relationships between variables.

Identify Factors Contributing to Employee Attrition:

The fourth step is to identify the factors that contribute to employee attrition by analyzing the data. This involves using graphical analysis or other statistical techniques to identify significant predictors of turnover intentions, such as job satisfaction, job involvement, and environmental satisfaction.

Develop Recommendations:

The final step is to develop recommendations for the organization based on the analysis of the data and the identification of the most effective retention strategies. This involves providing specific recommendations for the organization to implement to reduce employee attrition and improve employee retention.

Conclusion:

In conclusion, the IBM HR Analytics Employee Attrition project involves analyzing data to identify the factors that contribute to employee turnover and the most effective retention strategies. The methodology involves data collection, cleaning, analysis, and the identification of factors contributing to employee attrition and retention strategies. By using this methodology, organizations can identify the factors contributing to employee turnover and implement effective retention strategies to reduce employee attrition and improve employee retention. The results of the analysis can provide valuable insights into the organization's HR practices and inform the development of HR policies and programs that address employee attrition. The IBM HR Analytics Employee Attrition project is a useful tool for organizations looking to improve their employee retention rates and create a more engaged and productive workforce.

Historical Retrospect

The entity of the study for the IBM HR Analytics Employee Attrition project without using machine learning is the concept of employee turnover and retention in organizations. This concept has a long and complex history that dates back to the early industrial era, when the growth of large factories and organizations led to the development of complex management systems and hierarchical structures. The concept of employee turnover and retention became particularly important during the mid-20th century, when the rise of human resource management (HRM) as a distinct field of study and practice led to a greater emphasis on employee management and development.

During the early 20th century, the concept of employee turnover and retention was largely ignored by management theorists and practitioners. However, this began to change during the mid-20th century, when a number of studies were conducted that identified the negative impact of high employee turnover rates on organizational performance and productivity. These studies led to the development of a number of theories and frameworks that sought to explain the causes of employee turnover and to identify strategies for reducing it.

One of the most influential early theories of employee turnover was developed by the sociologist William H. Whyte Jr., who argued that employees left their jobs due to a lack of social ties and emotional attachments to their workplaces. This theory was later expanded upon by other sociologists, who identified a number of other factors that contribute to employee turnover, including job satisfaction, organizational commitment, and work-family conflict.

During the 1960s and 1970s, a number of studies were conducted that sought to identify the most effective retention strategies for organizations. These studies identified a number of effective retention strategies, including training and development programs, flexible work arrangements, and employee recognition programs. However, it was not until the 1980s and 1990s that the concept of employee retention became widely accepted as a key management issue.

During the late 20th century and early 21st century, the field of HRM continued to evolve, with a greater emphasis on strategic HRM and the development of new technologies and techniques for analyzing employee turnover and retention. The rise of data analytics and machine learning has led to the development of new methods for analyzing and predicting employee turnover, with a focus on identifying the most effective retention strategies for specific groups of employees.

Despite these developments, the concept of employee turnover and retention remains a complex and multifaceted issue that requires ongoing research and analysis. The IBM HR Analytics Employee Attrition project is an important tool for organizations looking to understand the factors that contribute to employee turnover and to develop effective retention strategies. By analyzing data and identifying the most effective retention strategies, organizations can create a more engaged and productive workforce and improve their overall performance and productivity.

Data Analysis

Introduction:

This chapter mainly focuses on the Findings obtained from the research and how they were used for progressing in the research work. It also includes the Data Analysis with covering the both Qualitative and Quantitative aspects of the research. Finally, it highlights the result obtained from the research in the same manner how they would be presented to the user. Also, this includes how the Methodological Assumptions were used to complete the research work and build the final working analytical application.

When the, research was started, with the study on the previous research work and case studies, we found that various methods and tools were developed for predicting the attrition, but it was for making the company ready for the new hire as a replacement of the old one. Sometimes, company lost the valuable employee and had to spend a good amount of money on hiring new employee and on new hire's training. There was no tool developed which performed the automated decision of retention. HR Managers needed to take the decision manually by involving in the discussion with the line managers and employee's reporting managers to know about the capability of the employee and decide whether to retain the employee or not and what are the factors which can help them to retain a valuable employee. But, with the help of an analytical application, it performs the two main type of analysis – predictive and prescriptive analysis. With the predictive analytics it predicts the employee attrition and the prescriptive analysis helps HR Managers to decide on valuable employees and know the factors most effective in retaining the employees.

Data Acquisition and Cleaning:

Data Acquisition: We use IBM HR Analytics Employee Attrition & Performance data, which is created by IBM data scientists. It has 1470 rows x 27 columns and contains numeric and categorical data types in columns. We loaded the dataset in csv format and read it in the Jupyter notebook after importing necessary libraries.

Data Specification: The dataset has 1470 rows and 27 columns. Rows are observations from each employee and columns are from different features which are obtained in order to explain the employee attrition. The features data types consist of 18 integers and 9 objects. List of attributes are presented below.

Integer	Object
Age, Distance from Home, Employee Number, Environment Satisfaction, Hourly Rate, Job Involvement, Job Satisfaction, Monthly Income, Num Companies Worked, Performance Rating, Relationship Satisfaction, Total Working Years, Training Times Last Year, Work Life Balance, Years at Company, Years In Current Role, Years Since Last Promotion, Years With Current Manage	Attrition, Business Travel, Department, Education Field, Gender, Job Role, Marital Status, Over 18, Over Time,

We searched for missing values in every feature of dataset, Total 231 null values were found from different objects. We observed 5 random sample records in the dataset to grasp the general intuition about whole picture. Besides that, we explored the statistical attributes of each feature such as their mean, median and mode. This research also gave us a general impression about unique and top values for each attribute in addition to their frequencies in the dataset. We made double checks on some of features in order to make sure that everything is good to go. Those results were also okay.

Univariate analysis:

- 1. Attrition:** In the company, there are 1470 employees. 237 employees who compose 16% of the total number of employees left the company for some reasons. Besides that, 1233 employee is currently continuing to work in the same company as shown in figure 1.

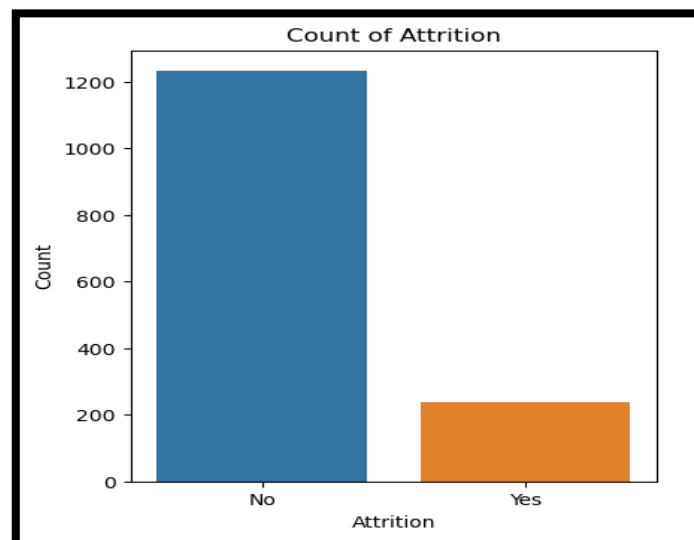


Fig 1

- 2. Age:** As per the below graph (fig 2) In 18-21 age group, young employees are more likely to leave the company. If we evaluate overall attrition number in the company, 26-35 age group's attrition number is the highest comparing to other age groups. 35-60 age group generally prefers to secure their job in the same company.

- 3. Gender:** Male employees are more likely to leave the company than female employees as seen in fig 3.

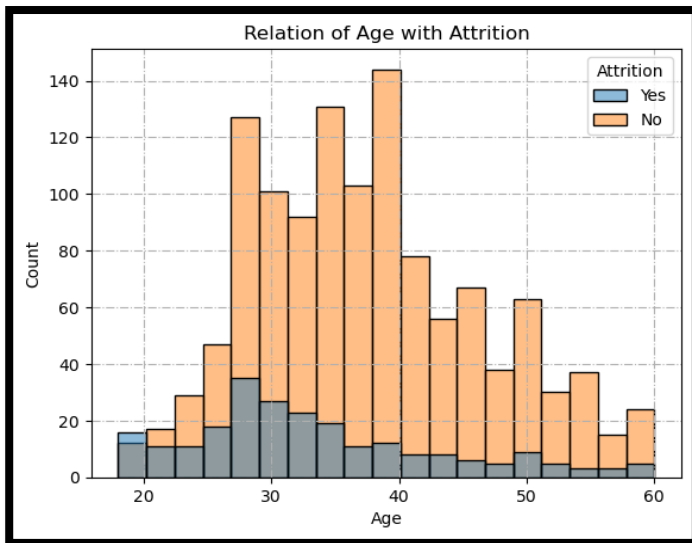


Fig 2

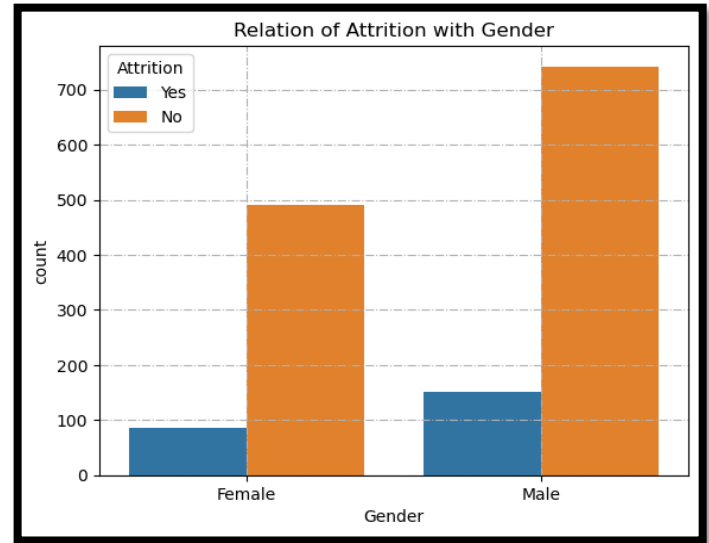


Fig 3

- 4. Marital Status:** As show in fig 4 single employees are more likely to leave the company. They have the highest attrition number and compose of the 50.6% employees who left the company. Married and Divorced employees are the followers of Single employees in the attrition number of the company respectively.
- 5. Department:** There are three departments in the company. As per graph (fig 5) Research & Development Department has the most attrition number in the company. In numbers, it is equal to 133, Actually, this attrition is a big number for company, but compared with other departments, Research & Development Department has the lowest attrition rate in itself as an individual department. Sales Department has mostly been affected by the attrition. Because 92 out of 237 employees left the organization. This is the highest number compared to the other two departments. Human Resources Department follows the Sales Department in terms of being affected by attrition itself. 12 out of 237 employees left the organization. But this is not that huge number in terms of whole attrition in company.

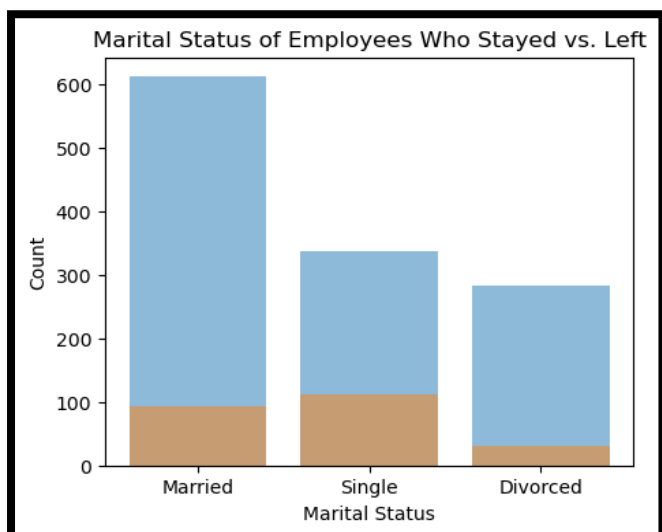


Fig 4

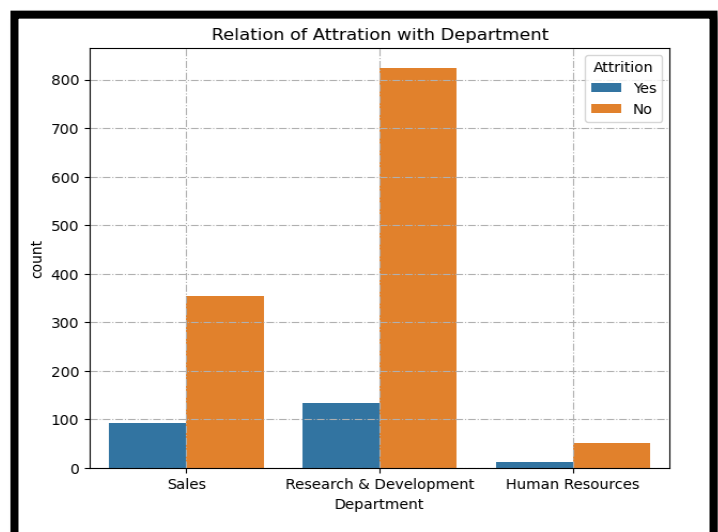


Fig 5

- 6. Performance Rating:** Performance rating has two category such as 3 and 4. Not surprisingly, performance rating 3 group has the highest attrition number as per graph (fig 6) 200 out of 237 employees left the organization.
- 7. Job Satisfaction:** In high job satisfaction, surprisingly employees leave the company most and their attrition composes 21.94% of company's attrition. From below graph (fig 7), We assume that job satisfaction should not be the main reason for employees to leave the company. As it may be expected, in low job satisfaction, employees leave the company more than other groups except high satisfaction. They compose 27.8% of all attrition in the company.

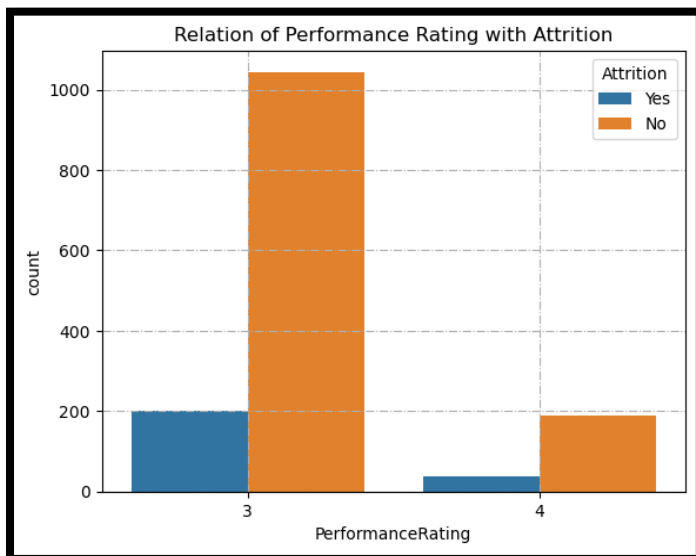


Fig 6

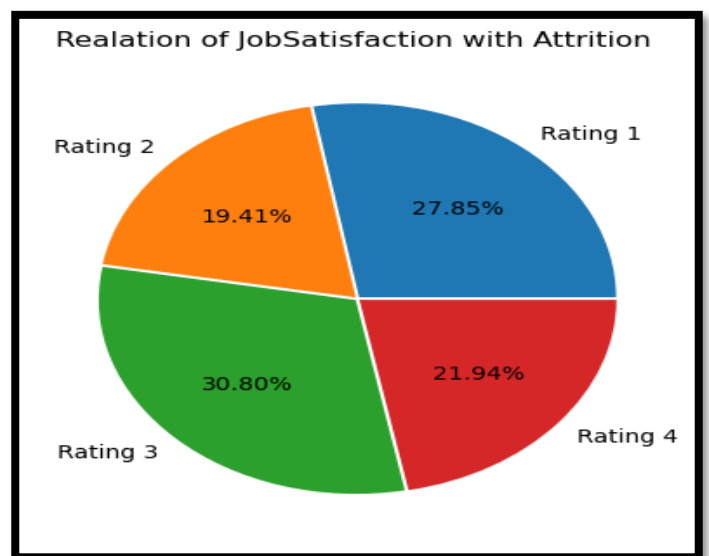


Fig 7

- 8. Business Travel:** As shown in below graph (fig 8) In the company, most of the employee travel rarely or don't travel according to their job description. The rest of the company employees which is 19.9% of them must travel frequently (277 out of 1470). The highest attrition number with 156 belongs to the employees who travels rarely. That is approximately 15% of employees in that group (156 out of 1043). But when you put this number overall attrition, it makes up 65.8% of all attrition in the company (156 out of 237). if we look at the attrition percentage of relevant travel group, the employees who are traveling frequently are in the danger zone. Because they have the highest attrition proportion, which is 24.9%, in their individual travel group (69 out of 277). That group's attrition rate composes of the 29.1% of overall attrition in the company (69 out of 237) Employees who don't travel in their current role have the lowest attrition rate, which is 8%.
- 9. Distance from Home:** Employees whose homes are 1-3 miles far away from the company compose approximately 1/3 of the whole company employee and their attrition rate is 28.7% of all company (68 out of 237). Also, employees whose homes are 10+ miles far away from the company compose approximately the other 1/3 of the whole company employee and their attrition rate is 39.2% of all company (93 out of 237). 6 Attrition rate within its own distance group seems to increase as the distance from home increases. (Refer fig 9)

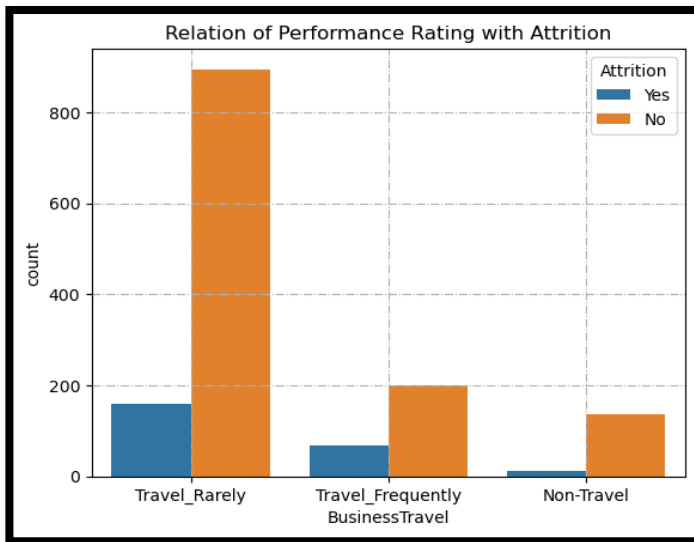


Fig 8

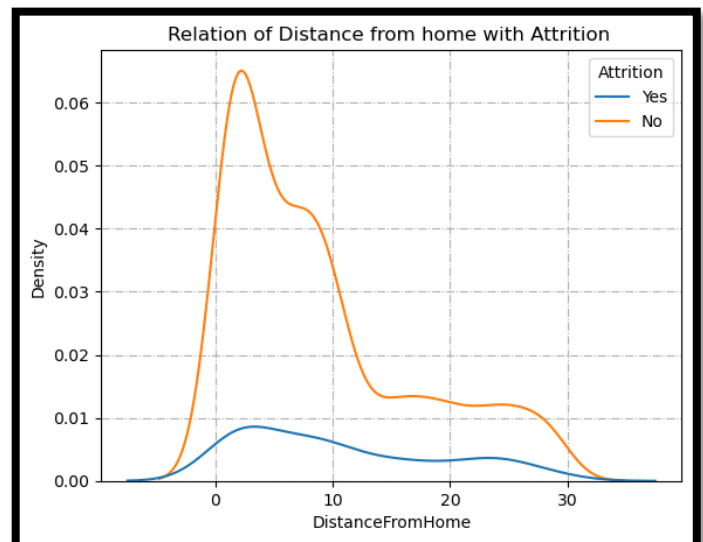


Fig 9

10. Monthly Income: The median salary of Sales Representatives, Laboratory Technicians and Human Resources Employees tends to be on the lower side. 2000-3000 dollars monthly income level, there is a high attrition and it compose the 40% of attrition in the company. 1000-2000 dollars monthly income level, there is a high attrition in its own income group level, which is 54.5%. 12 As the monthly income increase, it is observed that there is a decrease in attrition. But, in 9000-11000 dollars monthly income level, there is a rise in attrition of its own monthly income group level. (Refer fig 10 & 11)

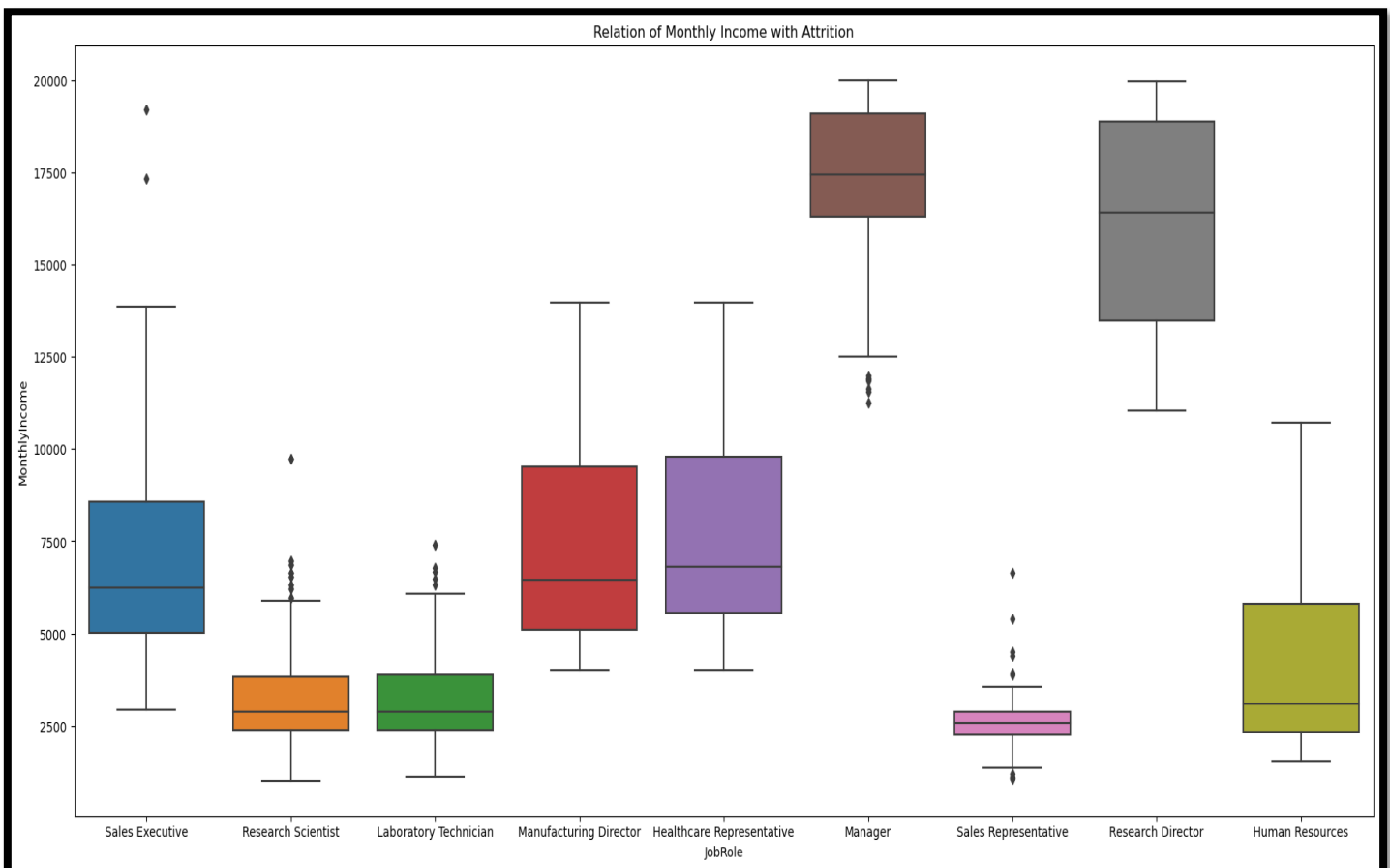


Fig 10

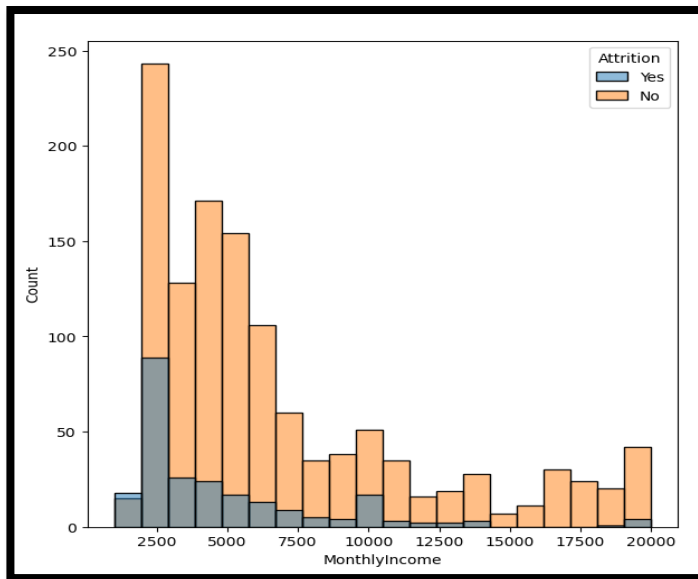


Fig 11

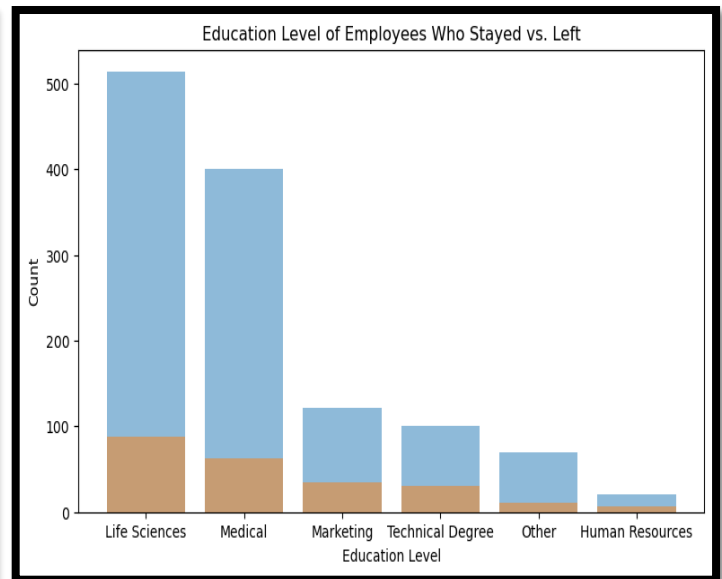


Fig 12

11. Education Field: As per graph (fig 12) Employees who have Life Science education level have the most attrition number which makes up the 37.5% of all attrition (89 out of 237). But that composes only 14.7% of attrition within Life Sciences field. Medical education level has the second highest attrition number which makes up the 13.57% of all attrition (63 out of 237). But that composes only 14.7% of attrition within Life Sciences field. Besides that, Human Resources, Technical Degree, and Marketing fields are mostly affected by the attrition respectively. Their approximately 22-26% employees left the company.

12. Promotion and Over Time: The Promotion graph getting spaced out towards the right-hand side, indicating a positive correlation between these two variables. Here we note two things. Firstly, there is a relatively higher percentage of people working overtime in the group of those who left, secondly, while things seem to be going in the right direction for the group of people who continue with firm, the opposite is happening in the other group (refer fig 13 & 14).

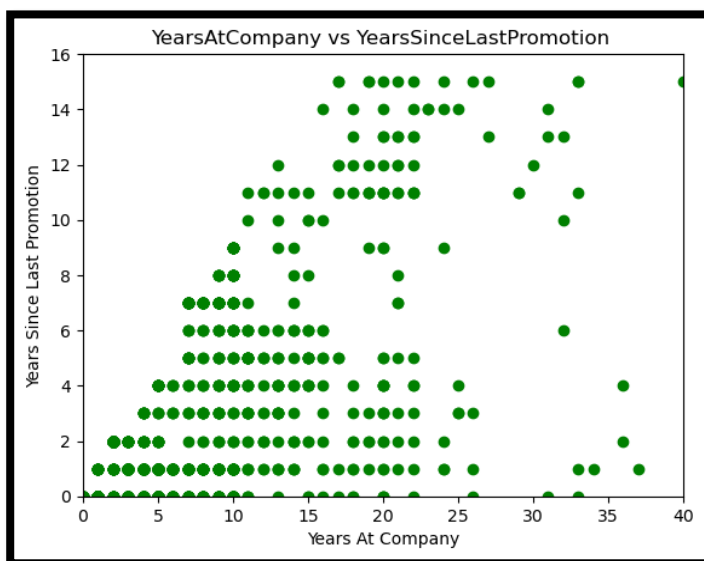


Fig 13

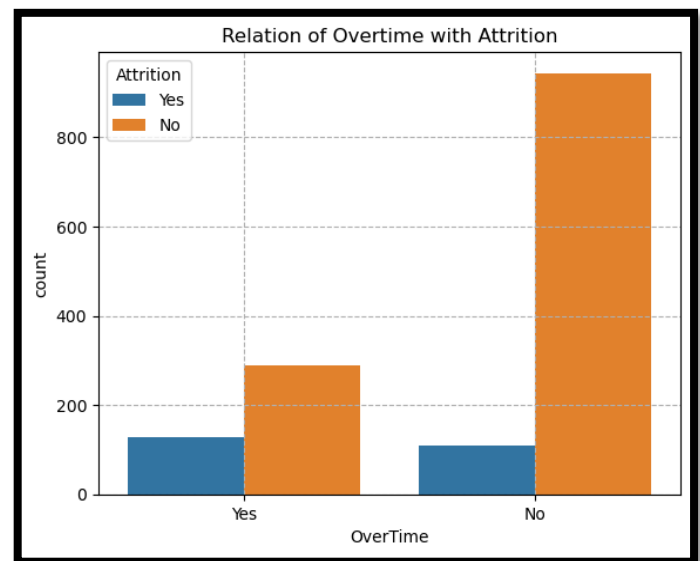


Fig 14

13. Relation with Current Manager: As per graph (fig 15) It is found that employee with new managers

quit he most.

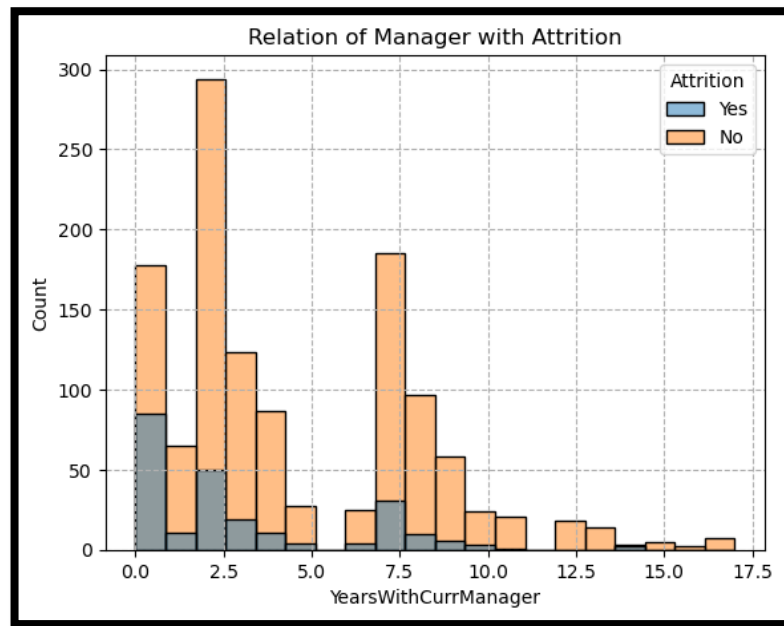


Fig 15

Summary

There are 1470 employees in the company and 16% of them left the company. We have some data about employees to examine the attrition reasons. To sum our exploratory data analysis;

1. Young employees (18-25 years old) compose the 8.4% of the company and they are more likely to leave the company than other age groups. As the employees get older, their attrition percentages drop.
2. Employees who travel rarely or don't travel according to their job description compose the 81.1% of entire company. 18.8% of the employees travels frequently and they have the highest attrition percentage (25%).
3. 30.3% of employees work in the Sales department and they have the highest attrition percentage (20.6%). Human Resources employees who compose the 4.3% of company, are the second highest attrition percentage (19%). Research and Development Department has 65.4% employee in the company and they have the lowest attrition percentage (13.8%).
4. As the distance between company and employees' homes increases, the attrition percentage increases. 11.6% of employees has below college degree and highest attrition percentage (18.2%).
5. Other attrition percentages according to education levels. 39% of employees has the bachelor's degree and 17.3% attrition. 19.2% of employees has the college degree and 15.6% attrition. 27.1% of employees has the master's degree and 14.6% attrition. 3.3% of employees has the doctorate degree and 10.4% attrition.
6. According to the education field, 1.8% of employees has Human Resources education field and highest attrition (25.9%). 8.9% employees have technical degree education field and 24.2% attrition. 10.8% of employees has Marketing education field and 22% attrition. Life Sciences, Medical and other education fields are affected respectively.
7. 61.2% of employees has environment satisfaction at very high/high level in the company. 19.3% of employees has low environment satisfaction and 25.3% attrition.
8. Male employees (60% of company) are more likely to leave the company than female employees.
9. 5.6% of employees works as Sales Representative and they have the highest attrition (39.8%). Laboratory Technician (17.6% of employees) and Human Resources (3.5% of employees) job role follows the Sales Representative attrition percentage by 23.9% and 23.07% respectively. Besides those job roles, Sale Executive (22.8% of the employees have 17.5% attrition) and Research Scientist (19.9% of employees have 16.1% attrition) have the higher attrition percentage than others. 61.3% of employees has high/very high job satisfaction in the company.
10. 19.7% of employees have low job satisfaction and highest attrition percentage (22.8%). As the job satisfaction increase, attrition percentage decreases. Medium (19% of employees) and high (30% of employees) job satisfaction has approximately same attrition percentage (16.5%).

11. 32% of employees are single and has the highest attrition percentage (25.5%).
12. 26.9% of employees have 1000-3000 dollars monthly income and highest attrition percentage (47.7%).
Other monthly income employees have been affected by attrition in the same percentage level except 9000-11000 dollars monthly income level. That group is 5.6% of employees and has 18.6% attrition.
13. The higher percent salary hike is, the more employees are likely and willingly to stay in the current company. Performance rating has two category such as 3 and 4. 84.6% of employees are in the performance rating 3 group and 16.08% attrition. The rest of the employees are in performance rating 4 group and their attrition percentage (16.37%) is a little bit higher than previous group.
14. As the total working years increases, the attrition percentage decreases in the relevant experience groups. 70.6% of employees have 2-3 times training in the previous year and have 17.9% and 14.1% attrition respectively.
15. As the years at company and current role increase, attrition percentage decreases in the company. 39.5% of employees leaves the company before fulfilling one year since their last promotion. 24.3% of employees fulfilling one year since their last promotion has the 13.7 attrition percentage. 7 years since last promotion employee group (5.2%) has the highest attrition (21.1%) in its individual group.
16. Most of the employee (17.9%) quit the company before completing their first year with their current manager (32.3% attrition). Other group who leaves the company most is the ones who work two years with current manager (23.4% of employees with 14.5% attrition).

Conclusion and Recommendation

This chapter concludes the research work and recommends on the future advancements of the application and research.

1. Males at higher attrition risk comparatively than females.
2. 5.6% of employees works as Sales Representative and 17.6% of employees works as Laboratory Technician. They have 39.8% and 17.6% attrition percentage respectively. These two job roles should be questioned, and the company should find the reason(s) why these job roles face more attrition rate than all others and take necessary actions.
3. 28.3% of employees have over time work in the company and 30.5% of those employees leave the company. As it is also reflected in the model, employees working overtime are significantly more likely to resign. Therefore, the company should understand the reason why they are working overtime. Is it for too high workload or are employees' qualifications not enough to complete the scheduled tasks on time? Maybe there might be some other reasons behind that. Our recommendation will be to understand the reason(s) for overtime with detail research and take appropriate measures to reduce the factors behind this attrition factor.
4. 18.8% of the employees travels frequently, and they have the highest attrition percentage (25%). The company should question what makes traveling a burden on their employees. The company should balance the travel status and if necessary, there might be some adjustments on the job description in terms of traveling. The company may use some extra incentives to motivate their employees who are supposed to travel.
5. 32% of employees are single and has the highest attrition percentage(25.5%). The company should be aware of this important factor and have strategy to deal with this groups' performance.
6. If the year increases since the employees' last promotion, the attrition percentage also displays increase. Especially, 7 years since last promotion employee group (5.2%) has the highest attrition (21.1%) in its individual group. For that reason, the company review their promotion policy, and maybe define the company's expectations from their employees and make clear to all employees how and when they may be promoted.
7. Beside those factors above, there are some other indicators which keep employees in the company. These factors are stated below.
 - ✓ Total working years in the company
 - ✓ Years in current role
 - ✓ Environment satisfaction
 - ✓ Medical education field
 - ✓ Job satisfaction

- ✓ Job involvement
- ✓ Years with current manager
- ✓ Work life balance
- ✓ Age

The company should primarily try to increase the effectiveness of those factors. As a result, it will yield to the decrease in the attrition rate.

Bibliography

1. Abdul-Kadar Masum, L.-S. B. A.-K. A. a. K. H., 2015. Intelligent Human Resource Information System (i-HRIS): A Holistic Decision Support Framework for HR Excellence, Malaysia: IAJIT
2. Adhikari, A., 2009. Factors Affecting Employee Attrition: A Multiple Regression Approach. The IUP Journal of Management Research, VIII (5)
3. Alex Frye, C. B. M. S., 2018. Employee Attrition: What Makes an Employee Quit? SMU Data Science Review, 1(1).
4. https://esource.dbs.ie/bitstream/handle/10788/3497/msc_salunkhe_t_p_2018.pdf?sequence=1&isAllowed=y
5. Bidisha Lahkar Das, D. M. B., 2013. Employee Retention: A Review of Literature. IOSR Journal of Business and Management, 14(2), pp. 08-16.
6. IBM, 2018. Employee Attrition and Performance - IBM Analytics. [Online] Available at: <https://www.ibm.com/communities/analytics/watson-analytics-blog/hremployee-attrition/>
7. Rahul Yedida, R. R. R. V. R. J. A. D. K., 2006. Employee Attrition Prediction, s.l.: s.n.
8. <https://blogs.sap.com/2017/07/20/applying-predictive-analytics-to-manage-employee-turnover/>

Appendix A

This appendix provides additional information and details related to the IBM HR Analytics Employee Attrition project without using machine learning.

A. Data Collection and Pre-processing

The data for this project was obtained from a publicly available dataset on Kaggle.com. The dataset contained information on 1,470 employees, including their demographics, job satisfaction, work performance, and employment status. The data was cleaned and pre-processed using Python, and missing values were imputed using mean imputation.

B. Variables and Measures

The following variables were used in this study:

Age: The age of the employee in years

Attrition: Whether the employee left (Yes else No)

Business Travel: The frequency of business travel, measured as either "Travel Frequently", "Travel Rarely", or "non-Travel"

Department: The department in which the employee works, measured as either "Sales", "Research & Development", or "Human Resources"

Distance From Home: The distance from the employee's home to their workplace, measured in miles

Education Field: The field of education of the employee, measured as either "Life Sciences", "Medical", "Marketing", "Technical Degree", "Other", or "Human Resources"

Employee Number: A unique identifier for each employee

Environment Satisfaction: The level of satisfaction with the work environment, measured on a scale from 1 (Low) to 4 (High)

Gender: The gender of the employee, measured as either "Male" or "Female"

Hourly Rate: Hourly salary earned by the employee

Job Involvement: The level of involvement in job tasks, measured on a scale from 1 (Low) to 4 (High)

Job Level: The level of the employee's position within the organization, measured on a scale from 1 (Entry Level) to 5 (Executive)

Job Role: The role of the employee within their department, measured as either "Sales Executive", "Research Scientist", "Laboratory Technician", "Manufacturing Director", "Healthcare Representative", "Manager", "Sales Representative", "Research Director", "Human Resources", or "Technical Representative"

Job Satisfaction: The level of satisfaction with the employee's job, measured on a scale from 1 (Low) to 4

(High)

Marital Status: The marital status of the employee, measured as either "Single", "Married", or "Divorced"

Monthly Income: The employee's monthly salary in US dollars

Num Companies Worked: The number of previous companies the employee has worked for

Over Time: Whether or not the employee works overtime, measured as either "Yes" or "No"

Performance Rating: The employee's performance rating, measured on a scale from 1 (Low) to 4 (High)

Relationship Satisfaction: The level of satisfaction with the employee's relationships at work, measured on a scale from 1 (Low) to 4 (High)

Total Working Years: The total number of years the employee has worked

Training Times Last Year: The number of times the employee received training in the past year

Work Life Balance: The employee's level of work-life balance, measured on a scale from 1 (Low) to 4 (High)

Years At Company: The number of years the employee has worked at the current company

Years In Current Role: The number of years the employee has worked

Years Since Last Promotion: The number of years since the employee got their promotion

Years With Current Manager: The number of years spent with the current manager