eNABLING THE SMART WORKFORCE MANAGEMENT THROUGH ANALYTICS AT IBM

TERM PROJECT PHASE 1

GROUP 1C

- Gayatri Das

- Bhavya Priya Akula

- Yash Parmar

- Manish Reddy Rajawala

Contents

[**ENABLING THE SMART WORKFORCE MANAGEMENT THROUGH ANALYTICS AT IBM** 2](#_Toc160540814)

[**Executive Summary –** 2](#_Toc160540815)

[**Introduction: Company Background** 3](#_Toc160540816)

[**Products and Services** 3](#_Toc160540817)

[**Geographic Market Area** 3](#_Toc160540818)

[**Structure in Terms of Departments** 4](#_Toc160540819)

[**Structure in Terms of Business Units** 5](#_Toc160540820)

[**IBM’s Porter’s Five Forces Analysis** 5](#_Toc160540821)

[**Current status of the use of reporting and business intelligence/analytics systems in IBM** 6](#_Toc160540822)

[**Analytical Maturity of IBM-DELTA Model Analysis** 7](#_Toc160540823)

[**The Business Analytics Problem** 10](#_Toc160540824)

[**Data used for Analysis.** 10](#_Toc160540825)

[**The Business Analytics Solution** 11](#_Toc160540826)

[**Excel Preliminary Analysis:** 11](#_Toc160540827)

[**Visualization and Analysis using Tableau.** 14](#_Toc160540828)

[**Problem Analysis using SPSS.** 20](#_Toc160540829)

[**Analytical approaches that can be taken from this problem (DELTA)** 24](#_Toc160540830)

[**Conclusion** 24](#_Toc160540831)

[**References:** 25](#_Toc160540832)

# **ENABLING THE SMART WORKFORCE MANAGEMENT THROUGH ANALYTICS AT IBM**

# **Executive Summary –**

The primary objective for any type of organization, be it non-profitable, governmental, public, or profitable, is to be able to manage and introduce effective changes that bring in substantial improvement. Business analysis is a crucial aspect and plays a pivotal role in identifying various organizational issues and requirements for change. Business analysts of any company work across all levels of hierarchy and are often involved in everything from describing plans and policies to creating the innovative architecture of the organization. They can also play a leadership role by outlining the goals and requirements for new plans and developments or by providing support for the company's ongoing professional advancement.

Over the past few years, IBM has been actively expanding its analytics capabilities through hiring new employees, creating new software, and buying businesses that have analytics tools to complement their existing services. In order to assist businesses in making data-driven choices, IBM provides a range of business intelligence and analytics products. However, IBM's HR department is having trouble figuring out employee sentiment and the driving forces behind attrition, and its current reliance on unstructured employee data does not meet the criteria for validity and reliability needed for analysis and decision-making. The goal of our group project is to improve and help the HR Organization in forming a strategy using business data to retain their employees.

To address this issue, we have analyzed the IBM HR Analytics Attrition and employee performance Dataset, which is taken from Kaggle. The dataset contained information on 1,470 employees and included various demographic and job-related factors of employees. We used various analytical techniques, including excel, descriptive analytics, correlation, logistic regression, and decision-tree analysis techniques to identify the factors that are contributing to employee attrition rate. Data visualizations of the dataset were also created using Tableau to get a clearer and better understanding of the situation.

Our analysis revealed several key insights that could potentially help IBM improve its current employee retention strategies. Factors such as Age, distance from home, and job involvement were some significant predictors when it came to employee attrition. In this report we have also suggested ways such as increasing salaries, offering more training opportunities etc. which can help improve employee retention. The dashboards created using Tableau also helped us present these insights in a user-friendly and visually appealing format for the higher management.

# **Introduction: Company Background**

IBM is a renowned multinational technology corporation, originally established as The Computing-Tabulating-Recording Company (CTR) in 1911. The company has since undergone a name change and is now recognized as IBM (International Business Machines Corporation). With a comprehensive range of software, hardware, and consulting services for the technology available to Fortune 500 firms across the world, IBM has grown into one of the biggest technology corporations in the world.

With a rich legacy of technology industry innovation, IBM has earned its position as a pioneering organization. The company has made significant contributions to the development of new technologies, such as the first hard drive, automated teller machine (ATM), and personal computer. In recent years, IBM has concentrated its efforts on developing novel solutions based on cutting-edge technologies such as blockchain, cloud computing, machine learning, and Artificial Intelligence (AI). Overall, IBM's continued dedication to technological advancement and industry leadership is a testament to its legacy of excellence and global impact.

## **Products and Services**

IBM is a pioneer in artificial intelligence (AI), offering solutions for the manufacturing, healthcare, finance, and retail sectors. Additionally, IBM offers blockchain-based solutions to businesses in the financial, healthcare, and supply chain sectors, assisting them in improving their security and transparency while gaining a competitive edge over their rivals. IBM is at the forefront of the research and study of quantum computing through its cloud-based access platform, IBM Quantum, and through its IBM Security group. It also provides IoT solutions for a range of industries through IBM Watson IoT and hybrid cloud solutions using IBM Cloud Pak to integrate on-premises and cloud-based infrastructure.

## **Geographic Market Area**

Several important American cities, including New York, San Francisco, Austin, and Atlanta, are home to IBM offices. IBM is concentrated on offering technological solutions to numerous sectors, including healthcare, finance, and retail in the USA. Internationally, IBM is well-represented in a variety of areas, such as Europe, the Asia-Pacific, and Latin America. The diverse variety of industries covered by IBM's international activities includes banking, telecommunications, government, and manufacturing. The regional market reach of IBM is wide, and the company enjoys a solid reputation for giving enterprises and organizations all around the world cutting-edge and dependable technological solutions.

## **Structure in Terms of Departments**

IBM has several departments that are responsible for different aspects of the company's operations. These different departments of the organization are what make a business’s operation more efficient and productive. For a large company like IBM, managing several departments can sometimes be challenging.

**Research and Analysis** **–** The company currently possesses the largest research facility in the world and has created more than 10 global research labs that encompass the Asian, American, and EU regions.

**Service & Products –** The cloud computing, cognitive computing, commerce, the Internet of Things, mobile apps, and user security are just a few of the categories that come under the IBM services and products division. The company has also observed a rapid advancement in cloud data services, such as the Cloud Data Encryption Services (ICDES), which uses cryptographic splitting to safeguard users' data. IBM has also established a presence in several other cutting-edge technology domains, such as the gaming sector, IT outsourcing, and various public sectors.

**Marketing and Public Relations -** IBM has bolstered its marketing initiatives with a number of additional top-tier partners, such as the Masters Tournament, a renowned golf competition.

**Software and Storage -** Mainframes, servers, storage systems, and operating systems are just a few of the hardware and software items that this division creates and markets. Analytics products from IBM, including IBM Cognos Analytics, IBM Statistical Product and Service Solutions (SPSS), IBM Maximo Asset Management, and Db2, are part of the company's extensive software line. As a part of storage IBM provides a variety of hardware items, such as its Flash System all-flash arrays, Storwize systems, and other hybrid arrays, as well as Fiber Channel storage area network hardware, storage media, and tape solutions.

**Cloud and Security -** This division is in charge of IBM's cloud and artificial intelligence (AI) products, such as Watson and Red Hat. This division offers cybersecurity, risk management, and compliance services to IBM's clients in the area of security.

Diagram

Description automatically generated

**Figure: Various Departments in IBM**

## **Structure in Terms of Business Units**

The four business units that make up IBM’s financials are software, consulting, infrastructure, and financing.

**Software –** In order to assist clients with their data demands and to help them automate and secure their systems, IBM's software sector connects the company's software products with its hybrid cloud platform.

**Consulting –**To assist IBM's clients in digitally transforming their organizations, the sector designs and develops hybrid cloud infrastructures, streamlines workflows and business processes, and jointly develops products and solutions with them.

**Infrastructure –** Hybrid cloud solutions are offered by IBM's Infrastructure division. It offers clients infrastructure systems that aid in meeting the new standards for hybrid multi-cloud and enterprise AI workloads.

**Financing –** The IBM Finance unit aids customers in purchasing the company's IT hardware, software, and services. The sector often offers financing for goods or services that are essential to the client's company operations and complement IBM's AI strategy and hybrid cloud platform.

Chart, pie chart

Description automatically generated

**Figure: IBM Segment Breakdown**

## **IBM’s Porter’s Five Forces Analysis**

Based on this IBM Porter model table mentioned below we can infer that; competition is the strongest force affecting IBM's business. Therefore, it is recommended that the company focuses on improving its competitive advantage through **human resource development and breakthrough innovation**.

| **Five Forces** | **Intensity** | **Explanation** |
| --- | --- | --- |
| Competitive Rivalry or Competition | Strong | The information technology industry has a high level of competition, with many companies offering similar products and services, resulting in pricing pressure, market share competition, and profit erosion. |
| Bargaining Power of Buyers or Customers | Moderate | Customers have some bargaining power, but it is not significant enough to affect IBM's overall market share or pricing. |
| Bargaining Power of Suppliers | Moderate | Suppliers have moderate bargaining power, with a few dominant suppliers in the industry. IBM can switch suppliers, if necessary, but it may affect product quality or delivery time. |
| Threat of Substitutes or Substitution | Moderate | The threat of substitutes is moderate as there are few alternatives to the products and services provided by IBM. However, advancements in technology may create substitutes in the future. |
| Threat of New Entrants or New Entry | Moderate | The threat of new entrants is moderate as the industry has high barriers to entry due to the need for significant capital investments, economies of scale, brand recognition, and intellectual property. However, the industry is attractive, which may encourage new entrants. |

## **Current status of the use of reporting and business intelligence/analytics systems in IBM**

IBM has moved aggressively the past several years to build their analytics capabilities through hiring talent, developing software, and acquiring firms with analytics tools that enhance their portfolio. IBM itself has been the alpha and beta for the deployment of these new technologies across operations—from managing finance and marketing to optimizing their supply chain and attracting, retaining, and developing talent. They have found in every aspect of their organization, that smarter, more informed decisions can be made by embedding analytics into how they do business.

IBM offers a suite of Business Intelligence and Analytics tools to help organizations make data-driven decisions. The tools allow users to access, analyze and report on data from various sources. The IBM Order Management Reports feature provides insights into order management data. Users can create and schedule reports, as well as customize dashboards and alerts to monitor key performance indicators. The reports can be used to track orders, inventory levels, shipping, and fulfillment performance. Cognos Analytics is another IBM tool that helps businesses make smarter decisions through analytics. It provides a platform for creating and sharing interactive reports, dashboards, and visualizations. Cognos Analytics also allows for predictive modeling, data mining, and data exploration. Overall, IBM's Business Intelligence and Analytics tools enable organizations to gain insights into their data and make informed decisions to improve their business operations.

# **Analytical Maturity of IBM-DELTA Model Analysis**

The Delta model offers businesses a framework to evaluate existing situations and finding areas for growth. Organizations can create a strategic plan that enable them to stay ahead of the competition and continuously enhance their performance by concentrating on these five crucial areas.

**D-Accessible and high-quality Data**

IBM uses data warehouses for storing and managing substantial amounts of data. In fact, IBM offers its own data warehousing solutions, such as IBM Db2 Warehouse, which is a fully managed cloud data warehouse designed for analytics, machine learning, and AI workloads. IBM Db2 Warehouse is the relational database management system used by the organization. IBM Db2 warehouse is a client-managed, preconfigured data warehouse that runs in private clouds, virtual private clouds, and other container-supported infrastructures. (“IBM Db2 Warehouse”) (“IBM Db2 Warehouse”) This data warehouse is designed to provide the ideal solution when you must maintain the control of your data but want cloud-like flexibility. (“On-Premises Data Warehouse Software - SourceForge”) Db2 warehouse provides a unified, powerful data warehouse delivering access to structured and unstructured Information in real time. Db2 warehouse provides insights into your data with the Cubing Services feature, a multidimensional analysis server that enables OLAP applications to access up to a terabyte of base OLAP data without requiring multiple copies. (“Data warehousing and analytics”). (“Data warehousing and analytics - IBM”) IBM also has partnerships with other data warehousing providers, such as Snowflake and Microsoft Azure, to offer customers a variety of options for storing and analyzing their data. Additionally, IBM has expertise in building and implementing custom data warehousing solutions for clients across a wide range of industries.

**E-Enterprise Orientation**

IBM’s transformation to use analytics enterprise-wide began in 2004, continues till date. The focus is on how to become more smarter and agile with the use of analytics to solve business problems. IBM has leveraged its math department both to apply analytics internally and to add some of the learnings and benefits to its analytics products and solutions.

1. **Globally Integrated Enterprise Shared Services:** To support this transformation, globally integrated organizational units providing support services to all of IBM were formed and human resources, integrated supply chain finance, information technology, and marketing.

2. **Enterprise Transformation Initiatives:** A number of enterprise-wide initiatives were launched to drive radical, innovative transformation in the way IBM works with clients, business partners, and its own employees. Three Enterprise Transformation Initiatives, which are described in later chapters, are Development, Smarter Commerce™ Inside IBM, and Hardware Product Management Transformation.

The next evolution of IBM’s transformation is to migrate from a globally integrated enterprise to a smarter enterprise by optimizing the entire enterprise with the following technologies:

* Analytics to gain business insight for customers and the enterprise.
* Social media for business collaboration both inside and outside the enterprise.
* Mobile communications for pervasive connectivity.
* Cloud technologies for IT enablement.

Also, to make sure that its staff members have the abilities necessary to flourish in a world that is becoming more and more data-driven, IBM has also been investing in talent development and training programs.

Graphical user interface

Description automatically generated

**Figure 1-1 Nine levers: Capabilities that enable and enhance big data and analytics development, delivery, and value creation.**

**L-Analytical Leadership**

The leadership team of the organization has been attempting to eliminate silos and promote an innovative culture with an emphasis on encouraging agility and responsiveness to shifting market conditions**.** The analytical leaders of IBM found a strong link between organizations that excel in these nine levers (mentioned in the figure below) deriving the greatest value from data and analytics. They identified a progression of three steps—Enable, Drive, and Amplify—each with three of the levers:

* **Enable** forms the foundation for creating value from big data and analytics.
* **Drive** has the actions needed to create value by moving from analytics discovery to value creation.
* **Amplify** increases the amount of value realized by providing momentum to translate insights into actions that increase an organization’s bottom line.

**T-Technology**

IBM utilizes an Enterprise Resource Planning (ERP) framework as one of its consistent/coordinated data frameworks (IS). IBM involves SAP as its ERP framework, which helps in overseeing different business processes, including finance, deals, obtainment, and production network the board, among other. The SAP framework permits IBM to smooth out its tasks, further develop navigation, and gain better perceivability in its business processes. Moreover, IBM likewise utilizes other consistent or incorporated Information systems like CRM, HCM, and SCM frameworks to oversee different parts of its tasks. IBM has created and carried out its own ERP arrangement. The board framework is a cloud-based stage that gives start-to-finish permeability and control of request satisfaction across the whole production network.

All this ERP framework empowers IBM to oversee and improve its business activities even more productively by coordinating and smoothing out its business processes, including store network, stock administration, monetary administration, and client relationships.

**A-Analytical Techniques**

IBM started developing sales analytics tools using descriptive and predictive reporting tools, but over the past decade there has been a sea change in attitude. Now sales managers looking for more analytics support are using advanced BI software, social media analytics, entity analytics and cognitive computing which can take their organizations to the next level.

IBM’s analytical maturity can be defined as follows:

* **Customer Productivity:** IBM places a high priority on increasing customer productivity and in this end, a variety of products and services are available. Solutions from IBM such as analytics software, cloud computing services, and consultancy services assist clients in deriving value and insights from their data.
* **Innovation:** IBM has made large investments in cutting-edge technologies including blockchain, quantum computing, and artificial intelligence. A strong research division within IBM is dedicated to creating cutting-edge products and services that can handle challenging business problems.
* **Organizational Architecture:** IBM's structure is built to encourage communication and creativity amongst various divisions and business units.

Overall, IBM has a very high analytical maturity level, and the corporation keeps innovating and creating new products to assist businesses in maximizing the value of their data.

# **The Business Analytics Problem**

IBM has various business functions including human resources, finance, supply chain, sales, information technology, marketing, and service. Other functions—such as software development and hardware manufacturing—are more specific to a technology, although analogous functions exist in other industries (for example, manufacturing cars in the automotive industry).

**Main Issues with IBM’s HR Organization**

HR organization is looking at understanding the overall employee sentiment and reasons behind attrition. The key issues identified were pertaining to different activities of workforce management. Starting from recruitment, improving the engagement, identifying the early signs of attrition, improving the job-fit, getting an accurate view of employees’ sentiment about IBM.

The firm currently relies on unstructured employee data that is not meeting the validity, reliability requirements for analysis and decision-making. This leads to untimely response to an existing issue, improper measurement of employee behavior and skillset requirements at a broad level. This further impacts all other lifecycle events of an employee. Hence, the objective is to help IBM address its strategic human capital challenges and enhance the HR Organization’s capability in driving business strategy through business analytics.

## **Data used for Analysis.**

The dataset that will be used for this project is the IBM HR Analytics Attrition and employee performance Dataset, which is taken from Kaggle. This dataset contains information on 1,470 employees and is fictional in nature. It includes various demographic and job-related factors of employees. The dataset has a total of 35 variables, including the employee's age, gender, job role, monthly income, performance rating, and whether they have left the company (attrition).

**Application of Data Mining Techniques**

* ***Problem Identification and Data Analysis:*** Descriptive Analytics, Correlation, VIF, t-Test, Regression and Decision-tree analysis techniques will be used to analyze the IBM employee dataset to identify the factors contributing to current problems with respect to employee behavior. Data Visualizations of the dataset will be created using Tableau to get a clearer understanding of the existing situation.
* ***Business Process Improvement:*** Descriptive, predictive, and advanced analytical solutions such as Mini Social Polls, social media analytics, Entity analytics can be used for improving their existing process and meet the organization’s objective of improving the strategic capability of HR function in driving IBM’s business strategy.

# **The Business Analytics Solution**

In this section, we will present our business analytics solution for analyzing the employee retention problem at IBM. Using the employee dataset provided by IBM, we applied various analytical techniques, including basic descriptive and visual analytics, and advanced data mining techniques like regression modeling using SPSS Modeler. Our objective is to demonstrate how these analytical tools can help IBM make data-driven decisions to improve their employee retention rates.

Initially, we leveraged Excel to conduct descriptive analysis and create visualizations to gain an initial understanding of the factors affecting employee retention. We then used regression modeling with SPSS Modeler to identify the most significant factors that influence employee attrition. Additionally, we used Tableau to provide a comprehensive view of IBM's employee attrition and identify potential areas where strategic modifications to business processes can be made to address this issue of high employee turnover. Our analysis provides actionable insights to help IBM improve its overall performance and address the challenge of employee retention.

Our analysis has revealed several key insights that can help IBM improve its employee retention strategies. We have found that age, distance from home, and job satisfaction are significant predictors of employee attrition. Our analysis has also highlighted that increasing salaries and offering more training opportunities can improve employee retention.

We have created several dashboards using Tableau to present these insights in a user-friendly and visually appealing format. Our data analytics solution can help IBM identify and prioritize the areas where it can make changes to improve its employee retention rates.

Our business analytics solution demonstrates the power of data-driven approaches to help organizations make better decisions. We have shown how advanced data mining techniques can uncover insights that can be used to develop effective employee retention strategies. Our findings have significant implications for IBM, and we are confident that our insights will help IBM improve its employee retention rates and ultimately enhance its performance.

## **Excel Preliminary Analysis:**

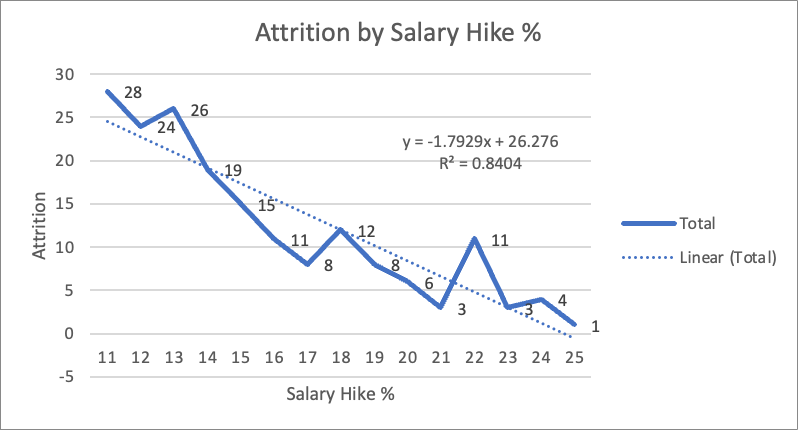
1. Attrition by Salary Hike

This line chart represents the relationship between salary hikes and employee attrition in IBM. The chart shows the percentage of salary hike on the X-axis and the sum of attrition (no. of employees leaving the firm) on the Y-axis.

Looking at the data, it appears that there is an inverse relationship between salary hike percentage and employee attrition. That is, as the salary hike percentage increases, the sum of attrition decreases. This can be seen by the fact that the highest attrition rate (28) occurs when the salary hike percentage is 11, and the lowest attrition rate (1) occurs when the salary hike percentage is 25.

This suggests that employees are more likely to stay with the company if they receive a higher salary hike. It also indicates that the company may be able to reduce employee turnover by offering higher salary hikes.

However, it is important to note that this chart only shows the relationship between salary hikes and attrition and does not necessarily prove causation. Other factors, such as job satisfaction, work-life balance, and career opportunities, may also play a role in employee turnover.



1. Attrition by No. of Previous Companies Worked by IBM employee.

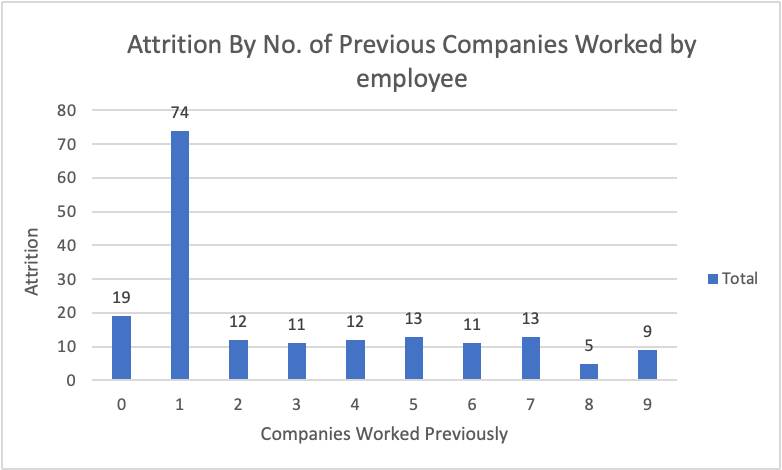
This bar chart represents the relationship between the number of companies that an IBM employee worked for previously and the sum of attrition. The chart shows the number of previous companies worked for on the X-axis and the sum of attrition on the Y-axis.

Based on the data, it seems that the highest attrition (74) occurs among employees who have worked for only one previous company. The attrition rate for employees who have never worked for any company before and started their first job at IBM is 19. For employees who have worked for two or more companies, the trend is almost horizontal, indicating that the number of companies worked for has limited influence on attrition.

One possible explanation is that these employees may be more open to new opportunities and less committed to IBM as an employer. This may be due to a lack of familiarity with the company culture or a lack of investment in the company's long-term goals. As a result, these employees may be more likely to leave IBM for other opportunities that arise.

Another possible explanation is that employees who have worked for only one company in the past or who are freshers may be more likely to encounter job dissatisfaction or culture mismatch. These employees may not have had the opportunity to develop a broader range of skills or work experience, which could limit their ability to find a suitable fit for IBM. As a result, they may be more likely to leave the company in search of a better fit.

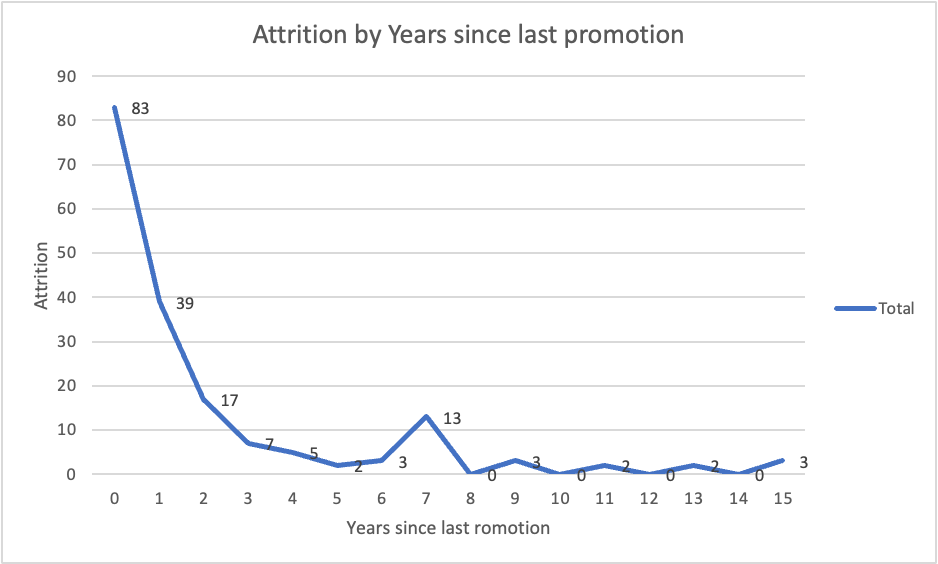
Overall, while the trend in the data suggests that the number of companies worked for previously has limited influence on attrition rates at IBM, there may be other factors at play that influence employee retention and turnover.



1. Attrition by Years since last Promotion

The data shows that employees who have not been promoted in the past year have a high attrition rate of 83. As the years since the last promotion increase, the attrition rate generally decreases, with 39 employees leaving after 1 year without a promotion, 17 leaving after 2 years, and so on. However, there are a few outliers in the data, such as the employee who left after 5 years without a promotion and the employee who left after 13 years without a promotion.

Based on the data provided, it could be speculated that some employees who have recently received a promotion might leverage it to enhance their career prospects and find higher paying jobs elsewhere. Conversely, it might be inferred that IBM employees who have been with the company for an extended period of time are more likely to be content with their work and less inclined to leave, regardless of whether they have received a promotion or not. However, it is important to note that these assumptions cannot be conclusively determined from the data alone, and other factors such as job satisfaction and professional development opportunities may also play a role in employee retention.



## **Visualization and Analysis using Tableau.**

Tableau is a powerful data visualization tool that enables us to transform complex datasets into engaging and insightful visualizations. As a part of our project, we have leveraged Tableau's features to create interactive dashboards that will allow the end users to explore and visualize the employee data of IBM, mainly focusing on the attrition rate of employees. The main goal here is to give the client organization that is IBM a visual representation along with, descriptive analytics as well as other data mining solution which they can use to benefit their organization and to make better decisions and extensively improve performance. For comprehending the employee attrition rate more thoroughly and pinpointing areas that may be improved to lower the attrition rate we developed interactive visualizations using Tableau that provided insights into employee attrition, using factors such as employee education, employee gender etc, to see the influence these different employee characteristics had on the attrition rates.

Visualizations

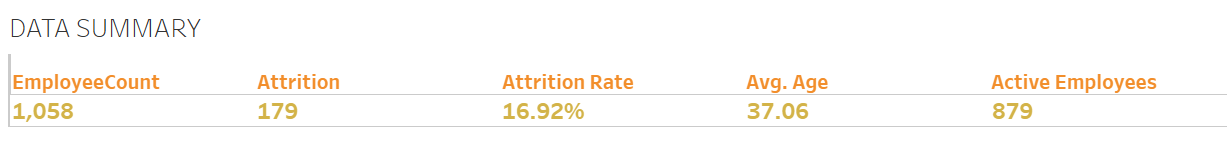


Figure: Data Summary (Created by the Author)

This is a data summary visualization that provides an overview of employee attrition rates in a fictional company. This displays the key metrics which are the number of employees, employees that have left (i.e., attrition), the attrition rate, the average age of the employees present in this data set no of active employees.

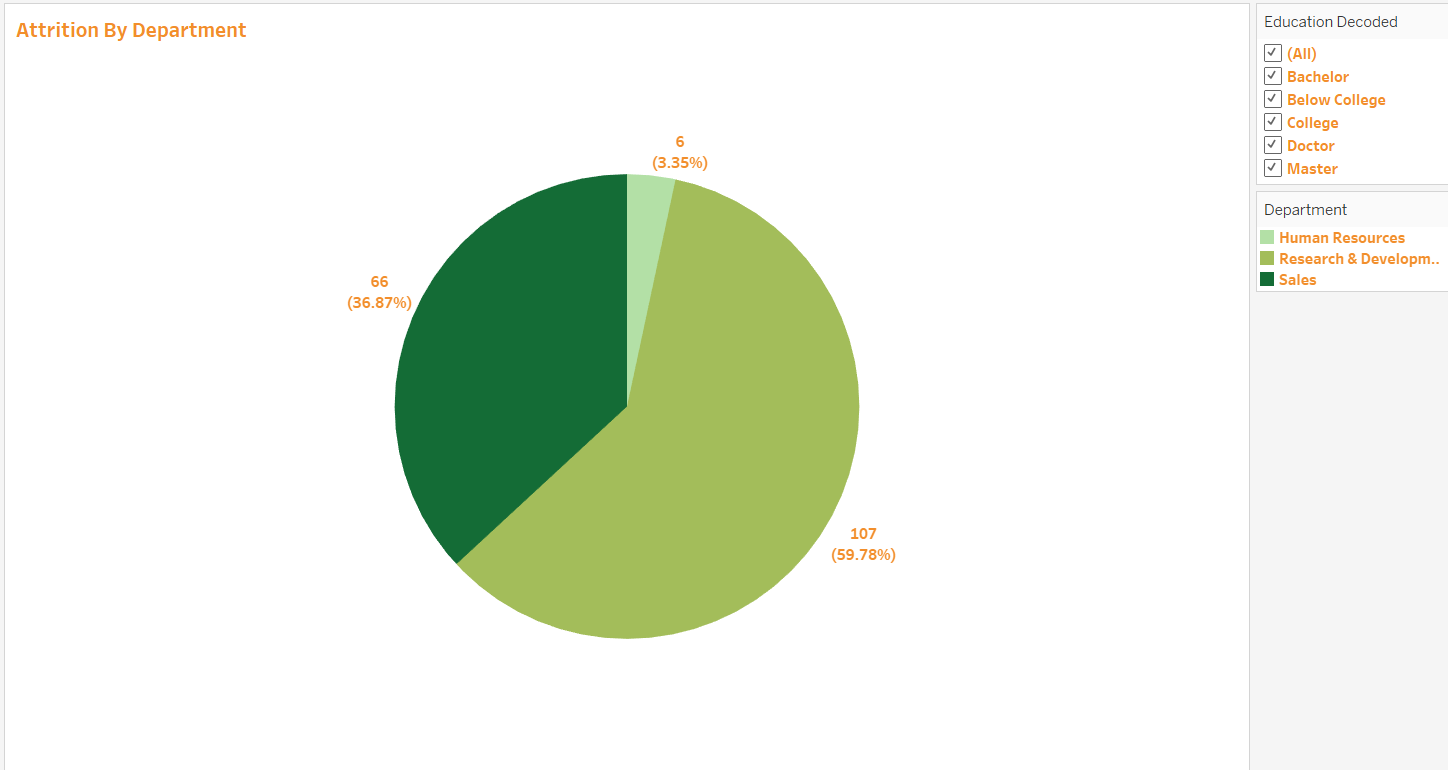


Figure: Attrition by Department (Created by the Author)

This visualization is in the form of a pie-chart displaying the attrition rate of employees across different departments in IBM, with each department represented by a slice of the pie. The slices are color-coded to indicate the various departments, as indicated by the right-hand legend. The labelling next to the slices tells us the number of employees that have left that department and the percentage is displayed under it.

Based on is visualization, it can be seen that majority of employees who left the company were from the Research & Development department. This department had the highest attrition rate among all the departments, hence indicating a potential issue in that department that may require higher management’s attention to identify the root cause and to develop effective retention strategies and improve overall employee morale.

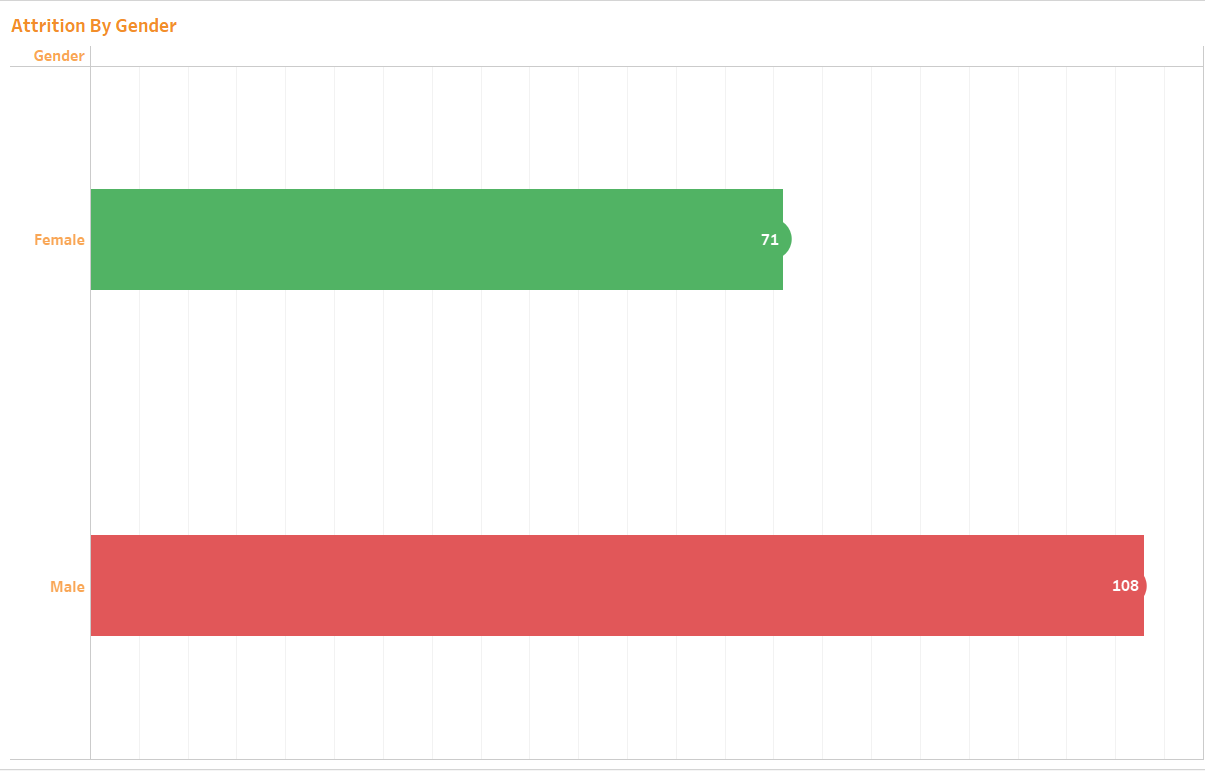


Figure: Attrition by Gender (Created by the Author)

This visualization is in the form of a horizontal bar chart, with separate bars for male and female employees. Each bar displays the number of employees that have left. The bars are color-coded to indicate the level of attrition, with red indicating a high attrition rate and green indicating a low attrition rate.

Based on this visualization we can say that there is a higher attrition rate amongst male as there is a clear disparity between the male and female attrition rate. This may indicate a need for HR policies that address issues affecting one gender more than the other.

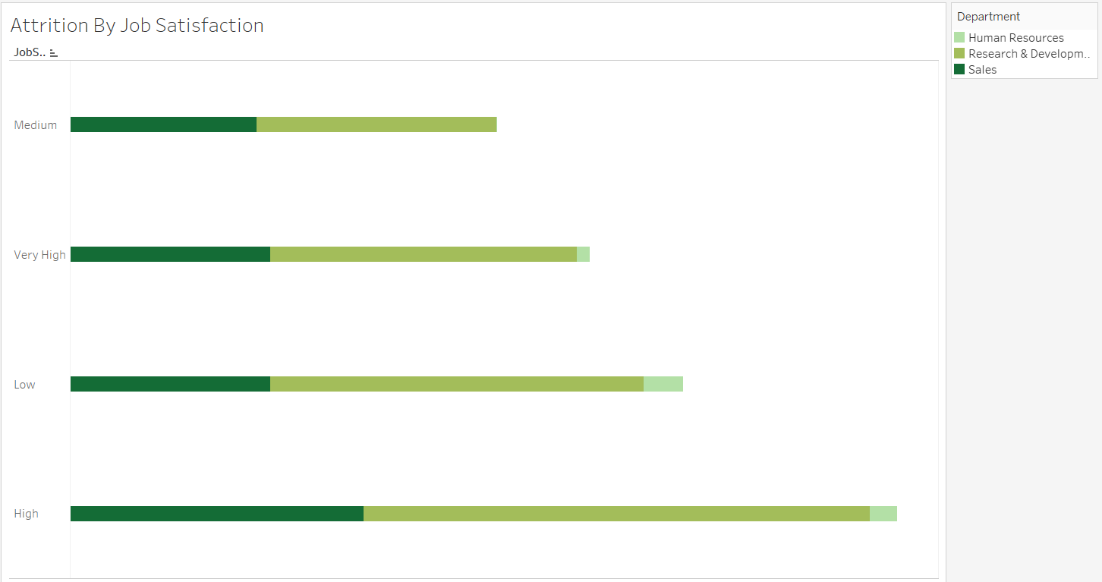


Figure: Attrition by Job Satisfaction (Created by the Author)

This displays the attrition rate of IBM employees based on their job satisfaction levels. The visualization is presented in the form of horizontal bar graph, with the attrition rates represented on the x-axis and the different educational levels represented on the y-axis. Further each bar is divided by different departments which can be seen by the colour variation in the bars.

Based on this visualization we can say that there is a positive correlation between job satisfaction levels and attrition rates, as higher attrition rates are observed for employees with higher job satisfaction levels. While it may seem counterintuitive, this can occur due to a variety of factors related to employee motivation, career goals, and workplace dynamics.

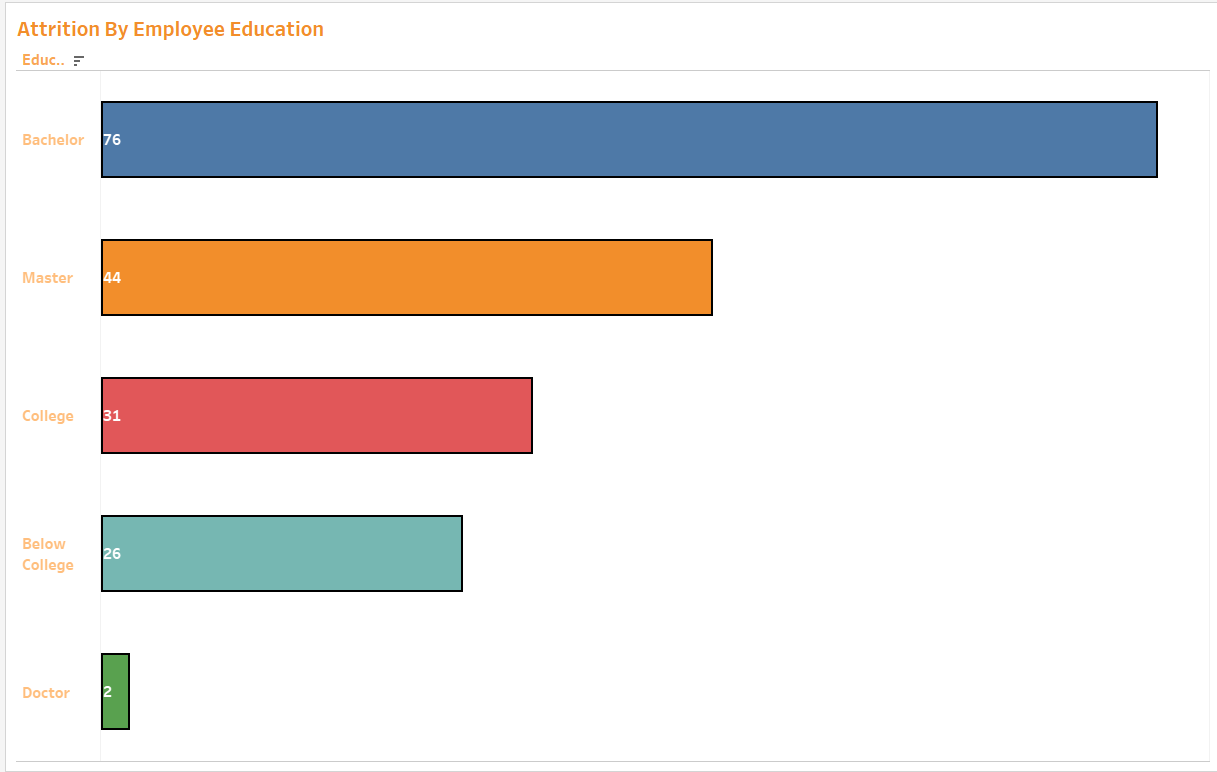


Figure: Attrition by Education Level (Created by the Author)

This visualization shows the attrition rates of employees based on their educational level. The data is presented in the form of a bar graph, with the attrition rates represented on the y-axis and the different educational levels represented on the x-axis. The graph shows that employees with a doctoral degree have the lowest attrition rate, followed by employees with college and below college educational level. On the other hand, employees with a bachelor’s have the highest attrition rates, followed by employees with a master's degree.

Here the correlation we can see is that specifically, employees with higher levels of education are more likely to leave their jobs than those with lower levels of education, except for employees with doctoral degrees.

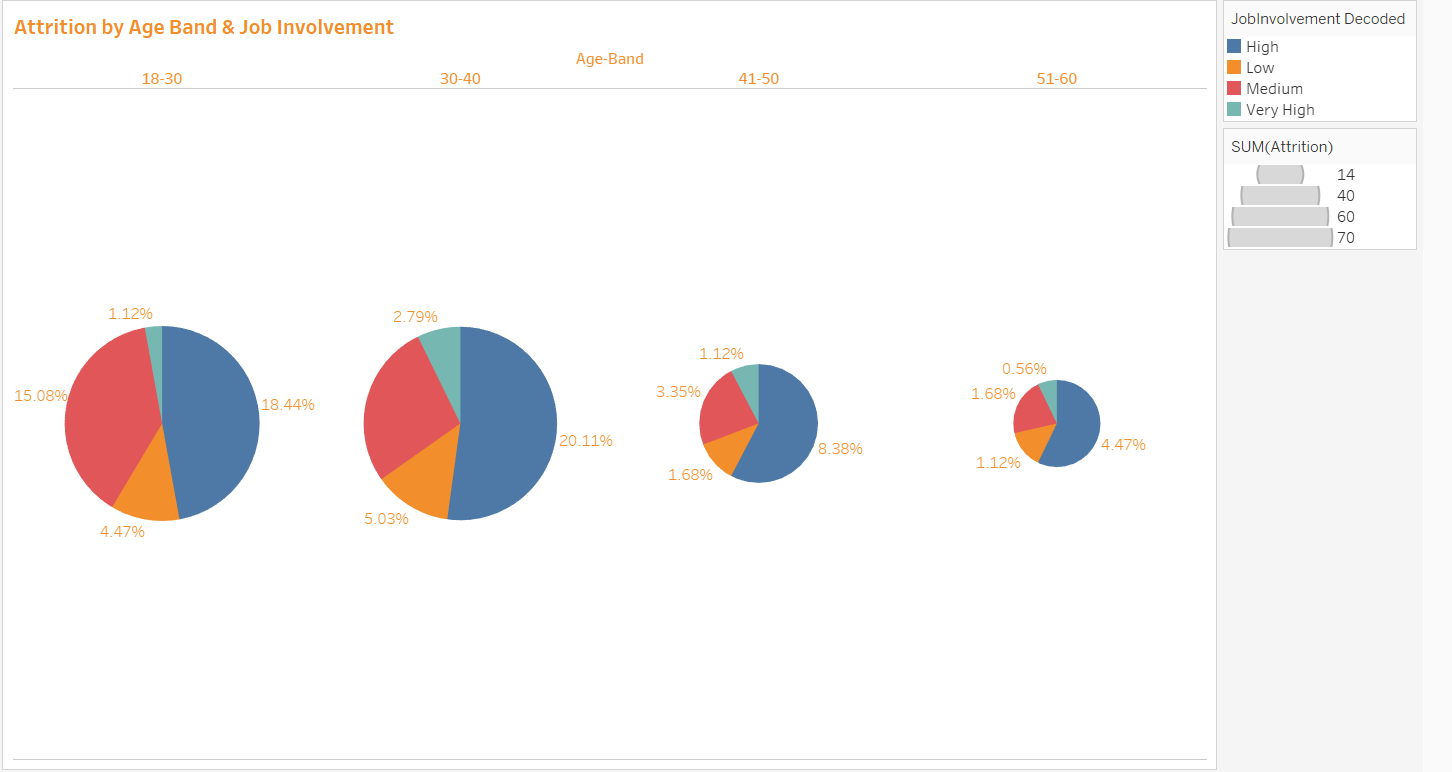


Figure: Attrition by Age Band and Job Involvement (Created by the Author)

This visualization shows the attrition rates by age band and job involvement of employees in an organization. The horizontal axis represents the different age bands ranging from 18-25 to over 60, while the vertical axis represents the job involvement levels ranging from low to high. The colors here indicate job involvement levels.

From this visualization it can be observed that the attrition rates are highest among employees in the age band of 18-25, which that younger employees may be more prone to leaving the organization, possibly due to factors such as lack of job security, lack of job satisfaction, or a desire to explore new opportunities.

Secondly, the attrition rates are generally higher for employees with high and medium job involvement levels, regardless of their age band. This suggests that employees who are highly involved in their jobs are also more likely to leave the organization, which could have negative implications for the company. One of the reasons is that employees with high job involvement may have higher expectations from their job and their employer and will be more likely to leave if those expectations are not met. They may have a strong emotional attachment to their work and the organization, which could lead to higher levels of job satisfaction but also make it harder for them to continue working in a role that they feel is not meeting their needs or expectations.

Dashboard –

This dashboard is a collation of all the visualizations explained above and provides a visual representation of key metrics related to employee attrition, such as gender, job involvement, employee satisfaction levels etc. By having this information readily available, the management can identify problem areas and take proactive measures to address them, ultimately reducing attrition rates and improving employee retention.

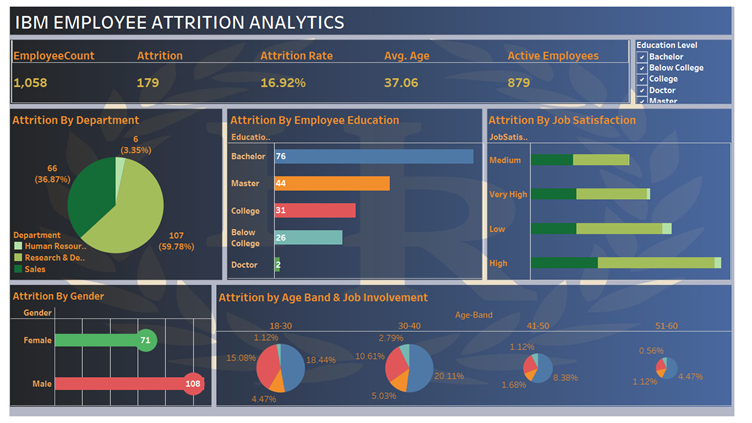


Figure: IBM Employee Attrition Analytics Dashboard (Created by the Author)

Suggestions –

Based on the insights found from the above visualizations provided, here are some managerial insights and recommendations:

1. In-depth analysis of the Research & Development department should be conducted to identify the root causes of high attrition rates, including factors like workload, job satisfaction, management style, and team dynamics and develop targeted retention strategies that are specific to the needs of the Research & Development department. This could involve offering additional professional development opportunities, creating a more positive and supportive work environment, or implementing policies and procedures that address specific concerns within the department.
2. Reassess the factors that contribute to job satisfaction, may be conduct surveys or focus groups to gain a better understanding of what factors are contributing to their decision to leave. Provide more opportunities for career growth to keep employees engaged and motivated.
3. Encourage open communication and offer incentives for long-term commitment. This can help to create a sense of loyalty and commitment among employees regarding the management.
4. Conduct exit interviews with employees who leave as this can provide valuable insights into any issues or concerns that employees may have and can help identify areas for improvement.

## **Problem Analysis using SPSS.**

Based on the analysis conducted, it was identified that attrition was one of the key issues IBM was experiencing, other issues being employee behavioral patterns not being properly captured. Hence, looking at the data type of attrition variable from the collected data, the below analysis was conducted to identify the underlying variables that causes attrition.

Steps performed in SPSS analysis.

* After cleaning up the dataset, by resetting the variables and data types of the clean data was uploaded to SPSS.
* The attrition was chosen as the target variable, and other relevant variables were set as Input variables.
* Using the filter node, we filtered out only those variables that were observed to be logically impacting attrition based on the business logic.
* Next, we have partitioned the dataset into training and testing datasets in 50:50 proportion.
* Since we have observed the dataset to be unbalanced, and as we did not have enough data to analyze the variable of interest, which is Attrition, we have balanced the dataset using balancing node.
* We have analyzed the data by running quest decision tree analysis, which resulted in the results below.

Diagram

Description automatically generated

Chart, box and whisker chart

Description automatically generated

Table

Description automatically generated

Since the model only resulted in accuracy of 52%, we went ahead and ran the logistic regression model for performing a better analysis of attrition.

The logistic regression model resulted in below results, which yielded in accuracy of 75.69% in Training dataset and 65.18% of accuracy in Testing dataset.

Table

Description automatically generated

Chart, line chart

Description automatically generated

**Futuristic Business Analytics Solutions**

Based on the above analysis, it was identified that attrition was the primary issue to be addressed by IBM. However, we have noticed that IBM has a huge business opportunity of “Enabling the smart workforce Management through Analytics” which can be achieved by implementing the below advanced business analytics solutions.

1. **Develop a Predictive Analytics Model through Clustering -** Using predictive analytics to proactively identify high-value skilled employees at risk of leaving could provide insights to proactively address the risk factors and improve retention. A predictive model can be designed a that could accurately predict workers at risk of leaving and then make recommendations on actions to take to retain them. The solution involves clustering the resource of interest by criteria such as country, business unit, and job category. Clustering refers to grouping a set of data so that data within each cluster is similar and data from different clusters is not similar. For each cluster, the voluntary attrition versus the compensation is determined along with the resulting ROI.

For example, some of the recommendations were to address concerns with compensation or the pace of promotions.

1. **Cognitive Computing- data -** Cognitive computing, computing systems that interact with people in new ways to provide insight and advice, is emerging just when we need it to help us uncover insight from the explosion of big data. The second era of computing, the programmable systems era, was characterized by general-purpose machines that are programmed by humans to perform a wide range of computational tasks. The third era, the cognitive era, is characterized by computing systems that are sensing, learning, reasoning, and interacting with people in new ways to provide insight and advice. Cognitive computing promises to provide invaluable assistance with the analysis of big data.

Examples include visualizing big data insights based on our questions, helping us explore data and uncover insights, and helping us detect anomalies in big data.

1. **Entity Analytics -** Entity Analytics refers to the practice of sorting through data and discovering data that relates to the same entity. It focuses on sorting through data and grouping together data that relates to the same entity. Entity analytics is a powerful technique for recognizing context and detecting like and related entities across large collections of data. Entity analytics is helpful in finding relationships between pieces of data and can incorporate new data to either confirm or negate previously found relationships. Additional methods can be applied to garner insight from unstructured data.
2. **Employee Behavioral pattern study using Social Media Analytics -** To get better insights than that provided by surveys by using social media to get a more accurate, more comprehensive, and more immediate view of what employees are thinking, can be known by making use of social media. This involves collecting data from only publicly available data and anonymizing all data. This, when properly implemented allows analytics, segmentation, and insights to be performed on the social media data without ever revealing the authors’ identities. This will give IBM HR the ability to use employee engagement surveys to assess the engagement of the organization at a specific point in time and to gather insights expressed on internal and external public social media to provide insights in “near real time.” When necessary, HR professionals can follow up with targeted surveys and focus groups to gain information on identified issues. Also, an employee’s demographic characteristics are augmented with the social media content, while the identity gets obfuscated. The text is then analyzed for sentiment to create interactive visualizations displaying the most-discussed topics and the sentiment around them, aggregated by demographic segments. To derive sentiment about a topic, the project can be started with the preconfigured set of sentiment terms and concepts in IBM Social Media Analytics and added to these terms, when necessary.

**Managerial Considerations/Issues in Business Analytics Solutions Implementation**

* For implementation of advanced analytical techniques like mentioned above the ideal team to do analytics is a collaboration between an experienced data scientist, a person steeped in the area of the business where the challenge needs to be solved, and an IT person with expertise in the data in that particular area of the business.
* Estimating an analytics project’s ROI involves both capturing the project costs and measuring the value.
* Relationships inferred from data today may or may not be present in data collected later.
* Difficult to make the end-user understand the technicality and complexity involved in newer forms of analytics such as entity analytics, cognitive computing etc.
* The accuracy of sentiment analysis must be continuously refined according to human language nuances (behavioral/subjective elements of the data)
* Customizing the predictive model to various geographical locations can get challenging, as the factors to be considered, standards of measurement etc. would vary based on the culture of the location.

## **Analytical approaches that can be taken from this problem (DELTA)**

**D - Accessible and high-quality data:** Conducting a thorough analysis of the data to identify trends and patterns in employee turnover. This can include analyzing employee demographics, job titles, performance ratings, and job satisfaction surveys. Use predictive analytics to identify high-risk employees and take proactive steps to retain them.

**E - Enterprise Orientation:** Encouraging a culture of transparency and open communication within the organization. Encourage employees to share their concerns and feedback and take action to address their needs. Develop retention strategies that are specific to each department and job role.

**L - Analytical Leadership:** Develop leadership training programs that emphasize the importance of employee engagement and retention. Train managers to be proactive in identifying and addressing employee concerns and hold them accountable for reducing attrition rates within their teams.

**T - Technology:** It can help improve employee engagement and retention by implementing technologies. This can include tools for measuring employee satisfaction, providing opportunities for career development, and facilitating communication and collaboration among team members.

**A - Analytical Techniques:** To ensure that the analytical tools and methods used to measure employee engagement and retention are reliable and accurate, and continuously monitor and evaluate the effectiveness of retention strategies and adjust them as needed based on new data and insights.

# **Conclusion**

Based on the analysis of the IBM employee dataset, we have created this report that explores the factors that contribute to employee attrition rates within IBM. The report highlights the importance of understanding the underlying causes of employee attrition and the negative impact it can have on the company.

Using data analysis tools such as Tableau, Excel, and SPSS, we have examined various factors that influence attrition rates, such as employee demographics, job satisfaction levels, etc. In this report we have emphasized the need for companies to prioritize employee retention and implement effective strategies to reduce attrition rates. We also highlighted the importance of improving working conditions, offering competitive compensation packages, and providing opportunities for professional growth and development, on the whole enabling the smart workforce management at the organization.

In conclusion, this report demonstrates how data analysis can help us provide valuable insights into the factors that contribute to employee attrition rates. By understanding these factors and implementing effective retention strategies, companies like IBM can reduce turnover rates, improve employee satisfaction, and ultimately, enhance their overall performance.

## **References:**

* <https://panmore.com/ibm-five-forces-analysis-porters-recommendations>
* <https://www.ibm.com/us-en/><https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset>
* <https://www.ibm.com/products/zos?lnk=flatitem>
* <https://www.ibm.com/docs/en/db2/10.5?topic=analytics-overview-db2-warehouse>
* <https://go-gale-com.esearch.ut.edu/ps/i.do?p=ITOF&u=tamp73569&v=2.1&it=r&id=GALE%7CA693024253&inPS=true&linkSource=interlink&sid=bookmark-ITOF>
* <https://en.wikipedia.org/wiki/IBM>
* <https://www.britannica.com/topic/International-Business-Machines-Corporation>
* <https://www.techtarget.com/searchitchannel/definition/IBM-International-Business-Machines>
* <http://www.orgcharting.com/ibm-org-chart/>
* <https://www.investopedia.com/how-ibm-makes-money-4798528>
* <https://www.ibm.com/downloads/cas/4ZR6LXE5>
* Dietrich, Brenda L.; Plachy, Emily C.; Norton, Maureen F. Analytics Across the Enterprise (IBM Press). Pearson Education.