MGT7222: Human Resource Analytics

Data-Driven Decisions: An HR Exploration of Employee Attrition and Recommendation in a pharmaceutical company

Name: Dhanush Mathighatta Shobhan Babu
Student ID: 40412492

Word Count: 1932

Table Of Content

Sl. No	Content	Page No
1.	1. Introduction and Background 1.1 Background 1.2 Overview of the Business Problem 1.3 Importance of the problem 1.4 Challenges faced 1.5 Analytical tasks and role of analytics 1.5.1 Data Exploration 1.5.2 Data Quality Assessment	1-4
2.	2. Methodology 2.1 Proposed solution design 2.2 Chosen model 2.3 Step-By-Step Flowchart or Infographic 2.4 Applicability and Efficacy of the Method	5-6
3.	3. Findings 3.1 Visualizations and Insights 3.1.1 Overtime VS Attrition 3.1.2 Age VA Attrition 3.1.3 Gender/Marital Status VS Attrition 3.1.4 Job Level VS Attrition 3.1.5 Gender VS Attrition 3.1.6 Attrition VS Job role 3.1.7 Department VS Attrition 3.1.8 Attrition VS Percent Salary Hike 3.1.9 Key Performance Indicators 3.2 Choice of chart type 3.2.1 Bar Chart 3.2.2 Line Chart 3.2.3 Pie Chart 3.3 Use of Parameters and Calculations	7-14
4.	4. Recommendations and Discussions 4.1 Recommendation to reduce attrition	15-16
5.	5. Future work 5.1 Follow-Up and Future work	17
6.	6. Referencing	18-19

Table of Figures

Sl. No	Content	Page No
1.	Figure 1.2	1
2.	Figure 2.2.1	5
3.	Figure 2.3.1	6
4.	Figure 3.1.1.1	7
5.	Figure 3.1.2.1	7
6.	Figure 3.1.3.1	8
7.	Figure 3.1.4.1	9
8.	Figure 3.1.5.1	9
9.	Figure 3.1.6.1	10
10.	Figure 3.1.7.1	11
11.	Figure 3.1.8.1	11
12.	Figure 3.1.9.1	12
13.	Figure 3.4.1	13

1.Introduction

1.1 Background

The firm in focus is a big pharmaceutical corporation that is facing severe problems as a result of high levels of workforce turnover. To monitor Human Resource Metrics, the organisation now depends on standard monthly reports delivered in Excel files. Recognising the limits of this strategy, they are eager to use more complex analytics tools, such as Tableau, to acquire a better grasp of the data and solve the problem of high attrition.

1.2 Overview of the Business Problem

According to Jain and Nayyar (2018), attrition refers to a decrease in the workforce size resulting from several factors such as voluntary resignations, mortality, and retirement, among others or Employee attrition refers to the occurrence of employees voluntarily or involuntarily leaving an organisation (Mozaffari et al., 2022).Currently in the company ,the overall attrition rate is 19.15 %.The report is presented to the Board members of the company and the main objective is to find the prominent factors affecting attrition

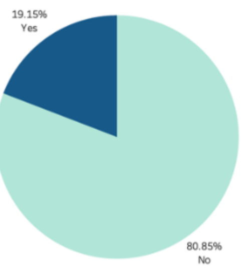

Fig 1.2: Overall Attrition

1.3 Importance of the problem

Employee churn is a significant concern for organisations across several sectors. A high attrition rate can result in financial losses and have a negative impact on the knowledge base of a firm as well as the degree of employee engagement. In addition, it can also have an effect on productivity and the timely attainment of organisational objectives. According to the study conducted by (Mozaffari et al.2022). Past research indicates that high staff attrition rates might have a detrimental impact on organisational efficiency (Han, 2020).

1.4 Challenges Faced

Employee attrition can cause a variety of operational issues, including knowledge loss, workflow disruption, financial costs, and strain on human resources. Teamwork, job happiness, client relationships, and service quality can all suffer as a result. Hiring replacements comes at a considerable cost to the organisation, including the price of hiring, training, and conducting interviews (Alduayj & Rajpoot, 2018).

1.5 Analytical Tasks and the Role of Analytics

The initial analytical step involves integrating the provided data into Tableaufor the purpose of data exploration and assessment of data quality. Subsequently, through conducting exploratory analysis, significant patterns within the data can be identified. Based on these findings, a comprehensive report should be generated to improve the decision-making process within the business. (Hoelscher & Mortimer, 2018)

1.5.1 Data Exploration

The data is obtained from the HRIS (Human Resource Information system) database. We have divided our data into two categories Numerical and categorical. The below table illustrates data exploration:

Categorical Data

Sl.No	DATA VARIABLE NAME	DISTINCT VALUE	No. Of Distinct Values
1.	ATTRITION	Yes and No	2
2.	BUSINESS TRAVEL	Rarely, Frequently, None	3
3.	DEPARTMENT	HR, Sales and R & D	3
4.	EDUCATION	BelowCollege,College, Bachelor, Master,Doctor	5
5.	EDUCATION FIELD	HR,Life Science, Marketing, Medical Science,Others,Technical	7
6.	ENVIRONMENT SATISFACTION	Low,Medium, High,Very High	4
7.	GENDER	Male and Female	2
8.	JOB INVOLVEMENT	Low, Medium, High, Very high	4
9.	JOB LEVEL	Junior to Senior	5
10.	JOB ROLE	Sales Executive, Research Scientist, Laboratory Technician, Manufacturing Director, Healthcare, Representative, Manager, Sales Representative, Research Director, Human Resources	10
11.	JOB SATISFACTION	Low, Medium, High, Very high	4
12.	MARITAL STATUS	Single, Married and Divorced	3
13.	OVERTIME	Yes and No	2
14.	PERFORMANCE RATING	Low, Good, Excellent, Outstanding	4
15.	RELATIONS SATISFACTION	Low, Medium, High, Very High	4
16.	STOCK OPTION LEVEL	None, Low, Medium, High	4
17.	WORK LIFE BALANCE	Bad, Good, Better, Best	4

Table 1.5.1.1 - categorical data analysis

Numerical Data

Sl.No	DATA VARIABLE NAME	MINIMUM	MEAN	MEDIAN	MAXIMUM
1.	AGE	7	36.97	36	85
2.	BILLABLE RATE	102	802.9	802	1499
3.	DISTANCE FROM HOME(MILES)	1	9.2002	7	29
4.	DOB	1936-09-02	-	-	2003-09-12
5.	HOURLY RATE	30	65.87	66	100
6.	ID	1	-	-	1467
7.	MONTHLY INCOME ($)	1009	6505	4908	19999
8.	MONTHLY RATE	2094	14323	14255	26999
9.	NUMCOMPANIES WORKED	0	2.695	2	9
10.	PECENT SALARY HIKE	11	15.21	14	25
11.	TOTAL WORKING YEARS	0	11.32	110	94
12.	TRANING TIMES LAST YEAR	0	2.801	3	6
13.	YEARS AT COMPANY	0	7.01	5	40
14.	YEARS IN CURRENT ROLE	0	4.32	3	18
15.	YEARS SINCE LAST PROMOTION	0	2.192	1	15
16.	YEARS WITH CURRENT MANAGER	0	4.123	3	17

Table 1.5.1.2 – numerical data analysis

1.5.2 Data Quality Assessment

The issues with data quality are as follows:

DATA VARIABLE NAME	DATA TYPE	OUTLIER	ISSUE
AGE	Numerical	7,77,85	It is possible to call these data points "outliers" because people under the age of 18 are not allowed to work by law, and the average retirement age is around 60 years old. Because of this, it is very unlikely that someone between the ages of 77 and 85 would be actively working for a company.
DEPARTMENT	Categorical	HR, Human Resources, R & D, Resource and Development	Despite the repetition, the indicated categories remain identical. Consequently, it is imperative to classify them under the same category, namely HR and R&D or Human Resource and Resource and Development.
TOTAL WORKING YEARS	Numerical	94	The present observation may be classified as an outlier on account of the exceedingly unusual occurrence of an individual having accumulated 94 years of job experience.

Table 1.5.2.1 – data quality issues

2. Methodology

2.1 Proposed solution design

An industry-neutral process model for data mining is called CRISP-DM. Six iterative phases make up this process, which goes from business knowledge to deployment (Schröer *et al.,* 2021). We will be employing certain phases of the CRISP-DM methodology, specifically focusing on descriptive analysis.

2.2 Chosen Methods

The following steps are adapted from CRISP-DM:

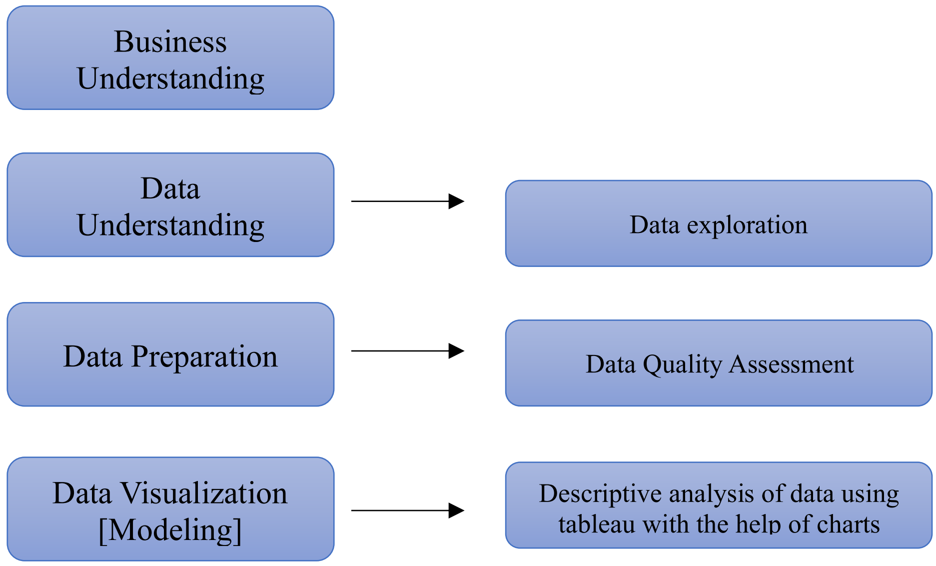

Fig 2.2.1

1. **Business Understanding**: An evaluation of the business environment is necessary in order to obtain a comprehensive understanding of the resources that are now available and those that are needed. The identification of the objective is a crucial element in this stage (Schröer *et al.,* 2021).

2. **Data Understanding**: The acquisition of data from various sources, the subsequent examination and depiction of the data, and the evaluation of its integrity are fundamental activities during this stage (Schröer *et al.,* 2021). In this particular instance, we obtained the data from the Human Resource Information System (HRIS) and subsequently classified it into two distinct data groups, as previously indicated.

3. **Data Preparation**: Data selection should be carried out through the establishment of criteria for inclusion and exclusion. Poor data integrity can be remedied through data cleansing (Schröer *et al.,* 2021). The data quality issues in table 3.1.2 have been identified and emphasized.

4. **Data Visualization**: Data visualisation is a process that aids in understanding and processing numerical data by presenting it in a visual framework, transforming both large and small data sets into comprehensible visual representations (Islam & Jin, 2019).

2.3 Step-by-Step Flowchart or Infographic

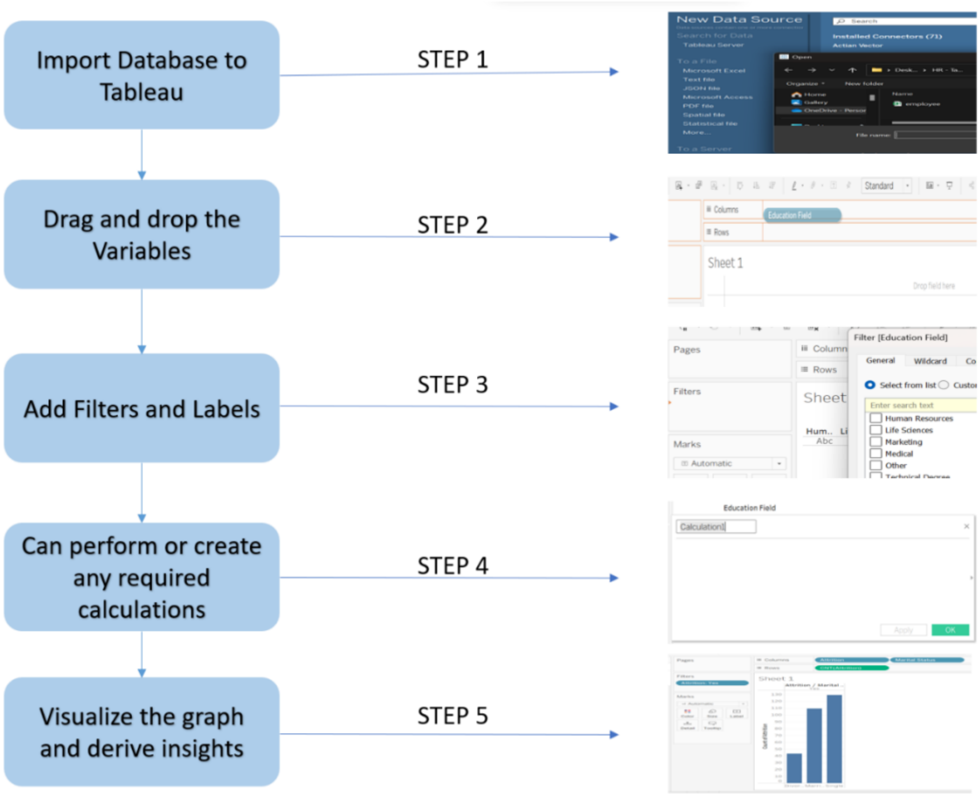

Fig 2.3.1

2.4Applicability and Efficacy of the Methods

CRISP-DM has many data mining benefits. First, it provides a methodical framework for successful project development, establishing and adhering to all necessary processes and obligations. Correlation-based feature selection helps CRISP-DM find and choose relevant characteristics. This improves classification model accuracy (Source: https://typeset.io/questions/what-are-the-advantages-of-using-crisp-dm-sldtuc7jl9)

.

3. Findings

3.1 Visualization and Insights

3.1.1 Overtime VS Attrition

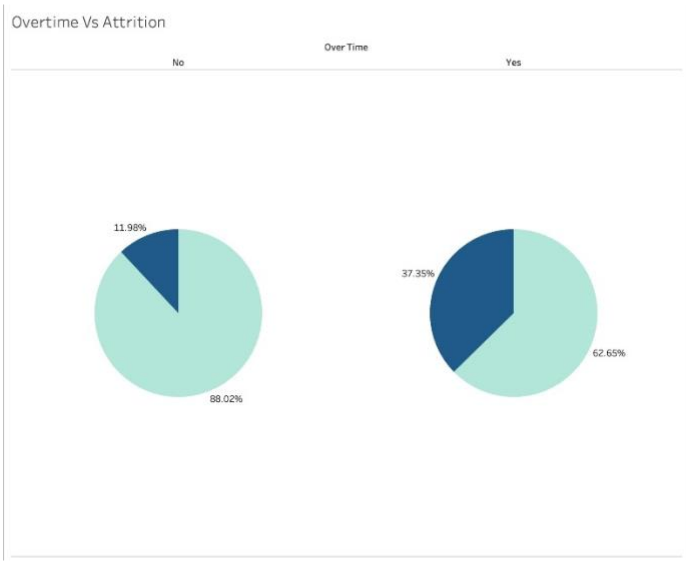

Fig 3.1.1.1

This visualisation uses pie charts. Our study examined overtime and attrition. Employees that have worked extra have much lower attrition rates, according to the visualisation. The data shows that 11.98% of non-overtime employees have left the organisation, whereas 37.35% of overtime employees did. Since overtime workers leave the organisation more often, this strongly shows that overtime affects attrition.

3.1.2 Age VS Attrition

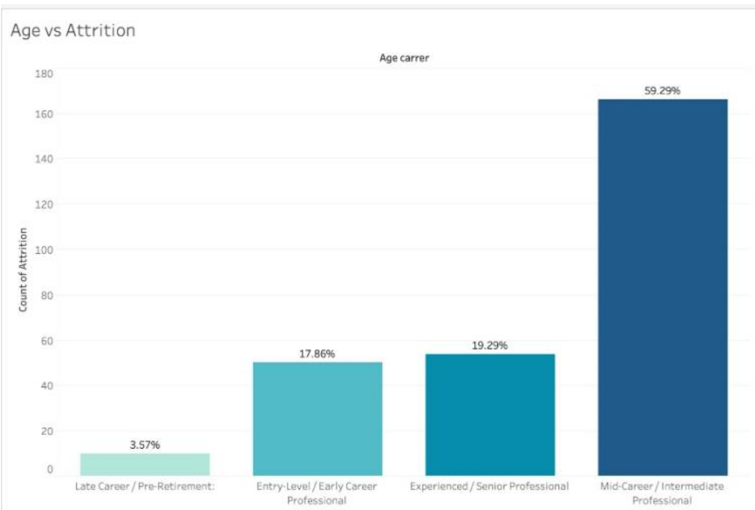

Fig 3.1.2.1

The bar graph presented herein displays the categorization of the parameter "Age" into four distinct categories. Individuals at the entry-level or early career stage, often ranging from 18 to 25 years of age. The mid-career or intermediate professional stage often encompasses individuals between the ages of 26 and 40. Experienced and senior professionals within the age range of 41 to 55 years old.
In the late career/pre-retirement stage, individuals typically range in age from 56 to 60 years old. It is evident from the data that a significant proportion of individuals classified as Mid-Career/Intermediate Professionals are departing from the organisation, accounting for around 59.29% of the total.

3.1.3. Gender/Marital Status Vs attrition

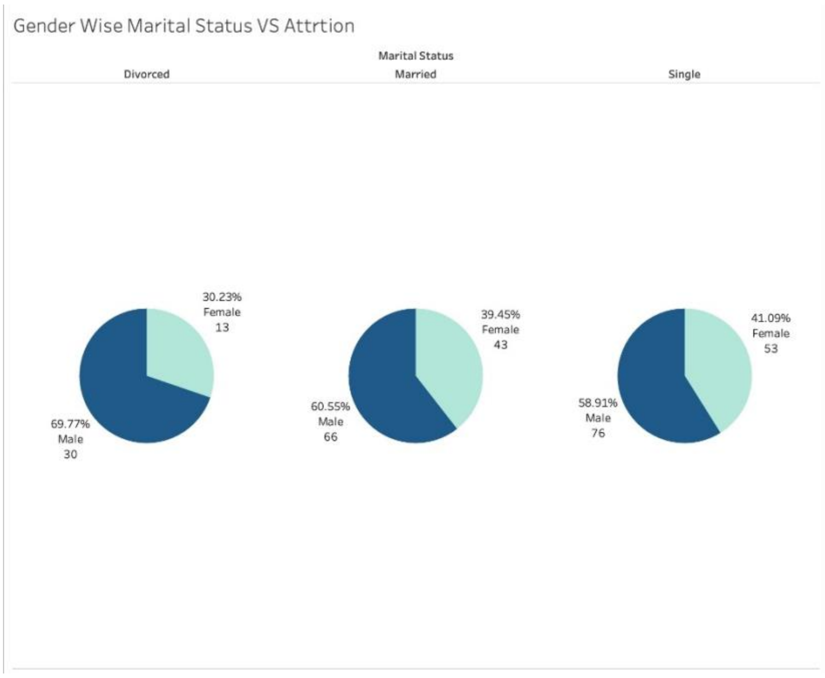

Fig 3.1.3.1

A multivariate analysis was conducted, examining three variables: Attrition, Gender, and Marital Status. The concept of attrition has been employed as a means of filtration. The visualisation reveals that the primary trend seen is the departure of male employees from the company, regardless of their marital status. Furthermore, it is noteworthy that the largest group of individuals departing the company consists of single males, with a count of 76.

3.1.4 Job Level VS Attrition

There exist five distinct employment levels, denoted numerically from one to five. Specifically, level one corresponds to junior positions, while level five corresponds to senior positions. The visualisation illustrates that individuals classified as Job level 1, denoting junior positions, exhibit the highest rate of attrition within the organisation, accounting for 50.89% of the total.

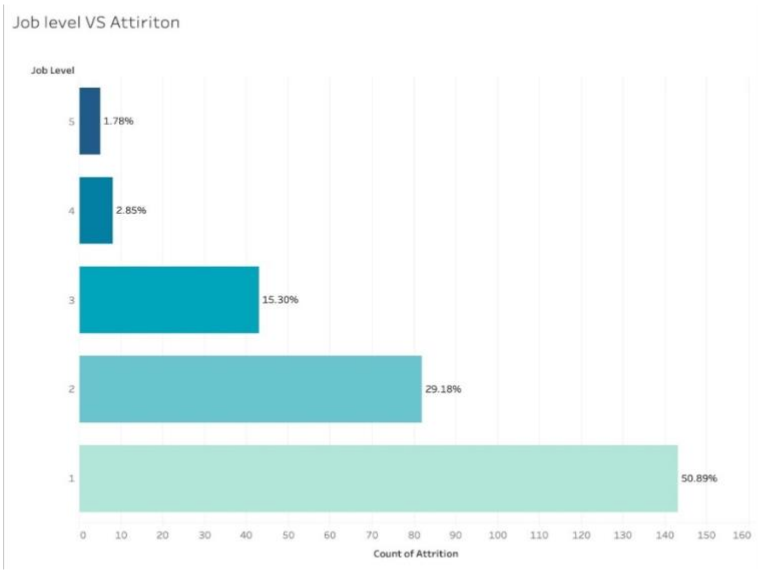

Fig 3.1.4.1

3.1.5 Gender VS Attrition

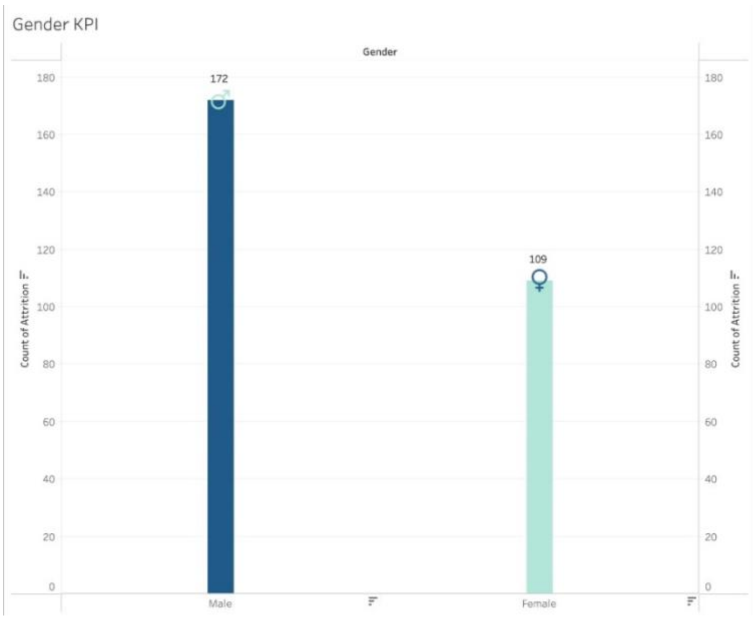

Fig 3.1.5.1

In this visualisation, a bar chart has been employed to represent data. The chart incorporates vector symbols to visually represent male and female individuals. Based on the visualisation, it can be inferred that out of a total of 281 employees who have departed from the organisation, males outnumber females. Specifically, there were 172 male employees who left, while the count for female employees who left stands at 109.

3.1.6 Attrition VS Job role

Attrition Vs Job role

Job Role	% of Total Count of Attrition along Job Role	Count of Attrition
Research Director	0.71%	2.0
Manager	1.78%	5.0
Healthcare Representative	3.20%	9.0
Manufacturing Director	3.56%	10.0
Human Resources	4.27%	12.0
Sales Representative	11.74%	33.0
Research Scientist	16.73%	47.0
Laboratory Technician	22.06%	62.0
Sales Executive	35.94%	101.0

Fig 3.1.6.1

This visualisation utilised a textual tabular format to ascertain the occurrence rate of attrition inside different occupational positions. The findings indicate that the sales executive role demonstrates the highest rate of attrition, with a total count of 101, or approximately 35.94% of the overall attrition.

3.1.7 Department VS Attrition

Based on the presented visualisation, it has been ascertained that the Sales department demonstrates the greatest attrition rate among the three departments, namely HR, R&D, and Sales, with a percentage of 30.56%. Following this, the Human Resources (HR) department exhibits an attrition rate of 19.05%, whereas the Research and Development (R&D) department demonstrates the lowest attrition rate of 13.87%.

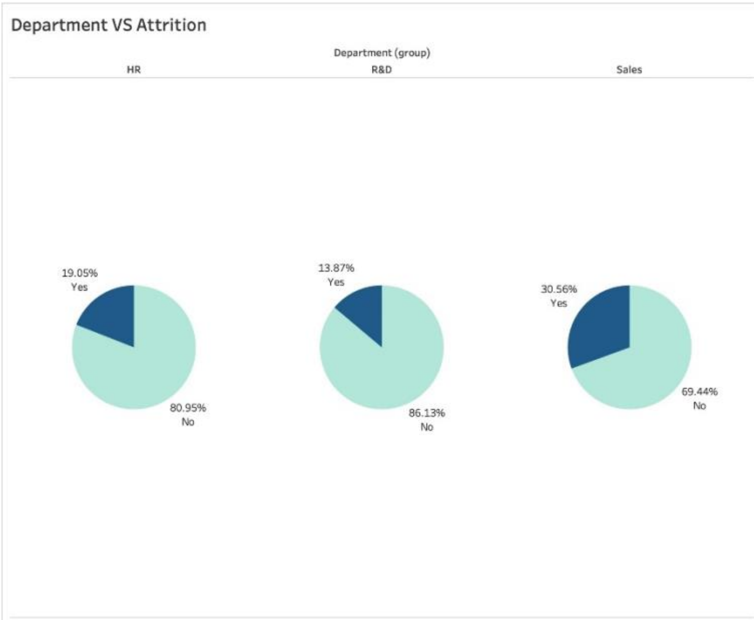

Fig 3.1.7.1

3.1.8 Attrition VS Percent Salary Hike

Fig 3.1.8.1

The graph shows a negative link between attrition and wage increases across all firm departments. This association is not sector-specific, emphasising the urgency of solving this issue due to its widespread impact. When compensation increases are minor, attrition is more likely. Salary increases are a major driver in employee attrition (Joseph et al. (2021)).

3.1.9 Key Performance Indicators

KPI

Count of employee	Attrition_ count	Active employee	Attrition rate	Avg. Age
1,467	281	1,186	19.15%	37

Fig 3.1.9.1

There are 5 KPI's :

- Total Count of Employees
- Count of Attrition
- Number of Active Employees
- Attrition rate
- Average age

3.2 Choice of Chart

3.2.1 Bar Chart

Bar charts are particularly efficient in visually representing numerical data that can be easily categorised, allowing for the rapid identification of patterns and trends within the dataset (Hardin et.al,2014).

3.2.2 Line Chart

Line charts are commonly utilised alongside bar and pie charts as one of the most frequently employed sorts of charts. Line charts are used to establish a connection between discrete numerical data points. The outcome is a concise and direct method for representing a series of values in a visual manner (Hardin et.al,2014).

3.2.3 Pie Chart

Pie charts are commonly employed to visually represent the relative proportions or percentages of information (Hardin et.al,2014).

3.3 Use of Parameters and Calculations

There are now three computed variables available: one for the attrition rate, one for the attrition count, and one for the number of active Employees.

The following formulas were used in this study:

1)Attrition count:

IF [Attrition] = 'Yes' THEN 1 ELSE 0 END

2) No of active employees:

COUNT([employee]) - SUM ([Attrition_count])

3)Attrition rate:

SUM ([Attrition_count]) /COUNT([employee])

3.4 Dashboard design explanation

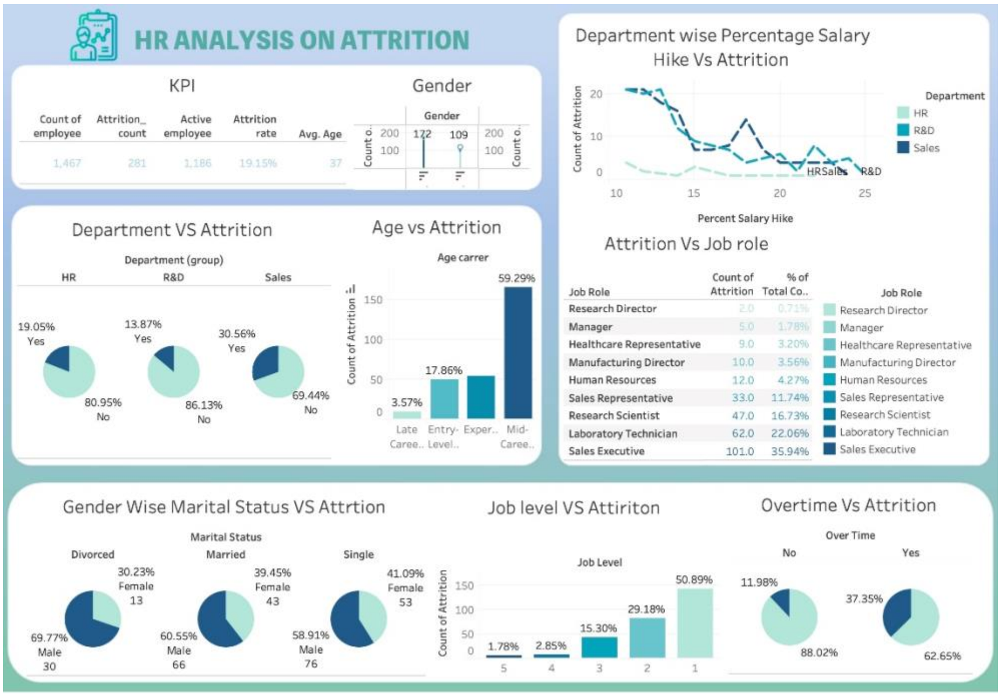

Fig 3.4.1

The dashboard provides a comprehensive summary of the primary elements that contribute to attrition, as well as the present state of the firm with regards to attrition. The dataset comprises descriptive and comparative information pertaining to attrition.

4. Recommendations and Discussions

4.1 Recommendation to reduce attrition,

Overtime - by implementing a system of monitoring and regulating overtime hours, organisations can effectively mitigate the risk of employee burnout.
Marital Status - by promoting a culture of open communication and actively seeking input from married employees, the organisation aims to enhance overall performance and well-being.
Job role - by fostering collaboration between sales and marketing divisions, the level of stress can be reduced, while simultaneously providing enhanced compensation packages and opportunities for professional development.
Job level - Effective retention techniques that encourage, recognise, and sustain motivation among new hires are needed to reduce employee churn. By offering incentives and rewards, companies can encourage employees to dedicate themselves to the company's goals.

4. 2 Literature based Recommendations

According to Al-Suraihi et al. (2021), businesses can use eight retention methods to retain employees. The company must follow all recommendations. According to Al-Suraihi et al. (2021), the following techniques were found.

1. Organisational Management: A structured internal job posting system helps individuals make meaningful contributions in their areas of competence, supporting professional growth and development.
2. Employee Training: Skills training is needed for employees.
3. Implement rigorous training classes that encompass the latest technical improvements and keep staff informed about industry trends.
4. Job involvement involves actively monitoring and increasing employee project engagement to help solve disengagement.
5. Recognising and praising employees' contributions can boost job satisfaction and workplace culture.
6. awards and Benefits: Identify and incentivize top performers with immediate awards, employee recognition, and performance-based compensation.
7. Job engagement involves acknowledging and appreciating individuals' unique traits and strategically assigning them to projects that maximise their skills and abilities. Implementing this technique is crucial to building and maintaining staff dedication and passion.

8. Workplace Environment: Encourage open communication, productivity, and a positive company culture.
9. Active involvement and teamwork in the workplace are greatly enhanced by employee participation. Through the integration of job engagement approaches with these efforts, firms can enhance employee participation in all aspects of the business.

5. Future work

5.1 Follow-Up and future work

Advanced analytics has many benefits, according to Hanna et al. (2021).
Forecasting accuracy is higher with advanced analytics than with BI tools. Increased precision lowers uncertainty.
Accurate forecasts accelerate decision-making. This can accelerate executive decision-making and enhance success replication confidence.
Advanced analytics helps stakeholders make strategy, consumer, market, and company decisions based on data.
Risk management may benefit from advanced analytics. Advanced analytics can help firms estimate more correctly and eliminate costly errors.

References

1. Alduayj, S.S. and Rajpoot, K. (2018) 'Predicting employee attrition using machine learning', 2018 International Conference on Innovations in Information Technology (IIT) [Preprint]. doi: Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8605976 [Accessed on: 8th November 2023].

2. Han, J.W. (2020) 'A review of antecedents of employee turnover in the hospitality industry on individual, team and organizational levels', International Hospitality Review, 36(1), pp. 156–173. doi: Available at: https://www.emerald.com/insight/content/doi/10.1108/IHR-09-2020-0050/full/html[Accessed on: 28th October 2023].

3. Hardin, M., Hom, D., Perez, R. and Williams, L., (2012). Which chart or graph is right for you?. Tell Impactful Stories with Data. Tableau Software. Doi: Available at: https://static1.squarespace.com/static/555cc0bbe4b0425d05f56579/t/559dbc40e4b0d1590e8a3ea0/1436400704467/CCSS+Math-which_chart_v5%5B1%5D.pdf [Accessed on: 6th November 2023].

4. Hoelscher, J. and Mortimer, A. (2018) 'Using Tableau to visualize data and drive decision-making', Journal of Accounting Education, 44, pp. 49–59. doi: Available at: Available at : https://www.sciencedirect.com/science/article/abs/pii/S0748575117301951?fr=RR-2&ref=pdf_download&rr=824e4b4c588f4922 [Accessed on: 4th November 2023].

5. Islam, M. and Jin, S. (2019) 'An overview of data visualization', 2019 International Conference on Information Science and Communications Technologies (ICISCT) [Preprint]. doi: Available at: https://ieeexplore.ieee.org/abstract/document/9012031 [Accessed on: 2nd November 2023].

6. Jain, R. and Nayyar, A. (2018) 'Predicting employee attrition using XGBoost machine learning approach', 2018 International Conference on System Modeling & Advancement in Research Trends (SMART) [Preprint]. doi: Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8746940 [Accessed on: 4th November 2023].

7. Mozaffari, F. et al. (2022) 'Employee attrition prediction in a pharmaceutical company using both machine learning approach and qualitative data', Benchmarking: An International Journal [Preprint]. doi: Available at: https://www.emerald.com/insight/content/doi/10.1108/BIJ-11-2021-0664/full/html?skipTracking=true [Accessed on: 2nd November 2023].

8. Peter Dunn (2022) Understanding the advantages and disadvantages of Advanced Analytics, Technology Solutions That Drive Government. Available at: https://fedtechmagazine.com/article/2023/03/understanding-advantages-and-disadvantages-advanced-analytics#:~:text=The%20Disadvantages%20of%20Advanced%20Analytics,collecting%2C%20offl [Accessed on: 30 October 2023].

9. Schröer, C., Kruse, F. and Gómez, J.M. (2021) ‘A systematic literature review on applying CRISP-DM process model’, Procedia Computer Science, 181, pp. 526–534. doi: Available at: https://www.sciencedirect.com/science/article/pii/S1877050921002416 [Accessed on: 8th November 2023].

10. What are the advantages of using CRISP-DM?: 4 answers from research papers (no date) SciSpace - Question. Available at: https://typeset.io/questions/what-are-the-advantages-of-using-crisp- dm-sldtuc7jl9 [Accessed on: 15 November 2023].

11. Al-Suraihi, W.A. et al. (2021) ‘Employee turnover: Causes, importance and retention strategies’, European Journal of Business and Management Research, 6(3), pp. 1–10. doi: Available at: https://www.ejbmr.org/index.php/ejbmr/article/view/893 [Accessed on: 8th November 2023].