

DATA VISUALISATION USING TABLEAU AND R SOFTWARE

EXECUTIVE SUMMARY

Not every organization or every job is right for every person. Employees are leaving the workforce faster than they are hired, and it is often outside the employer's control. Thus, it is impossible to prevent employee attrition. While some employees leave the company when they reach retirement age, many others depart for a variety of reasons, including poor job satisfaction rate, low compensation, and a hostile work environment. Attrition data can reveal a wealth of information about how the organisation is operating.

The study reveals the company's staff attrition rate. Through the analysis, it has been determined that the workplace environment for employees is comfortable and safe. The company's ability to grow depends on the attrition of its valuable personnel, who serve as the foundation of the business. Therefore, the business needs to fill the gap by addressing employee needs and expectations for their jobs and other relevant areas that aid in better employee retention. The business may also decide to enhance the bonus amount, provide better transportation and lodging options, and introduce new incentive programmes to encourage employees to work harder and provide the finest work possible.

Even in the most ideal workplaces, attrition happens. Attrition may give HR professionals the chance to re-evaluate and realign their recruitment and hiring tactics with the company's goals and values, despite the fact that it is frequently unpleasant to let go of employees. Additionally, attrition can make it possible to hire top talent or people who are better suited for particular occupations.

The core reason of this attrition could be mismatch in expectations of organization and expectations of employees from each other. This study aimed to analyse the attrition reasons as well as understand the expectation of employees from the organization.

Table of Contents

1. INTRODUCTION	6
1.1 ATTRITION	6
1.1.1 TYPES OF ATTRITION	6
1.1.2 REASONS FOR ATTRITION	7
1.1.3 IMPACTS OF EMPLOYEE ATTRITION	7
1.1.4 MANAGING EMPLOYEE ATTRITION	7
2. PROJECT PROBLEM STATEMENT	8
2.1 OBJECTIVES FOR THE STUDY	8
2.2 NEED FOR THE STUDY	8
3. PROJECT TOOLS AND TECHNIQUES	9
3.1 TOOLS	9
3.1.1 TABLEAU	9
3.1.2 R PROGRAMMING	9
3.2 TECHNIQUES	9
3.2.1 BAR GRAPH	9
3.2.2 PIE CHART	9
3.2.3 STACKED BAR CHART	10
3.2.4 AREA CHART	10
3.2.5 LINE CHART	10
3.2.6 DASHBOARD	10
3.2.7 DESCRIPTIVE STATISTIC	10
3.2.8 CORRELATION ANALYSIS	10
3.2.9 ONE WAY ANOVA	11
3.2.10 TWO WAY ANOVA	11
3.2.11 CHI - SQUARE TEST	11
3.2.12 LOGISTIC REGRESSION	11
4. DATA ANALYSIS - TABLEAU	12
4.1. Employee Attrition Vs Education/Age/Job Travel/Distance from Home/Marital Status:	12
4.2 Factors Affecting Employee Attrition:	13
5. PROJECT INFERENCE - TABLEAU:	14
5.1 EMPLOYEE ATTRITION BY GENDER DISTRIBUTION:	14
5.2 EMPLOYEE ATTRITION BY MARTIAL STATUS:	14
5.3 EMPLOYEE ATTRITION BY EDUCATION:	14
5.4 EMPLOYEE ATTRITION BY AGE:	15
5.5 EMPLOYEE ATTRITION BY DISTANCE FROM HOME:	15
5.6 EMPLOYEE ATTRITION BY DEPARTMENT:	15
5.7 EMPLOYEE ATTRITION BY JOB ROLE:	15
5.8 EMPLOYEE ATTRITION BY PROMOTION:	16

5.9 EMPLOYEE ATTRITION BY SALARY HIKE:	16
5.10 EMPLOYEE ATTRITION BY TRAINING TIME:	16
5.11 EMPLOYEE ATTRITION BY YEARS WITH CURRENT MANAGER:	17
6. DATA ANALYSIS AND INFERENCE – R STUDIO	18
6.1 DESCRIPTIVE ANALYSIS	18
6.2 CORRELATION	18
6.2.1 AGE AND ATTRITION	18
6.2.2 MONTHLY INCOME AND ATTRITION	19
6.2.3 DISTANCE FROM HOME AND ATTRITION	19
6.3 CHI-SQUARE	20
6.3.1 GENDER AND ATTRITION	20
6.3.2 DEPARTMENT AND ATTRITION	20
6.3.3 MARTIAL STATUS AND ATTRITION	21
6.4 ONE-WAY ANOVA	21
6.4.1 ATTRITION AND JOB ROLE	21
6.4.2 ATTRITION AND PERCENTAGE SALARY HIKE	22
6.5 TWO WAY ANOVA	22
6.5.1 ATTRITION, EDUCATION AND DEPARTMENT	22
6.5.2 ATTRITION, DEPARTMENT AND MARITAL STATUS	23
6.6 LOGISTIC REGRESSION	24
7. FINDINGS AND RECOMMENDATIONS	25
7.1 FINDINGS	25
7.2 RECOMMENDATIONS	26
8. CONCLUSION	27
APPENDIX	28

1. INTRODUCTION

Employee retention is crucial to your company's success and the attrition rate is the metric that provides insight into how well a company is retaining its employees.

The employee attrition analytics is specifically focused on identifying why employees leave, what factors have influenced them to leave, and what might have prevented them from leaving based on the data provided. It can be used to help organizations understand and design the interventions that will be most effective in reducing unwanted attrition.

1.1 ATTRITION

Attrition is the term used to describe the loss of workers due to retirement, resignation, or death. **Employee attrition** is the gradual reduction in employee numbers. Employee attrition happens when the size of your workforce diminishes over time. This means that employees are leaving faster than they are hired.

The rate of employee attrition can be a good indicator of the health of an organization. A high rate of attrition can be a sign that employees are unhappy with their jobs or that the organization is not a good place to work. A low attrition rate can be a sign that employees are happy with their jobs and that the organization is a good place to work.

Employee attrition can be caused by a number of factors, including poor working conditions, low pay, and job dissatisfaction. Attrition can also be caused by external factors such as the economy or changes in the industry.

Organizations can use a number of strategies to reduce employee attrition, including improving working conditions, providing training and development opportunities, and increasing pay and benefits.

1.1.1 TYPES OF ATTRITION

There are two types of employee attrition:

Voluntary attrition: When an employee leaves the company willingly, it is called voluntary attrition. Employment attrition can be voluntary due to many reasons, such as better job opportunities, earning more money, or being closer to home.

Involuntary attrition: When an employee is forced to leave the company against their will, it is called involuntary attrition. Involuntary attrition can be caused by many reasons, such as being fired, laid off, or taking a leave of absence.

1.1.2 REASONS FOR ATTRITION

There are many reasons why employees leave their jobs, but some of the most common reasons include dissatisfaction with:

- Work environment
- job itself
- Company culture
- Salary or benefits
- Management
- Work/life balance
- Location
- Company's direction
- Career growth potential

1.1.3 IMPACTS OF EMPLOYEE ATTRITION

Employee attrition can have many impacts on a company. Some of the most common impacts are:

1. Loss of productivity
2. Loss of morale
3. Increase in costs
4. Loss of customers
5. Damage to the company's reputation

1.1.4 MANAGING EMPLOYEE ATTRITION

There are many ways to reduce employee attrition. Some of the most common ways are:

1. Improve salaries
2. Improve working conditions
3. Provide career growth opportunities
4. Improve management
5. Improve work-life balance

2. PROJECT PROBLEM STATEMENT

A large company named XYZ, employs, at any given point of time, around 4000 employees. However, every year, around 15% of its employees leave the company and need to be replaced with the talent pool available in the job market.

Data about various aspects like Gender, Travel, Department, Distance from home, Education Level, Education Field, Job Level, Job Role, Martial status, Monthly income, Percentage of salary hike of each employee, Training, Years served in the company and under current manager, Last promotion, Total working years etc., about each employee is collected to understand the pattern of employee attrition in XYZ.

The management believes that this level of attrition (employees leaving, either on their own or because they got fired) is bad for the company, because of the following reasons -

- ✓ The former employees' projects get delayed, which makes it difficult to meet timelines, resulting in a reputation loss among consumers and partners.
- ✓ A sizeable department has to be maintained, for the purposes of recruiting new talent
- ✓ More often than not, the new employees have to be trained for the job and/or given time to acclimatise themselves to the company

The project thus aims to identify and analyse the rate of employee attrition through statistical tools.

2.1 OBJECTIVES FOR THE STUDY

1. To understand which factors affect attrition the most.
2. To study about the attrition trend
3. To understand the changes that must be made in the workplace in order to curb retention.

2.2 NEED FOR THE STUDY

1. To help the businesses to make informed decisions
2. To help businesses improve rate of retention of employees
3. To know what changes are to be made to the workplace

3. PROJECT TOOLS AND TECHNIQUES

Tools and techniques in research are the statistical methods of collection, analysis, interpretation, presentation, and organization of data. Statistics provides numerous tools and techniques to analyze the data and interpret the results of the analysis.

3.1 TOOLS

3.1.1 TABLEAU

Tableau is a software company that offers collaborative data visualization software for organizations working with business information analytics. Organizations use Tableau to visualize data and reveal patterns for analysis in business intelligence, making the data more understandable. It can be used to create charts, graphs, and other visual representations of data. Tableau products query relational databases, online analytical processing cubes, cloud databases, and spreadsheets to generate graph-type data visualizations. The software can also extract, store, and retrieve data from an in-memory data engine.

3.1.2 R PROGRAMMING

R is a language and environment for statistical computing and graphics. R provides a wide variety of statistical and graphical techniques, and is highly extensible. R can produce well-designed publication-quality plots, including mathematical symbols and formulae where needed. It is a free and open-source software environment that is available for Windows, MacOS, and Linux. R is a popular language for statistical analysis, data science and data analysis.

3.2 TECHNIQUES

3.2.1 BAR GRAPH

A bar graph is a graphical representation of information. It is used when x axis categorical and y axis is continuous variable. It uses bars that extend to different heights to depict value. Bar graphs can be created with vertical bars, horizontal bars, grouped bars or stacked bars. Bar graphs are commonly used in business and financial analysis to display often complicated data. They can convey information quickly and effectively.

3.2.2 PIE CHART

A pie chart (or a circle chart) is a circular statistical graphic, which is divided into slices to illustrate numerical proportion. In a pie chart, the arc length of each slice is proportional to the quantity it represents. It is used to represent the percentage distribution of the data and it is

applicable for categorical variables only. Pie charts are very widely used in the business world and the mass media.

3.2.3 STACKED BAR CHART

A stacked chart is a form of bar chart that shows the composition and comparison of a few variables, either relative or absolute, over time. Also called a stacked bar or column chart, they look like a series of columns or bars that are stacked on top of each other. Stacked charts are an incredibly effective tool for comparisons. They are designed to compare total values across categories.

3.2.4 AREA CHART

An area chart or area graph displays graphically quantitative data. It is based on the line chart. The area between axis and line are commonly emphasized with colours, textures and hatchings. Commonly one compares two or more quantities with an area chart. It is used to represent the distribution of area with respect to geographic location and a demographic factor.

3.2.5 LINE CHART

A line chart is a type of chart that provides a visual representation of data in the form of points that are connected in a straight line. A line chart is one of the simplest methods to understand any financial data and trading data. It is used to find the trend and patterns.

3.2.6 DASHBOARD

The dashboard chart is created to display data in a graphical format. It is a graphical representation of data that can be used to track and monitor the performance of a business or organization. The dashboard chart can be used to track the performance of a variety of metrics, including sales, marketing, customer satisfaction, and employee productivity.

3.2.7 DESCRIPTIVE STATISTIC

Descriptive statistics as the name implies is used to describe a dataset. It helps understand the details of your data by summarizing it and finding patterns from the specific data sample. They provide absolute numbers gotten from a sample but do not necessarily explain the rationale behind the numbers and are mostly used for analysing single variables.

3.2.8 CORRELATION ANALYSIS

correlation analysis calculates the level of change in one variable due to the change in the other. A high correlation points to a strong relationship between the two variables, while a low correlation means that the variables are weakly related. It is used to find the relation between two variables.

3.2.9 ONE WAY ANOVA

One-Way ANOVA ("analysis of variance") compares the means of two or more independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different. In ANOVA the independent variable is taken as categorical data and the dependent variable is taken as continues data.

3.2.10 TWO WAY ANOVA

A two-way ANOVA is used to estimate how the mean of a quantitative variable changes according to the levels of two categorical variables. Use a two-way ANOVA when you want to know how two independent variables, in combination, affect a dependent variable. In ANOVA the independent variable is taken as categorical data and the dependent variable is taken as continues data.

3.2.11 CHI - SQUARE TEST

Chi-square is a statistical test used to examine the differences between categorical variables from a random sample in order to judge goodness of fit between expected and observed results. It is also known as χ^2 test.

3.2.12 LOGISTIC REGRESSION

It is used to find the cause-and-effect relationship between independent variables on the dependent variable, in which the dependent variable is binary categorical and independent variable is categorical and continuous variable.

4. DATA ANALYSIS - TABLEAU

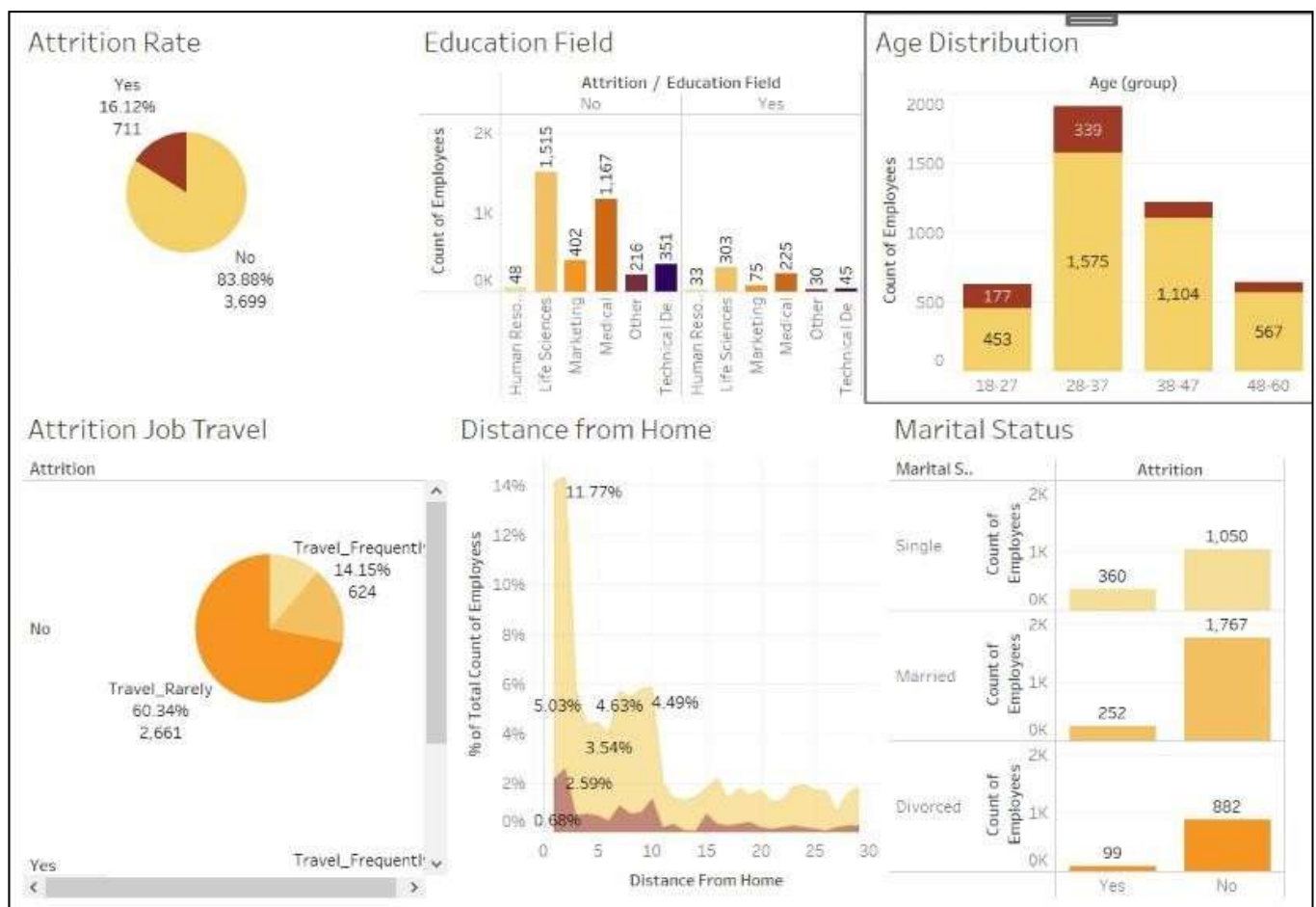
4.1. Employee Attrition Vs Education/Age/Job Travel/Distance from Home/Marital Status:

Using Tableau, the analysis is made between Education, Age, Job Travel, Distance from Home, Marital Status.

Employee Attrition Vs Age/Education/Marital Status is analyzed using a Bar Graph.

Employee Attrition Vs Job Travel is analyzed using Pie Chart.

Employee Attrition Vs Distance from Home is analyzed using Area Chart.

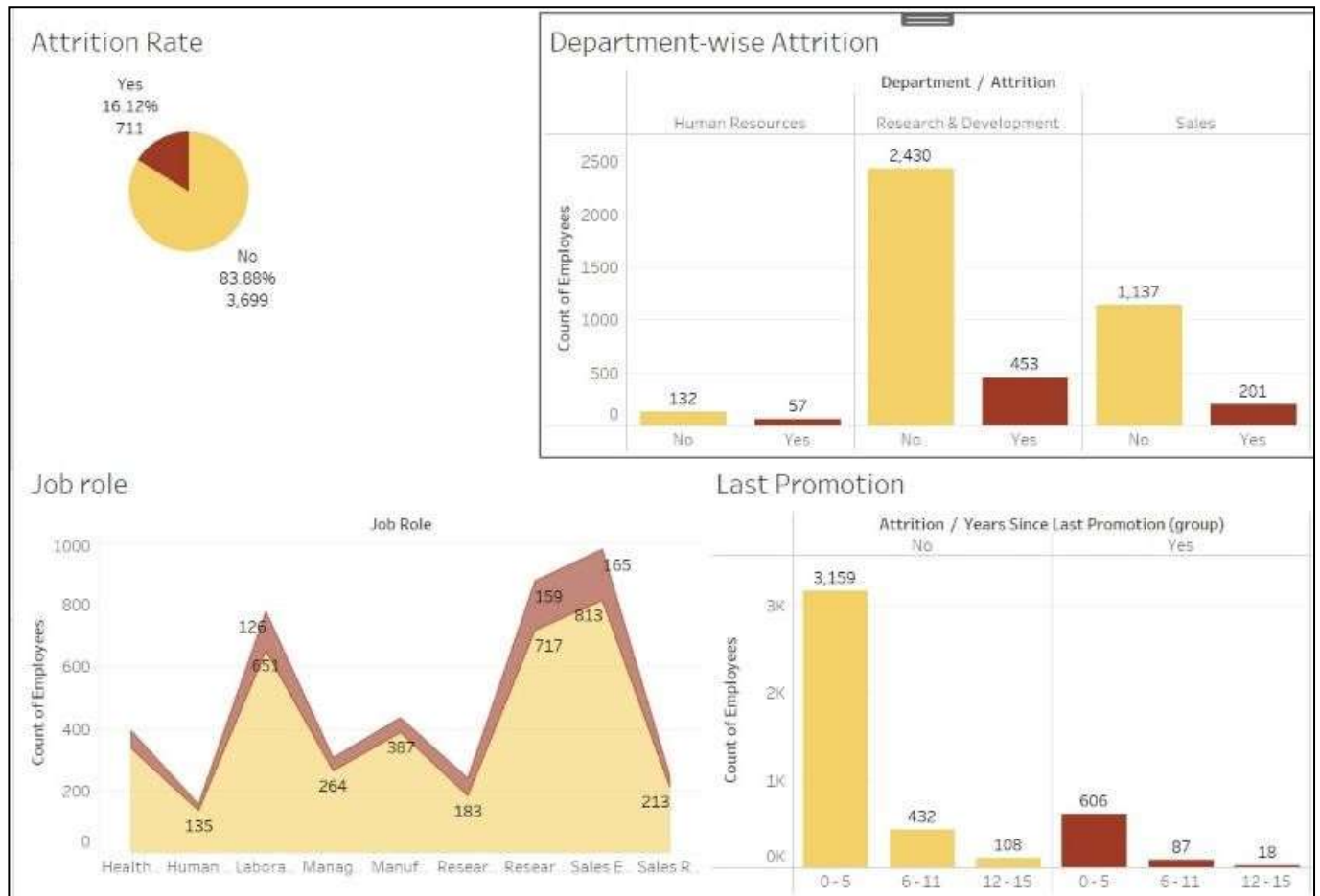


4.1 Dashboard- Employee Attrition Vs Education/Age/Job Travel/Distance from Home/Marital Status

4.2 Factors Affecting Employee Attrition:

Mostly Bar Graph is used to express the factors affecting the Employee Attrition.

- ❖ In Research and Development Department 2430 employees are willing to leave the organization and 453 employees are willing to leave the organization.



4.2 Dashboard - Employee Attrition Vs Department/Job Role/Promotion

- ❖ 3159 employees who got promoted within 0-5 years are not willing to leave the organization but 606 employees even if got promoted within 0-5 years are willing to leave the organization.

5. PROJECT INFERENCE - TABLEAU:

5.1 EMPLOYEE ATTRITION BY GENDER DISTRIBUTION:

Compared to female employees more male employees have left the organization. 60% of the male employees have left the organization while only 40% female employees have left the organization.

5.2 EMPLOYEE ATTRITION BY MARTIAL STATUS:

We can observe that majority of the people who left the organization are single. We can also observe that married people are the majority who do not want to left the organization.



5.1 STORY- Employee Attrition Vs Gender distribution/Marital Status

5.3 EMPLOYEE ATTRITION BY EDUCATION:

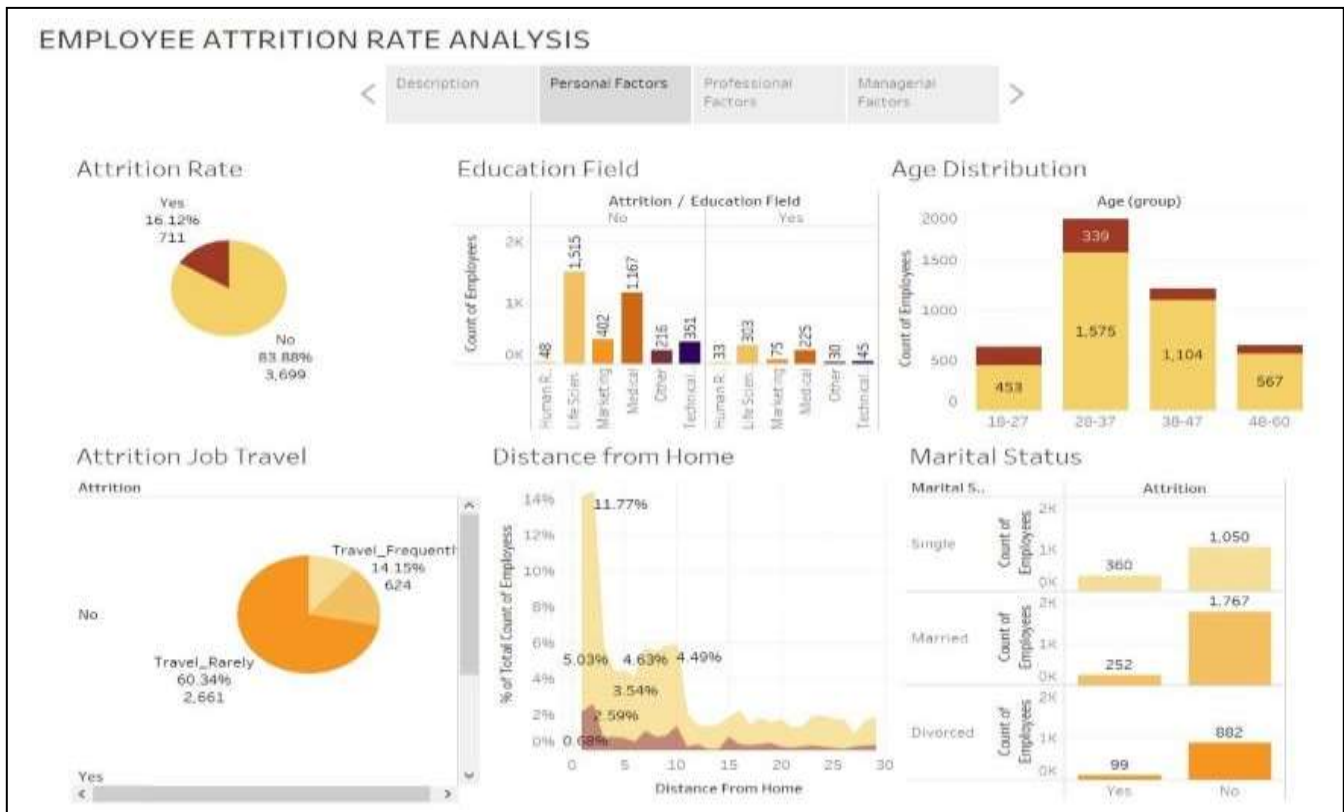
We can clearly see that employees whose field of study is life science are having high attrition rate and are 1515 in number, that is they are the ones mostly willing to leave the organization.

5.4 EMPLOYEE ATTRITION BY AGE:

1575 employees aged between 28-37 are not willing to leave the organization. Employees aged between 28-37 are mostly not willing to leave the organization and so the organization should focus on that age group employees.

5.5 EMPLOYEE ATTRITION BY DISTANCE FROM HOME:

Employees who live within 5kms from the workplace tends to leave the job very rarely and so it is better for the organization to focus on the employees who are residing near to the work place.



5.2 STORY- Employee Attrition Vs Educational field/Age/Distance from home

5.6 EMPLOYEE ATTRITION BY DEPARTMENT:

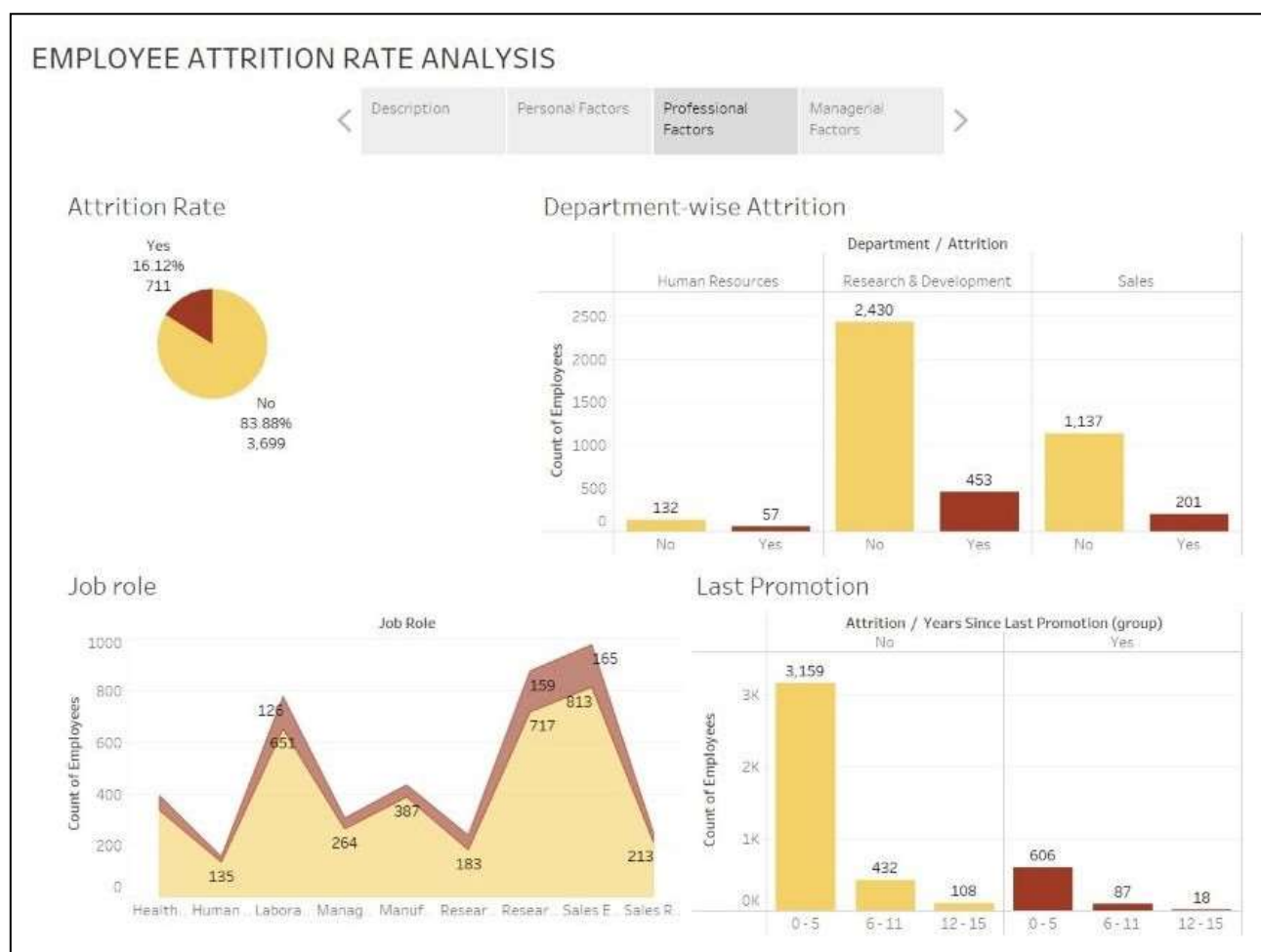
Employees from Research and Development Department who are working in the organization are 2430 and who left the organization are 453. Employees from research and development department are not willing to leave the organization at the same time when compared to other departments they are the ones having high attrition rate. But comparatively people not willing to leave the organization are more.

5.7 EMPLOYEE ATTRITION BY JOB ROLE:

Employees from Sales Department who are working in the organization are 813 and who left the organization are 165. Employees from Sales department are not willing to leave the organization at the same time when compared to other departments they are the ones having high attrition rate. But comparatively people not willing to leave the organization are more.

5.8 EMPLOYEE ATTRITION BY PROMOTION:

Employees who got promotion within the first 5 years are 3159 and are more likely to stay in the organization. So, the organization should focus on promotions and select right candidate.



5.3 STORY- Employee Attrition Vs Department/Job Role/Promotion

5.9 EMPLOYEE ATTRITION BY SALARY HIKE:

2325 employees whose salary increased by 11-15% are not willing to leave the organization. The company should give a salary hike between 11-15% for making the employees to stay in the organization.

5.10 EMPLOYEE ATTRITION BY TRAINING TIME:

282 employees who are trained for 2 years are most willing to leave the organization.

5.11 EMPLOYEE ATTRITION BY YEARS WITH CURRENT MANAGER:

255 employees who did not even work for one year under a manager are most liking to leave the organization.



5.4 STORY- Employee Attrition Vs Years with Current Manager/Training Time Last Year/Percentage Salary Hike

6. DATA ANALYSIS AND INFERENCE – R STUDIO

6.1 DESCRIPTIVE ANALYSIS

```
> myvar<-c('Age','DistanceFromHome','MonthlyIncome','PercentSalaryHike','YearsAtCompany')
> sapply(dataset[myvar],mean)
      Age DistanceFromHome MonthlyIncome PercentSalaryHike YearsAtCompany
36.923810      9.192517      65029.312925      15.209524      7.008163
> sapply(dataset[myvar],sd)
      Age DistanceFromHome MonthlyIncome PercentSalaryHike YearsAtCompany
9.133301      8.105026      47068.888559      3.659108      6.125135
> sapply(dataset[myvar],min)
      Age DistanceFromHome MonthlyIncome PercentSalaryHike YearsAtCompany
18      1      10090      11      0
> sapply(dataset[myvar],max)
      Age DistanceFromHome MonthlyIncome PercentSalaryHike YearsAtCompany
60      29      199990      25      40
```

6.1 Descriptive analysis using R

A descriptive analysis of some attributes like Age, Distance from home, Monthly Income, Percent Salary Hike and Years at Company is done.

Mean, Standard Deviation (SD), Minimum and Maximum of the attributes mentioned above are:

	Age	Distance from home	Monthly Income (in Rs.)	Percent Salary hike	Years at company
Mean	37 years	9 km	Rs. 65029	15.20%	7 years
SD	9.133301	8 km	Rs. 47068.88	3.65%	6.12 years
Min	18 years	1 km	Rs. 10090	11%	0 years
Max	60 years	29 km	Rs. 199990	25%	40 years

6.2 CORRELATION

6.2.1 AGE AND ATTRITION

```
> cor.test(Age,Attrition)

Pearson's product-moment correlation

data: Age and Attrition
t = -10.707, df = 4408, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.1878379 -0.1303018
sample estimates:
cor
-0.159205
```

6.2.1 Correlation- Age and Attrition

A correlation test is taken between two variable Age and Attrition.

The value for the test is **-0.159205** and the value is lesser than zero, thus it can be concluded that Age and Attrition are negatively correlated.

6.2.2 MONTHLY INCOME AND ATTRITION

```
> cor.test(Attrition,MonthlyIncome)

Pearson's product-moment correlation

data: Attrition and MonthlyIncome
t = -2.0709, df = 4408, p-value = 0.03843
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.060636003 -0.001662294
sample estimates:
cor
-0.03117628
```

6.2.2 Correlation- Monthly Income and Attrition

A correlation test is taken between two variable Monthly income and Attrition.

The value for the test is **-0.03117628** and the value is lesser than zero, thus it can be concluded that Age and Attrition are negatively correlated.

6.2.3 DISTANCE FROM HOME AND ATTRITION

```
> cor.test(Attrition,DistanceFromHome)

Pearson's product-moment correlation

data: Attrition and DistanceFromHome
t = -0.64604, df = 4408, p-value = 0.5183
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.03923439 0.01979106
sample estimates:
cor
-0.009730141
```

6.2.3 Correlation- Distance from home and Attrition

A correlation test is taken between two variable Distance From Home and Attrition.

The value for the test is **-0.09730141** and the value is lesser than zero, thus it can be concluded that Distance from home and Attrition are negatively correlated

6.3 CHI-SQUARE

6.3.1 GENDER AND ATTRITION

```
> chitable<-table(Gender,Attrition)
> chitable
      Attrition
Gender    0     1
  0 1494  270
  1 2205  441
> print(chisq.test(chitable))

      Pearson's Chi-squared test with Yates' continuity correction

data:  chitable
X-squared = 1.3499, df = 1, p-value = 0.2453
```

6.3.1 Chi-square- Gender and Attrition

Hypothesis

H_0 -There is no significant association between Gender and Attrition

H_1 - There is significant association between Gender and Attrition

The p-value for the test is **0.2453**. Since, the p-value is greater than 0.05, null hypothesis (H_0) is accepted and alternate Hypothesis (H_1) is rejected.

This means that there is no significant association between Gender and Attrition.

6.3.2 DEPARTMENT AND ATTRITION

```
> chitable<-table(Attrition,Department)
> chitable
      Department
Attrition    1     2     3
  0 2430 1137  132
  1  453  201   57
> print(chisq.test(chitable))

      Pearson's Chi-squared test

data:  chitable
X-squared = 29.09, df = 2, p-value = 4.821e-07
```

6.3.2 Chi-square- Department and Attrition

Hypothesis

H_0 -There is no significant association between Department and Attrition

H_1 - There is significant association between Department and Attrition

The p-value for the test is **4.4821e-07** OR 0.0000004821. Since, the p-value is lesser than 0.05, null hypothesis (H_0) is rejected and alternate Hypothesis (H_1) is accepted.

This means that there is significant association between Department and Attrition.

6.3.3 MARTIAL STATUS AND ATTRITION

```
> chitable<-table(Attrition,MaritalStatus)
> chitable
      MaritalStatus
Attrition    1      2      3
      0  882 1767 1050
      1   99  252  360
> print(chisq.test(chitable))

      Pearson's Chi-squared test

data:  chitable
X-squared = 138.49, df = 2, p-value < 2.2e-16
```

6.3.3 Chi-square- Martial Status and Attrition

Hypothesis

H₀-There is no significant association between Martial status and Attrition

H₁- There is significant association between Martial status and Attrition

The p-value for the test is **2.2e-16** or = 0.000000000000000022. Since, the p-value is lesser than 0.05, null hypothesis (H₀) is rejected and alternate Hypothesis (H₁) is accepted.

This means that there is significant association between Martial Status and Attrition.

6.4 ONE-WAY ANOVA

6.4.1 ATTRITION AND JOB ROLE

```
> oneway=aov(Attrition~JobRole)
> summary(oneway)

      Df Sum Sq Mean Sq F value Pr(>F)
JobRole    1    0.6  0.6192   4.581 0.0324 *
Residuals 4408 595.8  0.1352
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6.4.1 One-way Anova- Attrition and Job Role

Hypothesis

H₀-There is no significant association between Attrition and Job role

H₁- There is significant association between attrition and Job role

The F-value for the test is 0.0324. Since, the f-value is less than 0.05, null hypothesis (H_0) is rejected. Alternate Hypothesis (H_1) is accepted.

This means that there is significant difference between attrition and job role.

6.4.2 ATTRITION AND PERCENTAGE SALARY HIKE

```
> oneway=aov(Attrition~PercentSalar1Hike)
> summary(oneway)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PercentSalar1Hike	1	0.6	0.6312	4.67	0.0307 *
Residuals	4408	595.7	0.1351		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6.4.2 One -way Anova- Attrition and Percentage Salary Hike

Hypothesis

H_0 -There is no significant association between Attrition and Job role

H_1 - There is significant association between attrition and Job role

The F-value for the test is 0.0307. Since, the f-value is less than 0.05, null hypothesis (H_0) is rejected. Alternate Hypothesis (H_1) is accepted.

This means that there is significant difference between attrition and job role.

6.5 TWO WAY ANOVA

6.5.1 ATTRITION, EDUCATION AND DEPARTMENT

```
> twoway=aov(Attrition~Department+Education)
> summary(twoway)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Department	1	1.0	1.0306	7.631	0.00576 **
Education	1	0.1	0.1449	1.073	0.30032
Residuals	4407	595.2	0.1351		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6.5.1 Two-way Anova- Attrition, Education and Department

Hypothesis

H_0 -There is no significant association between Attrition, Education and department

H_1 - There is significant association between attrition, Education and Department.

The F-value for the test is **0.00576**. Since, the f-value is less than 0.05, null hypothesis (H_0) is rejected. Alternate Hypothesis (H_1) is accepted.

This means that there is significant association between attrition, education and department.

6.5.2 ATTRITION, DEPARTMENT AND MARITAL STATUS

```
> twoway=aov(Attrition~Department+MaritalStatus)
> summary(twoway)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Department	1	1.0	1.0306	7.630	0.00577	**
MaritalStatus	1	0.1	0.0693	0.513	0.47393	
Residuals	4407	595.3	0.1351			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

6.5.2 Two-way Anova- Attrition, Department and Marital Status

Hypothesis

H_0 -There is no significant association between Attrition, Department and Marital status

H_1 - There is significant association between Attrition, Department and Marital Status.

The F-value for the test is **0.00577**. Since, the f-value is less than 0.05, null hypothesis (H_0) is rejected. Alternate Hypothesis (H_1) is accepted.

This means that there is significant difference between attrition, Department and marital status.

6.6 LOGISTIC REGRESSION

```
> Attrition<-glm(Attrition~Age+BusinessTravel+Department+DistanceFromHome+Education+EducationField+EmployeeCount+EmployeeID+Gender+Job
level+JobRole+MaritalStatus+MonthlyIncome+NumCompaniesWorked+PercentSalaryHike+StandardHours+StockOptionLevel+TotalWorkingYears+Traini
ngTimesLastYear+YearsAtCompany+YearsSinceLastPromotion+YearsWithCurrManager)
> summary(Attrition)

Call:
glm(formula = Attrition ~ Age + BusinessTravel + Department +
    DistanceFromHome + Education + EducationField + EmployeeCount +
    EmployeeID + Gender + JobLevel + JobRole + MaritalStatus +
    MonthlyIncome + NumCompaniesWorked + PercentSalaryHike +
    StandardHours + StockOptionLevel + TotalWorkingYears + TrainingTimesLastYear +
    YearsAtCompany + YearsSinceLastPromotion + YearsWithCurrManager)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.43160  -0.20779  -0.12729  -0.01029   1.05310

Coefficients: (2 not defined because of singularities)
(Intercept)          2.164e-01  5.559e-02  3.893 0.000101 ***
Age                -4.379e-03  8.206e-04  -5.336 9.97e-08 ***
BusinessTravel     -5.956e-04  8.045e-03  -0.074 0.940967
Department         2.270e-02  9.460e-03  2.404 0.016240 *
DistanceFromHome   -7.311e-05  6.581e-04  -0.111 0.911548
Education          -8.243e-03  5.207e-03  -1.583 0.113512
EducationField     -6.620e-04  4.015e-03  -0.165 0.869067
EmployeeCount      -1.261e-06  4.173e-06  -0.302 0.762568
EmployeeID         -1.473e-02  1.091e-02  1.350 0.176057
Gender             -3.369e-03  4.885e-03  -0.695 0.486988
JobLevel           4.855e-03  2.167e-03  2.240 0.025183 *
JobRole            7.337e-02  7.344e-03  9.991  <. 2e-16 ***
MaritalStatus      -2.221e-07  1.143e-07  -1.957 0.053430 .
MonthlyIncome      1.442e-02  2.371e-02  0.601 1.30e-09 ***
NumCompaniesWorked 1.966e-03  1.460e-03  1.347 0.177941
PercentSalaryHike  -8.120e-03  1.271e-03  -1.296 0.195036
StandardHours      -5.072e-03  1.180e-03  -4.268 2.01e-05 ***
StockOptionLevel   -1.670e-02  4.162e-03  -4.026 5.70e-05 ***
TotalWorkingYears  2.760e-03  1.710e-03  1.607 0.108212
TrainingTimesLastYear 1.097e-02  2.310e-03  4.750 2.54e-07 ***
YearsAtCompany     1.504e-02  2.340e-03  6.429 2.00e-11 ***
YearsSinceLastPromotion 1.504e-02  2.340e-03  6.429 2.00e-11 ***
YearsWithCurrManager 1.504e-02  2.340e-03  6.429 2.00e-11 ***
---
Signif. codes:  0. '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.1242904)

Null deviance: 550.37 on 4409 degrees of freedom
Residual deviance: 545.51 on 4309 degrees of freedom
AIC: 3342.5

Number of Fisher Scoring iterations: 2
```

6.6 Logistic Regression Analysis

Logistic regression is done between various attributes like Age, Business Travel, Department, Distance from home, Education, Education Field, Employee count, Employee ID, Gender, Job level, Job Role, Marital Status, Monthly Income, Number of companies worked, Percentage Salary hike, Standard hours, Stock Option Level, Total Working Years, Training Times Last year, Years at Company, Years in Last Promotion, Years with current Manager and ATTRITION.

Negatively correlated attributes to Attrition- Business Travel, Department, Distance from home, Education, Employee Field, Employee ID, Gender, Job Level, Job Role, Monthly Income, Percentage Salary Hike, Stock Option Level and Years at company.

Positively correlated attributes to attribution- Age, Martial Status, Number of companies Worked, Total Working years, Training times last year, Years since last promotion, and Years with current managers.

7. FINDINGS AND RECOMMENDATIONS

7.1 FINDINGS

Out of a total employee count 4410, 3699 (84%) were retained in the company and 711 (16%) exited the company.

PERSONAL FACTORS

- Distance from Home – When employees lived farther away from the office, attrition rates peaked. Therefore, the rate of employee attrition for this organisation is not necessarily influenced by distance.
- People aged 28 to 37 experienced the greatest attrition (339 employees), followed by those aged 18 to 27.
- Unmarried employees were the ones who left the company most . Then it was followed by married and divorced employees.
- Employee attrition was highest among those with a background in the life sciences and medicine, who claimed that these abilities were not necessary for the job and chose to leave rather than stay in a position where they couldn't fully apply their education.

PROFESSIONAL FACTORS

- Employees that were promoted within the last 5 years were more likely to exit the organization. This might be an indicator of lower-level employees wanting to exit despite being promoted to a slightly higher level.Sales executives constituted the largest amount of exiting employees

MANAGERIAL FACTORS

- Compared to those who received wage increases of 16–25%, employees who received raises of 11–15% were more likely to leave the organisation. Employees were more likely to leave their existing bosses after less than a year of service.

7.2 RECOMMENDATIONS

- Maintaining a low employee attrition rate is important to both employees and employers
- So , basically there are both personal and professional factors that affect attrition rate of the employees
- Some Other suggestion include hiring the right talent for the right jobs.

The inference drawn from the data analysis shows that a number of workplace factors directly affect an employee's decision to remain on the job or depart. However, the company is unable to alter all of these contributing aspects because many of these crucial components are personal and unrelated to the workplace. For instance, the firm is powerless to alter Marital Status, despite the fact that it significantly affects attrition rate. To reduce turnover and promote retention, the company can make decisions about specific factors including the education level of potential hires, how personnel are divided among different departments, how often and how quickly promotions occur, as well as wage increases.

8. CONCLUSION

The objective of this investigation is to assemble data that was gathered from workplace personnel and other personal aspects.

Some of the Major factors that affect the employees attrition rates are :

- Department
- Marital status
- Job role
- Percentage salary hike
- Education

In those there are some which have direct influence over the attrition rates and some are non – influential factors.

Directly influential factors are:

- Marital status
- Education field
- Department
- Years since last promotion.
- Salary hike

Non – influential factors are:

- Distance from home
- Training times last year.

APPENDIX

Since some variable have values in character form so they were converted into numeric form as follow:

- ✓ **Attrition-** Attrition1, Yes =1, No=0
- ✓ **BusinessTravel** - BusinessTravel, “Non-Travel” = 1, “Travel_Rarely” = 2, “Travel_Frequently” =3
- ✓ **Department** -Department1, “Human Resources” = 1, “Research & Development” = 2, “Travel_Frequently” = 3
- ✓ **EducationField** -EducationField1, “Human Resources” = 1, “Life Sciences” =2, “Marketing” = 3, “Medical” = 4, “Other” = 5, “Technical Degree” = 6
- ✓ **Gender** - Gender1, “Female” = 1, “Male” = 2
- ✓ **JobRole** - JobRole1, “Human Resources” = 1, “Healthcare Representative” = 2, “Laboratory Technician” = 3, “Manager” = 4, “Manufacturing Director” = 5, “Research Director” = 6, “Research Scientist” = 7, “Sales Executive” =8, “Sales Representative” = 9
- ✓ **MaritalStatus-** MaritalStatus1 “Divorced” = 1, “Married” =2, “Single” = 3