Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

J-COMPONENT REPORT

TEAM MEMBERS:

Harini S-20BCE1832

FACULTY: Dr.Joshan **Date: 08-04-2023**

BONAFIDE CERTIFICATE

Certified that this project report entitled “Attrition of Employees” is a Bonafide work of Harini S -20BCE1832, who carried out the J-component under my supervision and guidance. The contents of this Project work, in fuller or in parts, have never been submitted to any other Institute or University for award of any degree or diploma and the same is certified.

Dr. Joshan

SCOPE, VIT, Chennai

ACKNOWLEDGEMENT

I wish to express our sincere thanks and deep sense of gratitude to our project guide, Dr.Joshan Professor, School of Computer Science Engineering, for his consistent encouragement and valuable guidance offered to me in a pleasant manner throughout the course of the project work. I express my thanks to our HOD Dr.P.Nithyanandam for his support throughout the course of this project. I also take this opportunity to thank all the faculty of the school for their support and their wisdom imparted to me throughout the course. I thank our parents, family, and friends for bearing with me throughout the course of my project and for the opportunity they provided me in undergoing this course in such a prestigious institution.

Abstract:

Attrition is a growing threat to service-based companies. It is the gradual reduction of a workforce by employees leaving and not being replaced rather than by redundancy. The abstract of this project is to visualise the factors that may be the reason for the attrition and predict the future attrition from the visualised data.

Objective of project:

Attrition is a major risk to service-providing organizations where trained and experienced people are the assets of the company. In order to avoid attrition, we will be looking into some data analysis techniques and visualize every possible outcome using Power BI. The main aim is to identify the factors which influence the attrition of employees.

Methods used in your project:

I have used Power BI to visualise the data with the help of various charts such as stacked-bar chart, pie-chart, etc., With the help of those numbers, we will be able to predict the future attrition rate and find the root cause for the attrition.

Outcome of project:

With the help of Power BI and the csv data file, we will have the current attrition rate and the factor that leads to maximum attrition. By predicting the future attrition rate, we will be able to prevent the gradual decrease in the employee attrition rate.

Scope of project

1. **Introduction**

My aim is to study employee attrition in an organisation, analyse the reasonings and predict future attrition. I used an existing database containing employee information and their attrition rates. After reading the given CSV file I transformed the data to remove null values, create the table, etc. Then after, I performed data visualization in which I created various tables like 'count of age', 'count of the gender of employees', 'Business Travel and attrition', 'dependency of overtime attrition', 'dependency of RelationshipSatisfaction',etc. I also created pie-charts for analysis of various fields like 'education', 'business travel', 'job role', 'employee marital info', etc.

2.Review of literature

A lot of studies have been made on attrition prediction analysis in the literature. The major focus was on predicting employee attrition. Researchers have applied machine learning classification models like logistic regression, random forests, support vector machine, and others to analyse the attributes that impact the attrition rate.

Shashikala S, Sushma Ravindra (2013), articulated that the groundwork in the study is in Indian Industry where attrition is present everywhere not in specific sector. Mainly there are few organizations like IT, Telecom which faces high attrition. The study portrays that compensation does not lead to employee loyalty but also salary/wages below the market therefore what are the techniques that can be used to control attrition.

N. Malati, Dr. Prakash Sharma (2013), stated that the attrition has been associated with high costs be in the form of direct or indirect. Due to many Institutions, the availability of competent and qualified faculty is very difficult and change the institutions frequently. Therefore, the study understands the relationship between direct cost incurred by the institutions and the faculty attrition and the study revealed that direct cost is decreasing with increasing attrition rates.

Rajesh Verma, Aanchal Aggarwal (2012), articulated that the last decade saw an outbreak in the growth and development of the Indian Economy. It was accompanied by the revolution in all fronts and a radical change in the way life insurance business was done. It is expected that Indian Insurance sector will be amongst the top 3 in the world by 2020. To reach up to this level the insurance companies have to take steps to reduce the increasing rate of attrition. As this study says that the attrition rate increasing in the sales force in the Insurance sector, the author made focus on the reasons for attrition among the sales force and how to manage the attrition rate.

Ankita Srivastava, Yogesh Tiwari, Hradesh Kumar (2011), stated that the main focus of the study is to identify the root cause of attrition and analyzing the level of employee motivation, satisfaction and involvement to generate employee sustenance in the organization. They delivered strong recommendations, which will eventually be valuable to the organizations to retain their employees for a long term.

Ms. Shaveta Gupta, Dr. Sukhmani, Ms. HarsimranKaur (2011), articulated that talent is the critical success factor to any organization. The research concentrates on retaining the employees and the cost of attrition. Retaining the employees involves understanding the intrinsic motivators as individuals differ a lot. Attempts to find out approaches and strategies to retain the employees and also the cost pertained to attrition

3.Materials and Methods

3.1 Info about models

The modules involved in this project are 1. Dataset collection 2. Attribute analysis 3. Data pre-processing

1. Dataset collection: The first step of a prediction system involves data collection and splitting the data into training and testing datasets. Employee datasets contain all the necessary attributes of employees to predict employee attrition such as Employee id, Satisfaction Level, Last Evaluation, Number of Projects, Average Monthly Hours, Time spent in Company (Years), Departments, Left, Promotion last 5 years, salary level.

2. Attribute analysis: The attributes present in the dataset are analyzed through charts, stacked bar chart and pair plots through data visualization techniques. The attributes which majorly affect employee attrition are analysed. Attributes like employee ID which have no significance are eliminated in the process.

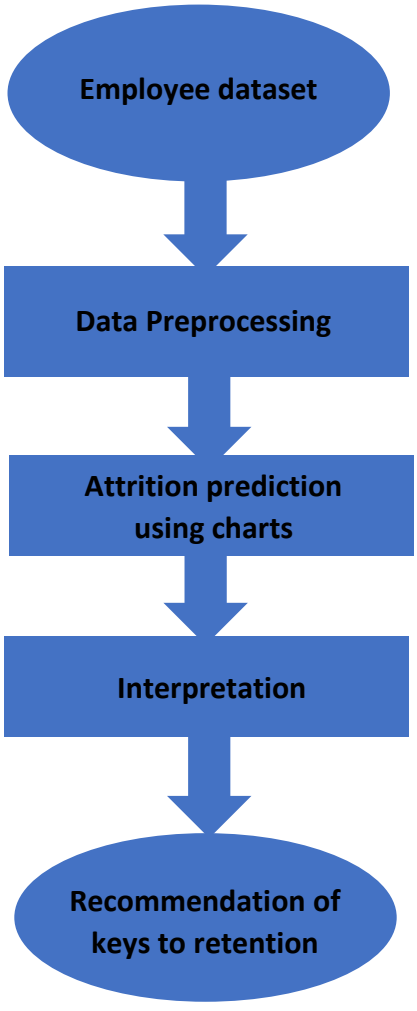
3. Data pre-processing: In visualisation, data pre-processing has a huge role as it pre-processes the data into the required format. This prevents from resulting in misleading outcomes. Data pre-processing deals with redundant data, imbalanced data, missing values, and categorical values. Label encoding is applied to deal with dummy variables.

3.2 dataset

The dataset is available here.

Employee Attrition | Kaggle

3.3 architecture and explanation



EXPLANATION:

The dataset was taken from Kaggle and data pre-processing is basically removing the null values, empty rows and redundant rows and columns. The attrition was then predicted using suitable chart models with different attributes. For some variable, new data has to be created so that we will get the efficient visualisation.

Then the interpretation was made from the chart models and hence I was able to find the main attrition causing factor. Then suitable measures can be taken in order to bring down the attrition rate.

4. Proposed works

4.1 novelty

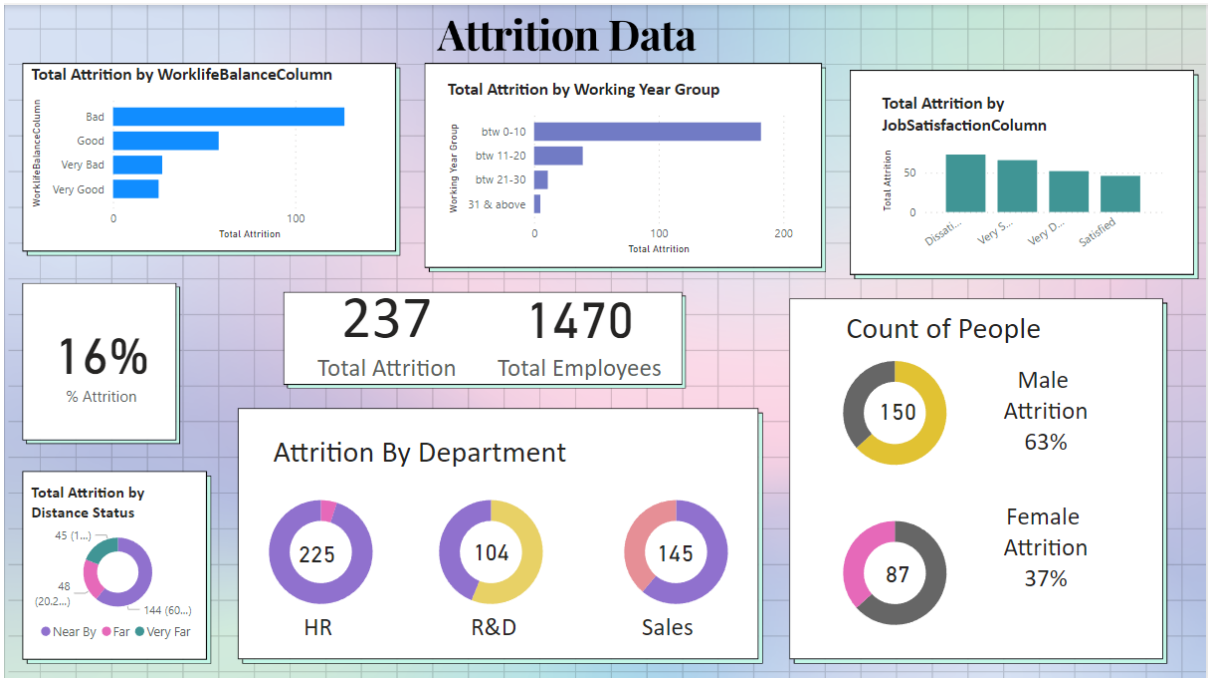
This project is different from the projects available in online since it uses only data visualisation tool (Power BI) to visualise the data and does not use any high-level machine learning algorithms for analysing.

4.2 Project contributions

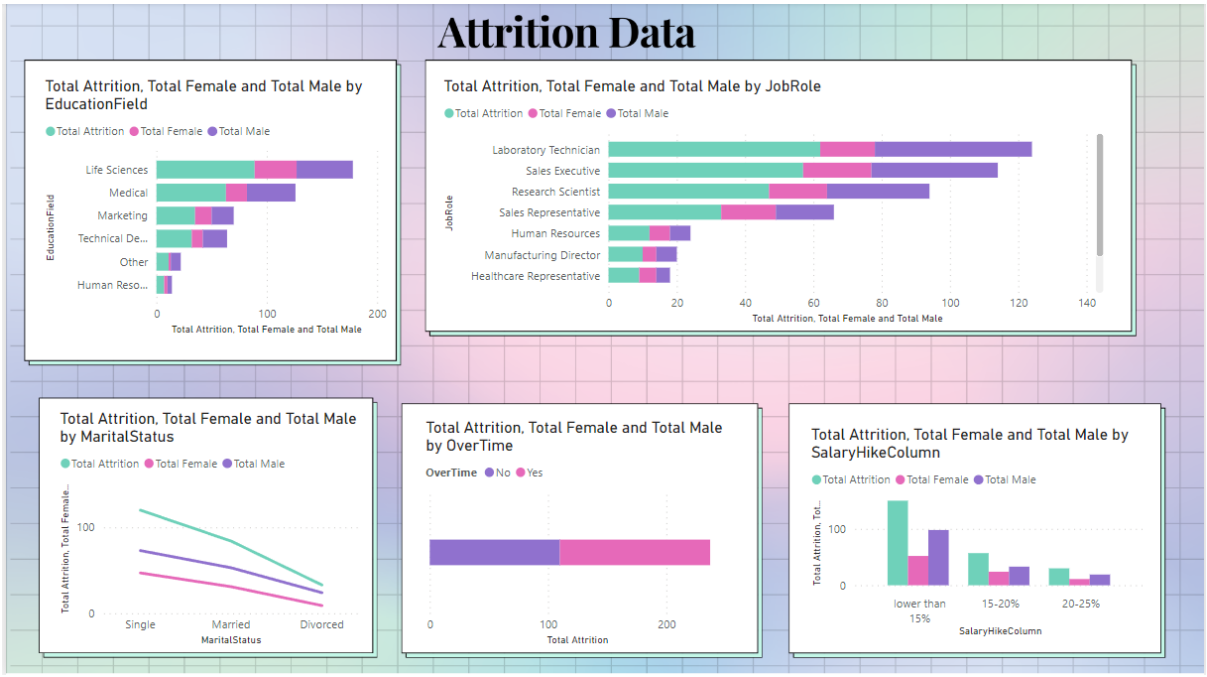
(No team members)

5. Results and discussion

5.1 results



The **total attrition rate** is found to be 16%, that is, out of 1470 **total employees**, 237 has left the office. The various factors that caused this attrition has been analysed.

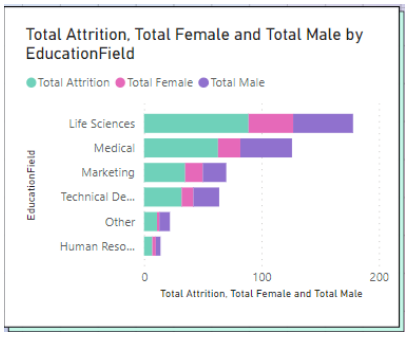


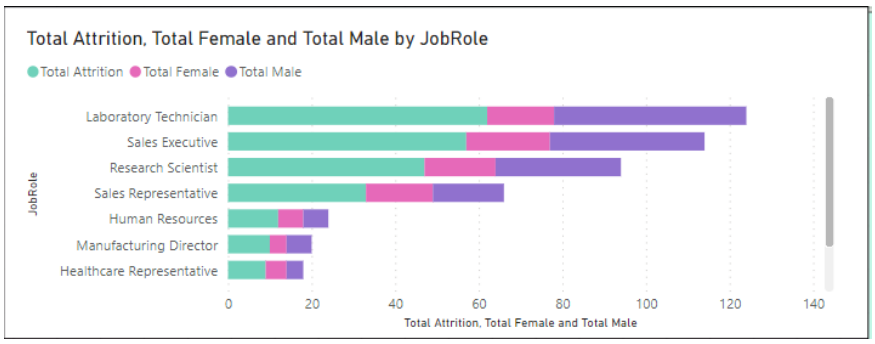
Here the visualisation dealt with the **job role**, **education field**, **Marital Status**, employees who worked **over time** and using **salary hike column**.

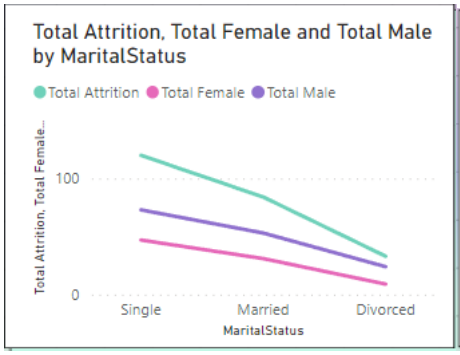
The salary hike column was made using the **percent salary hike**. A certain range was taken and names were given like lower than 15%, 15-20%, 20-25%.

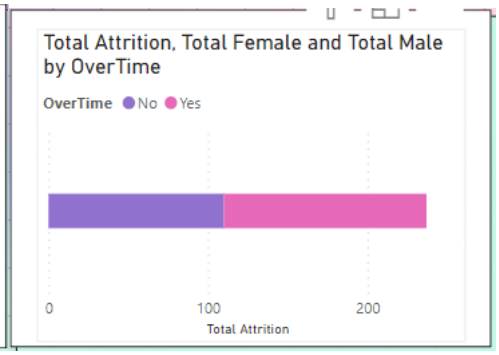
Here the attrition factor was viewed with the help of **monthly income**. Here Monthly Income Column was again made from **Monthly Income data** where the range has been split and the attributes name were given as Average pay, Minimum pay and Highest pay. And the **growth** was also taken in the similar manner from **Working Years Column**.

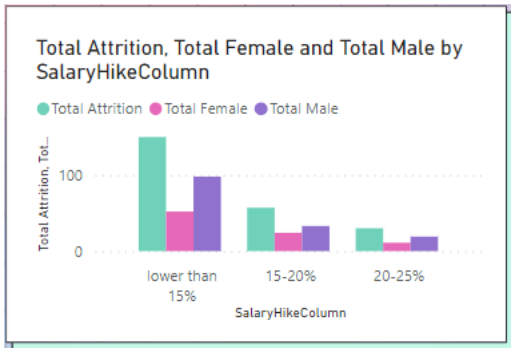
5.2 figures, comparisons tables

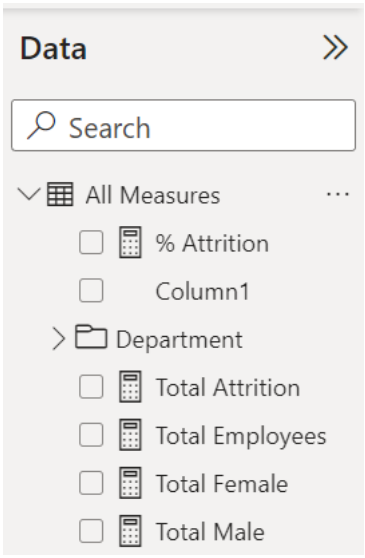












5.3 explanation

Findings of the Study

- It was found that employees **marital status** plays an important role in the attrition rate.
- It was identified that employees prefer a change of job due to ***lack of growth*** opportunities in their job.
- It was noticed that even when good income is provided based on their ***designation***; employees still shift to another organization if the prospects are higher.
- It was analysed that if there is no good understanding and support between superior and subordinate then employees would prefer to change to an organization with ***better environment***.

- It was found that if the ***income*** and the benefits are ***low*** then employees will not retain in the organization rather join an organization where their value are recognized.
- If the ***workload*** is high for the employees, it is noticed that the satisfaction is lowered thus leading to employee leaving the organization.
- It was identified that ***low age*** people do not retain in the organization for a long period of time.
- It was noticed that if there is low income and less job security, the ***stress (work-life balance)*** of the employee increases therefore employee will look for organization with a secure job and high income.

6. Conclusion

In conclusion, the unique contribution of this analysis is that I investigated the attributes which majorly contribute to attrition. The scope of this attrition project is to develop it into a retention project. Attrition is inevitable; it will always prevail; it can only be minimized.

- Intrinsic factors are equally and sometimes more important than extrinsic factors while controlling attrition.
- Effective leadership; to a great extent, may be helpful to control attrition.
- Attrition does not always have a negative impact on the organization.

github link: DataVisualisation_Attrition_Github

7. References

1. (PDF) Employee Attrition Analysis ML model (researchgate.net)

2. Analyze Employee Performance and Attrition Data in Oracle Analytics Cloud

3. Keramati, A.; Jafari-Marandi, R.; Aliannejadi, M.; Ahmadian, I.; Mozaffari, M.; Abbasi, U. Improved churn prediction in telecommunication industry using data mining techniques. Appl. Soft Comput. 2014, 24, 994–1012. [Google Scholar] [CrossRef]

4. Nagadevara, V. Early Prediction of Employee Attrition in Software Companies-Application of Data Mining Techniques. Res. Pract. Hum. Resour. Manag. 2008, 16, 2020–2032. [Google Scholar]

5. Rombaut, E.; Guerry, M.A. Predicting voluntary turnover through Human Resources database analysis. Manag. Res. Rev. 2018, 41, 96–112. [Google Scholar] [CrossRef]

Appendix:

```
% Attrition = DIVIDE([Total Attrition],[Total Employees],0)
```

16%

% Attrition

```
Total Attrition = CALCULATE([Total Employees],FILTER(Data,Data[Attrition]="yes"))
```

237

Total Attrition

```
Total Employees = COUNTROWS(Data)
```

1470

Total Employees

```
Total Female = CALCULATE([Total Attrition],FILTER(Data,Data[Gender]="Female"))
```


```
Total Male = CALCULATE([Total Attrition],FILTER(Data,Data[Gender]="Male"))
```