# Capstone Project - 1

Attrition Analysis of IBM HR Data

# Employee Attrition



- Attrition plays a huge role in the performance of organizations.

- To manage workforce it is necessary to understand relation between attrition and job satisfaction

Sukanya

# Dataset Source

IBM HR Analytics Employee Attrition & Performance dataset from Kaggle

**url:**
https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset/download?datasetVersionNumber=1

# Attrition % form the given dataset

The attrition % from the given set of data is 16.1%, which means for every 100 employees, approximately 16 members leave the organization.



Attrition Proportion in the dataset
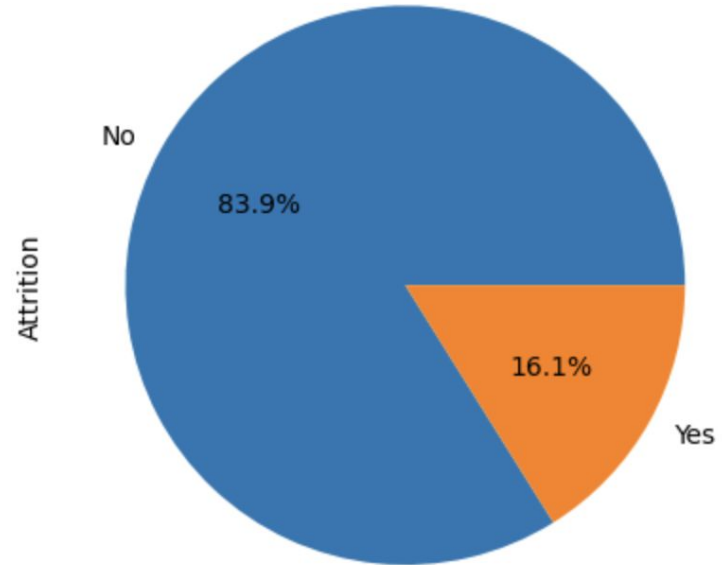
# Basic Analysis

- The dataset has 35 columns and 1470 clean records

- Job satisfaction is ordinal data where value 1 is low, 2 is medium, 3 is high and 4 is very high.

- Attrition column is nominal

- There are no null values

```python
#reading into dataframe
hr_data=pd.read_csv("/content/HR-Employee-Attrition.csv")
hr_data.head(5)
```

```python
[8]  hr_data.shape

     (1470, 35)
```

```python
[7]  hr_data.columns

     Index(['Age', 'Attrition', 'BusinessTravel', 'DailyRate', 'Department',
            'DistanceFromHome', 'Education', 'EducationField', 'EmployeeCount',
            'EmployeeNumber', 'EnvironmentSatisfaction', 'Gender', 'HourlyRate',
            'JobInvolvement', 'JobLevel', 'JobRole', 'JobSatisfaction',
            'MaritalStatus', 'MonthlyIncome', 'MonthlyRate', 'NumCompaniesWorked',
            'Over18', 'OverTime', 'PercentSalaryHike', 'PerformanceRating',
            'RelationshipSatisfaction', 'StandardHours', 'StockOptionLevel',
            'TotalWorkingYears', 'TrainingTimesLastYear', 'WorkLifeBalance',
            'YearsAtCompany', 'YearsInCurrentRole', 'YearsSinceLastPromotion',
            'YearsWithCurrManager'],
           dtype='object')
```

Sukanya

# Why Job Satisfaction

"

"Job satisfaction is a critical determinant in mitigating attrition, as content and fulfilled employees are more likely to remain committed to their organizations. As expressed by Daniel H. Pink in his book 'Drive: The Surprising Truth About What Motivates Us' (Riverhead Books, 2009), when individuals find purpose, autonomy, and mastery in their work, it significantly enhances their satisfaction and reduces the likelihood of attrition."

The statistics

Sukanya

# Hypothesis Testing

**Null Hypothesis(H0):**There is no significant relation between job satisfaction and attrition

**Alternate Hypothesis(Ha):**There is significant relation between job satisfaction and attrition

**Statistical Testing:** Chi-squared test

**Result:** Chi Reject Null hypothesis, as p value is less than 0.05 significance factor
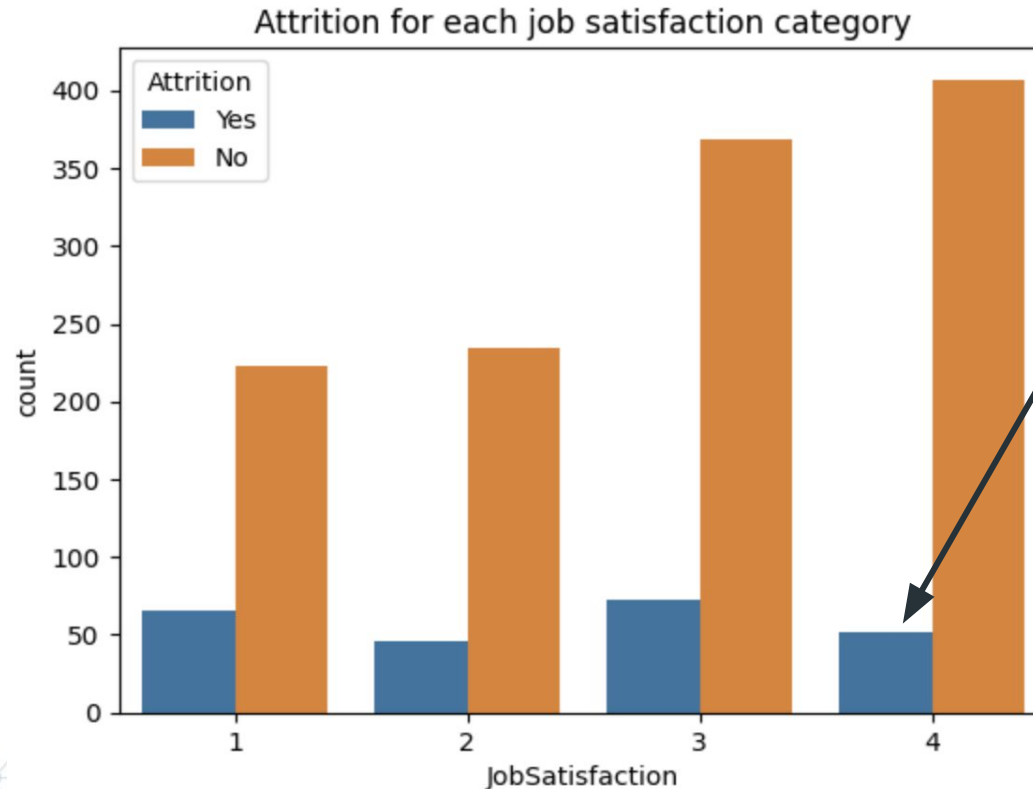
```
#chi square test
from scipy.stats import chi2_contingency

chi_sq_test, p_value, x,y = chi2_contingency(temp_df)

print("Chi square test value", chi_sq_test)
print("P value", p_value)
```

```
Chi square test value 17.505077010348
P value 0.0005563004510387556
```

# Attrition Visualized from the dataset



Attrition for each job satisfaction category

Attrition rate is lower in the high job satisfaction category

The same is inferred by the statistical testing

Sukanya

# Conclusion

- The presence of significant relationship between job satisfaction from the given dataset is backed by our statistical testing.

- This research can be adopted to suit various organizations and corporations

- This correlation indicates the necessity of organizations to involve in employee engagement activities for better employee retention.