IBM Employee Attrition Analysis

Group 17
Steven Chen, Simran Barnwal, Sai Srikar Vaidyula, Yuchao Du, Xiangchen Zhao

Motivation and Objective

- Attrition or **loss of employees** is a huge loss to companies in terms of
  - Training and searching for new employees
  - Loss of good and trained employees

- **Importance:** Analysis on employee attrition could help companies to **intervene on time** and remedy the situation

Goal: Find out factors that may cause employee attrition in a company

Dataset

The IBM dataset contains **1500 employees'** information

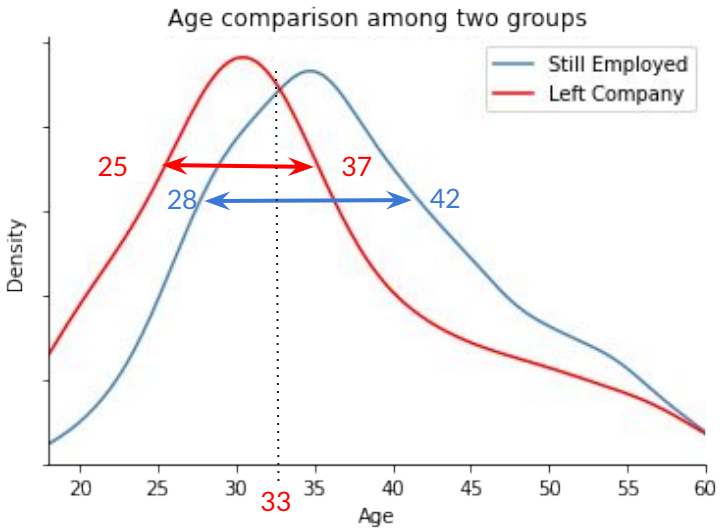
- **Key column**: Attrition (Yes/No)
- **Numerical columns**: Age, Monthly Income, Distance from home, etc.
- **Categorical columns**: Overtime, Job Level, Employee Satisfaction, etc.

	Attrition	Age	DistanceFromHome	Gender	MonthlyIncome	OverTime	JobRole	Department
0	1	41	1	Female	5993	Yes	Sales Executive	Sales
1	0	49	8	Male	5130	No	Research Scientist	Research & Development
2	1	37	2	Male	2090	Yes	Laboratory Technician	Research & Development

Methodology

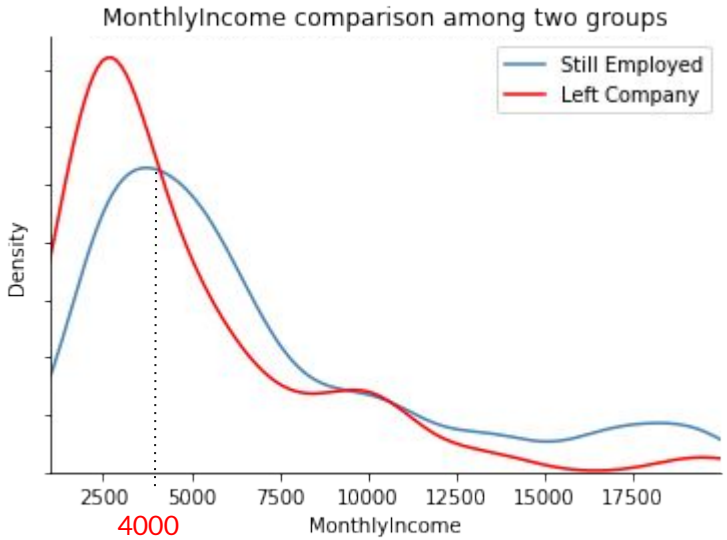
- **Setup hypotheses**
  - Attrition rate is related to **employee's age**; **monthly income**; **distance**.
  - Attrition increases if **income** does not increase along with **work years** as expected.
- **Test the hypotheses by visualizing the graphs**
  - For each variable (Age, Income, etc.), plot their distribution within two groups
  - Are the variable distributions of the two groups(left; remain) same or different?
- **Confirm what we found using statistical tests**
  - Use two sample **t test** and calculate the **p-value**
  - We can confirm the difference between two groups by rejecting the null hypothesis

Does attrition rate negatively relate to Age?



Younger employees are more likely to leave the company!

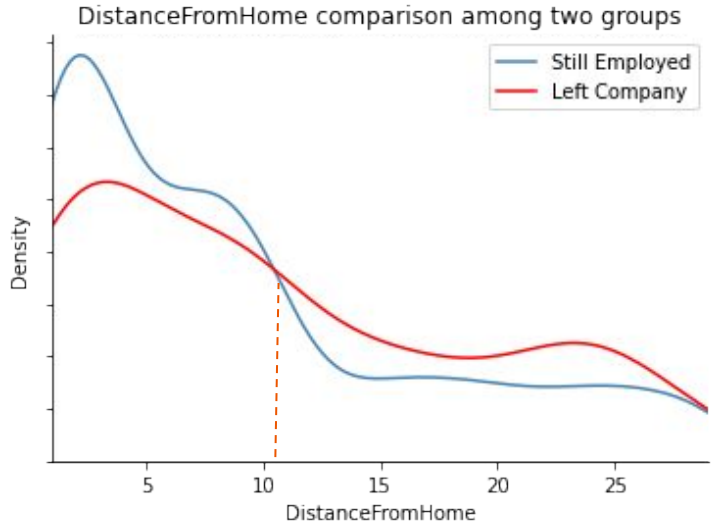
Does attrition rate negatively relate to Monthly Income?



Employees with relatively **low monthly income** are more **likely** to **leave** the company.

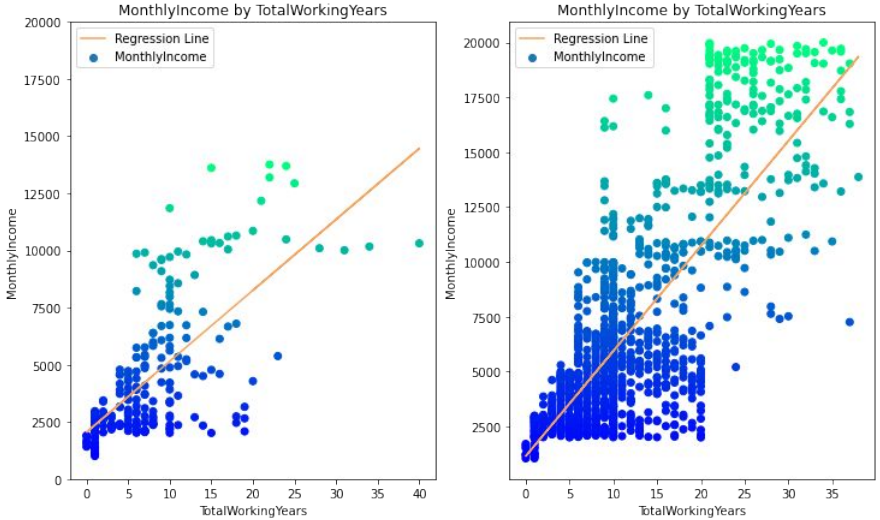
This is very intuitive!

Does attrition rate positively relate to distance from home?



The result is also consistent with our intuition that **long distance from home results in high attrition rates.**

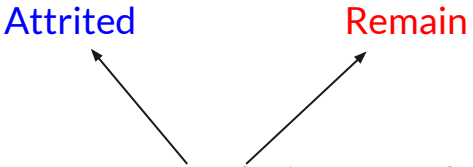
Combine monthly income and total working years



Company should **pay their employee with more work experience** more in order to keep them in company.

Slope: Attrited: $377/year of experience Remain: $479/year of experience

Two Sample T-tests



A two sample t-test is used to test whether the means of two populations are the same.

- **Significance Level**: 0.05
- **Null Hypothesis:** Age(Monthly Income/Distance) has **SAME** means among two groups (two-sides)
- **Alternative Hypothesis:** Age(Monthly Income/Distance) has the **DIFFERENT** means among two groups

Validate what we have visualized

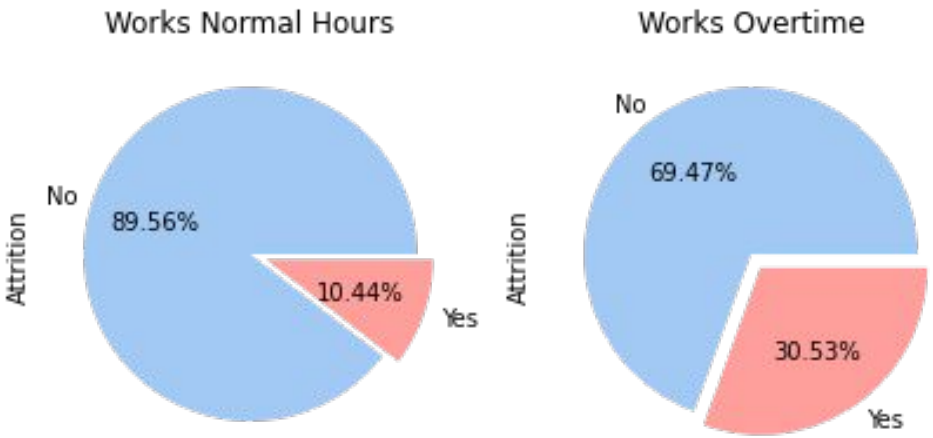
	p_value
Age	0.000
MonthlyIncome	0.000
DistanceFromHome	0.002
IncomeGrowthPerWorkyear	0.000

We **reject** the Null hypothesis when:

P value < significance level 0.05

Thus, these factors are indeed important for attrition

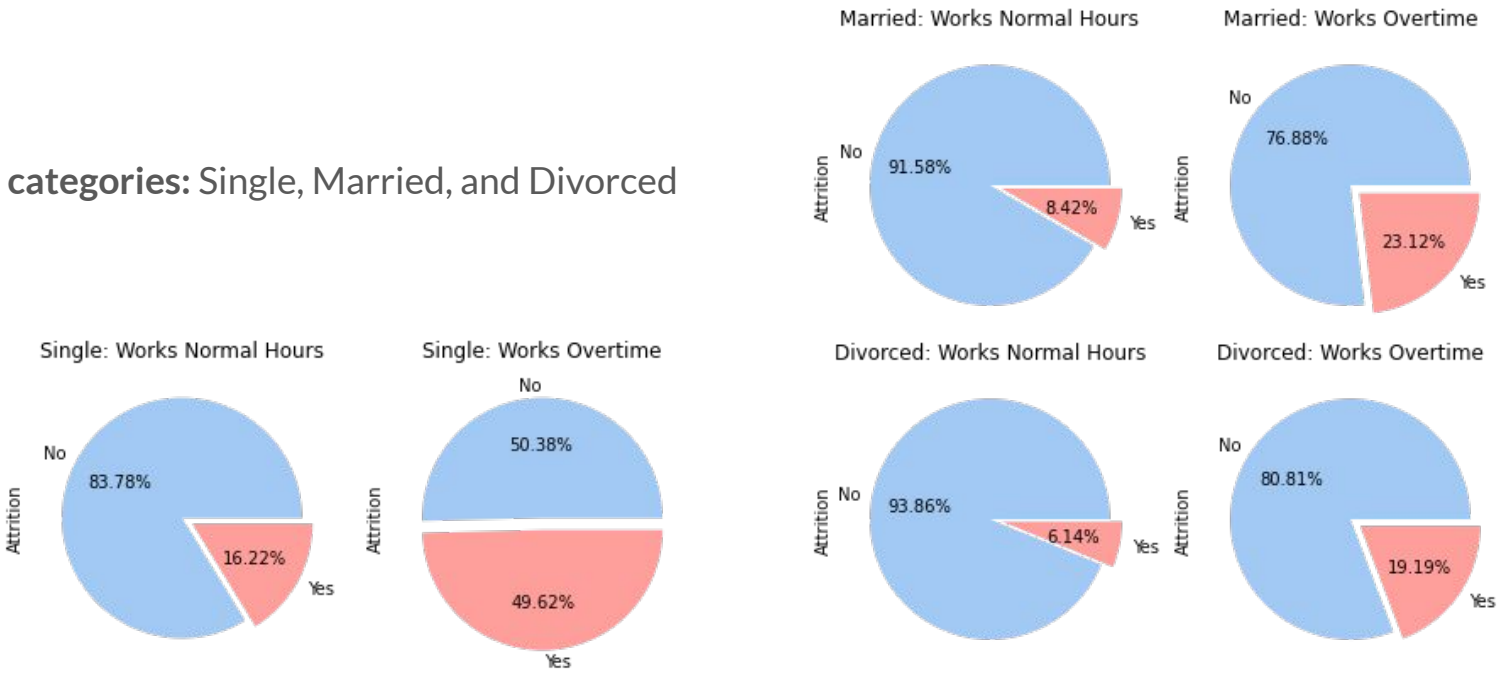
Overtime vs Attrition



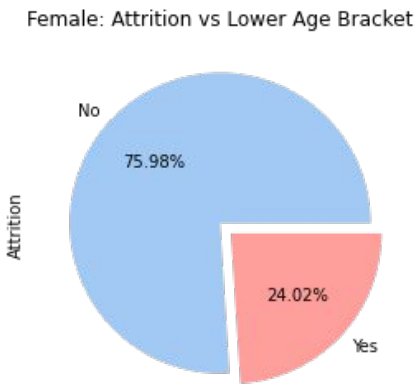
Employees working overtime are **three times** more likely to leave the company!

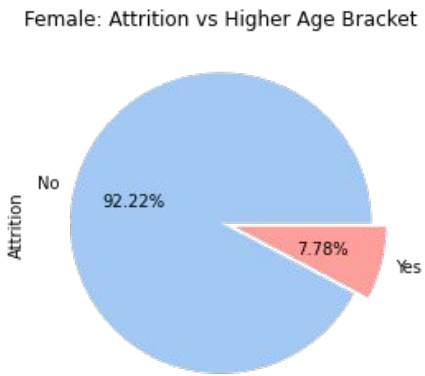
Deep into Overtime vs Attrition w.r.t. Marital Status

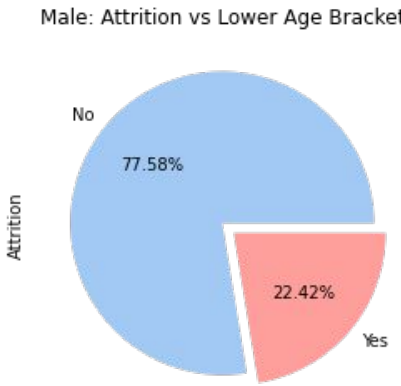
3 categories: Single, Married, and Divorced

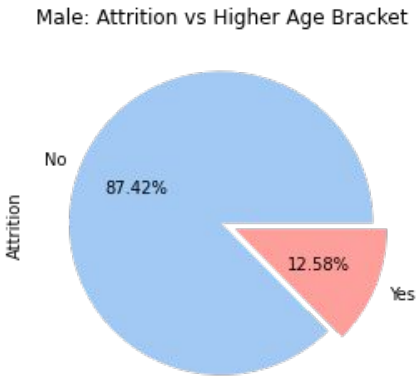


Gender Analysis: Age vs Attrition









Female employees

Male employees

Attrition in male employee is greater in the higher age bracket than female employees

Comments and Conclusion

- Working conditions are a major factor in attrition (almost 3 times!)
- In the higher age bracket, female employees tend to stay in the same company compared to their male counterparts
- Income, Age and distance from work are related to attrition as expected from common knowledge
- We can estimate a threshold ratio of income and work experience in order to keep their employee in the company

References

https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset

https://resources.workable.com/tutorial/effective-exit-interview