IBM HR Analytics – Employee Attrition Analysis

This project focuses on **Employee Attrition Prediction & Exploratory Data Analysis (EDA)** using the **IBM HR Dataset**. The goal is to identify key factors influencing employee turnover and derive data-driven insights that can help organizations improve retention strategies.

Project Overview

Employee attrition is a major concern for companies. In this project, we explored:

- What factors contribute to employee attrition?
- How salary, job role, age, experience, work-life balance affect retention.
- Which features have the strongest correlation with attrition?
- Visual insights to understand patterns & trends in employee turnover.

Tools & Libraries Used

Library	**Usage**
Python	Core programming
Pandas	Data cleaning & manipulation
NumPy	Numerical operations
Matplotlib	Data visualization
Seaborn	Statistical plots
Scikit-Learn (if used)	Preprocessing / ML model

Dataset

Dataset used: **IBM HR Analytics Employee Attrition Dataset**
Source: Kaggle / IBM HR Data

Example Features:

1. Age,
2. Gender,
3. Department,
4. JobRole,
5. MonthlyIncome,

6. JobSatisfaction,
7. Education,
8. DistanceFromHome,
9. Attrition,
10. Overtime,
11. WorkLifeBalance, etc.

Data Preprocessing

✓ Handled missing values & duplicates
✓ Converted categorical data to measurable form
✓ Feature analysis & statistical summary
✓ Outlier detection using boxplots

Exploratory Data Analysis (EDA)

Key visualizations include:

- Attrition by Age group
- Salary vs Attrition relationship
- Job role-wise attrition count
- Heatmap correlation between features
- Work-life balance & overtime effect
- Department vs Attrition comparison

Insights Obtained (Example)

- Employees doing overtime have higher attrition rate.
- Attrition rate is high in lower salary brackets.
- Job satisfaction directly affects employee turnover.
- Younger employees tend to switch more often.
- R&D department shows notable attrition compared to others.

Conclusion

This project demonstrates a complete **end-to-end HR Analytics workflow** involving data cleaning, visualization, and interpretation of attrition trends. Helpful for organizations to improve retention strategies and employee satisfaction.

Connect / Contribution

Feel free to contribute!
If you like this project, ⭐ star the repo 😊