Data Forecasting

Answer Predictive Questions

For HR Data Analysis

DATA DYNAMOS

Harnessing the Power of Data to Drive Innovation

The dataset appears to have a wealth of HR-related columns. Key fields include:

- **Demographics: Gender, Age, Ethnicity, MaritalStatus**
- **Job Details: Department, DistanceFromHome, BusinessTravel, HireDate**
- **Performance and Tenure: PerformanceRating, YearsAtCompany, YearsInMostRecentRole, YearsSinceLastPromotion, YearsWithCurrManager**
- **Compensation: Salary, StockOptionLevel**
- **Other: OverTime, Attrition**

Predictive Questions for Analysis

Based on this structure, Here are some predictive questions tailored to HR data analysis that you might consider based on the data you have:

Predictive Analysis:

1 .Can we predict the likelihood of attrition based on employee attributes?

2 .Are there any leading indicators of high-performing employees?

Why It Matters:

- **Proactive Decision-Making:** Predicting attrition or identifying potential high performers allows the company to act before problems escalate.
- **Strategic Planning:** Aligns HR strategies with business goals by forecasting workforce trends and needs.

1) Employee Turnover

Which employees are likely to leave the organization based on factors like job satisfaction, education level, or performance rating?

To predict which employees are likely to leave the organization based on factors like job satisfaction, education level, or performance rating, we can use a **logistic regression model** or any other classification model. However, since the dataset provided does not explicitly include columns like "Job Satisfaction" or "Performance Rating," we will need to make some assumptions or use the available columns (e.g., Attrition, Education, YearsAtCompany, OverTime, etc.) to build a predictive model.

EmployeeID	FirstName	LastName	Gender	Age	predicted attrition probability
C25C-C434	Jillayne	Goshawke	Female	35	63.9%
7805-8A7C	Agna	Grinnov	Non-Binary	28	63.9%
9E8E-F2FF	Dix	Presley	Female	24	63.9%
F877-3BF9	Cozmo	Duncklee	Non-Binary	21	63.9%
FD95-9EEB	Sasha	Incogna	Male	30	63.9%
DEBE-1622	Bartolemo	Balma	Male	28	63.6%
95D8-5DC7	Udall	Dowson	Male	48	63.6%
B34E-95FF	Samuele	Creek	Male	25	63.6%

The model has an overall accuracy of **66.8%**, but it's struggling to correctly identify employees who leave (low recall for class **1**). This suggests more factors may influence attrition beyond job satisfaction, education level, and performance ratings.

2) Tenure Prediction

How long is an employee expected to stay at the company?

To estimate how long an employee is expected to stay at the company , analyze the **YearsAtCompany** column from the Employee dataset. calculate the average tenure and also use regression modeling to predict expected tenure based on various factors.

Employee Tenure Insights:

- **Average tenure:** 4.56 years
- **Median tenure:** 4.0 years
- **Predicted expected tenure (based on regression model):** 5.73 years

This suggests that, on average, employees stay at the company for about **4.5 to 5.7 years**, depending on job satisfaction, education level, and performance ratings.

3) Attrition Prediction

What is the likelihood that an employee will leave the company based on their data (e.g., age, job role, years of service, salary, etc.)?

To predict the likelihood of an employee leaving based on various factors (e.g., age, job role, years of service, salary, etc.), I'll:

1. **Expand the dataset** to include more relevant features (age, job role, salary, years at company, etc.).
2. **Train a predictive model** (Random Forest or Logistic Regression) to estimate attrition probability.
3. **Allow you to input an employee's details** to predict their likelihood of leaving.

Attrition Prediction Model Performance:

- **Accuracy:** 97.6%
- **Precision (No Attrition):** 98.9%
- **Precision (Attrition):** 95.1%
- **Recall (No Attrition):** 97.6%
- **Recall (Attrition):** 97.7%

This means our model can **reliably predict** if an employee is likely to leave based on their profile.

4) Performance Prediction

What are the key factors that predict high or low employee performance ratings?

To determine the key factors that predict **high or low employee performance ratings**:

1. **Identify performance-related features** (e.g., job satisfaction, education level, experience, manager rating, etc.).
2. **Train a predictive model** (Random Forest or Linear Regression) to analyze which factors have the most influence.
3. **Extract feature importance** to rank the most significant predictors of employee performance.

Key Factors Predicting Employee Performance (Manager Rating)

SelfRating **(34.4%)** → Employees who rate themselves highly tend to receive better manager ratings.

Salary **(20.9%)** → Higher salaries are linked to better performance ratings, possibly due to motivation or job level.

Age **(13.2%)** → Older employees tend to have more experience, positively impacting their ratings.

Job Satisfaction **(10.4%)** → Employees satisfied with their jobs perform better.

Years at Company **(8.7%)** → Longer tenure contributes to higher performance, possibly due to experience.

Job Role **(6.8%)** → Some roles receive systematically higher ratings than others.

Education Level **(5.4%)** → Higher education has a minor impact on ratings.

5) Job Satisfaction Prediction

a) Which attributes (e.g., education level, performance, tenure) have the greatest impact on employee satisfaction levels?

To determine which attributes impact employee satisfaction levels:

1. Use "JobSatisfaction" as the target variable.
2. Analyze key predictors like education level, performance, tenure, salary, and age.
3. Train a Random Forest model to identify the most important factors.

Key Factors Affecting Employee Satisfaction

1. **Salary (30.5%)** → Higher salaries strongly correlate with higher job satisfaction.
2. **Age (18.9%)** → Older employees tend to report higher satisfaction, possibly due to stability.
3. **Years at Company (12.3%)** → Longer tenure contributes to higher satisfaction, indicating loyalty or familiarity.
4. **Manager Rating (11.5%)** → Employees with supportive managers tend to be more satisfied.
5. **Job Role (9.7%)** → Some roles naturally have higher satisfaction than others.
6. **Self Rating (9.0%)** → Employees confident in their performance report higher satisfaction.
7. **Education Level (8.1%)** → Education has the least impact on satisfaction.

b) What is the level of job satisfaction of an employee based on factors like travel, salary, management, etc.?

The SatisfiedLevel sheet provides a mapping of SatisfactionID to different levels of job satisfaction (e.g., Very Dissatisfied to Very Satisfied).

Key Findings on Job Satisfaction:

- **Salary** has a very weak positive correlation (**0.0053**) with job satisfaction.
  This suggests salary alone does not strongly impact satisfaction.

- **Years with Current Manager** (**0.0011**) and **Years at Company** (**0.00087**) show almost no correlation, indicating tenure and manager relationships may not significantly influence satisfaction.

- **Stock Option Level** has a weak negative correlation (**-0.0032**), meaning stock benefits have little effect on job satisfaction.

6) Promotion Readiness

Can we identify employees who are most likely to succeed if promoted, based on their past performance and other attributes? (What is the likelihood of an employee getting a promotion in the future based on their performance and years of service?)

To determine the likelihood of an employee getting promoted based on performance and years of service

- **Performance Rating** (from the PerformanceRating sheet)
- **YearsSinceLastPromotion** (from the Employee sheet)
- **YearsAtCompany** (experience)
- **JobRole** (to check if certain roles have higher promotion trends)

Insights on Promotion Likelihood:

- **Years at Company (0.84 correlation)**: The strongest predictor—employees with more experience are more likely to have been promoted before.
- **Manager Rating (0.02 correlation)**: A weak positive correlation, meaning manager ratings have little impact on the time since the last promotion.
- **Work-Life Balance (0.016 correlation)** and **Self-Rating (0.013 correlation)**: Minimal influence on promotions.
- **Job Satisfaction (-0.001 correlation)**: No significant relationship, meaning satisfaction levels don't predict promotions.

7) Retention Strategies

Which groups of employees are at the highest risk of leaving, and what strategies could reduce their likelihood of leaving?

To identify employees at high risk of leaving, analyze attrition trends based on key factors like:

- **OverTime** (workload)
- **Salary** (compensation)
- **YearsAtCompany** (tenure)
- **JobSatisfaction** (happiness at work)
- **WorkLifeBalance** (stress levels)

Groups at Highest Risk of Leaving:

- **Low Salary:** Employees who left had an average salary of **$81,956**, much lower than those who stayed (**$125,856**).
- **Short Tenure:** Those who left had **~2.5 years** at the company, compared to **7.4 years** for those who stayed.
- **Job Satisfaction & Work-Life Balance:** Nearly the same for both groups, meaning **these factors alone don't strongly predict attrition**.

Strategies to Reduce Attrition:

1. **Competitive Compensation**: Increase salaries for newer employees to match industry standards.
2. **Career Growth Opportunities**: Since shorter tenure employees leave more often, provide **fast-track promotions** or **clear career paths**.
3. **Better Onboarding & Engagement**: Support new hires in the **first 2-3 years** to ensure they feel valued.

8) Impact of Overtime on Attrition

Does working overtime increase the likelihood of an employee leaving the company?

analyze whether working overtime increases the likelihood of an employee leaving the company by checking the correlation between **OverTime** and **Attrition.**

Impact of Overtime on Employee Attrition:

- **Employees who do NOT work overtime:**
  - 89.56% stay with the company.
  - 10.43% leave.
- **Employees who work overtime**:
  - 69.47% stay.
  - 30.53% leave.

Employees who work overtime are about 3 times more likely to leave the company compared to those who don't (30.53% vs. 10.43%). This suggests that excessive overtime may contribute to employee burnout and attrition.

9) Salary Prediction Based on Job Role

What is the expected salary of an employee based on their department, experience, and job role?

analyze how department, experience (years at the company), and job role affect an employee's salary. Let me first check the relevant data.

Expected Salary Insights:

By Department:

- **Human Resources**: ~$119,699
- **Sales**: ~$119,118
- **Technology**: ~$109,655

By Job Role (Top & Bottom 3):

- **Highest Paid Roles**:
  - HR Manager: ~$449,331
  - Analytics Manager: ~$346,484
  - Manager: ~$317,531
- **Lowest Paid Roles**:
  - Recruiter: ~$37,648
  - Sales Representative: ~$40,656
  - Software Engineer: ~$51,967

By Experience (Years at Company):

- Salary peaks around 8–10 years (~$134,665 – $145,605).
- New employees (0 years) start around $91,418.

DataDynamos