User Story - HR Dashboard

As an HR manager, I want a comprehensive dashboard to analyze human resources data, providing both summary views for high-level insights and detailed employee records for in-depth analysis

Summary View

The summary view should be divided into three main sections: Overview, Demographics, and Income Analysis

Overview

The Overview section should provide a snapshot of the overall HR metrics, including:

- Display the total number of hired employees, active employees, and terminated employees.
- Visualize the total number of hired and terminated employees over the years.
- Present a breakdown of total employees by department and job titles.
- Compare total employees between headquarters (HQ) and branches (New York is the HQ)
- Show the distribution of employees by city and state.

Demographics

The Demographics section should offer insights into the composition of the workforce, including:

- Present the gender ratio in the company.
- Visualize the distribution of employees across age groups and education levels.
- Show the total number of employees within each age group.
- Show the total number of employees within each education level.
- Present the correlation between employees's educational backgrounds and their performance ratings.

Income

The income analysis section should focus on salary-related metrics, including:

- Compare salaries across different education levels for both genders to identify any discrepancies or patterns.
- Present how the age correlate with the salary for employees in each department.

Employee Records View

- Provide a comprehensive list of all employees with necessary information such as name, department, position, gender, age, education, and salary.
- Users should be able to filter the list based on any of the available columns.

Data Generation

Generate python script to generate a realistic dataset of 8950 records for human resources.

The dataset should include the following attributes:

1. Employee ID: A unique identifier.
2. First Name: Randomly generated.
3. Last Name: Randomly generated.
4. Gender: Randomly chosen with a 46% probability for ‘Female’ and a 54% probability for ‘Male’.
5. State and City: Randomly assigned from a predefined list of states and their cities.
6. 6.Hire Date: Randomly generated with custom probabilities for each year from 2015 to 2024.
7. 7.Department: Randomly chosen from a list of departments with specified probabilities.
8. Job Title: Randomly selected based on the department, with specific probabilities for each job title within the department.
9. Education Level: Determined based on the job title, chosen from a predefined mapping of job titles to education levels.
10. Performance Rating: Randomly selected from ‘Excellent’, ‘Good’, ‘Satisfactory’, ‘Needs Improvement’ with specified probabilities.
11. Overtime: Randomly chosen with a 30% probability for ‘Yes’ and a 70% probability for ‘No’.
12. Salary: Generated based on the department and job title, within specific ranges.
13. Birth Date: Generated based on age group distribution and job title requirements, ensuring consistency with the hire date.
14. Termination Date: Assigned to a subset of employees (11.2% of the total) with specific probabilities for each year from 2015 to 2024, ensuring the termination date is at least 6 months after the hire date.
15. Adjusted Salary: Calculated based on gender, education level, and age, applying specific multipliers and increments.
16. Be sure to structure the code cleanly, using functions where appropriate, and include comments to explain each step of the process