

Human Resources Dataset Analysis

Analytical Workflow

1. Data Understanding

- Identify all available files and the content of each one.
- Determine the type of each variable (numerical – categorical – textual).
- Define potential relationships between files (e.g., EmployeeID as a key).

2. Initial Data Exploration

- Display the first few rows of each file.
- Check the number of rows and columns.
- Inspect data types for each column.
- Detect missing or abnormal values.

3. Data Merging

- Identify common columns across datasets.
- Combine the files into one unified DataFrame for analysis.
- Ensure no duplicates appear after merging.

4. Data Cleaning

- Handle missing values (remove or impute).
- Fix or standardize column names if needed.
- Convert incorrect data types (e.g., strings to numbers).
- Remove duplicate rows.

5. Data Preparation

- Create new columns if needed for analysis.

- Encode categorical values into numerical form when necessary.
- Address outliers if they affect accuracy.

◆ 6. Descriptive Analysis

- Calculate mean, median, min, max for key variables.
- Analyze distribution of education levels.
- Analyze performance ratings and satisfaction levels.
- Count employees by category or group.

◆ 7. Analytical Questions

- Formulate key questions to explore relationships in the data.
- Investigate correlations between variables (e.g., satisfaction vs performance).

◆ 8. Data Visualization

- Create charts to illustrate key findings.
- Compare variables visually.
- Use bar charts, pie charts, histograms, or line charts as needed.

◆ 9. Insights & Findings

- Extract the most important results from the analysis.
- Highlight key observations and detected patterns.

◆ 10. Recommendations

- Suggest improvements or actions based on the findings.
- Connect insights to the project goals.



Question Analysis

- What is the relationship between education level and performance rating?
- Do employees with higher performance scores have higher satisfaction levels?
- How is satisfaction distributed across different education levels?
- Is there a correlation between final rating level and job satisfaction?
- What is the average performance score for each education level?
- Does education level influence the likelihood of receiving a high rating?
- What percentage of employees falls into each satisfaction category?
- How does performance vary across departments or age groups (if available)?
- Is there a link between performance rating and promotion or retention?
- What are the most influential factors affecting job satisfaction?
- Do employees with higher education achieve better ratings?
- Are there noticeable differences based on years of experience (if available)?
- What is the overall distribution of performance ratings in the organization?
- Does satisfaction level vary across different rating categories?
- How do education, performance, and satisfaction interact together?