Green-Destinations-Attrition-Rate-Analysis

Calculate the **attrition rate** from a dataset using pandas. It includes loading the data, exploring it, handling any missing values, and finally calculating the attrition rate.

Attrition Rate Analysis

This code will:

1. Load the dataset.
2. Explore it to ensure the necessary column (`Attrition`) exists.
3. Handle any missing values.
4. Calculate the attrition rate.
5. Save the summary as a CSV file for documentation.

Below is the code and explanation of each line:

```
import pandas as pd

# Step 1: Load the dataset

file_path = '/Users/monty/Downloads/greendestination.csv'  # Update the file path

try:

    data = pd.read_csv(file_path)

except FileNotFoundError:

    raise FileNotFoundError(f"File not found. Please check the file path: {file_path}")



# Step 2: Data Exploration

# Display the first few rows to understand the structure of the data

print("Dataset Overview:")

print(data.head())



# Check for missing values

print("\nMissing Values:")
```

```
print(data.isnull().sum())


# Ensure the 'Attrition' column exists
if 'Attrition' not in data.columns:
    raise ValueError("The dataset does not contain an 'Attrition' column.")


# Step 3: Handle Missing Data (if any)
# If there are missing values in the Attrition column, drop them (optional)
data = data.dropna(subset=['Attrition'])


# Step 4: Calculate Attrition Rate
# Total number of employees
total_employees = len(data)


# Number of employees with Attrition = 'Yes'
attrition_count = data[data['Attrition'] == 'Yes'].shape[0]


# Attrition rate formula
attrition_rate = (attrition_count / total_employees) * 100


# Step 5: Display Results
print("\nTotal Employees:", total_employees)
print("Employees with Attrition (Yes):", attrition_count)
print(f"Attrition Rate: {attrition_rate:.2f}%")
```

```
# Step 6: Save Results to a File (Optional)
result = {
    "Total Employees": total_employees,
    "Employees with Attrition (Yes)": attrition_count,
    "Attrition Rate (%)": round(attrition_rate, 2)
}

# Convert results to a DataFrame and save as a CSV file
result_df = pd.DataFrame([result])
output_path = 'attrition_rate_summary.csv'
result_df.to_csv(output_path, index=False)
print(f"\nAttrition summary saved to {output_path}")
```

Conclusion:

The dataset contains a column called `Attrition`, which indicates whether an employee has left the organization ("Yes") or not ("No"). To calculate the **% attrition rate**, we can:

1. Count the total number of employees.
2. Count the number of employees with `Attrition = "Yes"`.
3. Calculate the percentage as:

   $$\text{Attrition Rate} = (\text{Number of Attritions} / \text{Total Number of Employees}) \times 100$$

The attrition rate is approximately **16.12%**. This means that about 16.12% of employees have left the organization based on the given dataset. Let me know if you'd like further analysis or a breakdown by department, age group, or other factors!