

System Requirements

Specification Index

For

Employee Data Analysis and Management using Pandas

(Topic: Data Import and Export)

Version 1.0

Employee Data Analysis Console

Project Abstract

This Employee Data Analysis Console is a Python application designed to perform various analysis tasks on employee data stored in CSV format. The application uses the Pandas library to load, filter, and analyze employee data efficiently. The primary functionalities of this system include viewing the first few records, summarizing data types and column names, identifying the employee with the highest age, filtering employees by department, categorizing employees into age groups, and exporting updated data into a new CSV file. This application is crucial for businesses to analyze their workforce, streamline decision-making processes, and better manage employee-related statistics.

Business Requirements:

- **Loading employee data:** The system should be able to load employee data from a CSV file.
- **Viewing employee data:** The system should display the first few records for quick inspection.
- **Employee Analysis:** The system should be able to analyze employee attributes like age, department, and job roles.
- **Data Modification:** The system should allow for adding new columns and exporting modified data.

Constraints:

Input Requirements

- **Employee Data:**
 - The employee data must be stored in a CSV file.
 - The CSV file should have columns such as **Name**, **Age**, **Department**, and other relevant employee attributes.

- The file must be accessible through the provided file path.

Data Operations:

- **Displaying First Rows:**
 - The system must return the first 5 rows from the loaded DataFrame.
- **Displaying DataFrame Information:**
 - The system must return column names and data types from the loaded DataFrame.
- **Finding Highest Age Employee:**
 - The system must identify and return the employee with the highest age.
- **Filtering Sales Department:**
 - The system must filter employees by the **Sales** department and return the relevant subset.
- **Adding Age Category:**
 - The system must categorize employees into different age groups: 'Young', 'Mid-aged', and 'Old'.
 - The system must add an 'Age Category' column based on age range bins.

Output Constraints:

- 1. Data Display:**
 - The system must display the first 5 rows of the DataFrame.
 - It must show DataFrame column names and data types.
- 2. Employee Analysis:**

- The system must show the employee with the highest age.
- The system must return a filtered list of employees working in the **Sales** department.

3. Modified Data Export:

- The system must export the updated DataFrame with the new "Age Category" column to a new CSV file.

Required Output Format:

- Display the first 5 rows of the employee data.
- Display the column names and data types.
- Show the name of the employee with the highest age.
- Show the filtered employee list for the Sales department.
- Export the updated employee data to a new CSV file.

Template Code Structure:

1. Initialization Function:

- `__init__(self, file_path)`
 - Load employee data into a Pandas DataFrame.

2. Display Functions:

- `display_head()`
 - Display the first 5 rows of the DataFrame.
- `dataframe_info()`
 - Display DataFrame column names and data types.

3. Analysis Functions:

- `highest_age_employee()`
 - Identify the employee with the highest age.
- `filter_sales_department()`
 - Filter and display employees in the Sales department.
- `add_age_category()`
 - Add an 'Age Category' column based on age range.

4. Export Function:

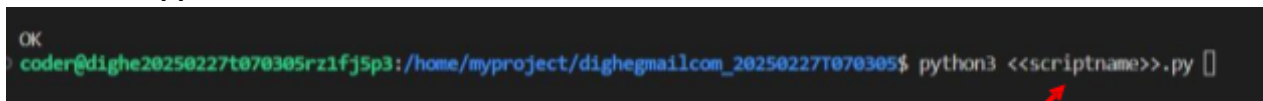
- `export_updated_csv(output_file="updated_employee_data.csv")`
 - Export the updated employee data to a new CSV file.

Execution Steps to Follow:

- All actions like build, compile, running application, running test cases will be through Command Terminal.
- To open the command terminal the test takers, need to go to Application menu (Three horizontal lines at left top) -> Terminal -> New Terminal
- This editor Auto Saves the code
- If you want to exit(logout) and continue the coding later anytime (using Save & Exit option on Assessment Landing Page) then you need to use **CTRL+Shift+B** -command compulsorily on code IDE. This will push or save the updated contents in the internal git/repository. Else the code will not be available in the next login.
- These are time bound assessments the timer would stop if you logout and while logging in back using the same credentials the timer would resume from the same time it was stopped from the previous logout.
- To setup environment:
You can run the application without importing any packages
- To launch application:
python3 mainclass.py
- To run Test cases:
python3 -m unittest
- Before Final Submission also, you need to use **CTRL+Shift+B** - command compulsorily on code IDE, before final submission as well. This will push or save the updated contents in the internal git/repository, and will be used to evaluate the code quality.

Screen shot to run the program


To run the application



A screenshot of a terminal window with a dark background. The prompt is 'coder@dighe20250227t070305rz1fj5p3: /home/myproject/dighegmailcom_20250227T070305\$'. The command 'python3 <<scriptname>>.py' is entered. A red arrow points from the text 'python3 mainclass.py' below to the '<<scriptname>>.py' part of the command in the terminal.

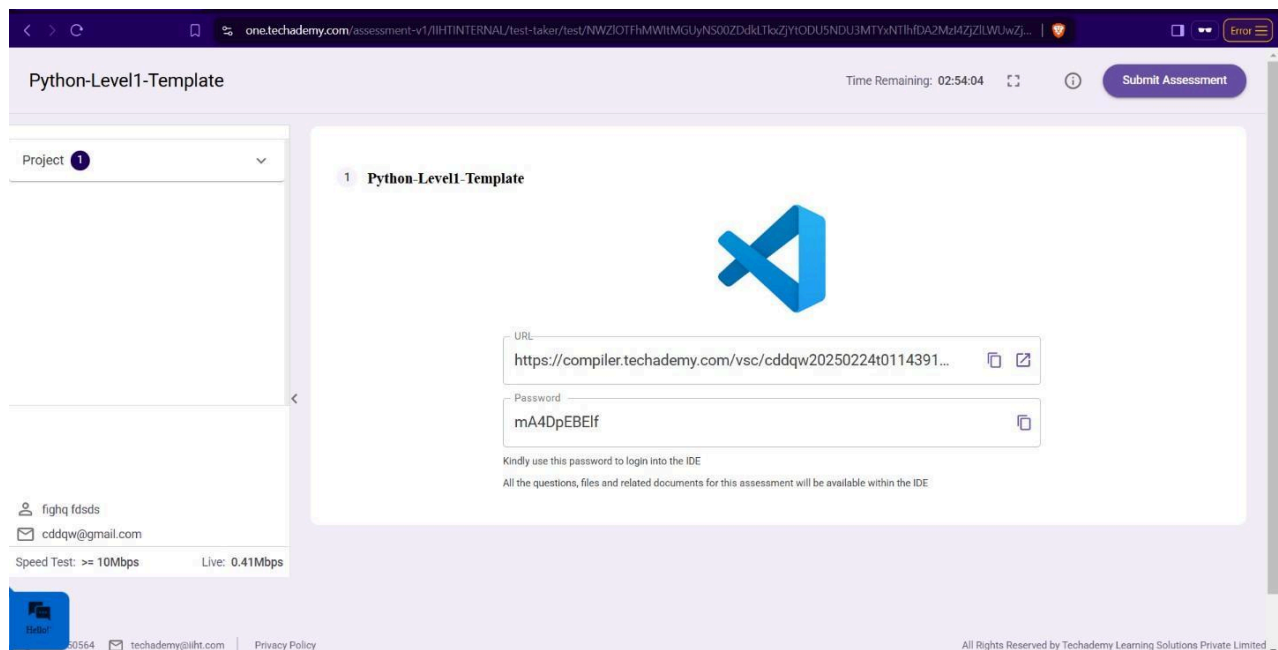
python3 mainclass.py

```
● coder@dighe20250227t070305rz1fj5p3:/home/myproject/dighegmailcom_20250227T070305$ python3 -m unittest
TestBoundary = Passed
.TestExceptional = Passed
.TestCalculateTotalDonations = Failed
.TestCalculateTotalStockValue = Failed
.TestCheckFrankWhiteDonated = Failed
```



To run the testcase

`python3 -m unittest`



- **Once you are done with development and ready with submission, you may navigate to the previous tab and submit the workspace. It is mandatory to click on “Submit Assessment” after you are done with code.**