**TRƯỜNG CAO ĐẲNG FPT POLYTECHNIC**

**🕮**

**Text

Description automatically generated**

**LAB 1**

MÔN HỌC : Lập Trình Pandas

MÃ MÔN : DAT 205

Báo Cáo Bài Tập môn Học

**Sinh viên thực hiện :**  *Trần Gia Lạc*

**MSSV :** *Ps31064*

**Giảng viên hướng dẫn :** Văn Công Khanh

**Mục Lục**

[MỤC TIÊU: 3](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191195)

[PHẦN I 3](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191196)

[Bài 1 (2 điểm) 3](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191197)

[Bài 2 (2 điểm) 5](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191198)

[PHẦN II 6](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191199)

[Bài 3 (2 điểm) 6](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191200)

[Bài 4 (2 điểm) 7](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191201)

[Bài 5: 8](file:///H:\Kho%20dữ%20liệu\PJ\PS31064_Trần%20Gia%20Lạc_%20DP18301_lab7.docx#_Toc157191202)

# MỤC TIÊU:

# PHẦN I

## Bài 1 (2 điểm)

- Tạo một Series từ một danh sách các số nguyên (ví dụ: [1, 3, 5, 7, 9]).

- Truy cập phần tử đầu tiên, phần tử cuối cùng và một phần tử ở vị trí bất kỳ trong Series.

- Thay đổi một phần tử trong Series.

- Tính tổng của tất cả các phần tử trong Series.

**Bài làm:**

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Tạo một series từ danh sách số nguyên  
   my\_series = pd.Series([1, 3, 5, 7, 9]) : sử dụng thư viện pandas để tạo
2. Truy cập phần tử đầu tiên  
   first\_element = my\_series[0]: tương tự như numpy, lấy mảng cần truy cập và đưa index
3. Truy cập phần tử cuối cùng  
   last\_element = my\_series[len(my\_series) - 1]: tương tự như ‘2)’.   
   => **len(my\_series) – 1** là chọn phần tử cuối cùng bằng cách tính chiều dài của mảng và trừ đi 1 vì điểm bắt đầu để truy cập phần tử là 0
4. Truy cập một phần tử ở vị trí bất kỳ (ví dụ: vị trí thứ 2)  
   random\_index = pd.Series(np.random.randint(0, len(my\_series)))  
   number\_index = my\_series[random\_index]  
   => đầu tiên **(np.random.randint(0, len(my\_series))** là chọn ngẫu nhiên 1 index trong mảng sau đó dùng **my\_series[random\_index]** để truy cập phần tử theo index đó
5. Thay đổi một phần tử trong Series (ví dụ: thay đổi phần tử thứ 3 thành 10)  
   my\_series[2] = 10: tương tự như numpy thay thế phần tử bằng index
6. Tính tổng của tất cả các phần tử trong Series  
   total\_sum = my\_series.sum(): dùng hàm tính tổng **sum()** để tính tổng các phần tử trong series

## Bài 2 (2 điểm)

- Tạo một DataFrame từ một từ điển có cấu trúc (ví dụ: {'Tên': ['Alice', 'Bob', 'Charlie'], 'Tuổi': [25, 30, 35]}).

- Truy cập một cột cụ thể bằng cách sử dụng tên cột.

- Truy cập một hàng cụ thể bằng cách sử dụng chỉ số hàng.

- Thêm một cột mới vào DataFrame.

- Tính trung bình của một cột cụ thể trong DataFrame.

**Bài làm:**

A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated

1. Tạo DataFrame từ từ điển có cấu trúc  
   data = {'Name': ['Alice', 'Bob', 'Charlie'], 'Age': [25, 30, 35]}: tạo **dictionary**  
   df = pd.DataFrame(data): tạo **dataframe** từ **dictionary**
2. Truy cập một cột cụ thể bằng tên cột  
   column\_name = 'Name' : đặt biến là tên cột trong **dictionary** ban đầu  
   specific\_column = df[column\_name]: tương tự như truy cập phần tử trong mảng thì lần này nó sẽ xuất ra nguyên cột **‘Name’**
3. Truy cập một hàng cụ thể bằng chỉ số hàng  
   row\_index = 2 : biến chứa index để chọn cột  
   specific\_row = df[row\_index-1:row\_index]: **[row\_index-1:row\_index]** đây là phép cắt trích ra hàng thứ 2
4. Thêm một cột mới vào DataFrame  
   ew\_column = [160, 170, 180] : tạo mảng dữ liệu  
   df['Length'] = new\_column : tạo ra cột **‘Length’** với dữ liệu lấy từ mảng trên
5. Tính trung bình của một cột cụ thể trong DataFrame  
   average\_age = df['Age'].mean(): tính trung bình cột cụ thể **‘Age’**

# PHẦN II

## Bài 3 (2 điểm)

- Tạo một DataFrame đơn giản chứa thông tin như tên và điểm số của một số học sinh.

- In ra tất cả các hàng trong DataFrame.

- Tính tổng điểm của tất cả học sinh.

- Tìm điểm số cao nhất và thấp nhất trong DataFrame.

- Tính điểm trung bình của các học sinh.

**Bài làm:** A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

1. Tạo DataFrame chứa thông tin của học sinh   
   data = {'Tên': ['Alice', 'Bob', 'Charlie', 'David'],  
    'Điểm số': [85, 90, 75, 80]}  
   df = pd.DataFrame(data)  
   print("DataFrame:")  
   => tạo dictionary chứa thông tin học sinh in ra
2. Tính tổng điểm của tất cả học sinh  
   total\_score = df['Điểm số'].sum(): dùng hàm **sum()** để tính tổng điểm
3. Tìm điểm số cao nhất và thấp nhất trong DataFrame  
   max\_score = df['Điểm số'].max()  
   min\_score = df['Điểm số'].min()  
   => dùng hàm **min()** và **max()** để tìm điểm cao nhất và thấp nhất
4. Tính điểm trung bình của các học sinh  
   average\_score = df['Điểm số'].mean(): dùng hàm **mean()** để tính điểm trung bình

## Bài 4 (2 điểm)

**Bài làm:**