**Lab 9 - Đáp án Câu hỏi thử thách**

**Câu 1.** Dữ liệu trong thư viện Pandas được tổ chức theo các class hay cấu trúc dữ liệu? Nêu đặc trưng của từng cấu trúc dữ liệu.

Dữ liệu trong Pandas được tổ chức theo DataFrame và Series.

DataFrame lưu dữ liệu dạng bảng, bao gồm tên cột, index, và dữ liệu trong bảng. Các dữ liệu trong bảng còn được chia thành các dòng, cột và ô.

Series lưu trữ dữ liệu dạng danh sách, bao gồm index và dữ liệu. Series là kiểu dữ liệu trả về khi truy vấn riêng một dòng hoặc một cột trong DataFrame.

**Câu 2.** Làm thế nào một text file như CSV có thể lưu trữ dữ liệu dạng bảng?

Các text file phân tách dữ liệu trong từng ô và từng dòng của bảng bằng các ký hiệu đã định trước. Đối với CSV, dữ liệu trong mỗi dòng của bảng được lưu trong một dòng của text file; mỗi ô trong cùng một dòng được tách nhau bởi dấu phẩy (,)

**Câu 3.** Dữ liệu trong DataFrame có thể được khởi tạo bằng:

1. Một dictionary, với key là số thứ tự từng dòng và values là dữ liệu trong từng dòng của bảng.
2. Một dictionary, với key là index từng dòng và values là dữ liệu trong từng dòng của bảng.
3. Một list chứa list, với mỗi list bên trong là dữ liệu của một dòng trong bảng.
4. Một list chứa list và một list chứa tên các cột. Mỗi list bên trong list chứa list là dữ liệu của một dòng trong bảng.

**Câu 4.** Chọn phát biểu đúng về kiểu dữ liệu trong Pandas.

1. Dữ liệu trong DataFrame có thể thuộc nhiều kiểu dữ liệu khác nhau.
2. Dữ liệu trong một cột của DataFrame có thể thuộc nhiều kiểu dữ liệu khác nhau.
3. Dữ liệu trong một Series có thể thuộc nhiều kiểu dữ liệu khác nhau.
4. Tất cả đều đúng.

**Câu 5.** Khi lưu DataFrame vào file, Pandas mặc định:

1. Lưu theo định dạng .csv
2. Lưu theo định dạng .xls
3. Lưu kèm index của DataFrame
4. Không lưu index của DataFrame