Phần n : Thống kê dữ liệu

* 1. Thống kê dữ liệu cơ bản

Thống kê dữ liệu bao gồm như : tập dữ liệu có bao nhiêu hàng và cột; trung bình; phân vị 25, 50, 75; phương sai và độ lệch chuẩn; giá trị lớn nhất nhỏ nhất của một thuộc tính trong tập dữ liệu; tần xuất giá trị của thuộc tính;...Để thực hiện những thống kê như trên thì có thể sử dụng Python kết hợp các thư viện như Pandas, Numpy,..

Mã nguồn thống kê cơ bản: [PTDL/thongkecoban.py at main · BAOTIN2004/PTDL · GitHub](https://github.com/BAOTIN2004/PTDL/blob/main/Do_An/thongkecoban.py)

Phần n+1: Ước lượng và kiểm định

* 1. Bài toán kiểm định

Có phát biểu nhận định rằng có sự khác nhau giữa lương trung bình của ngành “” với ngành “Data Engineer”. Hãy kiểm định nhận định đó mới mức ý nghĩa 5% ?

Mã nguồn giải quyết bài toán kiểm định: [PTDL/kiemdinh.py at main · BAOTIN2004/PTDL · GitHub](https://github.com/BAOTIN2004/PTDL/blob/main/Do_An/kiemdinh.py)