TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÀI TẬP LỚN/ĐỒ ÁN CUỐI KÌ MÔN ………………….**

**…tên đề tài…**

*Người hướng dẫn*: **TS NGUYỄN VĂN A**

*Người thực hiện*: **NGUYỄN THỊ B – MSSV**

**TRẦN VĂN C – MSSV**

Lớp **: 10050301**

Khoá  **: 17**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2014**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÀI TẬP LỚN/ĐỒ ÁN CUỐI KÌ MÔN ………………….**

**…tên đề tài…**

Người hướng dẫn: **TS NGUYỄN VĂN A**

Người thực hiện: **NGUYỄN THỊ B**

**TRẦN VĂN C**

Lớp **: 10050301**

Khoá  **: 16**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2014**

LỜI CẢM ƠN

Đây là phần tác giả **tự viết** ngắn gọn, thể hiện sự biết ơn của mình đối với những người đã giúp mình hoàn thành Luận văn/Luận án. Tuyệt đối không sao chép theo mẫu những “lời cảm ơn” đã có.

**ĐỒ ÁN ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là sản phẩm đồ án của riêng tôi / chúng tôi và được sự hướng dẫn của TS Nguyễn Văn A;. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình.** Trường đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày tháng năm*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Văn B*

*Trần Văn C*

PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN

**Phần xác nhận của GV hướng dẫn**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

**Phần đánh giá của GV chấm bài**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

TÓM TẮT

Trình bày tóm tắt vấn đề nghiên cứu, các hướng tiếp cận, cách giải quyết vấn đề và một số kết quả đạt được, những phát hiện cơ bản trong vòng 1 -2 trang.

MỤC LỤC

[LỜI CẢM ƠN i](#_Toc387692905)

[PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN i](#_Toc387692906)

[TÓM TẮT i](#_Toc387692907)

[MỤC LỤC 1](#_Toc387692908)

[DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ 1](#_Toc387692909)

[CHƯƠNG 1 – MỞ ĐẦU 1](#_Toc387692910)

[1.1 Tiểu mục cấp 1 1](#_Toc387692911)

[1.1.1 Tiểu mục cấp 2 1](#_Toc387692912)

[1.1.1.1 Tiểu mục cấp 3 1](#_Toc387692913)

[1.1.1.2 Tiểu mục cấp 3 tiếp theo. 1](#_Toc387692914)

[1.1.2 Tiểu mục cấp 2 tiếp theo 1](#_Toc387692915)

[1.2 Nội dung của chương này 1](#_Toc387692916)

[CHƯƠNG 2 – TỔNG QUAN 1](#_Toc387692917)

[1.1 Trình bày công thức toán học 1](#_Toc387692918)

[1.2 Trình bày một hình vẽ, sơ đồ 1](#_Toc387692919)

[CHƯƠNG 3 – CƠ SỞ LÝ THUYẾT / NGHIÊN CỨU THỰC NGHIỆM 1](#_Toc387692920)

[3.1 Chèn bảng: 1](#_Toc387692921)

[3.2 Viết tắt 1](#_Toc387692922)

[3.3 Trích dẫn 1](#_Toc387692923)

[3.3.1 Tài liệu tham khảo và cách trích dẫn 1](#_Toc387692924)

[3.3.2 Qui định của Khoa Công nghệ thông tin 1](#_Toc387692925)

CHƯƠNG 1 – TỔNG QUAN VỀ THƯ VIỆN THỐNG KÊ TRONG PYTHON

*Tổng quan*: khảo sát các hàm trong thư viện thống kê của Python về cách sử dung,tham số đầu vào,ý nghĩa của của kết quả trả về

*Các thư viện cần chuẩn bị*: statistics

* 1. Hàm mean()

- Tổng quan:

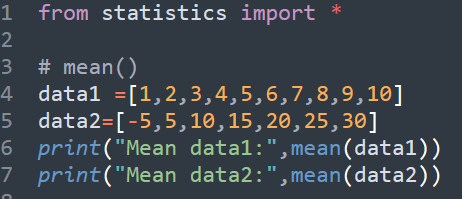
* Hàm mean() dùng để tính trung bình cộng của một dãy số truyền vào.Trung bình cộng trong thống kê là một đại lượng mô tả thống kê, được tính ra bằng cách lấy tổng giá trị của toàn bộ các phần tử trong tập chia cho số lượng các phần tử trong tập.
* Giá trị trung bình của mẫu đưa ra một ước tính không chệch về giá trị trung bình của tổng thể thực, để khi lấy trung bình trên tất cả các mẫu có thể, sẽ mean(sample)hội tụ về giá trị trung bình thực của toàn bộ tổng thể. Nếu dữ liệu đại diện cho toàn bộ tổng thể thay vì một mẫu, thì mean(data)tương đương với việc tính giá trị trung bình tổng thể thực sự.

- Cách sử dung:mean(data)

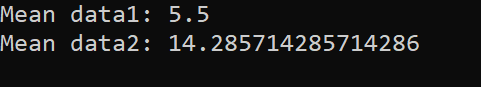
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Trung bình cộng của dãy số truyền vào.

- Ví dụ:



-Kết quả:



* 1. Hàm fmean()

- Tổng quan:

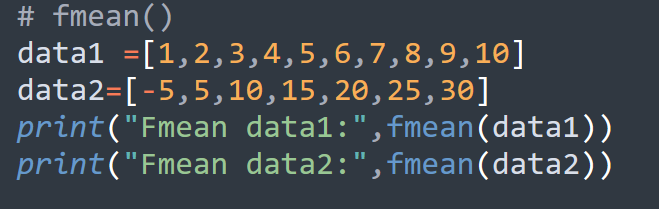
* Hàm fmean() dùng để tính trung bình cộng của một dãy số truyền vào..
* Điểm khác biệt giữa fmean() và mean() là fmean() sử dụng kiểu dữ liệu float để tính toán, còn mean() sử dụng kiểu dữ liệu int và fmean() chạy nhanh hơn mean().

- Cách sử dung:fmean(data)

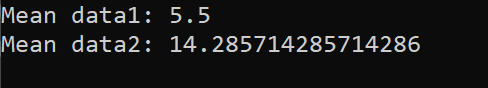
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Trung bình cộng của dãy số truyền vào dưới dạng dấu phẩy động.

-Ví dụ:



-Kết quả:



1.3 Hàm geometric\_mean()

- Tổng quan:

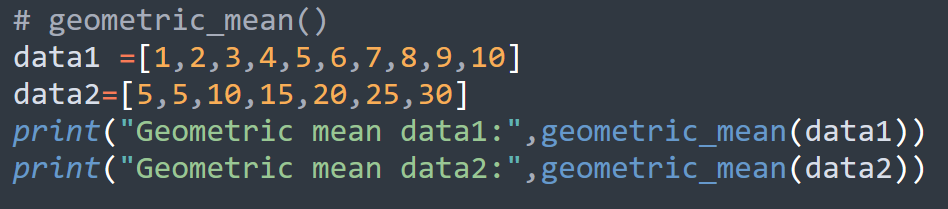
* Hàm geometric\_mean() dùng để tính trung bình hình học của một dãy số truyền vào.
* Trung bình hình học là một đại lượng thống kê được sử dụng để tính toán trung bình của các giá trị số dương.
* Trung bình hình học được tính bằng cách lấy tích của các giá trị số dương, sau đó lấy căn bậc hai của tích đó.
* Trung bình hình học biểu thị xu hươnhs trong tâm của tập dữ liệu, và nó có thể được sử dụng để tính toán tỷ lệ trung bình của các giá trị số dương.
* Trung bình hình học khác với trung bình cộng trong trường hợp các giá trị số âm, vì trung bình hình học không thể được tính toán với các giá trị số âm.

- Cách sử dung: geometric\_mean(data)

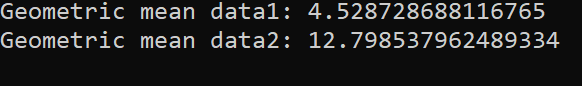
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Trung bình hình học của dãy số truyền vào.

-Ví dụ:



-Kết quả



1.4 Hàm harmonic\_mean()

- Tổng quan:

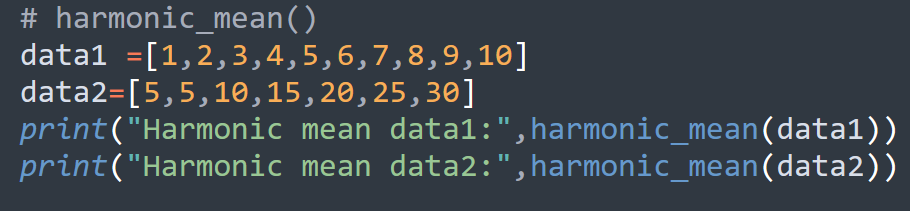
* Hàm harmonic\_mean() dùng để tính trung bình hàm của một dãy số truyền vào.
* Trung bình điều hoà là một đại lượng thống kê được sử dụng để tính toán trung bình của các giá trị số dương
* Trung bình điều hoà được tính bằng cách lấy nghịch đảo của các giá trị số dương, sau đó lấy nghịch đảo của trung bình của các giá trị số dương đó.

- Cách sử dung: harmonic\_mean(data)

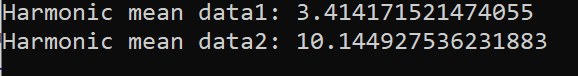
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Trung bình hàm của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.5 median()

- Tổng quan:

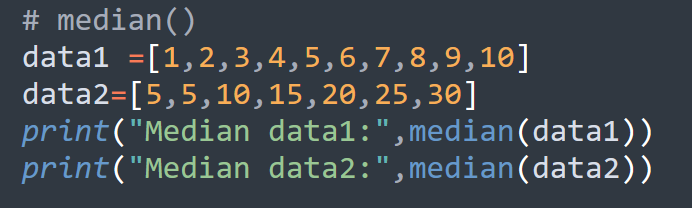
* Hàm median() dùng để tính giá trị trung vị của một dãy số truyền vào.
* Giá trị trung vị là một số tách giữa nửa lớn hơn và nửa bé hơn của một mẫu, một quần thể, hay một phân bố xác suất. Nó là giá trị giữa trong một phân bố, mà số các số nằm trên hay dưới con số đó là bằng nhau.
* Một nữa số lượng các số trong một phân bố là số lượng các số nhỏ hơn giá trị trung vị, và nửa còn lại là số lượng các số lớn hơn giá trị trung vị.
* Số trung vị của một danh sách là giá trị nằm giữa nếu số lượng các phần tử là số lẻ, hoặc là trung bình của hai giá trị nằm giữa nếu số lượng các phần tử là số chẵn.

- Cách sử dung: median(data)

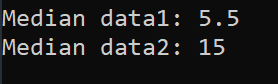
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Giá trị trung vị của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.6 median\_low()

- Tổng quan:

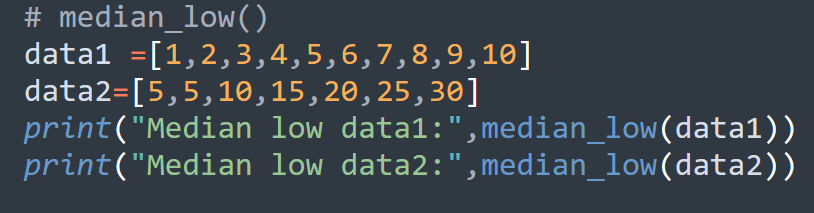
* Hàm median\_low() dùng để tính giá trị trung vị nhỏ nhất của một dãy số truyền vào.
* Giá trị trung vị nhỏ nhất là giá trị trung vị của một dãy số khi các số trong dãy số được sắp xếp theo thứ tự tăng dần.
* Nếu dãy số có số lượng phần tử là số lẻ thì giá trị trung vị nhỏ nhất là giá trị ở vị trí giữa của dãy số.
* Nếu dãy số có số lượng phần tử là số chẵn thì giá trị trung vị nhỏ nhất là giá trị ở vị trí giữa của dãy số.

- Cách sử dung: median\_low(data)

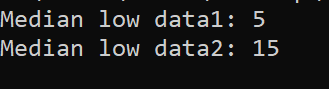
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Giá trị trung vị nhỏ nhất của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.7 median\_high()

- Tổng quan:

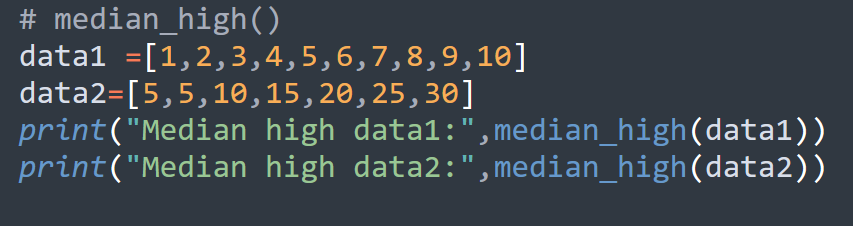
* Hàm median\_high() dùng để tính giá trị trung vị lớn nhất của một dãy số truyền vào.
* Giá trị trung vị lớn nhất là giá trị trung vị của một dãy số khi các số trong dãy số được sắp xếp theo thứ tự giảm dần.
* Nếu dãy số có số lượng phần tử là số lẻ thì giá trị trung vị lớn nhất là giá trị ở vị trí giữa của dãy số.
* Nếu dãy số có số lượng phần tử là số chẵn thì giá trị trung vị lớn nhất là giá trị ở vị trí giữa của dãy số.

- Cách sử dung: median\_high(data)

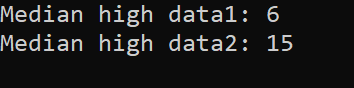
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Giá trị trung vị lớn nhất của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.8 median\_grouped()

- Tổng quan:

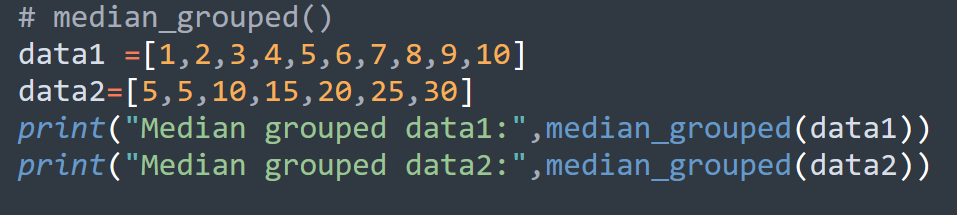
* Hàm median\_grouped() dùng để tính giá trị trung vị của một dãy số truyền vào.
* Giá trị trung vị là giá trị trung bình của giá trị trung vị nhỏ nhất và giá trị trung vị lớn nhất

- Cách sử dung: median\_grouped(data, interval=1)

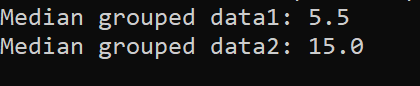
- Tham số đầu vào: data là một dãy số, interval là khoảng cách giữa các giá trị trong dãy số

- Kết quả trả về: Giá trị trung vị của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.9 Hàm mode()

- Tổng quan:

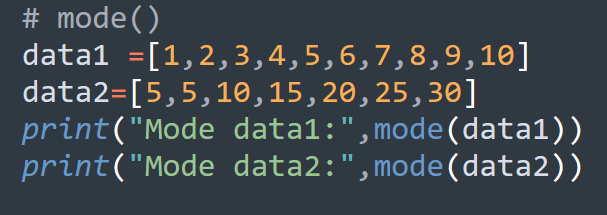
* Dùng để tính giá trị xuất hiện nhiều nhất của một dãy số truyền vào.
* Nếu có nhiều giá trị xuất hiện nhiều nhất thì hàm mode() sẽ trả về giá trị nhỏ nhất trong các giá trị đó.

- Cách sử dung: mode(data)

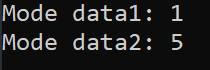
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Giá trị xuất hiện nhiều nhất của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.10 Hàm multimode()

- Tổng quan:

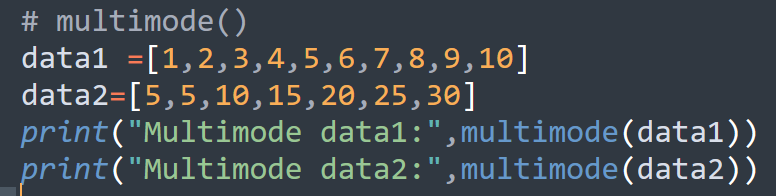
* Hàm multimode() dùng để tính giá trị xuất hiện nhiều nhất của một dãy số truyền vào.
* Nếu có nhiều giá trị xuất hiện nhiều nhất thì hàm multimode() sẽ trả về tất cả các giá trị đó.

- Cách sử dung: multimode(data)

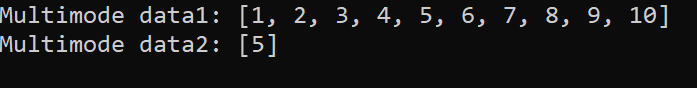
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Danh sách các giá trị xuất hiện nhiều nhất của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.11 Hàm quantiles()

- Tổng quan:

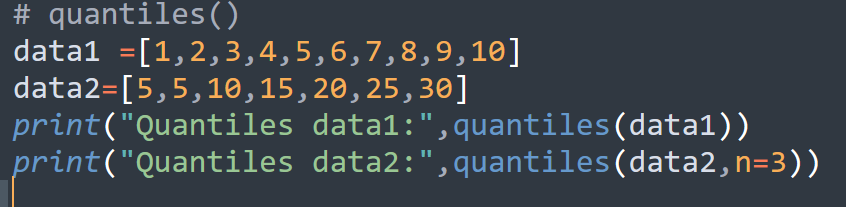
* Hàm quantiles() dùng để tính các giá trị phân vị của một dãy số truyền vào.
* Các điểm phân vị là các điểm mà dãy số được chia thành các phần bằng nhau.

- Cách sử dung: quantiles(data, n=4)

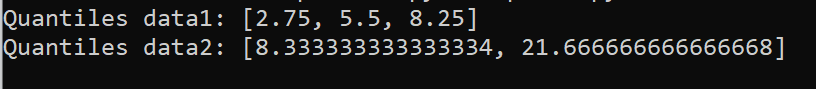
- Tham số đầu vào: data là một dãy số, n là số lượng phân vị cần tính

- Kết quả trả về: Danh sách các giá trị phân vị của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.12 Hàm pstdev()

- Tổng quan:

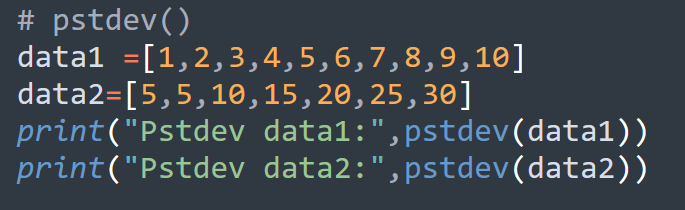
* Hàm pstdev() dùng để tính độ lệch chuẩn của một tổng thể..
* Độ lệch chuẩn là độ lệch trung bình của các giá trị trong dãy số truyền vào so với giá trị trung bình của dãy số đó.
* Là một đại lượng thống kê dùng để đo mức độ phân tán của một tập dữ liệu đã được lập thành bảng tần số.
* Có thể tính ra độ lệch chuẩn bằng cách lấy căn bậc hai của phương sai. Khi hai tập dữ liệu có cùng giá trị trung bình cộng, tập nào có độ lệch chuẩn lớn hơn là tập có dữ liệu biến thiên nhiều hơn. Trong trường hợp hai tập dữ liệu có giá trị trung bình cộng không bằng nhau, thì việc so sánh độ lệch chuẩn của chúng không có ý nghĩa.
* Độ lệch chuẩn còn được sử dụng khi tính sai số chuẩn. Khi lấy độ lệch chuẩn chia cho căn bậc hai của số lượng quan sát trong tập dữ liệu, sẽ có giá trị của sai số chuẩn.

- Cách sử dung: pstdev(data)

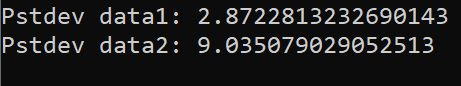
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Độ lệch chuẩn của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.13 Hàm pvariance()

- Tổng quan:

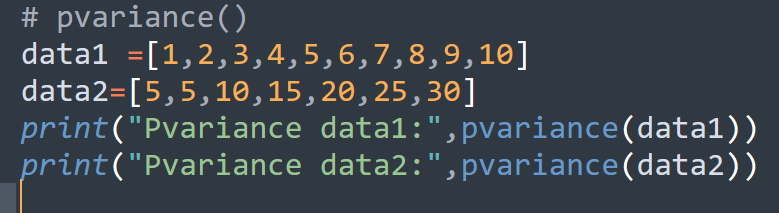
* Hàm pvariance() dùng để tính phương sai của một tổng thể.
* Phương sai là giá trị kỳ vọng của bình phương của độ lệch của X so với giá trị trung bình của nó.
* Có thể hiểu phương sai là "trung bình của bình phương khoảng cách của mỗi điểm dữ liệu tới điểm trung bình". Do đó, nó là giá trị trung bình của bình phương độ lệch.

- Cách sử dung: pvariance(data)

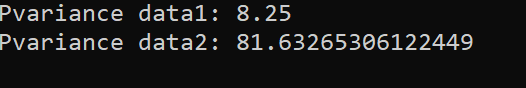
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Phương sai của dãy số truyền vào.

-Ví dụ:



-Kết quả



1.14 Hàm stdev()

- Tổng quan:

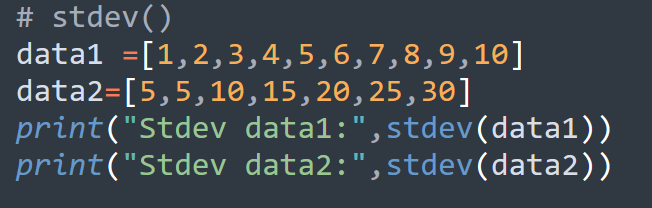
* Hàm stdev() dùng để tính độ lệch chuẩn của một mẫu.
* Độ lệch chuẩn là độ lệch trung bình của các giá trị trong dãy số truyền vào so với giá trị trung bình của dãy số đó.
* Độ lệch chuẩn là một đại lượng thống kê dùng để đo mức độ phân tán của một tập dữ liệu đã được lập thành bảng tần số.
* Độ lệch chuẩn còn được sử dụng khi tính sai số chuẩn. Khi lấy độ lệch chuẩn chia cho căn bậc hai của số lượng quan sát trong tập dữ liệu, sẽ có giá trị của sai số chuẩn.
* Có thể tính ra độ lệch chuẩn bằng cách lấy căn bậc hai của phương sai. Khi hai tập dữ liệu có cùng giá trị trung bình cộng, tập nào có độ lệch chuẩn lớn hơn là tập có dữ liệu biến thiên nhiều hơn.
* Trong trường hợp hai tập dữ liệu có giá trị trung bình cộng không bằng nhau, thì việc so sánh độ lệch chuẩn của chúng không có ý nghĩa. Độ lệch chuẩn còn được sử dụng khi tính sai số chuẩn. Khi lấy độ lệch chuẩn chia cho căn bậc hai của số lượng quan sát trong tập dữ liệu, sẽ có giá trị của sai số chuẩn.

- Cách sử dung: stdev(data)

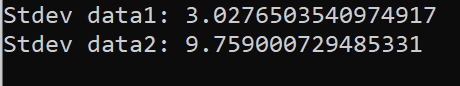
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Độ lệch chuẩn của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.15 Hàm variance()

- Tổng quan:

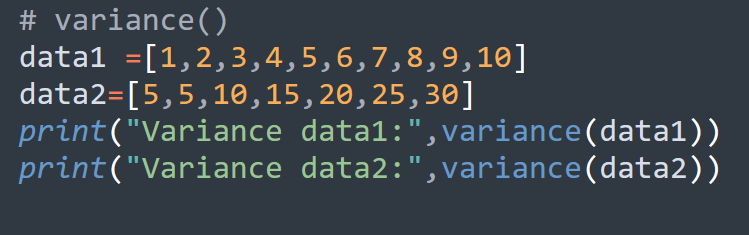
* Hàm variance() dùng để tính phương sai của một mẫu.
* Phương sai là giá trị kỳ vọng của bình phương của độ lệch của X so với giá trị trung bình của nó

- Cách sử dung: variance(data)

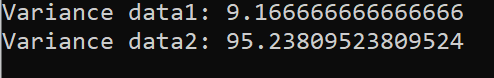
- Tham số đầu vào: data là một dãy số.

- Kết quả trả về: Phương sai của dãy số truyền vào.

-Ví dụ:



-Kết quả:



1.16 Hàm covariance()

- Tổng quan:

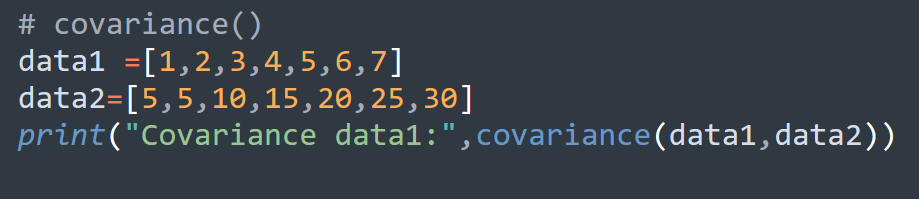
* Hàm covariance() dùng để tính hiệp phương sai.
* Hiệp phương sai là độ đo sự biến thiên cùng nhau của hai biến ngẫu nhiên.
* Hiệp phương sai được tính bằng cách phân tích các biến động về lãi suất (độ lệch chuẩn so với lợi nhuận kì vọng) hoặc bằng cách nhân mối tương quan giữa hai biến với độ lệch chuẩn của từng biến.

- Cách sử dung: covariance(data1, data2)

- Tham số đầu vào: data1, data2 là hai dãy số.

- Kết quả trả về: Hiệp phương sai của hai dãy số truyền vào.

-Ví dụ:



-Kết quả



1.17 Hàm correlation()

- Tổng quan:

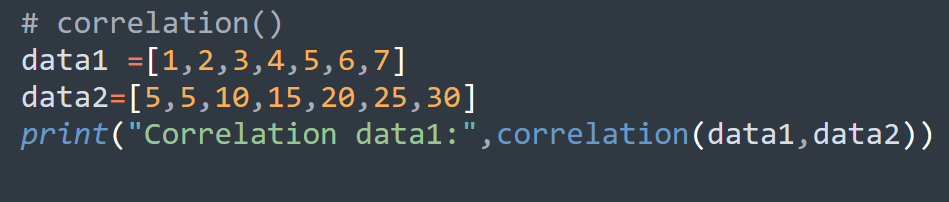
* Hàm correlation() dùng để tính độ tương quan giữa hai dãy số.
* Hệ số tương quan là một chỉ số đo lường của một số loại tương quan, nghĩa là mối liên hệ thống kê giữa hai biến số.Các biến có thể là hai cột của một bộ dữ liệu quan sát đã cho, thường được gọi là mẫu hoặc hai phần của một biến ngẫu nhiên đa biến số có phân phối đã biết trước.
* Có một số loại hệ số tương quan, mỗi loại lại có định nghĩa riêng, phạm vi sử dụng và đặc tính riêng. Tất cả đều giả định các giá trị nằm trong phạm vi chạy từ −1 đến +1, trong đó ± 1 biểu thị hai biến số có mối tương quan tuyệt đối có thể và 0 chỉ hai biến số không có liên hệ gì với nhau.
* Là công cụ phân tích, các hệ số tương quan thể hiện một số vấn đề nhất định, bao gồm khuynh hướng của một số loại yếu tố nhiễu bởi ngoại lai và khả năng được sử dụng tương đối để suy ra mối quan hệ nhân quả giữa các biến số.

- Cách sử dung: correlation(data1, data2)

- Tham số đầu vào: data1, data2 là hai dãy số.

- Kết quả trả về: Độ tương quan của hai dãy số truyền vào.

-Ví dụ:



-Kết quả



1.18 Hàm linear\_regression()

- Tổng quan:

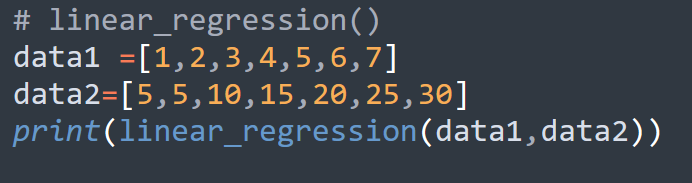
* Hàm linear\_regression() dùng để tính hệ số hồi quy tuyến tính.
* Hồi quy tuyến tính là một trong những phương pháp phân tích dữ liệu phổ biến nhất. Nó được sử dụng để dự đoán mối quan hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập. Hồi quy tuyến tính được sử dụng để dự đoán mối quan hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập. Hồi quy tuyến tính được sử dụng để dự đoán mối quan hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập. Hồi quy tuyến tính được sử dụng để dự đoán mối quan hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập.

- Cách sử dung: linear\_regression(data1, data2)

- Tham số đầu vào: data1, data2 là hai dãy số.

- Kết quả trả về: Hệ số hồi quy tuyến tính của hai dãy số truyền vào.

-Ví dụ:



-Kết quả



CHƯƠNG 2 – KHẢO SÁT THUẬT TOÁN HISTOGRAM EQUALIZATION ĐỂ XỬ LÝ ẢNH

*Tổng quan*: khảo sát các vấn đề liên quan đến thuật toán,công thức,thuật toán,..chương trình minh hoạ

*Thư viện cần thiết*:numpy,cv2,matplotlib.pyplot.

2.1 Tổng quan

- Thuật toán Histogram Equalization:

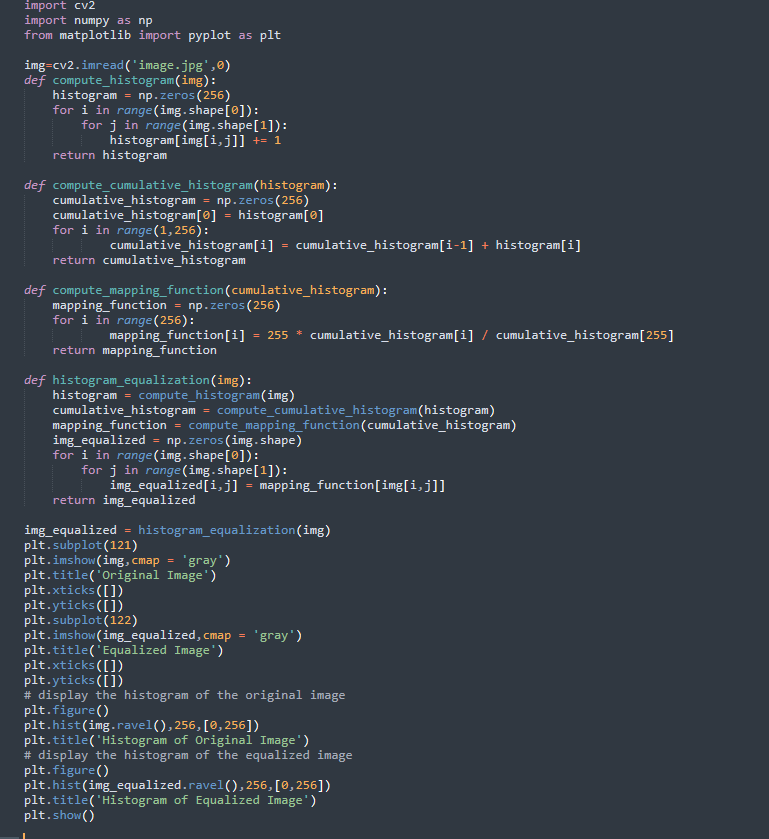
* Là kỹ thuật xử lý hình ảnh để cải thiện độ tương phản trong hình ảnh bằng cách trải rộng các giá trị cường độ xuất hiện nhiều nhất, kéo dài phạm vi cường độ của hình ảnh.
* Phương pháp này thường làm tăng độ tương phản của hình ảnh . Cho phép các khu vực có độ tương phản thấp hơn để có được độ tương phản cao hơn.

- Các bước để thực hiện thuật toán:

* Nhận hình ảnh đầu vào.
* Tạo biểu đồ cho hình ảnh
* Tìm cực tiểu cục bộ của hình ảnh
* Chia biểu đồ dựa trên cực tiểu cục bộ
* Có các mức xám cụ thể cho từng phân vùng của biểu đồ
* Áp dụng cân bằng biểu đồ trên mỗi phân vùng

2.2 Chương trình minh hoạ

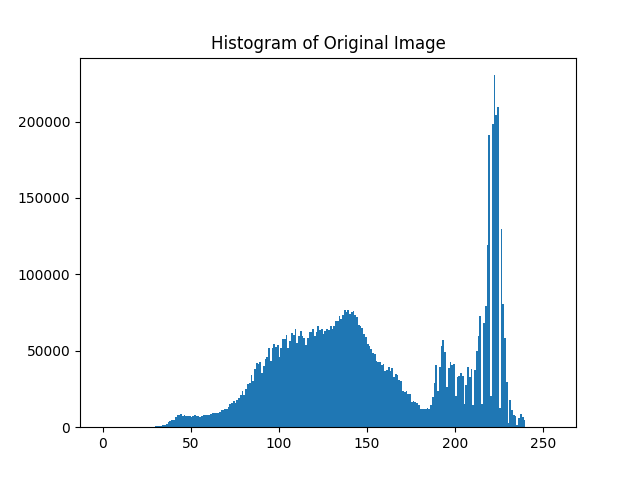
-Ví dụ:

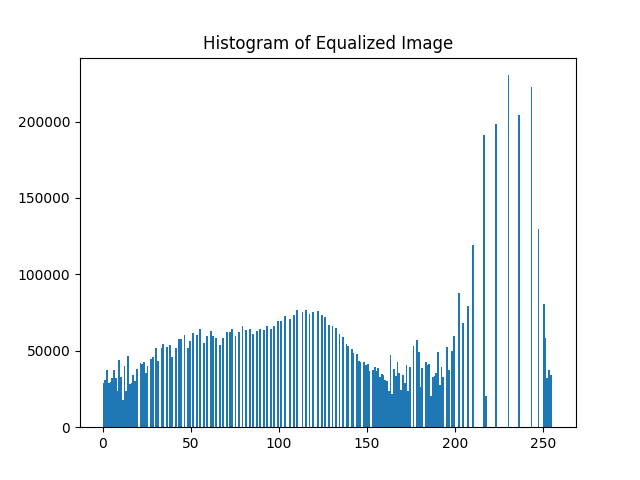


-Kết quả



Ảnh trước và sau khi xử lí





**TÀI LIỆU THAM KHẢO**

**Tiếng Việt**

1. Quách Ngọc Ân (1992), “Nhìn lại hai năm phát triển lúa lai”, *Di tuyền học ứng dụng*, 98(1), tr. 10-16.
2. Bộ nông nghiệp & PTNT (1996), *Báo cáo tổng kết 5 năm (1992-1996) phát triển lúa lai,* Hà Nội.
3. Nguyễn Hữu Đống, Đào Thanh Bằng, Lâm Quang Dụ, Phan Đức Trực (1997), *Đột biến –* *Cơ sở lý luận và ứng dụng,* Nhà xuất bản nông nghiệp, Viện khoa học kỹ thuật nông nghiệp Việt Nam, Hà Nội.
4. Nguyễn Thị Gấm (1996), *Phát hiện và đánh giá một số dòng bất dục đực cảm ứng nhiệt* *độ,* Luận văn thạc sĩ khoa học nông nghiệp, Viện khoa học kỹ thuật nông nghiệp Việt Nam, Hà Nội.

……….

1. Võ Thị Kim Huệ (2000), *Nghiên cứu chẩn đoán và điều trị bệnh…,* Luận án Tiến sĩ y khoa, Trường đại học y Hà Nội, Hà Nội.

**Tiếng Anh**

1. Anderson J.E. (1985), The Relative Inefficiency of Quota, The Cheese Case, *American* *Economic Review*, 75(1), pp. 178-90.
2. Borkakati R. P.,Virmani S. S. (1997), Genetics of thermosensitive genic male sterility in Rice, *Euphytica* 88, pp. 1-7.
3. Boulding K.E. (1955), *Economics Analysis*, Hamish Hamilton, London.
4. Burton G. W. (1988), “Cytoplasmic male-sterility in pearl millet (penni-setum glaucum L.)”, *Agronomic Journal* 50, pp. 230-231.
5. Central Statistical Oraganisation (1995), *Statistical Year Book*, Beijing.
6. FAO (1971), *Agricultural Commodity Projections (1970-1980)*, Vol. II. Rome.
7. Institute of Economics (1988), *Analysis of Expenditure Pattern of Urban Households in* *Vietnam,* Departement pf Economics, Economic Research Report, Hanoi.

**PHỤ LỤC**

Phần này bao gồm những nội dung cần thiết nhằm minh họa hoặc hỗ trợ cho nội dung luận văn như số liệu, biểu mẫu, tranh ảnh. . . . nếu sử dụng những câu trả lời cho một *bảng câu hỏi thì bảng câu hỏi mẫu này phải được đưa vào phần Phụ lục ở dạng nguyên bản* đã dùng để điều tra, thăm dò ý kiến; **không được tóm tắt hoặc sửa đổi**. Các tính toán mẫu trình bày tóm tắt trong các biểu mẫu cũng cần nêu trong Phụ lục của luận văn. Phụ lục không được dày hơn phần chính của luận văn