TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÀI TẬP LỚN**

**MÔN XÁC SUẤT THỐNG KÊ   
ỨNG DỤNG CHO CÔNG NGHỆ THÔNG TIN**

**Mathematical statistics functions -  
Histogram equalization algorithm**

*Người hướng dẫn*: **ThS NGUYỂN HUỲNH MINH DUY**

*Người thực hiện*: **MAI HOÀNG VIỆT - 51900847**

Lớp **: 19050402**

Khoá **: 23**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2021**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÀI TẬP LỚN**

**MÔN XÁC SUẤT THỐNG KÊ   
ỨNG DỤNG CHO CÔNG NGHỆ THÔNG TIN**

**Mathematical statistics functions -  
Histogram equalization algorithm**

Người hướng dẫn: **ThS NGUYỄN HUỲNH MINH DUY**

Người thực hiện: **MAI HOÀNG VIỆT - 51900847**

Lớp**: 19050402**

Khoá**: 23**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2021**

LỜI CẢM ƠN

Để hoàn thành bài tập lớn môn **Xác suất thống kê ứng dụng cho Công nghệ thông tin** này em đã nhận được những sự giúp đỡ, đóng góp ý kiến từ thầy cô và bạn bè.

Em xin gửi lời cảm ơn chân thành các thầy/cô trong khoa **Công nghệ thông tin** của trường **Đại học Tôn Đức Thắng** đã truyền đạt cho em những kiến thức đại cương cơ bản cũng như các môn chuyên ngành, giúp em nắm được những kiến thức cơ bản và tạo điều kiện cho các môn học tiếp theo và trong suốt quá trình học tập.

Em xin đặc biệt gửi lời cảm ơn đến thầy **Nguyễn Huỳnh Minh Duy**, giảng viên khoa Công nghệ thông tin, người đã tận tình giải đáp, hướng dẫn và giúp đỡ em trong quá trình hoàn thành bài tập lớn.

Trong quá trình hoàn thành bài tập lớn môn Xác suất thống kê ứng dụng cho Công nghệ thông tin, kiến thức của em còn hạn chế nên trong quá trình làm bài thì không thể tránh được xảy ra sai sót. Kính mong nhận được sự quan tâm, góp ý từ cô.

**EM XIN CHÂN THÀNH CẢM ƠN!**

**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 và được sự hướng dẫn của ThS Nguyễn Quốc Bình. 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 2021*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Mai Hoàng Việt*

PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN

**Phần xác nhận của GV hướng dẫn**

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

Tp. Hồ Chí Minh, ngày tháng năm 2021

(kí và ghi họ tên)

**Phần đánh giá của GV chấm bài**

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

Tp. Hồ Chí Minh, ngày tháng năm 2021

(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](#_Toc70027682)

[PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN iii](#_Toc70027683)

[TÓM TẮT iv](#_Toc70027684)

[MỤC LỤC 1](#_Toc70027685)

[CHƯƠNG 1 – MATHEMATICAL STATISTICS FUNCTIONS 2](#_Toc70027687)

[1.1 Các hàm tính giá trị trung bình và số đo của vị trí trung tâm 2](#_Toc70027688)

[1.1.1 Hàm statistics.mean(data) 2](#_Toc70027689)

[1.1.2 Hàm statistics.fmean(data) 3](#_Toc70027690)

[1.1.3 Hàm statistics.geometric\_mean(data) 4](#_Toc70027691)

[1.1.4 Hàm statistics.harmonic\_mean(data) 5](#_Toc70027692)

[1.1.5 Hàm statistics.median(data) 6](#_Toc70027693)

[1.1.6 Hàm statistics.median\_low(data) 7](#_Toc70027694)

[1.1.7 Hàm statistics.median\_high(data) 8](#_Toc70027695)

[1.1.8 Hàm statistics.median\_grouped(data, interval=1) 9](#_Toc70027696)

[1.1.9 Hàm statistics.mode(data) 10](#_Toc70027697)

[1.1.10 Hàm statistics.multimode(data) 11](#_Toc70027698)

[1.1.11 Hàm statistics.quantiles(data, \*, n=4, method = ‘exclusive’) 12](#_Toc70027699)

[1.2 Các hàm tính độ chênh lệch 13](#_Toc70027700)

[1.2.1 Hàm statistics.pstdev(data, mu = None) 14](#_Toc70027701)

[1.2.2 Hàm statistics.pvariance(data, mu = None) 15](#_Toc70027702)

[1.2.3 Hàm statistics.stdev(data, xbar = None) 16](#_Toc70027703)

[1.2.4 Hàm statistics.variance(data, xbar = None) 17](#_Toc70027704)

[CHƯƠNG 2 – HISTOGRAM EQUAIZATION ALGORITHM 18](#_Toc70027705)

[1.1 Khái niệm Histogram 18](#_Toc70027706)

[1.2 Histogram Equalization 22](#_Toc70027707)

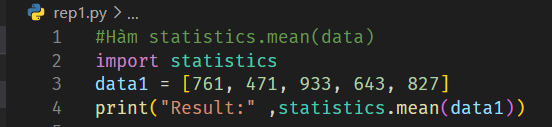
CHƯƠNG 1 – MATHEMATICAL STATISTICS FUNCTIONS

1.1 Các hàm tính giá trị trung bình và số đo của vị trí trung tâm

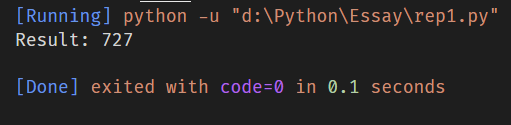
1.1.1 Hàm statistics.mean(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tính giá trị trung bình cộng của một tập dữ liệu cho trước. |
| Giá trị trả về | Giá trị trung bình cộng tổng thể của danh sách. |
| Ngoại lệ | Khi *data* rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ:**



* **Output:**

****

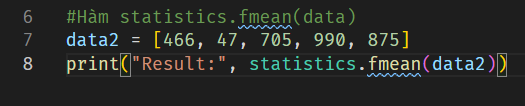
**- Cách sử dụng:**

Hàm statistics.mean(data) thực hiện bằng cách tính giá trị trung bình cộng các phần tử có trong tập hợp, sau đó trả về giá trị đó.

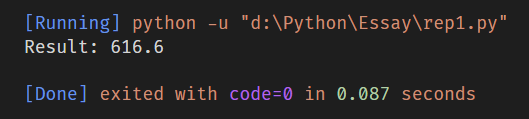
1.1.2 Hàm statistics.fmean(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tính giá trị trung bình của một tập dữ liệu cho trước. |
| Giá trị trả về | Một giá trị kiểu float, đại diện cho giá trị trung bình cộng mẫu. |
| Ngoại lệ | Khi *data* rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ:**

****

**- Output:**



* **Cách sử dụng:**

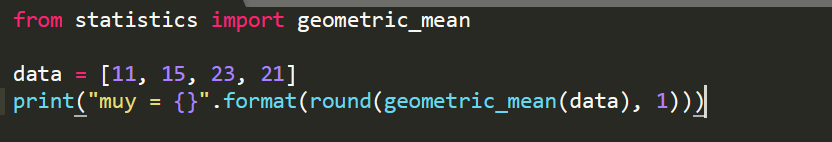
Hàm statistics.fmean(data) thực hiện bằng cách tính giá trị trung bình cộng các phần tử có trong tập hợp, sau đó trả về giá trị với kiểu dữ liệu là float.

Hàm này tương tự như hàm mean(data) nhưng tốc độ tính và hiển thị nhanh hơn.

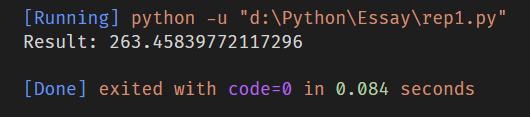
1.1.3 Hàm statistics.geometric\_mean(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tính giá trị trung bình nhân của một tập dữ liệu cho trước. |
| Giá trị trả về | Một giá trị kiểu float, đại diện cho giá trị trung bình nhân của mẫu. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng, *data* có chứa số 0 hoặc là giá trị âm thì StatisticsError sẽ được gọi. |

* **Ví dụ:**

****

* **Output:**



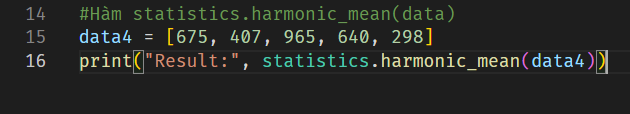
* **Cách sử dụng:**

Hàm statistics.geometric\_mean(data) thực hiện bằng cách tính giá trị trung bình nhân các phần tử có trong tập hợp, sau đó trả về giá trị với kiểu dữ liệu là float.

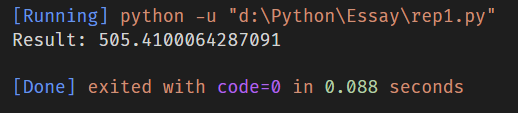
1.1.4 Hàm statistics.harmonic\_mean(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tính giá trị trung bình hài hòa của một tập dữ liệu. |
| Giá trị trả về | Một giá trị kiểu float, đại diện cho giá trị trung bình hài hòa của mẫu. Khi tập dữ liệu truyền vào có chứa số 0, kết quả trả về là 0. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng, *data* có chứa giá trị âm thì StatisticsError sẽ được gọi. |

* **Ví dụ :**



* **Output:**

****

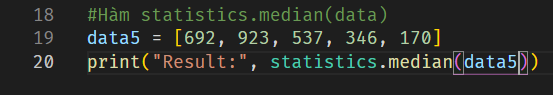
* **Cách sử dụng:**

Hàm statistics.harmonic\_mean(data) thực hiện bằng cách tính nghịch đảo của giá trị trung bình cộng của nghịch đảo các phần tử có trong tập hợp, sau đó trả về giá trị với kiểu dữ liệu là float.

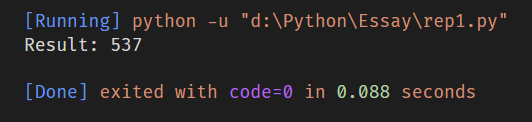
1.1.5 Hàm statistics.median(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tìm trung vị (giá trị giữa) của một tập dữ liệu cho trước. |
| Giá trị trả về | Trả về trung vị của một tập dữ liệu rời rạc. Nếu số lượng phần tử truyền vào là số chẵn, ta lấy trung bình cộng của cả hai, ngược lại ta lấy giá trị giữa. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ :**

****

* **Output:**

****

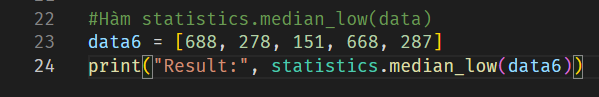
* **Cách sử dụng:**

Hàm statistics.median(data) sẽ sắp xếp các giá trị đó theo thứ tự tăng dần rồi bắt đầu thực hiện việc tìm ra trung vị của tập dữ liệu cho trước.

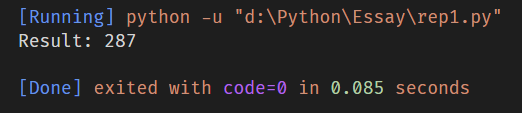
1.1.6 Hàm statistics.median\_low(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tìm trung vị thấp của một tập dữ liệu cho trước. |
| Giá trị trả về | Trả về trung vị thấp một tập dữ liệu rời rạc. Nếu số lượng phần tử truyền vào là số chẵn, ta lấy giá trị nhỏ hơn giữa hai phần tử, ngược lại giá trị ở giữa sẽ được trả về. Trung vị thấp luôn là một phần tử trong tập hợp. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ :**

****

* **Output:**



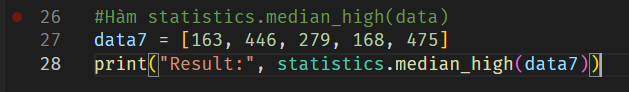
* **Cách sử dụng:**

Hàm statistics.median(data) sẽ sắp xếp các giá trị đó theo thứ tự tăng dần rồi bắt đầu thực hiện việc tìm ra trung vị thấp của tập dữ liệu cho trước.

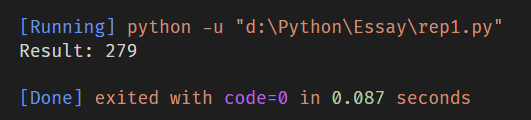
1.1.7 Hàm statistics.median\_high(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tìm trung vị cao của một tập dữ liệu cho trước. |
| Giá trị trả về | Trả về trung vị cao một tập dữ liệu rời rạc. Nếu số lượng phần tử truyền vào là số chẵn, ta lấy giá trị lớn hơn giữa hai phần tử, ngược lại giá trị ở giữa sẽ được trả về. Trung vị cao luôn là một phần tử trong tập hợp. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ :**



* **Output:**



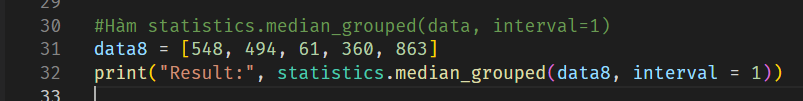
* **Cách sử dụng:**

Hàm statistics.median(data) sẽ sắp xếp các giá trị đó theo thứ tự tăng dần rồi bắt đầu thực hiện việc tìm ra trung vị cao của tập dữ liệu cho trước.

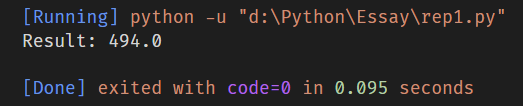
1.1.8 Hàm statistics.median\_grouped(data, interval=1)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán.  *interval*: mặc định là 1, đại diện cho độ rộng của nhóm. |
| Ý nghĩa | Dùng để tìm trung vị nhóm của tập dữ liệu cho trước. |
| Giá trị trả về | Trả về trung vị nhóm của một tập dữ liệu liên tục, tính toán ở phân vị thứ 50. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ:**



* **Output:**



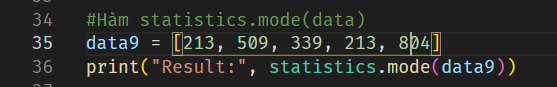
* **Cách sử dụng:**

Hàm này sẽ thực hiện nhóm các giá trị trong tập hợp với độ rộng bằng interval=1 (mặc định).

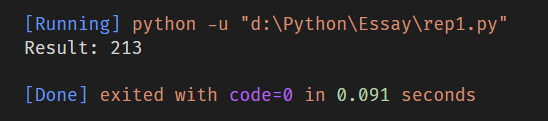
1.1.9 Hàm statistics.mode(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc không phải là số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tìm giá trị xuất hiện nhiều lần nhất trong tập dữ liệu cho trước. |
| Giá trị trả về | Trả về giá trị xuất hiện nhiều nhất trong tập dữ liệu rời rạc hoặc không phải là số. Nếu có nhiều kết quả xuất hiện như nhau thì trả về giá trị xuất hiện đầu tiên. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ:**



* **Output:**



* **Cách sử dụng:**

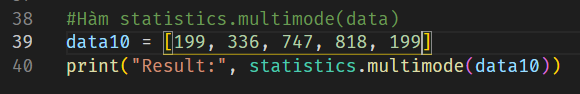
Hàm mode() sẽ trả về data point phổ biến nhất từ ​​số rời rạc cũng như dữ liệu không phải là số.

Đây là hàm thống kê duy nhất có thể được sử dụng với dữ liệu không phải là số.

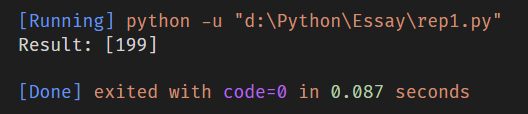
1.1.10 Hàm statistics.multimode(data)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc không phải là số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để tìm giá trị xuất hiện nhiều lần nhất trong tập dữ liệu cho trước. |
| Giá trị trả về | Trả về một danh sách giá trị xuất hiện nhiều nhất trong tập dữ liệu rời rạc hoặc không phải là số. Trả về rỗng nếu giá trị truyền vào rỗng. |

* **Ví dụ:**



* **Output:**



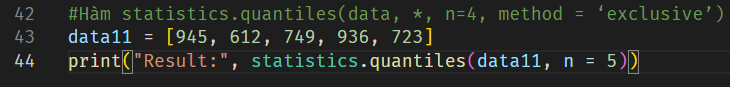
**- Cách sử dụng:**

Hàm multimode() sẽ trả về data list phổ biến nhất từ ​​số rời rạc cũng như dữ liệu không phải là số.

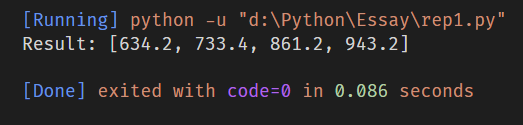
1.1.11 Hàm statistics.quantiles(data, \*, n=4, method = ‘exclusive’)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán. |
| Ý nghĩa | Dùng để chia tập dữ liệu thành các khoảng có xác suất bằng nhau. |
| Giá trị trả về | Trả về tứ phân vị của tập dữ liệu. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ:**



* **Output:**

****

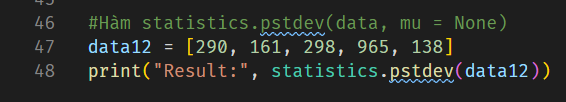
* **Cách sử dụng:**

1.2 Các hàm tính độ chênh lệch

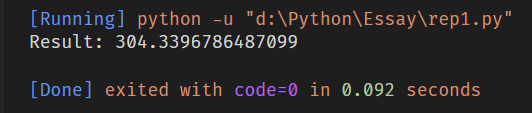
1.2.1 Hàm statistics.pstdev(data, mu = None)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán.  *mu:* mặc định là None, đại diện cho trung bình tổng thể của tập hợp cho trước. |
| Ý nghĩa | Dùng để tính độ lệch chuẩn của tổng thể. |
| Giá trị trả về | Trả về độ lệch chuẩn của tổng thể. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ:**

****

* **Output:**

****

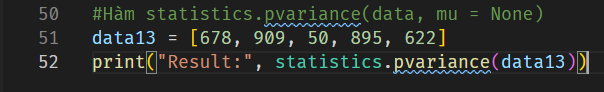
* **Cách sử dụng:**
  + Hàm sẽ thực hiện tính độ lệch chuẩn của tổng thể bằng cách lấy căn bậc hai của phương sai của tổng thể.
* Trước hết ta tính trung bình của tổng thể theo công thức:
* Sau đó ta tính phương sai của tổng thể:

= 304.3396786487099

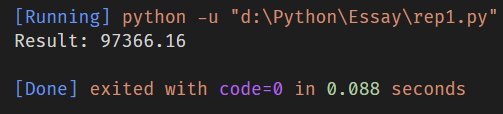
1.2.2 Hàm statistics.pvariance(data, mu = None)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán.  *mu:* mặc định là None, đại diện cho trung bình tổng thể của tập hợp cho trước. |
| Ý nghĩa | Dùng để tính phương sai của tổng thể. |
| Giá trị trả về | Trả về phương sai của tổng thể. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng thì StatisticsError sẽ được gọi. |

* **Ví dụ:**

****

* **Output:**

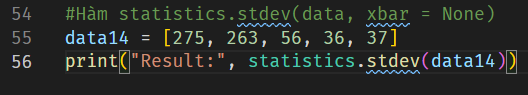
****

* **Cách sử dụng:**
* Trước hết ta tính trung bình của tổng thể theo công thức (1.1):
* Sau đó ta tính phương sai của tổng thể:

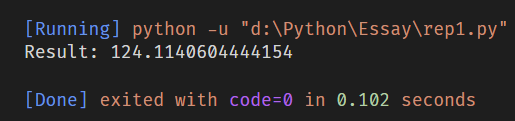
1.2.3 Hàm statistics.stdev(data, xbar = None)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán.  *xbar:* mặc định là None, đại diện cho trung bình mẫu của tập hợp cho trước. Khi *xbar* không được truyền vào hoặc None thì hàm sẽ tự động tính. |
| Ý nghĩa | Dùng để tính độ lệch chuẩn của mẫu. |
| Giá trị trả về | Trả về độ lệch chuẩn của mẫu. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng hoặc có ít hơn hai giá trị trong *data* truyền vào thì StatisticsError sẽ được gọi. |

* **Ví dụ:**

****

* **Output:**

****

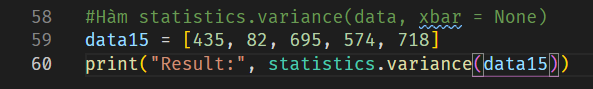
* **Cách sử dụng:**
* Hàm sẽ thực hiện tính độ lệch chuẩn của mẫu bằng cách lấy căn bậc hai của phương sai của mẫu.
* Trước hết ta tính trung bình của tổng thể theo công thức:
* Sau đó ta tính phương sai của mẫu:

= 124.1140604444154

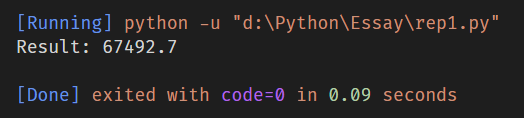
1.2.4 Hàm statistics.variance(data, xbar = None)

|  |  |
| --- | --- |
| Tham số truyền vào | *data*: là một danh sách các số hoặc một hàm thực hiện phép toán.  *xbar:* mặc định là None, đại diện cho trung bình mẫu của tập hợp cho trước. Khi *xbar* không được truyền vào hoặc None thì hàm sẽ tự động tính. |
| Ý nghĩa | Dùng để tính phương sai của mẫu. |
| Giá trị trả về | Trả về phương sai của mẫu. |
| Ngoại lệ | Khi tham số *data* truyền vào rỗng hoặc có ít hơn hai giá trị trong *data* truyền vào thì StatisticsError sẽ được gọi. |

* **Ví dụ:**

****

* **Output:**

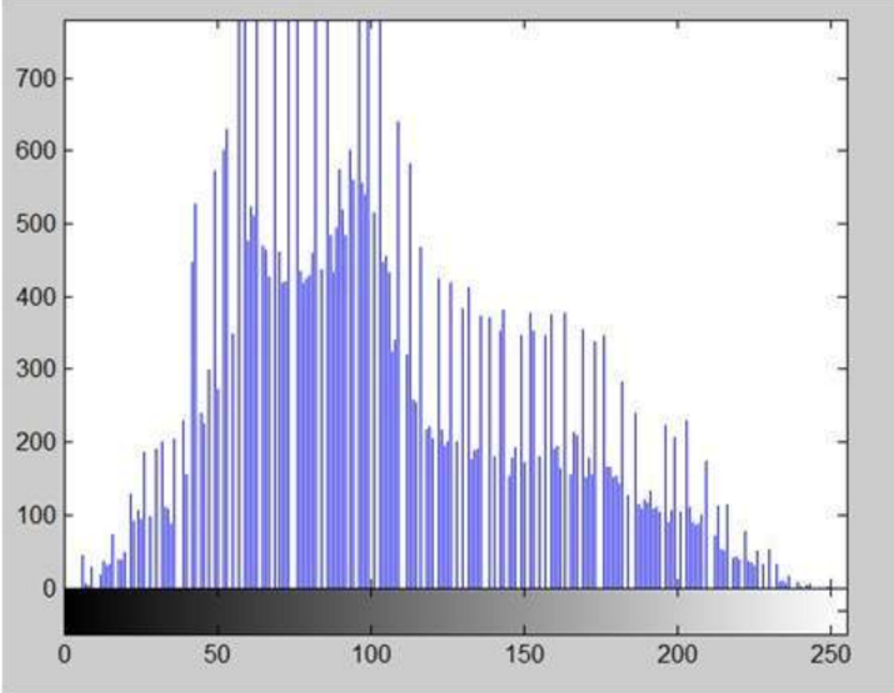


* **Cách sử dụng:**
* Hàm sẽ thực hiện tính độ lệch chuẩn của mẫu bằng cách lấy căn bậc hai của phương sai của mẫu.
* Trước hết ta tính trung bình của tổng thể theo công thức:
* Sau đó ta tính phương sai của mẫu:

= 67492.7

CHƯƠNG 2 – HISTOGRAM EQUAIZATION ALGORITHM

* 1. Khái niệm Histogram
* Biểu đồ phân bố tần số (Biểu đồ phân bố mật độ, biểu đồ cột) dùng để đo tần số xuất hiện của một vấn đề nào đó, cho thấy rõ hình ảnh sự thay đổi, biến động của một tập dữ liệu.
* Trong xử lý ảnh, biểu đồ Histogram của một ảnh là biểu đồ mô tả sự phân bố của các giá trị mức xám của các điểm ảnh (Pixel) trong một bức ảnh hoặc một vùng ảnh (Region).
* Trong một ảnh, giá trị của một điểm ảnh (Pixel) thường là từ 0-255.
* Trục tung (Oy) biểu diễn số lượng điểm ảnh (Pixel) của mức xám.
* Trục hoành (Ox) biểu diễn mức xám.
* Một biều đồ tốt là biểu đồ có số lượng điểm ảnh nhiều nhất ở vùng giữa (Độ sáng trung bình) và ít dần ra 2 vùng sáng tối (Ngọn núi).



* 1. Histogram Equalization

1. Khái niệm:

Histogram Equalization là sự điều chỉnh histogram về trạng thái cân bằng, làm cho phân bố (distribution) giá trị pixel không bị co cụm tại một khoảng hẹp mà được "kéo dãn" ra.

1. Công thức:

* Gọi hàm biến đổi ta cần xác định là K(i) với i ϵ [0,255]
* Bước 1: Thống kê số lượng pixel cho từng mức sáng, ta được histogram H(i)
* Bước 2:  Tính "hàm tích lũy" Z cho từng mức sáng theo công thức:

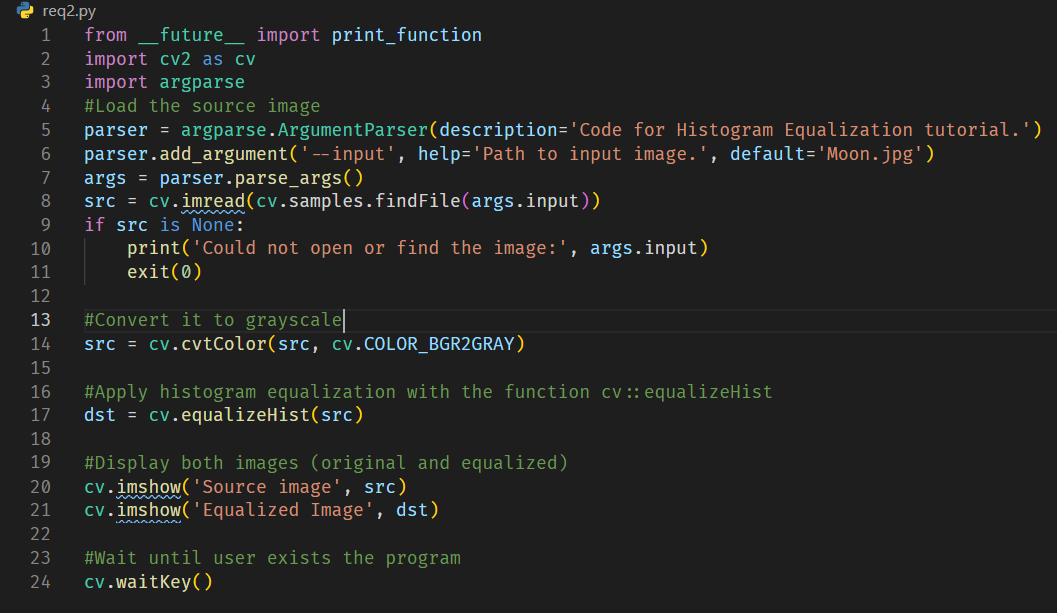
Z(i) =

* + Trong đó:

Z(i) chính là tổng số pixel có giá trị ⩽i

* + Z(i) chắc chắn là hàm đồng biến tăng
  + Ta cần biến đổi các giá trị pixel sao cho giá trị của Z sẽ phải tăng dần đều thay vì tăng đột ngột như trong ảnh gốc
* Bước 3: Hàm biến đổi K tại một mức sáng i được tính như sau:

1. Hiện thực Histogram Equalization



Input:



Output:

