**[Câu 3](#De_3): Chọn dữ liệu cho k biến (k ) (hoặc xử lý số liệu theo nhóm cho phù hợp) để lập bài toán so sánh về phân tích phương sai.**

**PHẦN 1: LẬP BÀI TOÁN**

Để đánh giá mức độ ứng phó của y tế các nước với đại dịch Covid-19, ta lấy dữ liệu tỉ lệ ca tử vong (đơn vị: %) trên tổng số ca mắc của các quốc gia theo từng châu lục, số liệu thu được đến ngày 22/5/2021 như sau:

**Nhóm 1:** *Tỉ lệ* ca tử vong trên tổng số ca bệnh do nCoV của các nước châu Phi

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2.69 | 2.22 | 1.26 | 1.52 | 1.23 | 0.13 | 1.59 | 0.88 | 1.37 | 3.51 |
| 3.77 | 1.31 | 0.63 | 2.52 | 1.32 | 5.81 | 1.34 | 0.36 | 3.62 | 1.51 |
| 0.61 | 2.93 | 0.84 | 0.69 | 1.81 | 1.81 | 3.01 | 3.97 | 1.7 | 1.94 |
| 3.36 | 3.6 | 2.4 | 1.31 | 1.76 | 1.18 | 1.45 | 3.58 | 1.25 | 1.31 |
| 0 | 1.54 | 2.75 | 0.36 | 1.92 | 5.25 | 3.42 | 1.08 | 7.36 | 4.13 |
| 0.94 | 3.65 | 0.8 | 1.36 | 4.1 |  |  |  |  |  |

**Nhóm 2:** *Tỉ lệ* ca tử vong trên tổng số ca bệnh do nCoV các nước châu Á

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 4.26 | 1.98 | 1.46 | 0.38 | 1.57 | 0.07 | 1.28 | 0.68 | 5.09 | 1.37 |
| 1.78 | 1.12 | 2.78 | 2.78 | 1.39 | 0.76 | 1.7 | 1.28 | 0.8 | 0.58 |
| 1.71 | 0.11 | 1.43 | 0 | 0.44 | 0.23 | 0.47 | 2.25 | 1.22 | 0 |
| 1.07 | 2.25 | 1.13 | 1.69 | 0.25 | 1.65 | 0.05 | 1.42 | 0.73 | 7.19 |
| 0.44 | 0.68 | 0.6 | 0.2 | 0.89 | 0 | 0.3 | 0.69 | 0.82 | 19.61 |

**Nhóm 3:** *Tỉ lệ* ca tử vong trên tổng số ca bệnh do nCoV của các nước châu Âu

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1.85 | 0.94 | 1.64 | 0.72 | 2.37 | 4.48 | 4.2 | 2.23 | 0.49 | 1.81 |
| 0.92 | 0.97 | 0 | 1.02 | 1.81 | 2.39 | 0 | 3.02 | 0 | 3.68 |
| 0.44 | 1.94 | 0 | 2.99 | 0 | 2.1 | 1.78 | 1.93 | 1.54 | 1.16 |
| 1.37 | 2.38 | 1.28 | 1.58 | 1.08 | 3.42 | 0.64 | 2.54 | 2.01 | 2.78 |
| 2.35 | 1.77 | 0.95 | 3.16 | 1.73 | 2.19 | 1.36 | 1.57 | 2.29 | 2.86 |
| 0 |  |  |  |  |  |  |  |  |  |

**Nhóm 4:** *Tỉ lệ* ca tử vong trên tổng số ca bệnh do nCoV của các nước bắc Mỹ

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 3.34 | 0 | 1.95 | 1.18 | 2.53 | 0 | 0 | 1.85 | 0 |
| 1.26 | 0.65 | 0 | 0 | 1.28 | 3.06 | 0 | 0.62 | 3.24 | 2.03 |
| 2.65 | 1.9 | 9.25 | 0 | 2.57 | .69 | 0 | 1.56 | 0.61 | 0 |
| 1.88 | 0 | 1.78 |  |  |  |  |  |  |  |

Với mức ý nghĩa 5% hãy so sánh tỉ lệ tử vong trên tổng số ca bệnh của 4 nhóm trên. Cho nhận xét. Giả sử tỉ lệ này của 3 nhóm tuân theo phân phối chuẩn và có cùng phương sai.

**PHẦN 2: CƠ SỞ LÝ THUYẾT**

**Dạng bài:** Kiểm định trung bình

**Phương pháp giải:** Phân tích phương sai 1 yếu tố

**Lý thuyết phân tích phương sai 1 yếu tố**:

Phép phân tích phương sai là so sánh trung bình của 2 hay nhiều nhóm dựa trên các giá trị trung bình của các mẫu quan sát từ các nhóm này, và thông qua kiểm định giả thiết để kết luận về sự bằng nhau của các trung bình tổng thể này.

**Phương pháp phân tích phương sai 1 yếu tố:**

Phép phân tích phương sai được dùng trong các trắc nghiệm để so sánh các giá trị trung bình của hai hay nhiều mẫu được lấy từ các phân số.

Đây có thể được xem như phần mở rộng các trắc nghiệm t hay z (so sánh hai giá trị trung bình).

Mục đích của sự phân tích phương sai một yếu tố là đánh giá sự ảnh hưởng của 1 yếu tố (nhân tạo hay tự nhiên) nào đó trên các giá trị quan sát, Yi(i=0,1,2,…,k)

**Điều kiện:**

Các giá trị được chọn 1 cách ngẫu nhiên và độc lập.

Mẫu N lớn hơn 30 hoặc tuân theo phân phối chuẩn.

Phương sai của các nhóm bằng nhau.

**Mô hình:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Yếu tố thí nghiệm | | | |  |
| 1 | 2 | ........ | k |
|  | Y11 | Y21 | ........ | Yk1 |  |
| Y12 | Y22 | ........ | Yk2 |
| .......... | ......... | ........ | ........... |
| Y1N | Y2N | ........ | YkN |
| Tổng cộng  trung bình | T1  \_ | T2  \_ | ........  ........ | Tk  \_ | T  \_ |

**Bảng Anova:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nguồn sai số | Bậc sai số | Tổng số bình phương | Bình phương trung bình | Gía trị thống kê |
| Yếu tố |  |  |  |  |
| Sai số |  |  |  |  |
| Tổng cộng |  |  |  |  |

Đặt giả thiết:

H0: μ1 = μ2 = ...... = μn ⭤ “Các giá trị trung bình bằng nhau”

H1: μi ≠ μj ⭤ “Ít nhất có hai giá trị trung bình khác nhau”

Tiêu chuẩn kiểm định: F = MSF/MSE

Biện luận: Nếu F < → Chấp nhận giả thiết H0.

**PHẦN 3: LỜI GIẢI TÍNH TAY**

**Giả thiết H0:** Tỉ lệ ca tử vong trên tổng số ca bệnh do nCov-19 của 4 châu lục là như nhau.

**Giả thiết H1:** Có ít nhất 2 châu lục mà tỉ lệ ca tử vong trên tổng số ca bệnh là khác nhau.

**Miền bác bỏ:** = ( = (2,65; +)

Trung bình mẫu từng nhóm:

1 = 2,14109

2 = 1,65260

3 = 1,72019

4 = 1,42061

Trung bình chung 4 mẫu:

Tổng chênh lệch bình phương cho từng nhóm:

Tổng chênh lệch bình phương trong các nhóm:

SSW = SS1 + SS2 + SS3 + SS4 = = 696,787

Tổng chênh lệch bình phương giữa các nhóm:

SSG = n1(1 - )2 + n2(2 - )2 +n3(3 - )2 + n4(4 - )2

= 12,41695

Phương sai giữa các nhóm:

MSG = = = 4,1390

MSW = = 3,76642

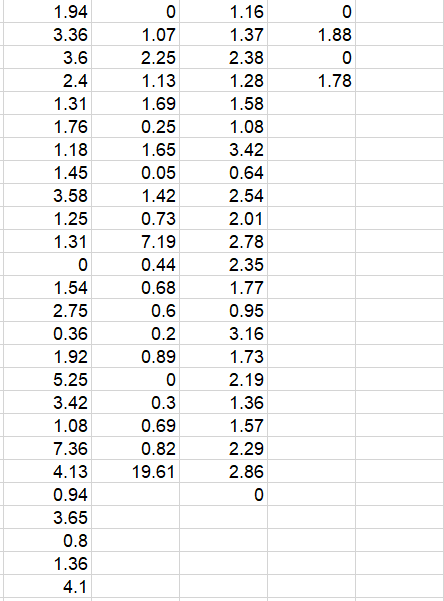
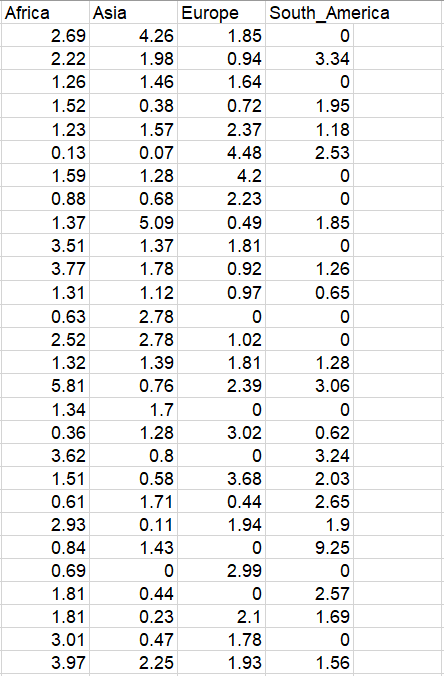
**Tiêu chuẩn kiểm định:** F = = = 1,0989

Vì nên chấp nhận giả thiết H0

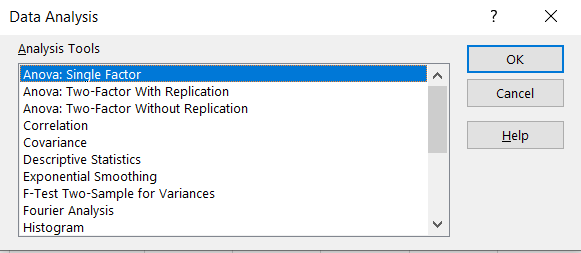
**Kết luận:** Tỉ lệ ca tử vong trên tổng số ca bệnh do nCov-19 của 4 châu lục là như nhau.

**PHẦN 4: GIẢI QUYẾT BÀI TOÁN TRÊN EXCEL**

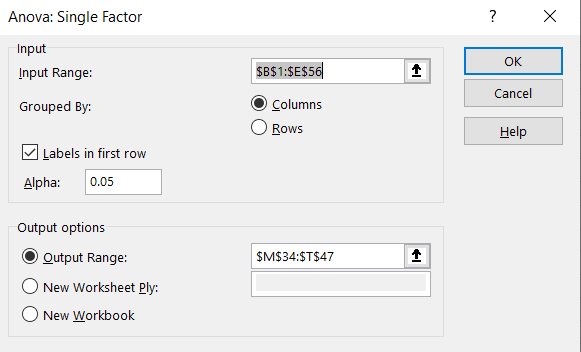
**Bước 1:** Nhập dữ liệu:

****

**Bước 2:** Sử dụng công cụ Data Analysis/Anova: Single Factor:



**Bước 3:** Chọn các mục như hình:



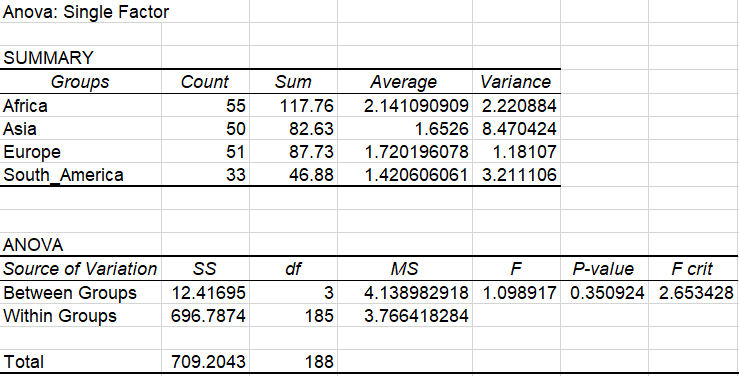
Input Range: Chọn địa chỉ tuyệt đối chứa dữ liệu

Output Range: Chọn vị trí xuất kết quả.

Chọn Labels in first row

Alpha: nhập 0.05 (mức ý nghĩa 5%)

Kết quả:



**Bước 4:** Biện luận:

**Giả thiết H0:** Tỉ lệ ca tử vong trên tổng số ca bệnh do nCov-19 của 4 châu lục là như nhau.

**Giả thiết H1:** Có ít nhất 2 châu lục mà tỉ lệ ca tử vong trên tổng số ca bệnh là khác nhau.

**Tiêu chuẩn kiểm định:** 1,09892

**Miền bác bỏ:** 2.65

Vì nên chấp nhận giả thiết H0

**Kết luận:** Vậy tỉ lệ tử trong trên tổng số ca nhiễm nCov-19 của 4 châu lục là như nhau. Dựa trên quả thu được, ta thấy tỉ lệ tử vong !..!.