

BÀI KIỂM TRA PYTHON

(ĐỀ 1)

Cho tập dữ liệu về các vận động viên tham gia Olympic. Tập tin "**athlete_events.csv**" chứa 271116 dòng và 15 cột. Mỗi dòng tương ứng với một vận động viên thi đấu trong một sự kiện Olympic cá nhân Olympic event (sự kiện vận động viên, athlete-events). Các cột là:

1. **ID** - Mã số duy nhất của mỗi vận động viên
2. **Name** - Tên vận động viên
3. **Sex** – Giới tính của vận động viên (M hoặc F)
4. **Age** – Tuổi của vận động viên
5. **Height** - Chiều cao của vận động viên (cm)
6. **Weight** – Cân nặng của vận động viên (kg)
7. **Team** – Tên đội
8. **Country** – Tên quốc gia của vận động viên
9. **Games** - Year và season
10. **Year** – Năm Olympic
11. **Season** – Mùa (Summer hoặc Winter)
12. **City** – thành phố đăng cai
13. **Sport** – Môn thể thao
14. **Event** – Sự kiện thể thao
15. **Medal** – Huy chương (Gold, Silver, Bronze, hoặc NA)

Yêu cầu:

1. Đọc dữ liệu từ tập tin CSV vào một DataFrame.
2. Loại bỏ bất kỳ dòng nào có giá trị thiếu (NaN) trong DataFrame.
3. Trung bình tuổi của các vận động viên tham gia Olympic là bao nhiêu?
4. Tổng số có bao nhiêu nam và nữ vận động viên tham gia Olympic?
5. Liệt kê 10 quốc gia có số lượng vận động viên tham gia Olympic nhiều nhất.
6. Vận động viên cao nhất và thấp nhất tham gia Olympic là ai, và họ thuộc quốc gia nào?
7. Tính trung bình cân nặng của nam và nữ vận động viên tham gia Olympic?
8. Sự kiện thể thao nào có số lượng vận động viên tham gia nhiều nhất?
9. Tìm năm Olympic có nhiều quốc gia tham gia nhất và in ra danh sách các quốc gia tham gia ở năm đó.
10. Tìm quốc gia có nhiều huy chương vàng nhất và số lần đạt huy chương vàng của quốc gia đó.
11. Trực quan hóa biểu đồ cột để thể hiện số lượng vận động viên tham gia Olympic theo từng năm.
12. Biểu đồ phân phối tuổi (biểu đồ Histogram) của vận động viên tham gia Olympic.
13. Tạo biểu đồ tròn để thể hiện phân phối huy chương (Gold, Silver, Bronze) trong tập dữ liệu.