**Tệp dữ liệu bao gồm 9 biến, với 3 nhóm chính:**

1. **Biến định danh và mô tả người dùng: id, full\_name, age, gender**
   1. **Id**

* Số lượng: 1 biến.
* Ý nghĩa: Mã định danh duy nhất cho mỗi người dùng trong dữ liệu.
* Loại dữ liệu: kiểu số nguyên (integer).
  1. **Full\_name**
* Số lượng: 1 biến.
* Ý nghĩa: Tên đầy đủ của người dùng.
* Loại dữ liệu: Biến chuỗi (string).
  1. **Age**
* Số lượng: 1 biến
* Ý nghĩa: Độ tuổi của người dùng. Thông tin này giúp phân khúc người dùng theo nhóm tuổi, hỗ trợ trong việc nhắm đến đối tượng quảng cáo phù hợp.
* Loại dữ liệu: Biến số học (numerical) - kiểu số nguyên (integer).
  1. **Gender**
* Số lượng: 1 biến
* Ý nghĩa: Giới tính của người dùng, có thể được mã hóa như "M" (nam) và "F" (nữ), hoặc bằng các số đại diện (ví dụ, 0 cho nữ và 1 cho nam).
* Loại dữ liệu: Biến phân loại (categorical) - thường là chuỗi (string) hoặc số nguyên (integer).

1. **Biến thiết bị và hành vi: device\_type, ad\_position, browsing\_history, time\_of\_day**
   1. **Device\_type**

* Số lượng: 1 biến
* Ý nghĩa: Loại thiết bị mà người dùng sử dụng để xem quảng cáo. Bao gồm "Mobile", "Desktop", hoặc "Tablet".
* Loại dữ liệu: Biến phân loại (categorical) - dạng chuỗi (string).
  1. **Ad\_position**
* Số lượng: 1 biến
* Ý nghĩa: Vị trí của quảng cáo mà người dùng thấy, bao gồm “top”, “bottom”, “side”. Biến này giúp đánh giá xem vị trí quảng cáo ảnh hưởng đến tỷ lệ nhấp chuột như thế nào.
* Loại dữ liệu: Biến phân loại (categorical) - dạng chuỗi (string).
  1. **Browsing\_history**
* Số lượng: 1 biến
* Ý nghĩa: Lịch sử duyệt web của người dùng, danh sách các loại trang web hoặc chủ đề mà người dùng đã truy cập gần đây bao gồm “Entertaiment”, “Social media”, “Education”, “Shopping”, “News”.. Biến này giúp hiểu rõ sở thích hoặc hành vi duyệt web của người dùng.
* Loại dữ liệu: Biến danh sách (list), vì một người dùng có thể truy cập nhiều trang khác nhau.
  1. **Time\_of\_day**
* Số lượng: 1 biến
* Ý nghĩa: Thời điểm trong ngày mà người dùng thấy quảng cáo. Bao gồm “morning”, “afternoon”, “evening”, “night”. Thông tin này giúp xác định các khoảng thời gian mà người dùng có xu hướng nhấp chuột cao hơn.
* Loại dữ liệu: kiểu dữ liệu phân loại (category)

1. **Biến kết quả: click**

* Số lượng: 1 biến
* Ý nghĩa: Kết quả nhấp chuột của người dùng với quảng cáo, được mã hóa là 1 nếu người dùng nhấp chuột vào quảng cáo và 0 nếu không. Biến này là mục tiêu trong các phân tích liên quan đến tỷ lệ nhấp chuột.
* Loại dữ liệu: Biến nhị phân (binary) - kiểu số nguyên (integer), chỉ có hai giá trị (0 hoặc 1).