1. **Lĩnh vực**: Tệp dữ liệu này thuộc lĩnh vực quảng cáo kỹ thuật số và phân tích hành vi người dùng. Nó chứa thông tin về các tương tác của người dùng với quảng cáo, giúp tìm hiểu thói quen và sở thích của họ khi duyệt web trên các thiết bị khác nhau.
2. **Nguồn thu thập**: Tệp dữ liệu này có thể được thu thập từ các website, các ứng dụng di động, các mạng xã hội hoặc các dịch vụ quảng cáo như GG ads.
3. **Cách thu thập:**
   1. **Cookie và mã theo dõi**: Dữ liệu có thể được thu thập bằng cách sử dụng cookie và mã theo dõi (tracking pixel) để ghi lại hành vi của người dùng trên website.
   2. **Thông tin từ thiết bị**: Thiết bị truy cập có thể được xác định qua User-Agent (thông tin về thiết bị và trình duyệt) để hiểu rõ loại thiết bị người dùng đang sử dụng.
   3. **Thời gian và vị trí**: Thời điểm truy cập và vị trí quảng cáo thường được thu thập tự động từ hệ thống khi người dùng truy cập vào trang chứa quảng cáo.
4. **Mục đích:**
   1. Dự đoán người dùng có nhấp vào quảng cáo hay không dựa trên thông tin người dùng, hành vi, cách hiển thị quảng cáo và thời gian trong ngày.
   2. Tối ưu hóa quảng cáo theo các tiêu chí:
      1. *Phân khúc khách hàng*: Dữ liệu giúp phân nhóm người dùng dựa trên độ tuổi, giới tính, thiết bị sử dụng, để quảng cáo hiệu quả hơn.
      2. *Phân tích hiệu suất quảng cáo*: Đo lường tỷ lệ nhấp chuột cho từng vị trí quảng cáo và thiết bị, từ đó điều chỉnh chiến lược quảng cáo.
5. Loại dữ liệu:
   1. Dữ liệu định danh: id, full\_name, age, gender
   2. Dữ liệu thiết bị: devide\_type
   3. Dữ liệu tương tác: ad\_position, browsing\_history, time\_of\_day
   4. Dữ liệu kết quả: click (0 hoặc 1)