

Thực hành mini-project -A/B testing-

TS.Tô Đức Khánh

01/04/2024

Thực hành A/B testing với hai projects sau đây.

1 Project 1 - Thương mại điện tử (E-commerce A/B Testing)

Đối với dự án này, bạn sẽ làm việc để hiểu kết quả của thử nghiệm A/B do một trang web thương mại điện tử thực hiện. Công ty đã phát triển một trang web mới nhằm cố gắng tăng số lượng người dùng “chuyển đổi”, nghĩa là số lượng người dùng quyết định trả tiền cho sản phẩm của công ty. Mục tiêu của bạn là nghiên cứu sổ ghi chép này để giúp công ty hiểu liệu họ có nên triển khai trang mới này, giữ lại trang cũ hay có thể chạy thử nghiệm lâu hơn để đưa ra quyết định.

Dữ liệu của 294478 người dùng được tổng hợp trong file dữ liệu `ab_test_commerce.csv`, các biến bao gồm:

- `id` - một số duy nhất xác định mỗi người dùng.
- `time` - thời gian hoạt động với trang web.
- `con_treat` - người dùng được đưa vào control group (`control` - sử dụng `old_page`) hay treatment group (`treatment` - sử dụng `new_page`).
- `page` - trang web mà người dùng được sử dụng.
- `converted` - tỷ lệ giao dịch.

Hãy load dữ liệu và thực hiện một bản báo cáo cho nhà quản lý và phát triển trang web, với các nội dung như sau:

1. Lưu ý rằng, dữ liệu này có sự sai khác giữa số lượng người trong hai nhóm control-treatment với số lượng người sử dụng `old_page`-`new_page`, do đó, cần hiệu chỉnh/làm sạch dữ liệu trước khi phân tích.
2. Bảng tóm tắt, khái quát về dữ liệu.
3. Trang web mới có thực sự tốt hơn trang web cũ.
4. Viết các nhận xét và kết luận.

2 Project 2 - DataCamp

Cookie Cats là trò chơi giải đố cực kỳ nổi tiếng trên thiết bị di động được phát triển bởi Tactile Entertainment. Đây là một trò chơi giải đố kiểu “kết nối ba” cổ điển trong đó người chơi phải kết nối các ô cùng màu để xóa bảng và giành được cấp độ. Nó cũng có tính năng ca hát mèo.

Khi người chơi tiến bộ trong trò chơi, họ sẽ gặp phải các cổng buộc họ phải đợi một thời gian trước khi có thể tiến bộ hoặc thực hiện mua hàng trong ứng dụng. Trong dự án này¹, chúng tôi sẽ phân tích kết quả của

¹<https://www.datacamp.com/projects/184>



Hình 1: Cookie Cats.

thử nghiệm A/B trong đó cổng đầu tiên trong Cookie Cats được chuyển từ cấp 30 lên cấp 40. Đặc biệt, chúng tôi sẽ phân tích tác động đến việc giữ chân người chơi.

Dữ liệu `cookie_cats.csv` cung cấp thông tin của 90,189 người chơi đã cài đặt trò chơi trong khi kiểm định A/B testing được tiến hành. Các biến là

- **userid** - một số duy nhất xác định mỗi người chơi.
- **version** - người chơi được đưa vào control group (`gate_30` - bắt đầu với cổng ở cấp 30) hay treatment group (`gate_40` - bắt đầu với cổng ở cấp 40).
- **sum_gamerounds** - số vòng trò chơi mà người chơi đã chơi trong 2 tuần đầu tiên sau khi cài đặt.
- **retention_1** - người chơi có quay lại và chơi sau 1 ngày cài đặt không?
- **retention_7** - người chơi có quay lại và chơi sau 7 ngày cài đặt không?

Hãy load dữ liệu và thực hiện một bản báo cáo cho nhà quản lý và phát triển game, với các nội dung như sau:

1. Bảng tóm tắt, khái quát về dữ liệu.
2. Tỷ lệ người chơi quay lại chơi game sau khi cài đặt.
3. Việc thay đổi cổng bắt đầu có làm thay đổi số lượng vòng trò chơi mà người chơi đã chơi hay không?
4. Tỷ lệ người chơi quay lại chơi game sau khi cài đặt, có bị ảnh hưởng bởi việc thay đổi cổng hay không?
5. Viết các nhận xét và kết luận.