A/B test on Search System Second Hand APP

Marc Roig Lama

Context

- Analysis of **Event Logging** dataset. (events_log.csv)

Tasks

- What is our **daily overall click through rate**? How does it vary between the groups?

- Which **results do people tend to try first**? How does it change day-to-day?

- What is our **daily overall zero results rate**? How does it vary between the groups?

- Choose a variable (**group A/B**) from the dataset and describe its relationship to **session length**.

Code Workflow Overview

Read Data
Using Pandas. Perform some timestamp processing and sort by values.
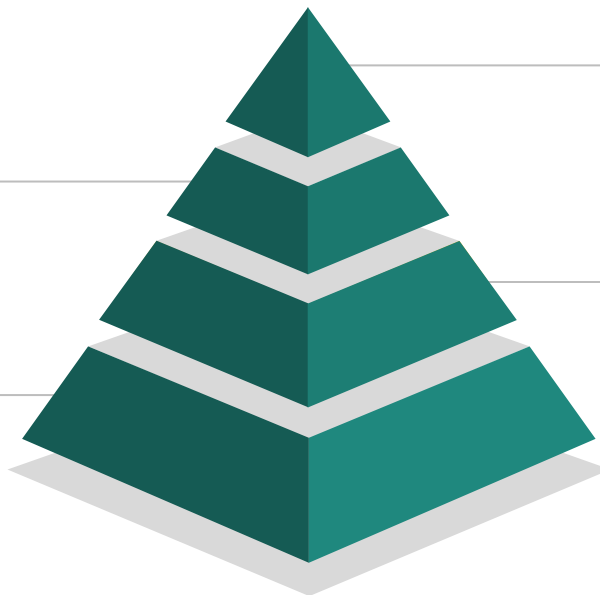
Clean Data
Delete page visits without link number. Remove sessions where the link visited is outside the number of search results. Sum up checkin logs into the last checkin to reduce size and erase redundant data.

Data Analysis
Get an overview for every column.

Metric Analysis
Compute metrics and visualizations to get conclusions.

Inferential statistics
Pending!
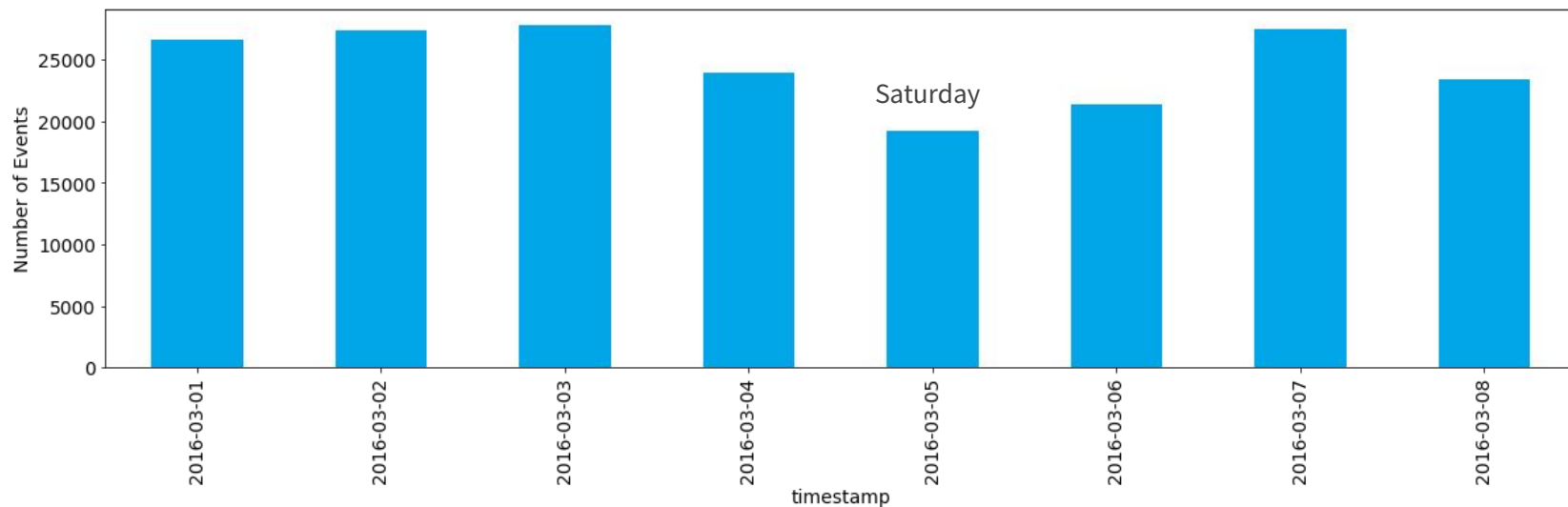
Data Overview

400k raw events

197k cleaned events

1 week

67 k user sessions

A/B ANALYSIS

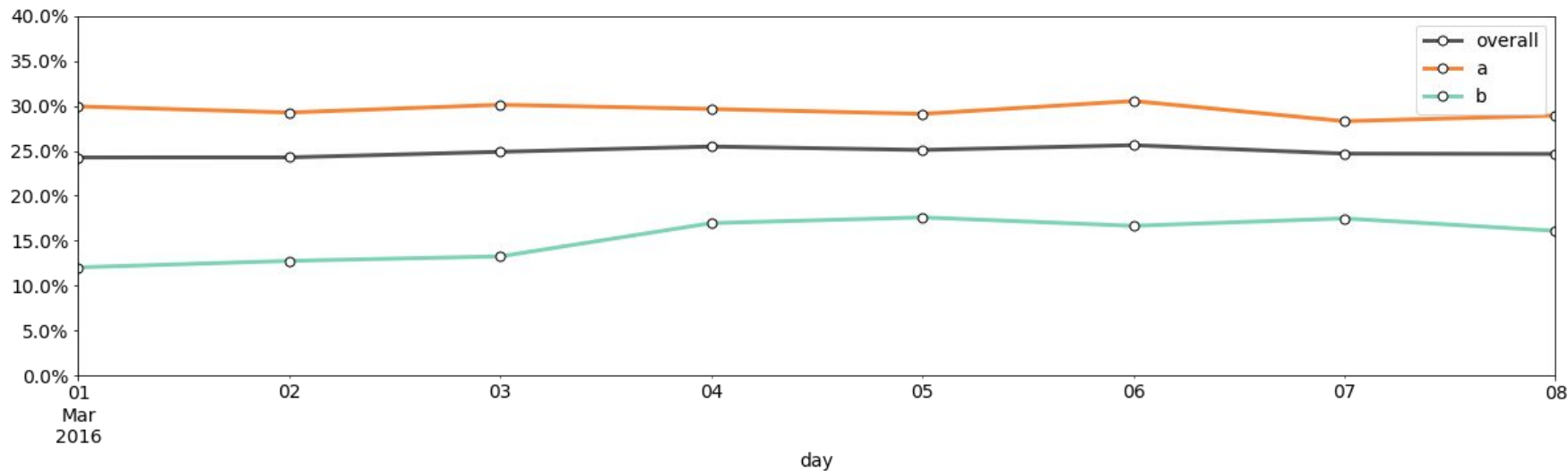Sessions by Group (%)

A
43,2%
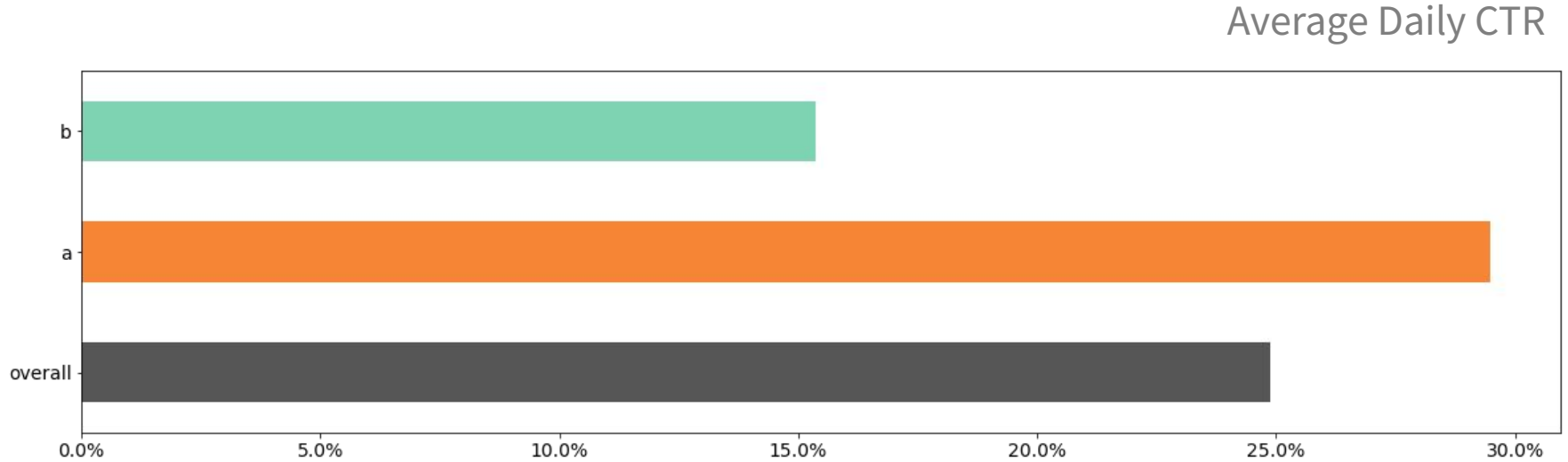29K

B
56,8%
38K
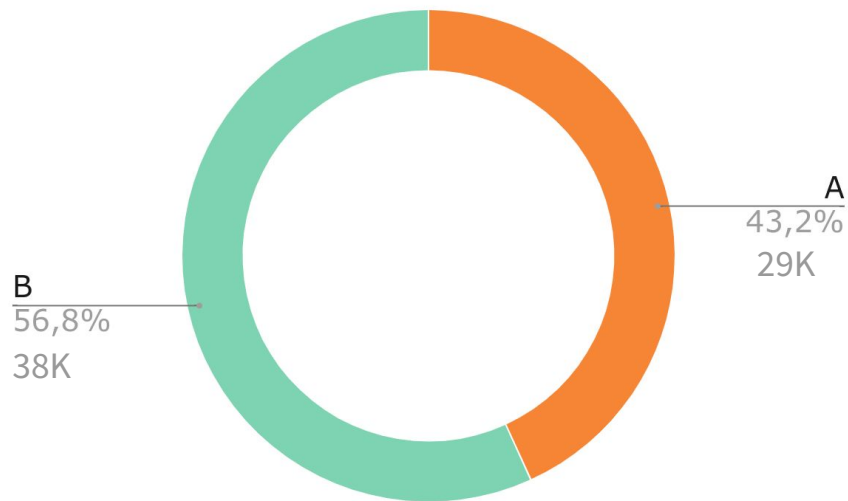
Let's explore 4 different metrics:

1. Clickthrough Rate (CTR)
2. Visited Pages
3. Zero Results Rate (ZRR)
4. Session Length

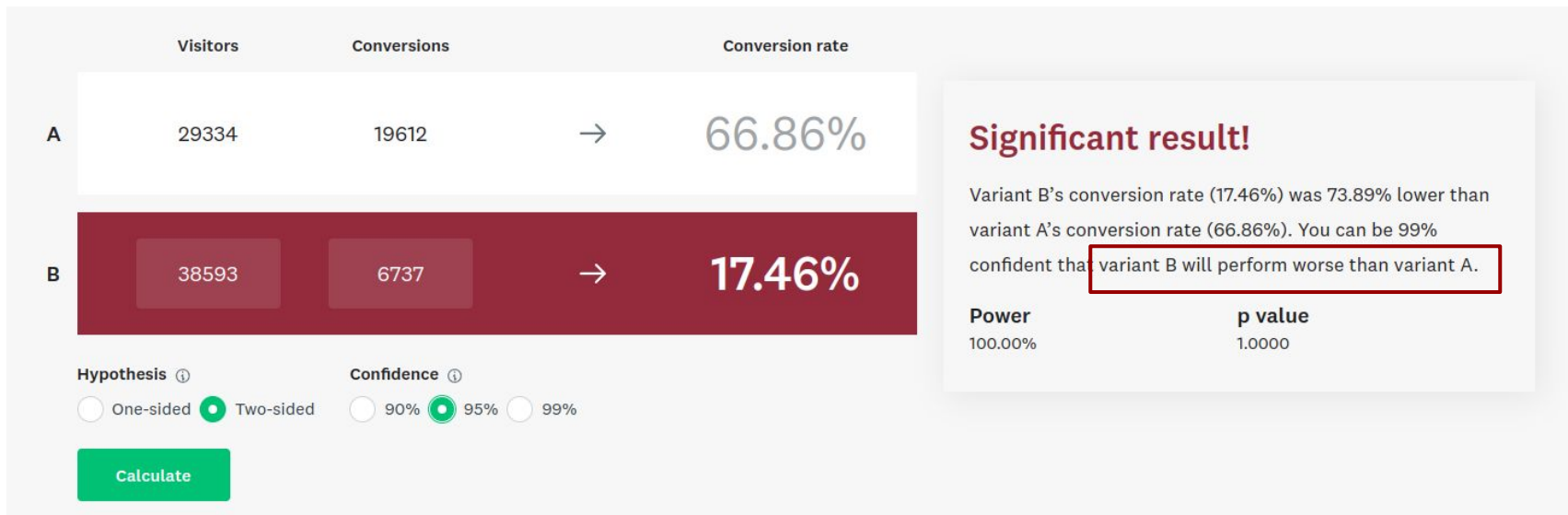# Clickthrough Rate (CTR)

Average Daily CTR

Clickthrough Rate (CTR)

- The daily overall CTR is **25%**. So 25% of the users that perform a search will visit a result on daily basis.
- On average, A users are visiting **14%** more links than B users.
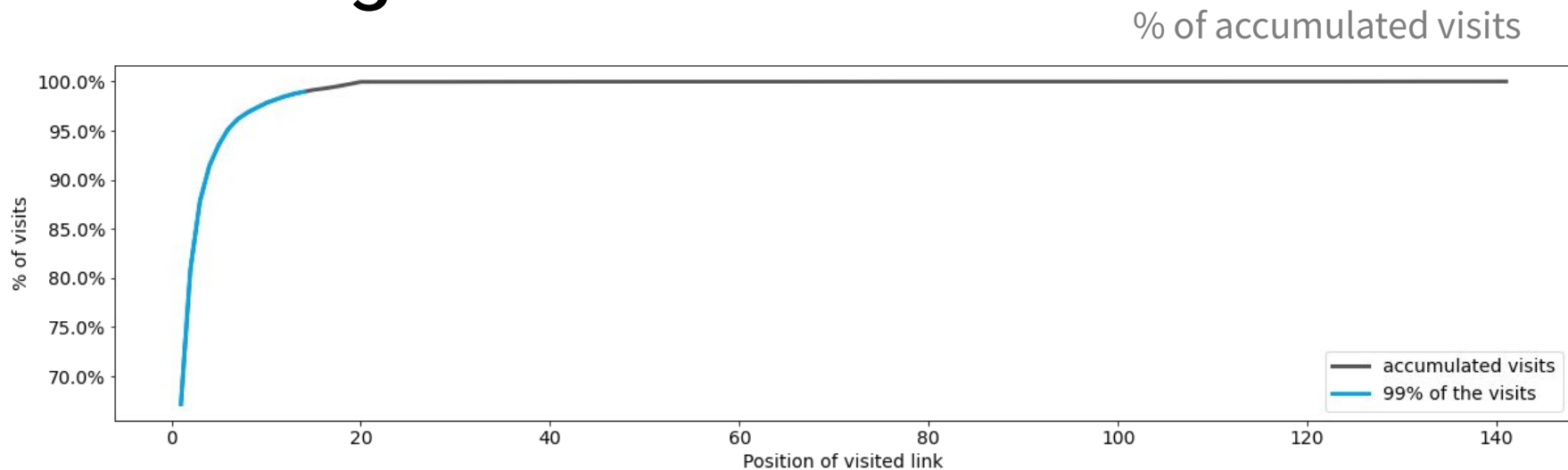
Statistical Significance on Visited Pages

Overall Conversion (Visited Pages over sessions) p-value.

	Visitors	Conversions		Conversion rate
A	29334	19612	→	66.86%
B	38593	6737	→	17.46%

Hypothesis ⓘ
○ One-sided ● Two-sided

Confidence ⓘ
○ 90% ● 95% ○ 99%

[Calculate]

Significant result!

Variant B's conversion rate (17.46%) was 73.89% lower than variant A's conversion rate (66.86%). You can be 99% confident that variant B will perform worse than variant A.
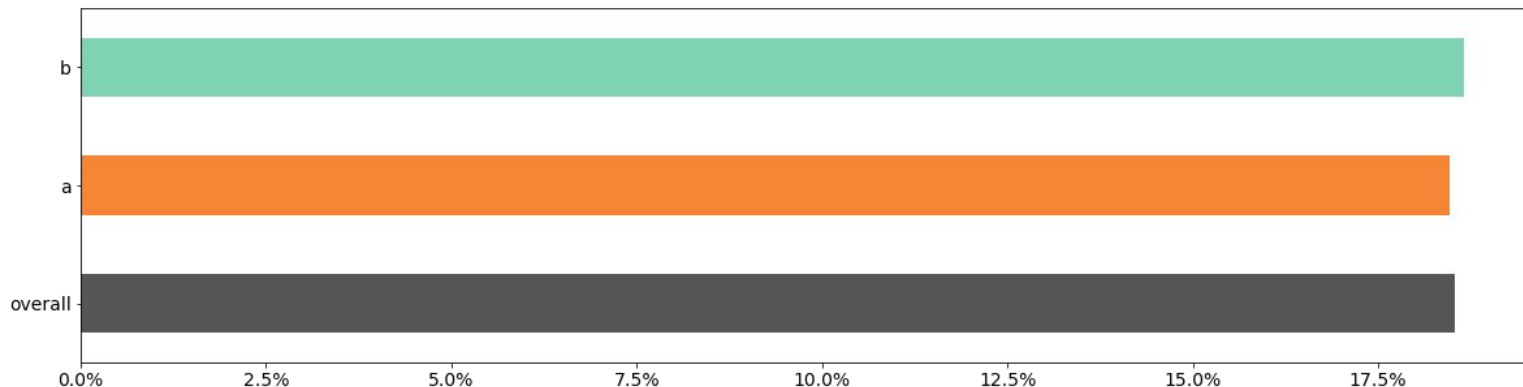
Power
100.00%

p value
1.0000

Visited Pages

- The 1st result is the most visited (**67%** of the sessions) at the first try.
- 90% of the sessions who opens a link area under the 4th result.
- The distribution does not change depending on the day.

Zero Results Rate (ZRR)

- The daily overall ZRR is **18%**. 18% of the searches doesn't return any result.
- No significant difference between groups.

Session Length

- **QUESTION: is the length of your session affected by the new feature?**
- NULL HYPOTHESIS: the session length is not affected by the new feature.
- ALTERNATIVE HYPOTHESIS: the session length is affected by the new feature.

Session Length Distribution
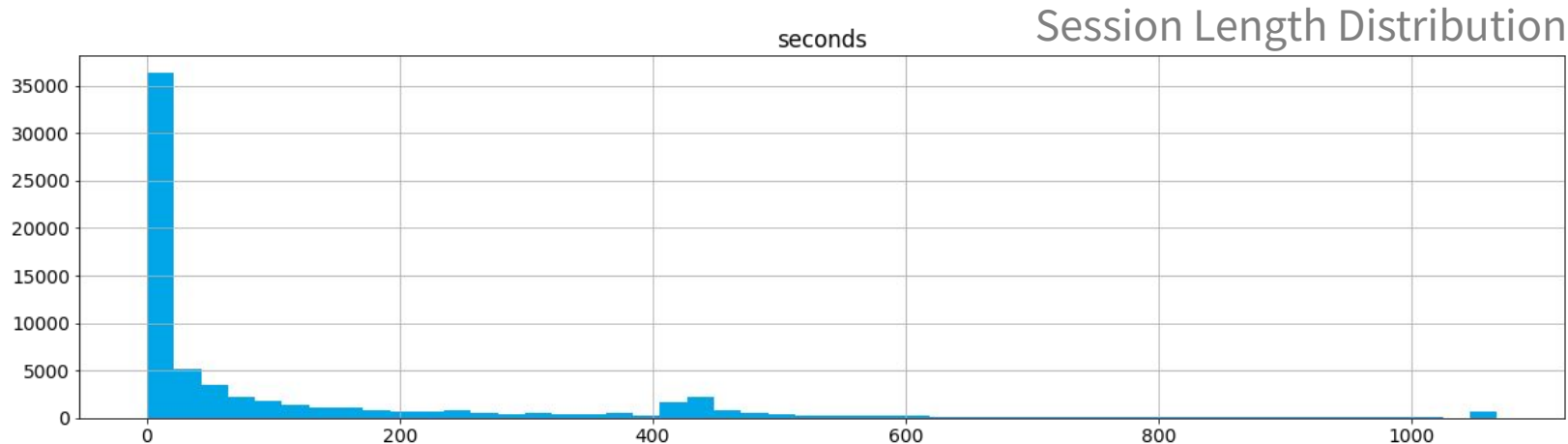
Session Length with 0 seconds sessions

Distribution by group
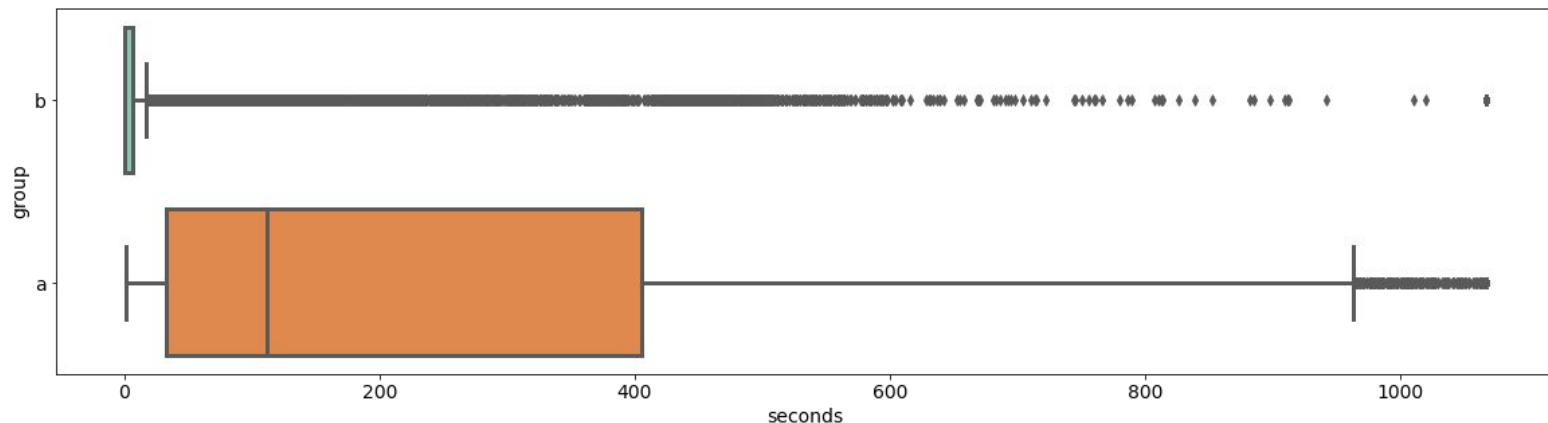
- A users will spend 225s on average.
- B users will spend 40s on average
- **72%** of the B sessions last 0s.
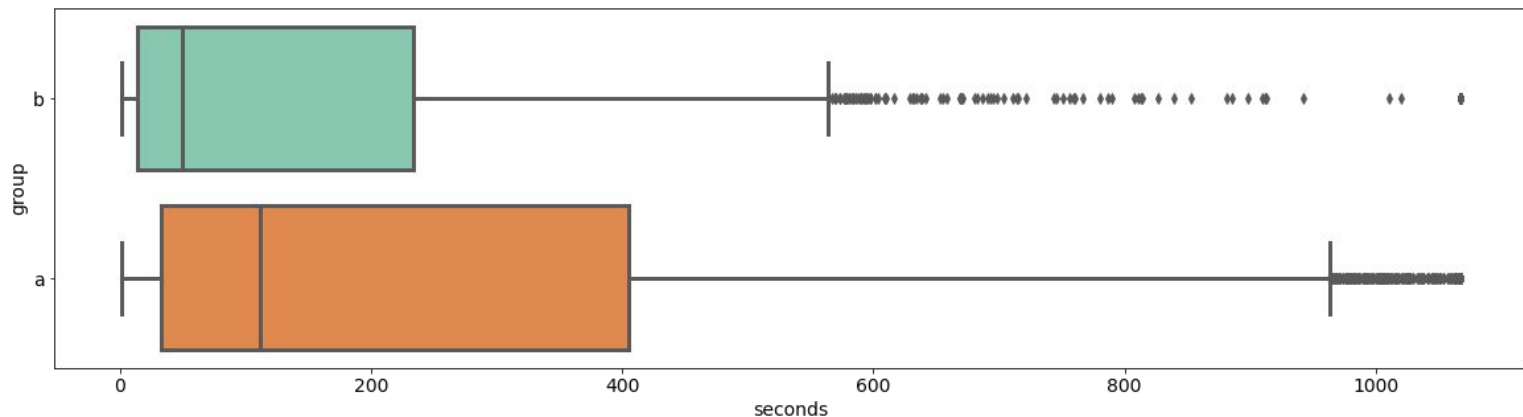- On average, A users will spend **185s** more than a B user on the app.

0 seconds sessions

If we consider 0 seconds sessions a bug in the logging system we can repeat the analysis **avoiding** those sessions on group B.

Session Length without 0 seconds sessions

Distribution by group

- A users will spend 225s on average.
- B users will spend 141s on average.
- On average, A users will spend **84s** more than a B user.

Conclusions

- Group A is performing better than B (+14%) in terms of CTR.

- ZRR is not affected by the new feature B.

- Group A is spending more time (+84s) on the app.

- If we are interested on **user engagement** (time) and **conversion** (clicks on ads) we should keep feature A.

- We have to make sure that this numbers are **statistically significant**: p-value test

Thank you!

in [Marc Roig](#) 👋

marcroiglama@gmail.com