[Mobile] UI Feature X A/B Experiment Analysis

Patricia Luk

Context

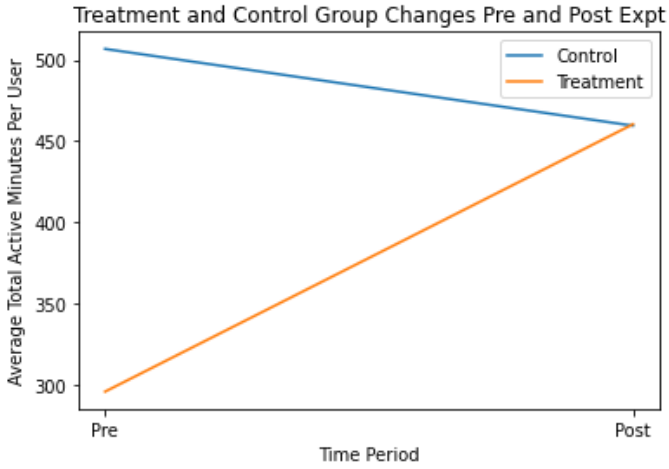
The Mobile team is exploring the potential impact of feature X. The primary metric of interest is engagement, defined as total active minutes per user. 12 subgroups were analyzed - a combination of gender (male, female, unknown) and user type (new user, reader, non-reader, contributor).

Findings

Treatment assignment was imbalanced on

- user type: treated users are more likely to be non-readers and new users and less likely to be readers and contributors[2]
- engagement: pre-treatment, the mean engagement of the treated group is about half that of the control group[1]

These may be correlated - readers and contributors may signal more highly engaged users. No gender imbalance was observed.



We observe

- statistically significant (95% confidence) increases in engagement in all subgroups except male new users and unknown new users[2]
- no statistically significant decreases in engagement in any subgroups[2]

Average Treatment Effect of the Treated (ATT) by Subgroup				
Gender	**User Type**	**t value**	**CI Lower**	**CI Upper**
Male	Non-reader	9.659	61.20	92.37
	Reader	11.17	740	1055
	New User	1.93	-0.43	56.7
	Contributor	3.60	923	3138
Female	Non-reader	9.34	49.2	75.381
	Reader	6.07	461	900.0
	New User	2.30	3.24	40.5
	Contributor	2.37	335	3661
Unknown	Non-reader	4.88	28.4	66.50
	Reader	5.75	545	1111
	New User	1.37	-10.0	56.5
	Contributor	2.53	629	5210

See appendix for detailed distributions by subgroup[1].

Recommendations

Because treatment assignment was imbalanced on user type and pre-treatment engagement, preliminary results are inconclusive although promising.
I recommend re-running the experiment with modification to the treatment assignment process to ensure covariate balance on pre-treatment engagement and user type.

Appendix

[1] Python notebook
[2] R script

Active Minutes Before and After Test: Distribution Close Look