MuscleHub A/B Test

To fitness test or not to fitness test . . . that is the question

Christa Seeberger March 2018

Background of the A/B test

The task: Determine whether a fitness test was encouraging or discouraging membership sign-up using data analytics.

The process for signing up:

1. Take a fitness test with a personal trainer
2. Fill out an application for the gym
3. Send in their payment for their first month's membership

How the test was conducted:

1. Randomly divided the 5004 visitors into A and B groups. A would take a fitness test, B would not
2. Using data analytics we wanted to show that B group would be more likely to purchase a membership
3. This presentation will show the outcome of the analytics and make a proposal

Research:

Customer interviews did not provide a clear preference for either A or B group

“I always wanted to work out like all of the shredded people on the fitness accounts I see on Instagram, but I never really knew how to start. MuscleHub’s introductory fitness test was super helpful for me! After taking the fitness test, I had to sign up and keep coming back so that I could impress my trainer Rachel with how much I was improving!” - Cora, 23, Hoboken

“When I walked into MuscleHub I wasn’t accosted by any personal trainers trying to sell me some mumbo jumbo, which I really appreciated. Down at LiftCity they had me doing burpees 30 seconds after I walked in the door and I was like “woah guys slow your roll, this is TOOOO much for Jesse!” I still ended up not signing up for a membership because the weight machines had all those sweat stains on them and you know, no thanks.” - Jesse, 35, Gowanes

“I took the MuscleHub fitness test because my coworker Laura recommended it. Regretted it.” - Sonny "Dad Bod", 26, Brooklyn

“I saw an ad for MuscleHub on BookFace and thought I'd check it out! The people there were suuuuuper friendly and the whole sign-up process took a matter of minutes. I tried to sign up for LiftCity last year, but the fitness test was way too intense. This is my first gym membership EVER, and MuscleHub made me feel welcome.” - Shirley, 22, Williamsburg

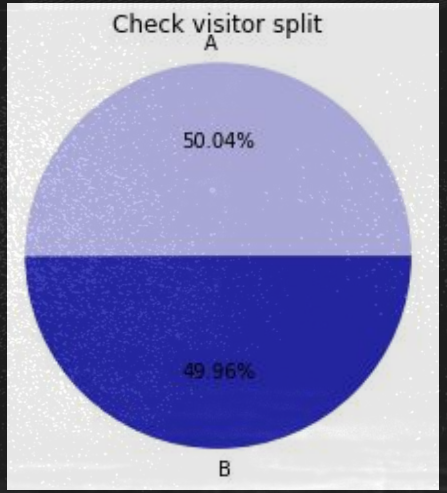
First steps of the analysis

1. Combine 4 data sets provided: Using SQL we left joined data from visits, fitness tests, purchases and applications to provide 5004 rows of data. This is a sample of the joined data sets:

	first_name	last_name	visit_date	fitness_test_date	application_date	purchase_date
0	Kim	Walter	7-1-17	2017-07-03	None	None
1	Tom	Webster	7-1-17	2017-07-02	None	None
2	Edward	Bowen	7-1-17	None	2017-07-04	2017-07-04
3	Marcus	Bauer	7-1-17	2017-07-01	2017-07-03	2017-07-05
4	Roberta	Best	7-1-17	2017-07-02	None	None
5	Joseph	Foley	7-1-17	None	None	None

2. Ensure we have split the data set equally: We imported Pandas to group the data and Matplotlib to create a pie chart and found the data to be equally represented:

	ab_test_group	first_name
0	A	2504
1	B	2500

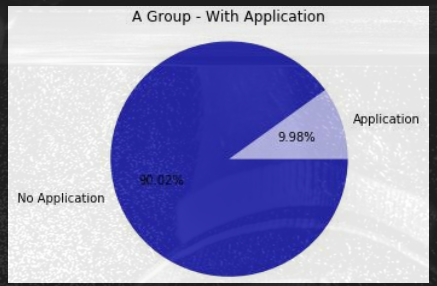


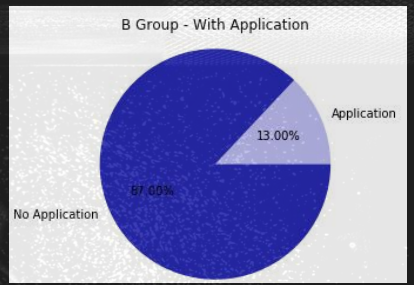
3rd step of the analysis:

Of groups A (took a fitness test) and B (did not take a fitness test) - who filled out an application?

We grouped A and B according to application and no application. With 13% applications in the B group it looks like not offering a fitness test creates more applications. However, we need to understand whether this difference is statistically significant.

is_application	ab_test_group	Application	No Application	Total	Percent with Application
0	A	250	2254	2504	0.09984
1	B	325	2175	2500	0.13000





As we are using categorical data, we use chi2 contingency to test the statistical significance. A p-value of less than 0.05 is required to prove the null hypothesis*. With a p-value of 0.0009 we can say the outcome IS significant.

P-value = 0.0009647827600722304

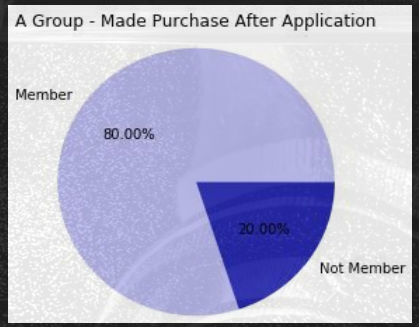
**The null hypothesis: a p-value of more than 0.05 means that there is no significant difference between datasets, i.e. the result is NOT significant*

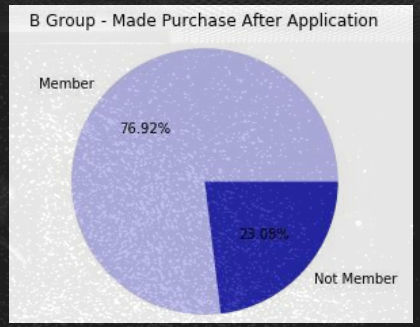
4th step of the analysis:

Of those who made an application, who became a member?

We grouped applications made by A and B by who became a member. In this case, it seems as if A (took a fitness test) acquire higher membership rates! However, again, we need to understand whether this difference is statistically significant.

is_member	ab_test_group	Member	Not Member	Total	Percent Purchase
0	A	200	50	250	0.800000
1	B	250	75	325	0.769231





Using the chi2 contingency test again, we can see that the p-value is 0.43. As this is higher than a p-value of 0.05 we can say the outcome is NOT significant and the null hypothesis is true.

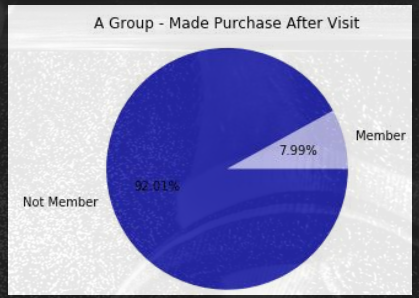
P-value = 0.43258646051083327

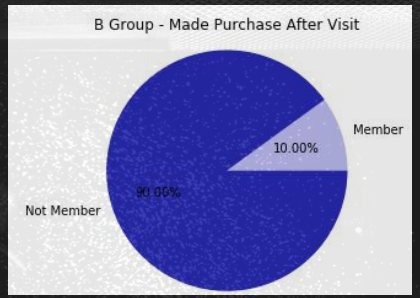
5th step of the analysis:

Of total visitors to the gym, who became a member?

Instead of grouping by applications we should rather group by all visitors to the gym to assess whether A (took a fitness test) became a member vs B (did not take a fitness test) became a member.

is_member	ab_test_group	Member	Not Member	Total	Percent Purchase
0	A	200	2304	2504	0.079872
1	B	250	2250	2500	0.100000

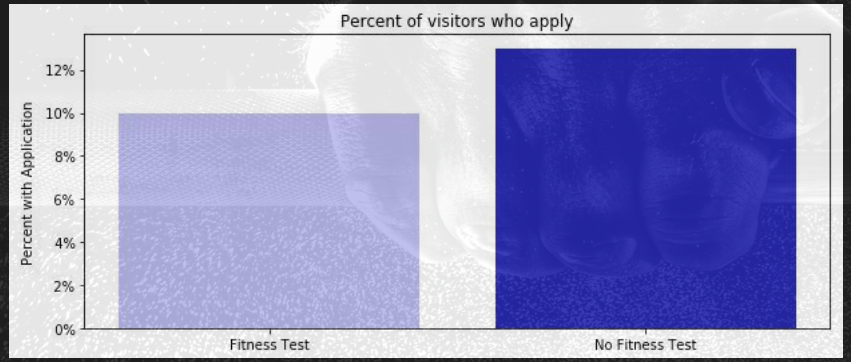


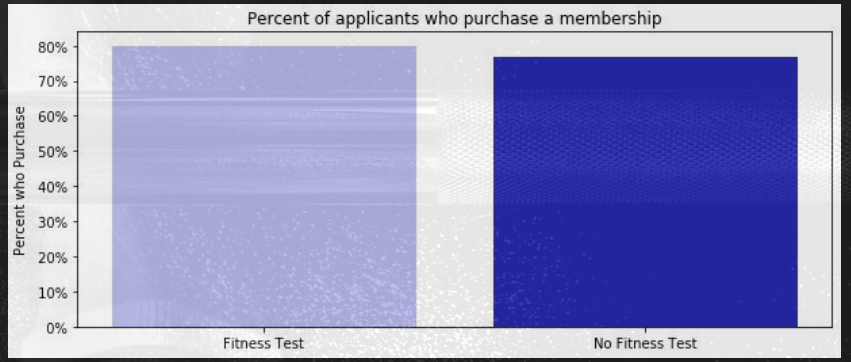


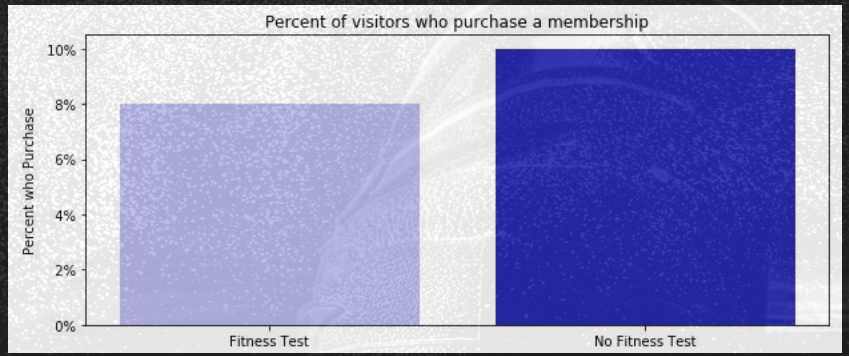
Using the chi2 contingency test again, we can see that the p-value is 0.014. As this is less than a p-value of 0.05 we can say the outcome IS significant and the null hypothesis is false.

P-value = `0.014724114645783203`

Summary graphs of the qualitative data







Recommendation for MuscleHub

1. Based on the data given, there is a clear preference for not having to take a fitness test, as do some of the interviews. We therefore recommend not enforcing a fitness test going forward.

2. We also understand that there was take-up on memberships when a fitness test was done. Therefore alternatives to a mandated fitness test could include:

    a. Offering an optional fitness test that shows a customer where and how improvements can be made in the most effective manner, which may improve the lifetime value of that customer;
    b. Offering an optional induction session for health and safety;
    c. Free training and menu planners for specific fitness categories;
    d. Investment in tracking fitness progress as a customer logs in during a training session.

These alternatives would offer additional data points for us to evolve this initial analysis and make further recommendations.