Description

For this test we wanted to use empirical data to prove whether presenting prospective members with a fitness test encouraged or discouraged eventual membership signup.

During our data gathering phase visitors will randomly be assigned to one of two groups:

- **Group A** will still be asked to take a **fitness test with a personal trainer**

- **Group B** will **skip the fitness test** and proceed directly to the application

Subsequently, we will analyze the membership data of each group and determine which approach was ultimately more effective. If possible, we will report back to MuscleHub our recommendation on which approach we believe is more effective and that they should continue to use.

Interviews

Janet, the manager, suspects that the fitness test intimidates prospective members based upon some feedback she gathered from interviews:

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
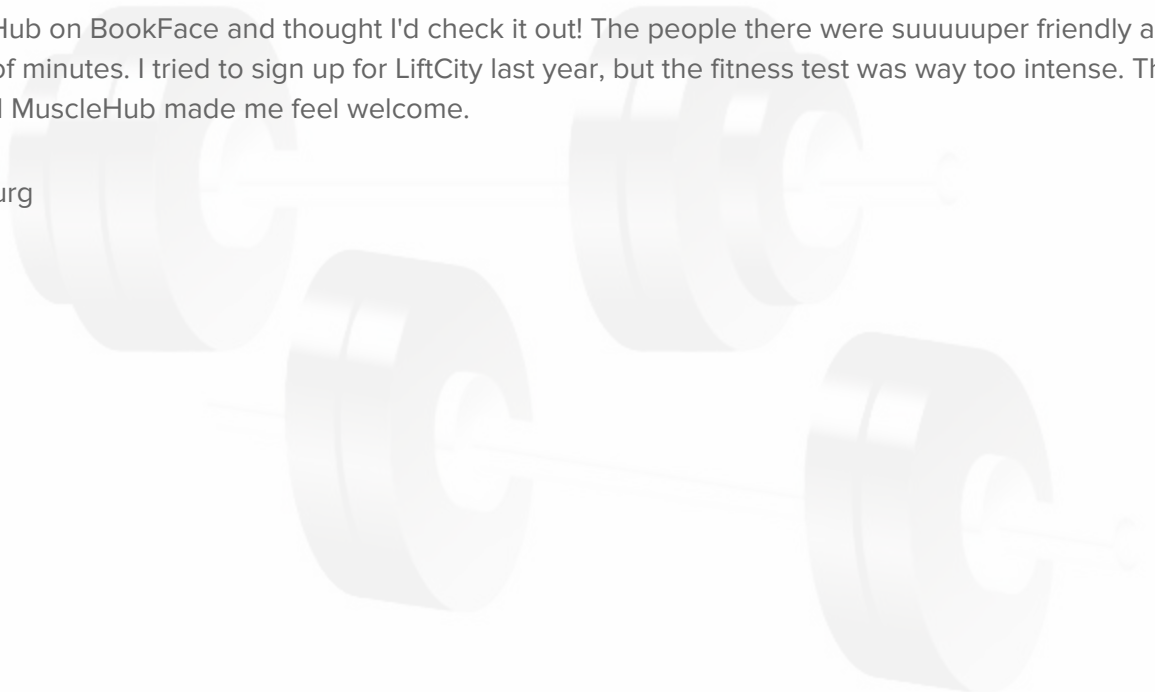
The Data

While the interviews are important, the small subset makes it hard to trust for making decisions. Data ultimately speaks much louder than words. For that reason we recorded 4 datasets

- **visits** contains information about potential gym customers who have visited MuscleHub
- **fitness_tests** contains information about potential customers in "Group A", who were given a fitness test
- **applications** contains information about any potential customers (both "Group A" and "Group B") who filled out an application. Not everyone in visits will have filled out an application.
- **purchases** contains information about customers who purchased a membership to MuscleHub.

Each dataset contains name, email, and date. The name and email allow us to join together the datasets. The dates in each dataset will let us know which event took place and when. Ultimately we will slice and dice the data, run some statistical tests and then, hopefully, we can determine which group was more effectively targeted.

Now let's take a look at the data in more detail. There are 5004 rows of **visit** data collected and represents all the people in our study. The other tables **fitness test**, **application**, and **purchases** may or may not have happened but all rows/people in those tables are also in the **visits** table. Thus we can left join the smaller subset tables to **visits** to give us a merged table that looks something like this:

	first_name	last_name	gender	email	visit_date	fitness_test_date	application_date	purchase_date
0	Kim	Walter	female	KimWalter58@gmail.com	7-1-17	2017-07-03	None	None
1	Tom	Webster	male	TW3857@gmail.com	7-1-17	2017-07-02	None	None
2	Edward	Bowen	male	Edward.Bowen@gmail.com	7-1-17	None	2017-07-04	2017-07-04
3	Marcus	Bauer	male	Marcus.Bauer@gmail.com	7-1-17	2017-07-01	2017-07-03	2017-07-05
4	Roberta	Best	female	RB6305@hotmail.com	7-1-17	2017-07-02	None	None

As can been seen in this sample, each row represents a person/visit and has name, email and visit date. Because fitness test, application, and purchase may or may not have taken place the date for any one is either populated with a date of that event or null or "None" (and did not happen) as represented in our data frame.

Now we should be able to group the data to get a better idea of what it all means.

Since a visitor either did or did not take a fitness test we can merely look for the presence of a *fitness_test_date* and set another column that represents our two group: **A** - did take the test and **B** - did not take the test.

Test Group A

50.04%

49.96%

Test Group B
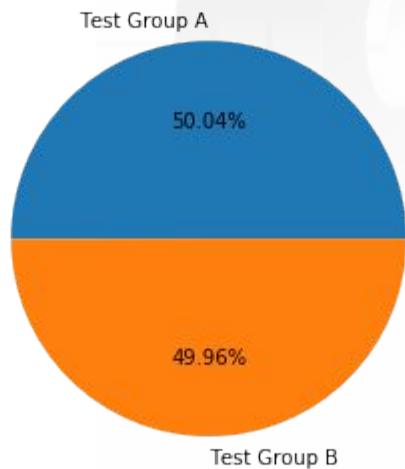
A count of each group show that the groups are very close to equally represented.

Now let's take a look at applications. Because Group B (did not take the fitness test) proceeded directly to the application step we could expect that their application rate would be higher than Group A. The data bears that out:
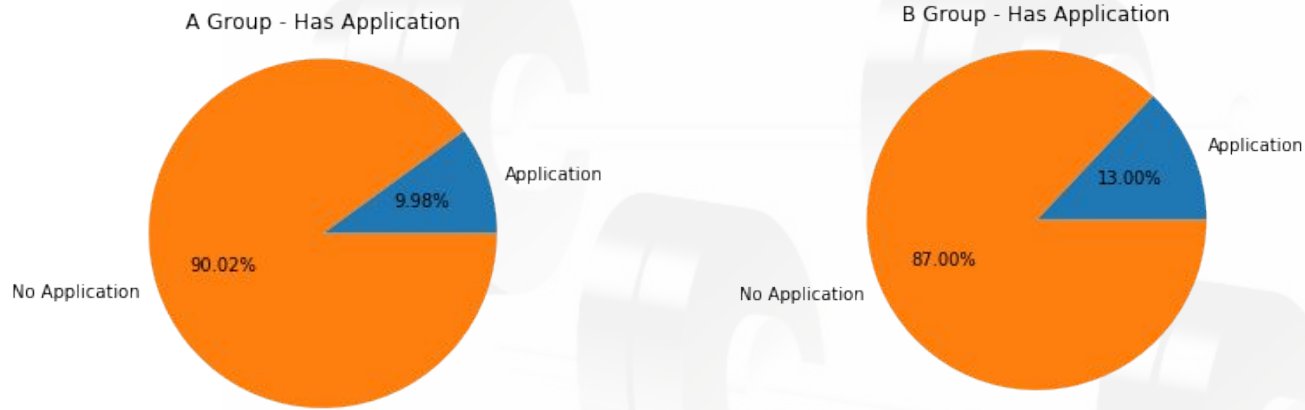
To know whether that difference is statistically important (the null hypothesis states that there is no difference between the datasets and thus any differences are significant **if** we get a p-value of less than .05) we can run a chi2 contingency test (because we are dealing with categorical data) on the data. The p-value of the test, 0.00096, proves that it is indeed significant. But the real question is which group ultimately ended up purchasing a membership.

If we look only at how many applications led to membership purchases then it might appear that Group A was more successful as 80% applications were converted vs. 77% for Group B. But doing a chi2 contingency test on the data shows that this result is statistically insignificant with a p-value of .43

Furthermore, what we really care about is how many **visitors** from each group were converted to memberships when all is said and done.   If we look at that data the results tell a slightly different story. Group B outperforms Group A here 10% to 8%. Our chi2 contingency test shows that this difference is statistically important with a p-value of .014

In Summary, the data proves, with some statistical significance that those visitors who were sent straight to fill out an application ultimately signed up at a rate 2% greater than those visitor who were given a fitness test.

Recommendation

Our recommendation is as follows:

- All things being equal, the fitness test seems to be a hindrance to sign-up as compared to giving people an application as soon as they visit. The data collected, as well as testimonials, support this conclusion in abundance. If the visitors can't be further categorized then we would highly recommend the latter approach.
- That being said, some people did seem to respond to the fitness test. If MuscleHub can identify the type of people who respond better to the fitness test it could be a good tool for converting those who are prone to that responding to that sales technique (perhaps they are already physically fit and looking for a differentiation in gyms facilities, personal trainers, and approach to fitness. If that is the case perhaps a fitness test can be used effectively in targeted cases.

Percent of visitors who apply
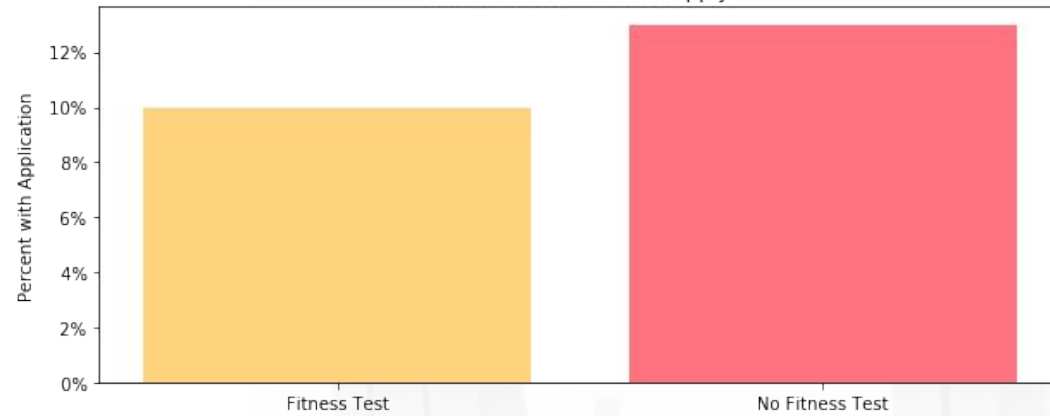
Percent of applicants who purchase a membership

Percent of visitors who purchase a membership

J.W. Goldsteen - <https://github.com/jonnygo55/>