MuscleHub A/B Test

Data Analysis Report

Capstone project by Alexei Vanyashin | Codecademy Intensive: Introduction to Data Analysis

Background. Current membership process

MuscleHub is a fancy gym. Currently, when a visitor to MuscleHub is considering buying a membership, he or she follows the following steps:

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

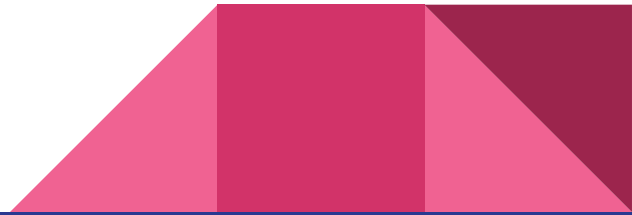
Quotes on Fitness Test

I took the MuscleHub fitness test because my coworker Laura recommended it. Regretted it.

- Sonny "Dad Bod", 26, Brooklyn

...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



Hypothesis: Fitness Test intimidates prospective members

Janet, the manager of MuscleHub, thinks that the fitness test intimidates some prospective members.

Let’s set up a Null Hypothesis and prove or reject it based on statistical tests.

Hypothesis: Fitness Test intimidates prospective members

Janet, the manager of MuscleHub, thinks that the fitness test intimidates some prospective members.

Let’s set up a Null Hypothesis and prove or reject it based on statistical tests.

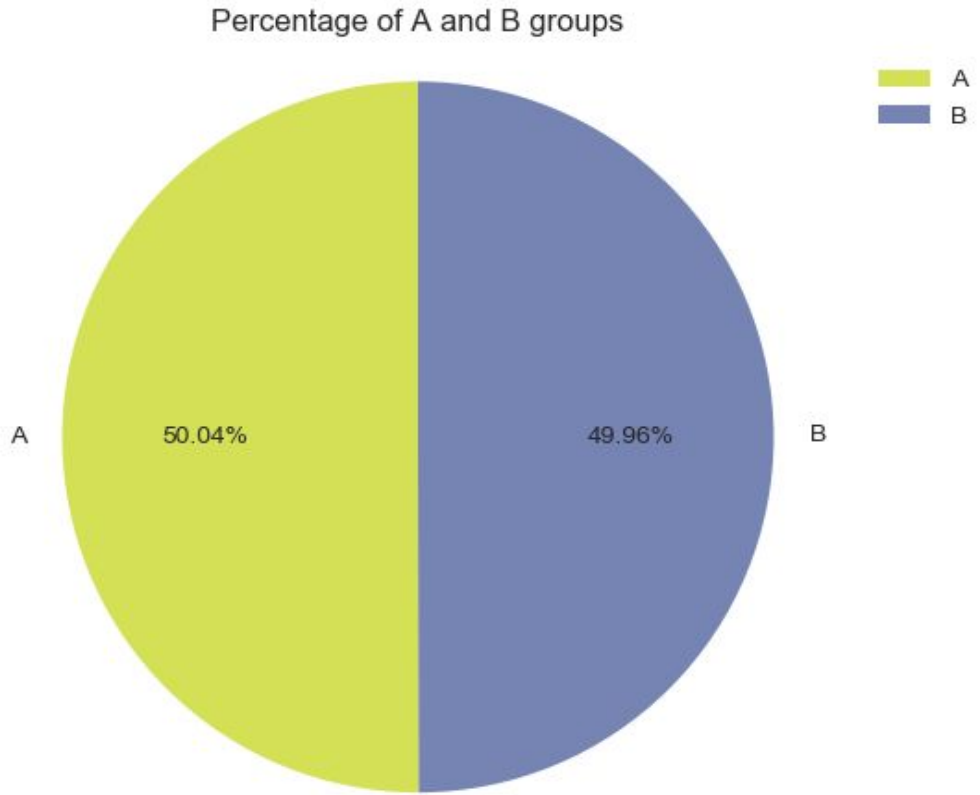
Null Hypothesis: Fitness Test is not a significant factor on prospective members’ behaviour

A/B Test Groups

To test the hypothesis Janet, PM, setup an A/B test. Visitors were randomly assigned to one of two groups:

- Group A was asked to take a fitness test with a personal trainer
- Group B skipped the fitness test and proceed directly to the application

Groups were split equally.



Dataset summary

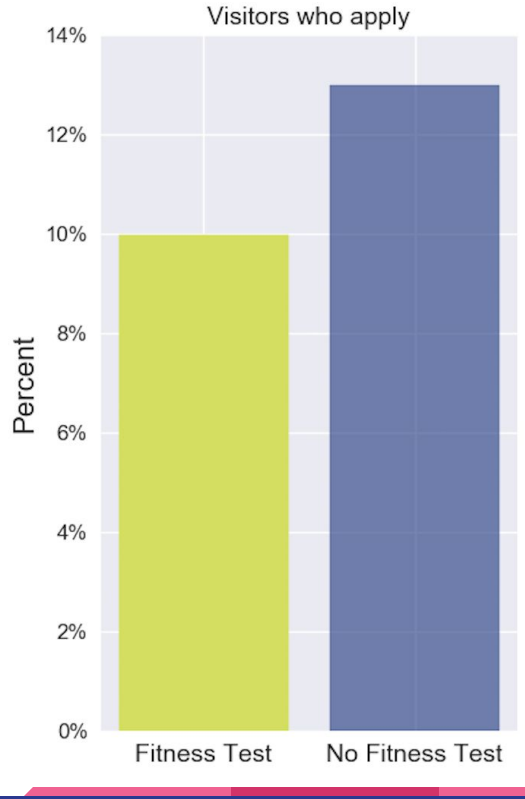
A collective user database called a `DataFrame` was created to analyze prospective customers' behaviour on all stages of the membership process

first_name	last_name	email	gender	visit_date	fitness_test_date	application_date	purchase_date	ab_test_group	is_application	is_member
Anne	Hensley	AnneHensley25@gmail.com	female	9-18-17	None	None	None	B	No Application	Not Member
Marilyn	Cardenas	MC3535@gmail.com	female	7-6-17	2017-07-09	None	None	A	No Application	Not Member
Amber	Head	AmberHead20@gmail.com	female	8-12-17	2017-08-16	None	None	A	No Application	Not Member
Mike	Irwin	Mike.Irwin@gmail.com	male	8-26-17	2017-08-29	None	None	A	No Application	Not Member
Rodney	Floyd	RodneyFloyd91@gmail.com	male	8-19-17	2017-08-25	None	None	A	No Application	Not Member
Allen	Mcleod	AMcleod1983@yahoo.com	male	8-23-17	2017-08-24	2017-08-23	2017-08-23	A	Application	Member
Ronald	Melendez	Ronald.Melendez@gmail.com	male	9-16-17	None	None	None	B	No Application	Not Member
Janice	Acevedo	JA9368@gmail.com	female	7-26-17	2017-07-27	None	None	A	No Application	Not Member
Diane	Haney	DHaney1996@gmail.com	female	8-19-17	2017-08-21	None	None	A	No Application	Not Member
Ida	House	IdaHouse42@gmail.com	female	8-21-17	None	None	None	B	No Application	Not Member

Step 1. How many signed Application?

Our data shows people who did not take Fitness Test (Group B) turned in more applications. Is the difference significant enough to prove Janet's assumption?

Test Group	Application	No Application	Total	% with Application
A	250	2254	2504	9.98%
B	325	2175	2500	13.00%



Hypothesis Test 1

Since we are dealing with Categorical Data (groups A, B) Chi-squared test is more appropriate. Chi-2 Test outputs a p-value. If the value is less than 0.05 the results are significant.

Test Input:

```
from scipy.stats import chi2_contingency
contingency = [[250, 2254],
               [325, 2175]]

chi2, pval, dof, expected = chi2_contingency(contingency)
```

Test Output:

```
Chi-2 test p-val: 0.0009647827600722304
Difference is significant? True
```

A significant difference means we can reject the Null Hypothesis.

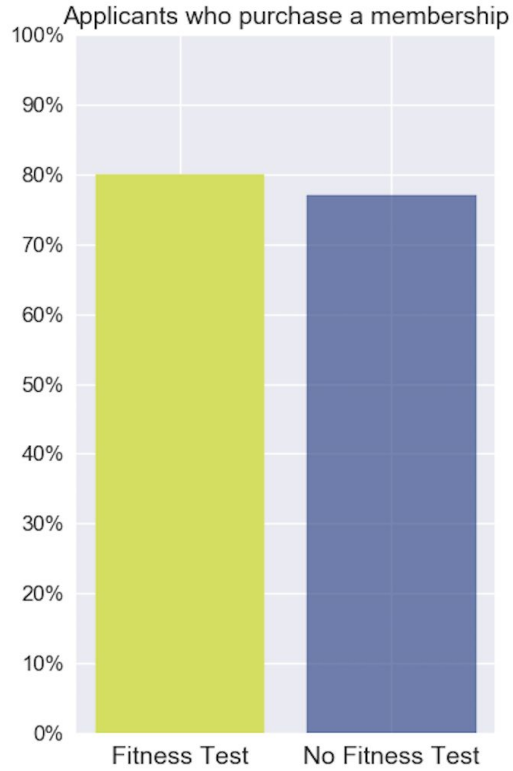
Null Hypothesis: Fitness Test is not a significant factor on prospective members' behaviour

R E J E C T E D

Step 2. Which group of applicants purchased more Memberships?

Out of people who picked an application those who took the fitness test were more likely to purchase a membership. Is this result significant?

Test Group	Member	Not Member	Total	% Purchase
A	200	50	250	80.00%
B	250	75	325	76.92%



Hypothesis Test 2

Again, for Categorical Data Chi-squared test is more appropriate.

Test Input:

```
contingency2 = [[200, 50],
                [250, 75]]

chi2, pval, dof, expected = chi2_contingency(contingency2)
```

Test Output:

```
Chi-2 test p-val: 0.43258646051083327
Result is significant? False
```

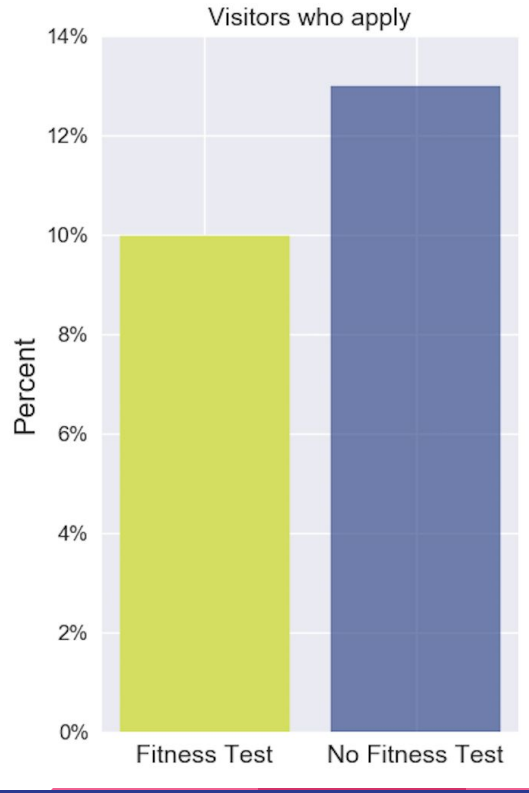
The difference is not significant, this is not a deal-breaker, or a ‘false positive’.

No significant difference
FALSE POSITIVE

Step 3. How many purchased Membership?

What really matters is how many visitors purchased a membership. People who did not take Fitness Test were more likely to purchase membership. Is this difference significant enough to reject the Null Hypothesis?

Test Group	Member	Not Member	Total	% Purchase
A	200	2304	2504	7.99%
B	250	2250	2500	10.00%



Hypothesis Test 3

Let's run Chi-squared test on our most important data. If the output p-value is less than 0.05 the result is significant.

Test Input:

```
contingency3 = [[200, 2304],
                [250, 2250]]

chi2, pval, dof, expected = chi2_contingency(contingency3)
```

Test Output:

```
Chi-2 test p-val: 0.014724114645783203
Result is significant? True
```

Our most important test shows a significant difference. We can now safely reject the Null Hypothesis and prove Janet's assumptions.

Null Hypothesis: Fitness Test is not a significant factor on prospective members' behaviour

R E J E C T E D

Summary. Qualitative Data

Who signed more Applications?

People who did not take Fitness Test (Group B) turned in significantly more applications.

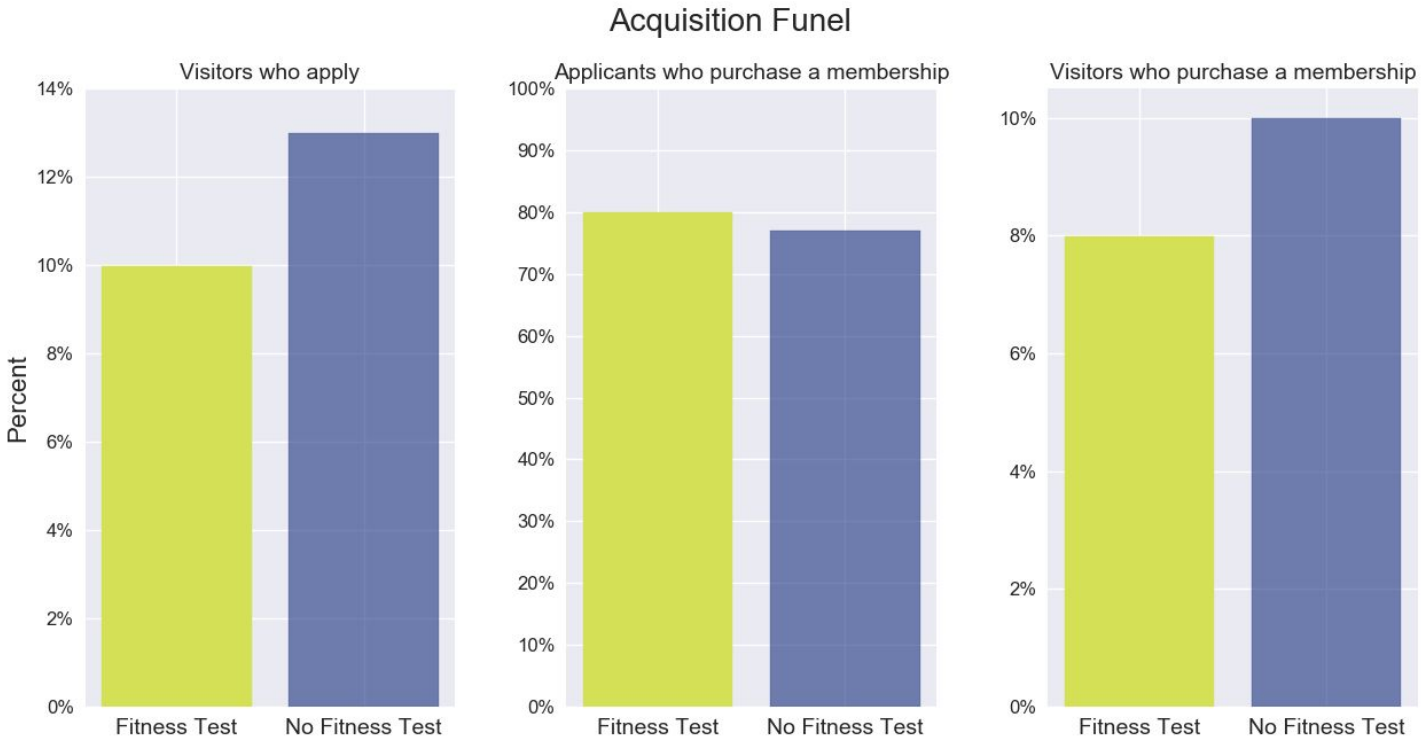
Which group of applicants purchased more Memberships?

Out of people who picked an application those who took the Fitness Test (Group A)were more likely to purchase a membership. However the difference between these two groups is insignificant.

Who purchased more Memberships?

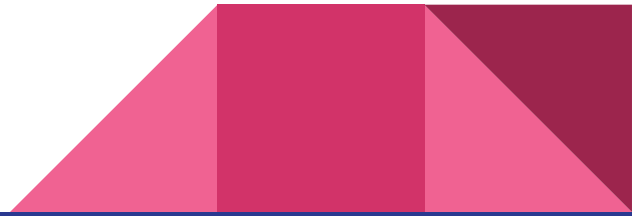
People who did not take Fitness Test (Group B) were more likely to purchase membership. Significantly.

Summary. Qualitative Data



Summary. Quantitative Data

1. On average MuscleHub attracts 1688 potential members per month.
2. 34 people (or 2%) are intimidated by the fitness test each month.
3. Applicants are 3% more likely to purchase membership after taking a fitness test.



Recommendations

MEMBERSHIP PROCESS **BEFORE:**

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

MEMBERSHIP PROCESS **AFTER:**

1. Fill out an application for the gym
2. Take a Fitness Test with a personal trainer (optional)
3. Send in payment for first month's membership

Data Analysis revealed that Fitness Test is in fact an intimidating factor on prospective members. It scares away 2% or 34 potential members per month.

Recommendation: instead of taking a Fitness Test visitors should fill out an application form.

Fitness Test isn't a significant factor on applicants behaviour, so it may be offered optionally after filling a form. In fact it may slightly boost purchases within the applicants group. Due to mixed feedback within visitors group Fitness Test should be totally optional.