Capstone Project - MuscleHub A/B Test

Hypothesis test

Reading the feedback in interview.txt, there may be the probability that potential members are discouraged from subscribing a membership to MuscleHub by fitness test.
So 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

To prove this hypothesis we will perform three different A/B test:

1. An A/B Test to evaluate if the visitor assigned to Group B will be more likely to fill an application.

2. Considering in both groups only the users who have filled the application, we will perform an A/B Test to evaluate if the visitors assigned to Group B will be more likely to purchase a membership to MuscleHub.

3. Finally we will perform the A/B Test including all users in each group (regardless of filling in the application).

In every step of the analysis it has been used the Chi Square test. It is the more suitable method because we are in the situation where half of users did the fitness test and the other half did not and the null hipothesis to verify is no significant difference between the groups (p value > 0.05).

Dataset summary 1/2

- visits contains information about potential gym customers who have visited MuscleHub

	index	first_name	last_name	email	gender	visit_date
0	0	Karen	Manning	Karen.Manning@gmail.com	female	5-1-17
1	1	Annette	Boone	AB9982@gmail.com	female	5-1-17
2	2	Salvador	Merritt	SalvadorMerritt12@outlook.com	male	5-1-17
3	3	Martha	Maxwell	Martha.Maxwell@gmail.com	female	5-1-17
4	4	Andre	Mayer	AndreMayer90@gmail.com	male	5-1-17

- fitness_tests contains information about potential customers in "Group A", who were given a fitness test

	index	first_name	last_name	email	gender	fitness_test_date
0	0	Kim	Walter	KimWalter58@gmail.com	female	2017-07-03
1	1	Tom	Webster	TW3857@gmail.com	male	2017-07-02
2	2	Marcus	Bauer	Marcus.Bauer@gmail.com	male	2017-07-01
3	3	Roberta	Best	RB6305@hotmail.com	female	2017-07-02
4	4	Carrie	Francis	CF1896@hotmail.com	female	2017-07-05

Dataset summary 2/2

- `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.

	index	first_name	last_name	email	gender	application_date
0	0	Roy	Abbott	RoyAbbott32@gmail.com	male	2017-08-12
1	1	Agnes	Acevedo	AgnesAcevedo1@gmail.com	female	2017-09-29
2	2	Roberta	Acevedo	RA8063@gmail.com	female	2017-09-15
3	3	Darren	Acosta	DAcosta1996@hotmail.com	male	2017-07-26
4	4	Vernon	Acosta	VAcosta1975@gmail.com	male	2017-07-14

- `purchases` contains information about customers who purchased a membership to MuscleHub.

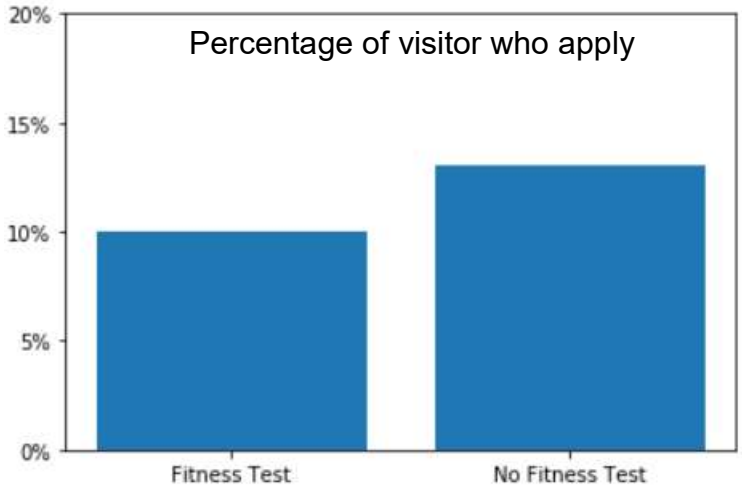
	index	first_name	last_name	email	gender	purchase_date
0	0	Roy	Abbott	RoyAbbott32@gmail.com	male	2017-08-18
1	1	Roberta	Acevedo	RA8063@gmail.com	female	2017-09-16
2	2	Vernon	Acosta	VAcosta1975@gmail.com	male	2017-07-20
3	3	Darren	Acosta	DAcosta1996@hotmail.com	male	2017-07-27
4	4	Dawn	Adkins	Dawn.Adkins@gmail.com	female	2017-08-24

A/B Test 1 results

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

```
print (pval)
```

```
0.000964782760072
```



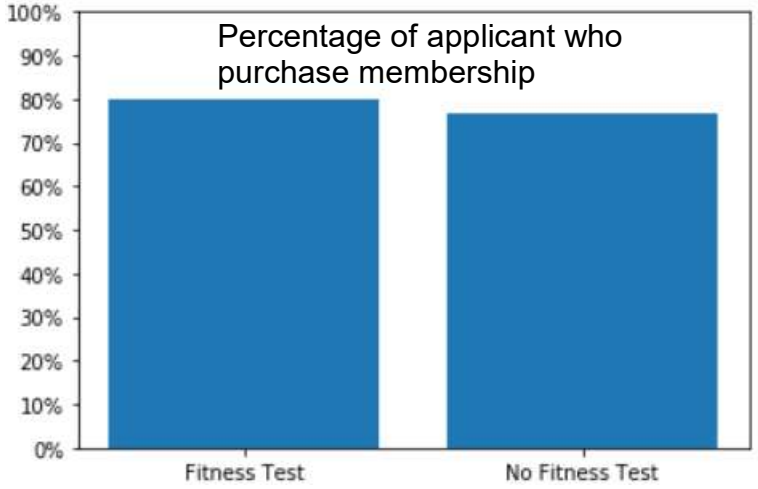
The result is significant.

It means that users who skip the fitness test are more likely to fill the application

A/B Test 2 results

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

```
print (pval)
0.432586460511
```



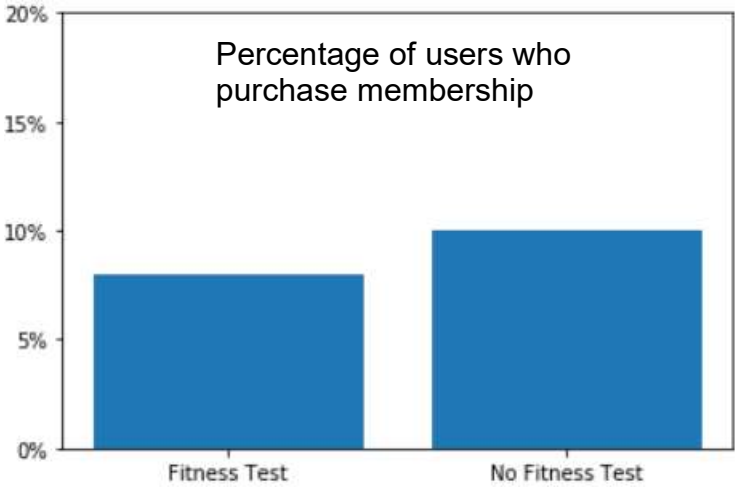
The result is not significant. There is no difference between the two groups

A/B Test 3 results

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

```
print (pval)
```

```
0.0147241146458
```



The result is significant.

It means that users who skip the fitness test are more likely to purchase a membership to MuscleHub

Conclusion

The results confirms the initial hypothesis according to which the fitness test may discourage from purchasing a membership to MuscleHub.

As a consequence, I would recommend modifying the test. It should be adapted to the characteristics of different people. For example, made easier for those who have never practiced fitness before.