

# Capstone Project: MuscleHub A/B Test

Author: Anne Luft, Germany

Date: 24 June 2018

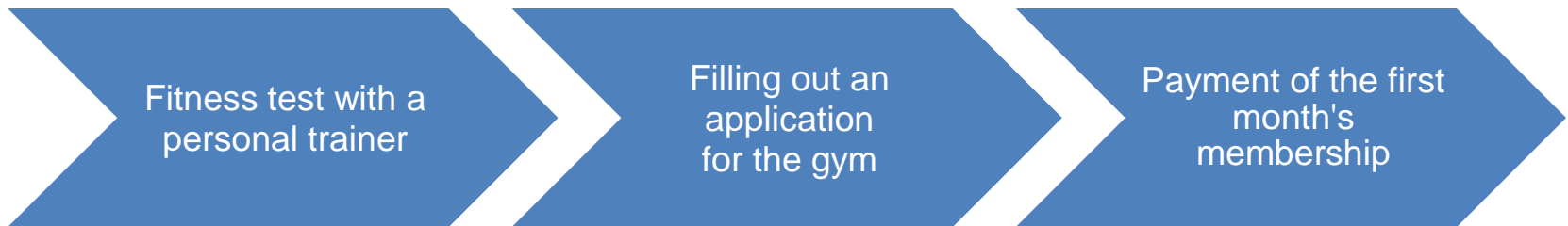
# Contents

- Initial Situation
- Description of A/B test
- Original Data
- Summary of Dataset
- Sanity Check
- Type of Hypothesis Test
- Results of Hypothesis Tests
- Summary of Qualitative Data
- Recommendation for MuscleHub

# Initial Situation

The fancy gym MuscleHub needs an A/B test to be run to optimize the process of winning new members.

For visitors to MuscleHub interested in a membership, following procedure has been determined:



## **Reason to perform an A/B test:**

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

# Description of A/B Test

## What is the A/B test about?

This A/B Test is comparing two approaches for winning new members to MuscleHub to see which one is more successful.

Accordingly, visitors will randomly assigned into two groups:

### Group A

- Visitors will follow the current procedure, starting with a fitness test with a personal trainer

### Group B

- Visitors will proceed directly to the application and skip the fitness test

## Janet's Hypothesis

More Visitors assigned to Group B than those to Group A will eventually become member of MuscleHub.

# Original Data

## Data Available

Data of visits, fitness tests, applications and purchases are stored into 4 data tables in an SQL Database.

	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

### Data Table “visits”

Data of potential gym customers who have visited MuscleHub

	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

### Data Table “fitness\_tests”

Data of potential gym customers who were given a fitness test

	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

### Data Table “applications”

Data of any potential customers who filled out an application. Not everyone in visits will have filled out an application.

	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

### Data Table “purchases”

Data of customers who purchased a membership to MuscleHub

# Summary of Dataset

## Consolidation of Original Data to one DataFrame

The original data of the just presented four data tables are consolidated to one DataFrame by using an SQL query:

	first_name	last_name	gender	email	visit_date	fitness_test_date	application_date	purchase_date	ab_test_group	is_application	is_member
0	Kim	Walter	female	KimWalter58@gmail.com	7-1-17	2017-07-03	None	None	A	No Application	Not Member
1	Tom	Webster	male	TW3857@gmail.com	7-1-17	2017-07-02	None	None	A	No Application	Not Member
2	Edward	Bowen	male	Edward.Bowen@gmail.com	7-1-17	None	2017-07-04	2017-07-04	B	Application	Member
3	Marcus	Bauer	male	Marcus.Bauer@gmail.com	7-1-17	2017-07-01	2017-07-03	2017-07-05	A	Application	Member
4	Roberta	Best	female	RB6305@hotmail.com	7-1-17	2017-07-02	None	None	A	No Application	Not Member

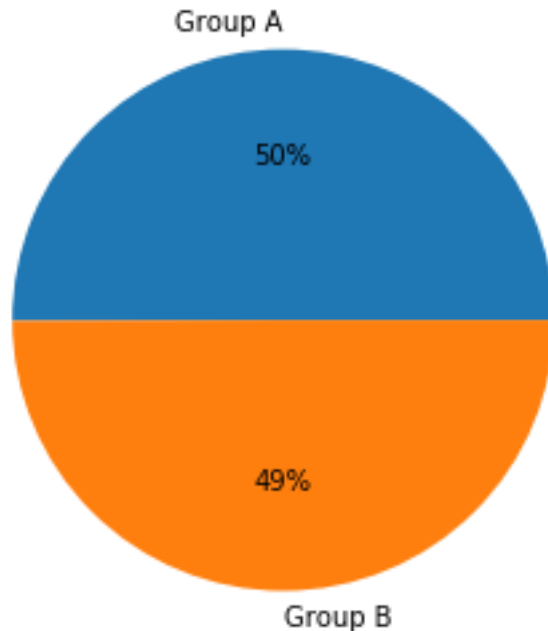
## Content of Table

- All visits starting at 7-1-17, when the A/B test took place
- Dates of fitness tests, applications and purchases, where available
- Additional Column “ab\_test\_group” to indicate to which group each visitor belongs: Group “A” performed a fitness test, group “B” applied without fitness test
- Additional Column “is\_application” with an entry for each visitor: “Application” when an application has been filled out, “No Application” when not
- Additional Column “is\_member” to list membership: “Member” or “No Member” depending on purchase of membership

# Sanity Check

Before analyzing the data, a sanity check is necessary: The size of the two groups “A” and “B” should be comparable.

A/B Test -  
Relative Distribution of Groups A and B



Group A consists of 2,504 Group B of 2,500 visitors, meaning the sanity check has been passed.

# Type of Hypothesis Test

## Type of Hypothesis Test

The reason for a hypothesis test after having performed the A/B test is:

Finding out whether the different observations really are the results of different conditions or just of random chance.

Janet's hypothesis was "More Visitors assigned to Group B than those to Group A will eventually become member of MuscleHub".

The null hypothesis we need to reject is that there's no significant difference between both groups. We reject that hypothesis (and confirm Janet's), if we get a p-value less than 0.05 or 5%.

The appropriate type of test is the Chi Square Test.

## Why Chi Square Test?

- There are two or more categorical datasets to be compared.
- Chi-Square only tests whether two individual variables are independent in a binary, "yes" or "no" format.
- Test results are numerical values, not percentages or ratios.



# Results of Hypothesis Tests

## Chi Square Test I

	ab_test_group	Application	No Application
0	A	250	2,254
1	B	325	2,175

**Contingency table:** Visitors who filled out an application

**P-Value:** 0.00096

**Result is significant.**

## Chi Square Test II

	ab_test_group	Member	Not Member
0	A	200	50
1	B	250	75

**Contingency table:** Visitors with application who purchased a membership

**P-Value:** 0.43259

**Result is not significant.**

## Chi Square Test III

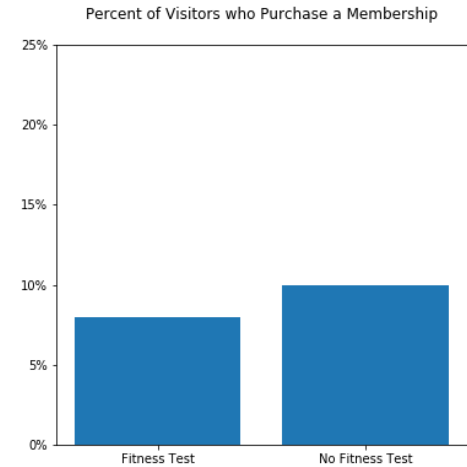
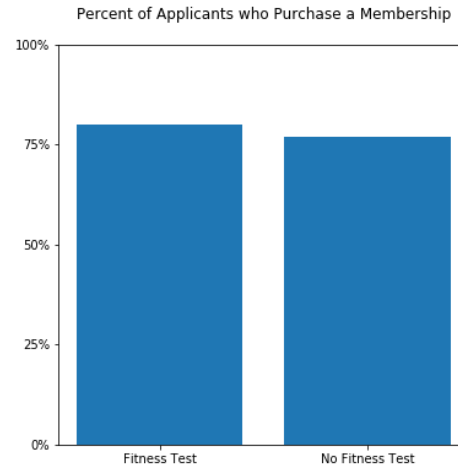
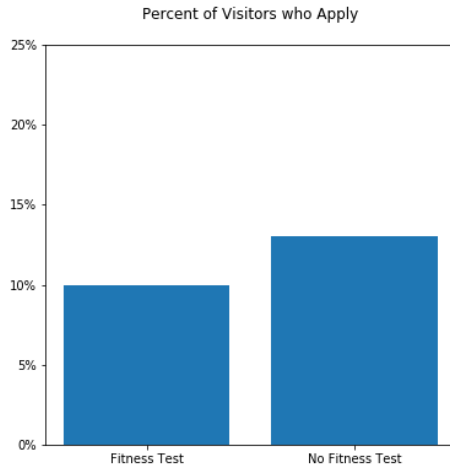
	ab_test_group	Member	Not Member
0	A	200	2,304
1	B	250	2,250

**Contingency table:** Visitors who purchased a membership

**P-Value:** 0.01472

**Result is significant.**

# Summary of Qualitative Data



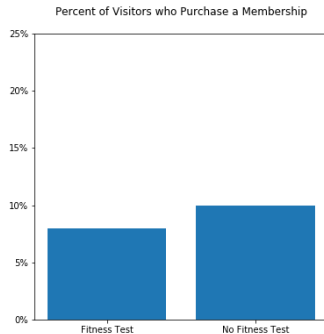
## Result of A/B Test

Janet's hypothesis "More Visitors assigned to Group B than those to Group A will eventually become member of MuscleHub" has been confirmed:

Of those visitors NOT doing a fitness test (Group "B")

- around 3% more applied.
- around 3% less of those who already applied purchased a membership. Since the result of the chi square test was not significant, this is to neglect.
- around 2% more purchased a membership.

# Recommendation for MuscleHub



## Key Result

Deciding is the number of visitors who could be acquired as member of MuscleHub.

Performing a fitness test with a personal trainer right after visiting obviously discourages a considerable number of potential customers.

## Recommendation

Taking the interview given by visitors into account, performing a fitness test appears to be appropriate and challenging particularly for visitors with ambitious training targets (e.g. Cora, 23, Hoboken).

Most of those who have given their views of their experiences with MuscleHub put the focus on enjoying themselves and feeling welcome at a gym. Cleanliness is according to Jesse, 35, Gowaness, also an issue.

My recommendation is

- Offering fitness tests individually as an optional choice and challenge.
- Offering special fitness classes for “mere mortals” to achieve basic training goals, for example “Become a member and secure your place in our introductory course!”.
- Attach importance to cleanliness and a feel-good atmosphere.