Codecademy:
Introduction to Data Analysis
Capstone Project

_____

Musclehub A/B Test

Michael Brown

mabrown1

Musclehub A/B Test
Problem Statement

_____

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 their payment for their first month's membership.

Janet, the manager of MuscleHub, thinks that the fitness test intimidates some prospective members, so she has set up an A/B test to test this hypothesis.

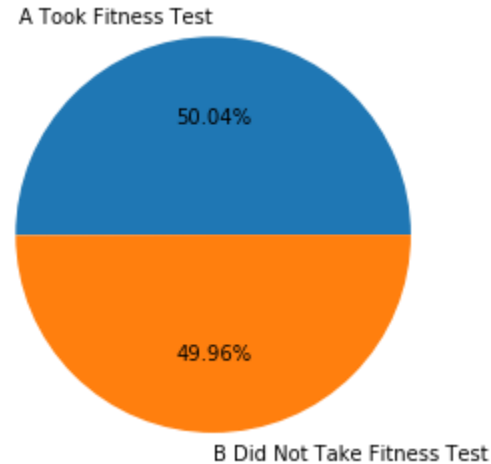Visitors will randomly be assigned to one of two groups:

- 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.

Musclehub A/B Test
Problem Statement

_____

- A/B Test Group Breakdown.

  - 2504 Users in A test group.
  - 2500 Users in B test group

- Janet's hypothesis is that visitors assigned to Group B will be more likely to eventually purchase a membership to MuscleHub.

- The end result of this analysis is to prove or disprove Janet's hypothesis that visitors assigned to Group B will be more likely to eventually purchase a membership to MuscleHub.

Musclehub A/B Test
Analysis Approach

_____

- Data is maintained in four data sets (visits, fitness_tests, applications, purchases) with common information (name, gender, email)

- Each data set contains specifics such as;
    - visits (visit_date)
    - fitness_tests (fitness_date)
    - applications (application_date)
    - Purchases (purchase_date)

- A master data set (df) combines the above information needed to prove or disprove the hypothesis and is analyzed in the following sequence:
    - Percent of visitors who apply.
    - Percent of applications who purchase a membership.
    - Percent of visitors who purchase a membership.

Musclehub A/B Test
Analysis Approach
Data Summary – Percent of visitors who apply.

_____

- Percent of visitors who picked up applications

Test Group	Pick up Application	Did Not Pick up Application	Total	% with Application
A	250	2254	2504	9.984
B	325	2175	2500	13.000

- Group A will take a fitness test with a personal trainer.

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

- To determine if there is a significant difference between the two data sets a Chi Square test is applied for multiple variables. Results are

$$p = 0.0009648 < 0.05$$

- The hypothesis that visitors assigned to Group B will be more likely to eventually purchase a membership to MuscleHub is rejected as there is a significant difference between the data sets.

Musclehub A/B Test
Analysis Approach
Data Summary – Percent of applicants who purchased a membership

_____

- Percent of applicants who picked up an application and purchased a membership.

Test Group	Member	Not a Member	Total	% with Application
A	200	50	250	80.000
B	250	75	325	76.923

- Group A will take a fitness test with a personal trainer.

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

- To determine if there is a significant difference between the two data sets a Chi Square test is applied for multiple variables. Results are

$$p = 0.432586 > 0.05 .$$

- The hypothesis that applicants assigned to Group B will be more likely to eventually purchase a membership to MuscleHub is not rejected as there is a no significant difference between the data sets.
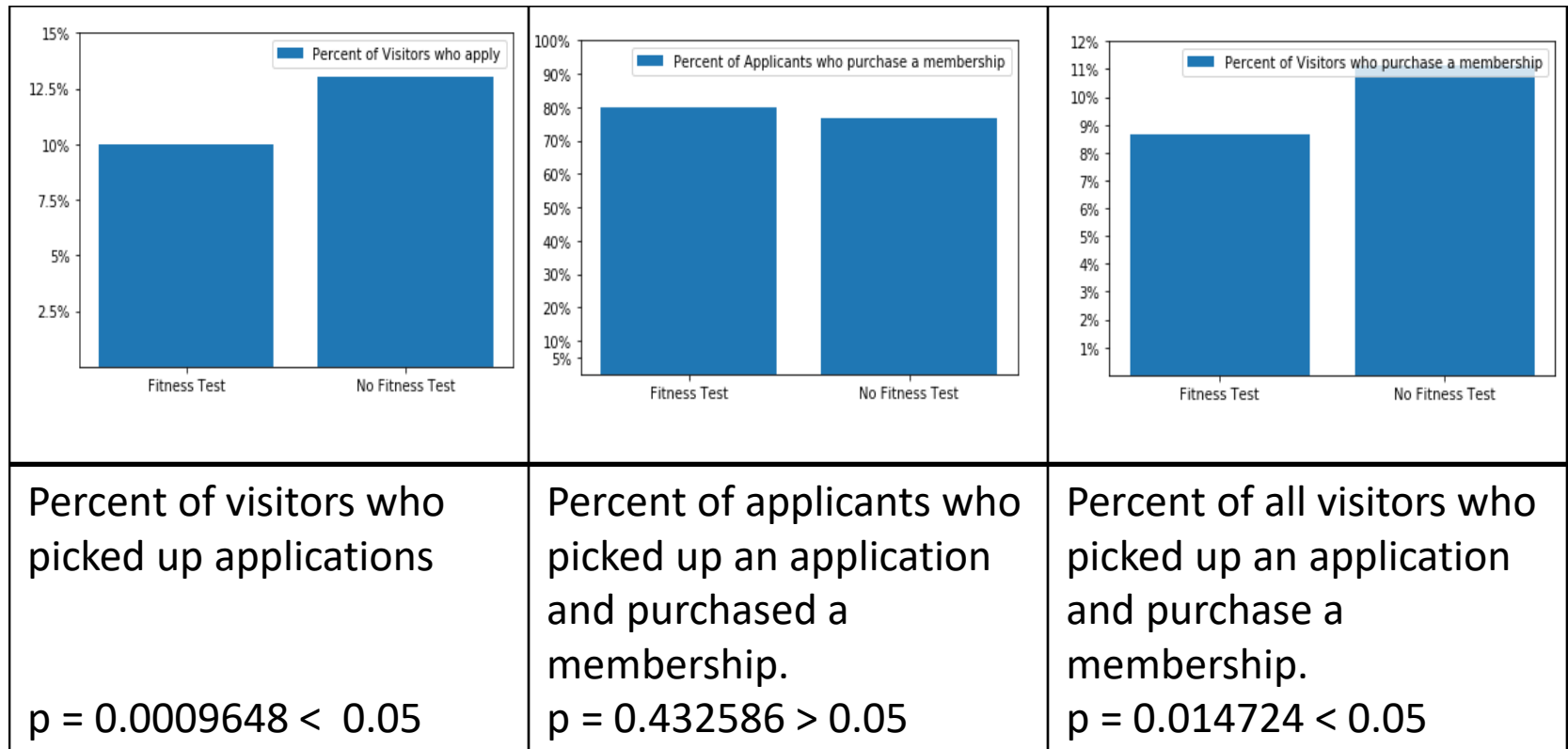
.

Musclehub A/B Test
Analysis Approach
Data Summary – Percent of visitors who purchase a membership
_____

- Percent of all visitors who picked up an application and purchase a membership.

Test Group	Member	Not a Member	Total	% with Application
A	200	2304	2504	8.681
B	250	2250	2500	11.111

- Group A will take a fitness test with a personal trainer.

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

- To determine if there is a significant difference between the two data sets a Chi Square test is applied for multiple variables. Results are

$$p = 0.014724 < 0.05 \; .$$

- The hypothesis that all visitors assigned to Group B will be more likely to eventually purchase a membership to MuscleHub is rejected as there is a significant difference between the data sets.

.

Musclehub A/B Test Summary

_____

Percent of visitors who picked up applications	Percent of applicants who picked up an application and purchased a membership.	Percent of all visitors who picked up an application and purchase a membership.
$p = 0.0009648 < 0.05$	$p = 0.432586 > 0.05$	$p = 0.014724 < 0.05$

Recommendation: Based on the data sets and Chi Square testing, the applicants who skipped the fitness test and picked up an application are more likely to purchase membership.