MuscleHub A/B Test

Codecademy Capstone Project
For Data Analysis Training

KOFFI ALEXIS

A/B test description

We've been hired to help MuscleHub, a fancy gym, run an A/B test !

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

- Take a fitness test with a personal trainer
- Fill out an application for the gym
- 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.

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

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

We will help to analyze the data.

Summary of the dataset

Janet of MuscleHub has a SQLite database, which contains several tables that will be helpful in this investigation:

- `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

- `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

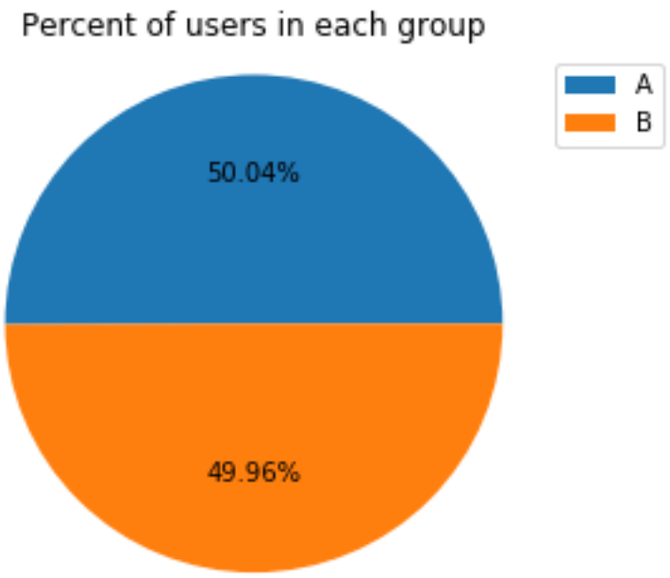
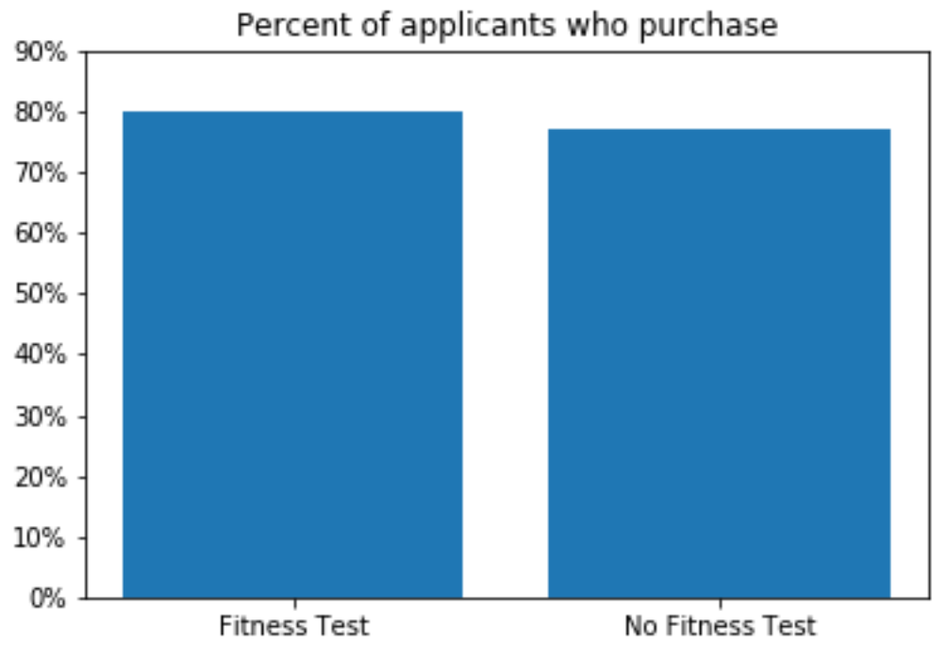
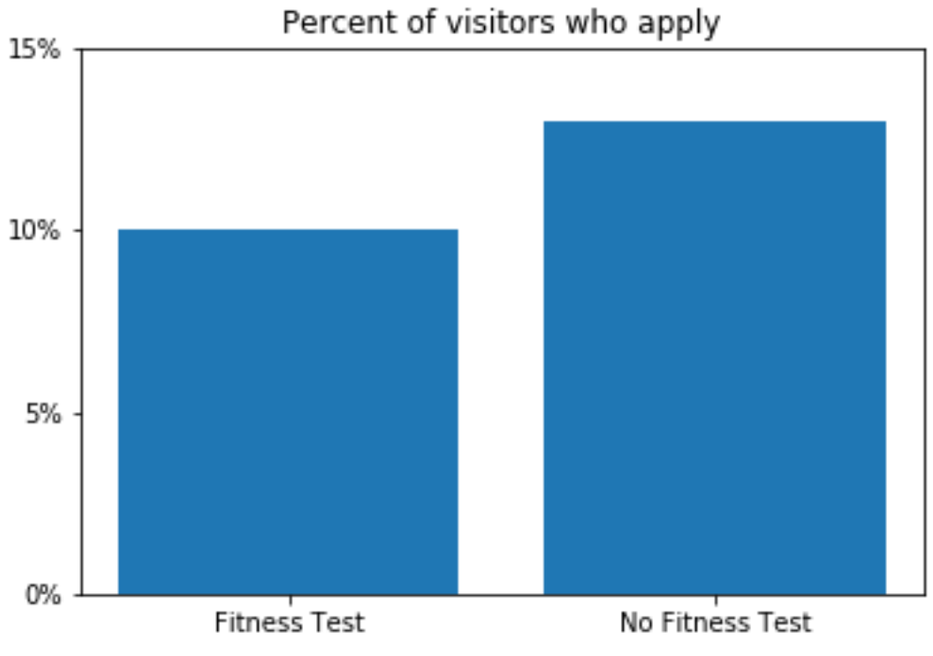
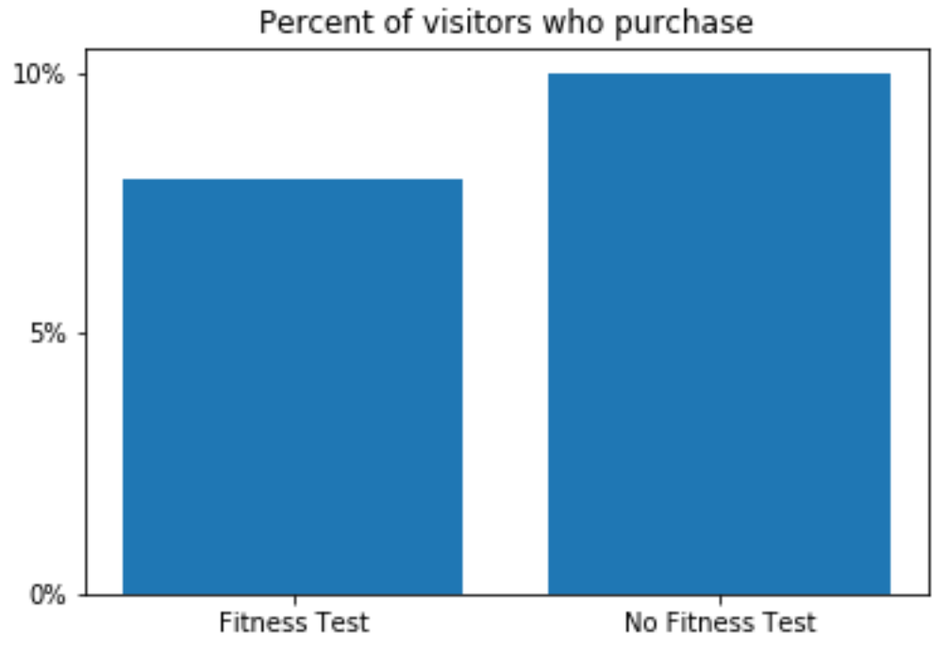
	index	first_name	last_name	email	gender	purchase_date
0	0	Roy	Abbott	RoyAbbott32@gmail.com	male	2017-08-18
1	1	Roberta	Acevedo	RA8063@gmail.com	female	2017-09-16
2	2	Vernon	Acosta	VAcosta1975@gmail.com	male	2017-07-20
3	3	Darren	Acosta	DAcosta1996@hotmail.com	male	2017-07-27
4	4	Dawn	Adkins	Dawn.Adkins@gmail.com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

Results of the hypothesis tests

Samples	Type of test	p value
People in each group who complete an application or not	Chi-Squared Test for two samples of categorical data	0.0009647827600722304
People in each group who picked up applications purchased memberships or not	Chi-Squared Test for two samples of categorical data	0.43258646051083327
All people memberships or not in each group who visited MuscleHub	Chi-Squared Test for two samples of categorical data	0.014724114645783203

ab_test_group	Application	No Application	Total	Percent with Application
A	250	2254	2504	0.09984
B	325	2175	2500	0.13000

ab_test_group	Member	Not Member	Total	Percent Purchase
A	200	50	250	0.800000
B	250	75	325	0.769231

ab_test_group	Member	Not Member	Total	Percent Purchase
A	200	2304	2504	0.079872
B	250	2250	2500	0.100000

Summary of the qualitative data