

Introduction to Data Analysis

Capstone project:

MuscleHub *A/B* Test

Yu-Ting Shen

Description of A/B test:

Target: does the fitness test affect the visitors to buy memberships?

A/B test:

- Group A: Take fitness test
- Group B: No fitness test

Hypothesis: Visitors in the group B is more likely to purchase memberships

Summary of dataset

There are 5004 rows and 8 columns in our raw dataset.

- 2504 visitors take fitness test
Group A: 2504 visitors
Group B: 2000 visitors
- 575 visitors submit application
Group A: 250 visitors
Group B: 325 visitors
- 450 visitors purchase memberships
Group A: 200 visitors
Group B: 250 visitors

Column name	Number of entry	Is null	Type
first_name	5004	non-null	object
last_name	5004	non-null	object
gender	5004	non-null	object
email	5004	non-null	object
visit_date	5004	non-null	object
fitness_test_date	2504	non-null	object
application_data	575	non-null	object
purchase_date	450	non-null	object

Results and explanation

Submit application:

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

It looks like more visitors in Group B submit applications.

Hypothesis: The difference is caused by chance.

By using the Chi2-test, the $p\text{-value} < 0.05$. So the result is not significant.

Results and explanation

Purchase memberships:

Group	Member	Not Member	Total	Percentage with Application
A	200	50	250	80.00%
B	250	75	325	76.92%

It looks like the visitors in Group A purchase memberships if they submit application.

Hypothesis: The difference is causing by chance.

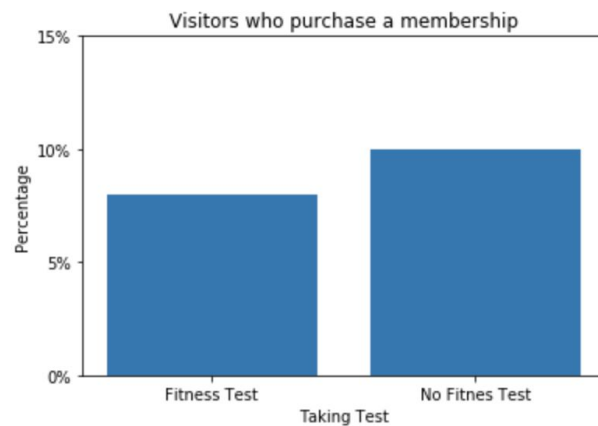
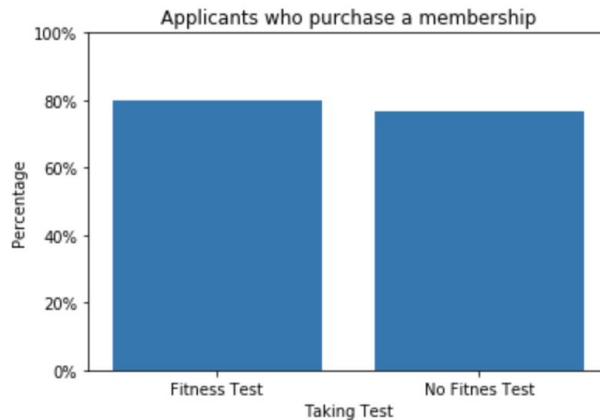
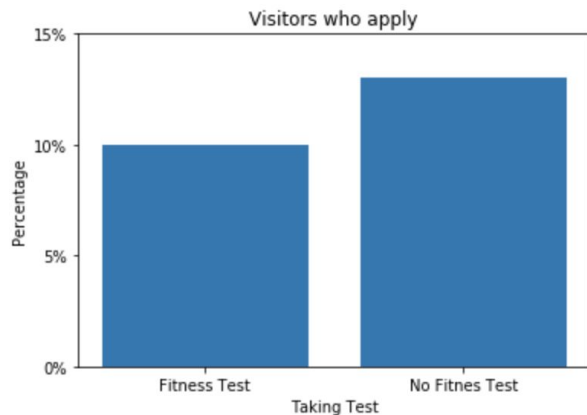
By using the Chi2-test, the p-value < 0.05 . So the result is not significant.

Summary

More visitors in Group B submit applications.

Similar amount of visitors in both groups purchase memberships if they submit applications.

More visitors in Group B purchase memberships.



Conclusion

More visitors would like to purchase memberships if there is no fitness test.