MuscleHub A/B Test

Which Group will be more likely to buy a membership?

Group A who will take the fitness test.

Or

Group B who will skip the test

A/B Test Summary

MuscleHub gym`s current process for visitors of the gym who intend on buying a membership are: a physical fitness test with a personal trainer, fill out an application and pay for the first months membership if joined. But the manager Janet believes that the fitness test intimidates prospective members. So I was hired to provide an A/B hypothesis analysis to test this hypothesis. There will be an A test group (who will be asked to take the fitness test) and a B test group (who will skip the test and proceed to the application) and each visitor within either test group will be randomly assigned. SQL, Pandas, Pyplot, and Numpy were used to either access (SQL) or analyze Janet's data from her database. Once the data was imported from her SQL database the investigation of the tables began. There were four tables within the database; purchase which contained information on customers who bought a membership, applications that contained information about customers who filled out an application, fitness test which contained information about which customers took the test, and visits which contained information about customers who just visited the gym.

A/B Test Summary Cont.

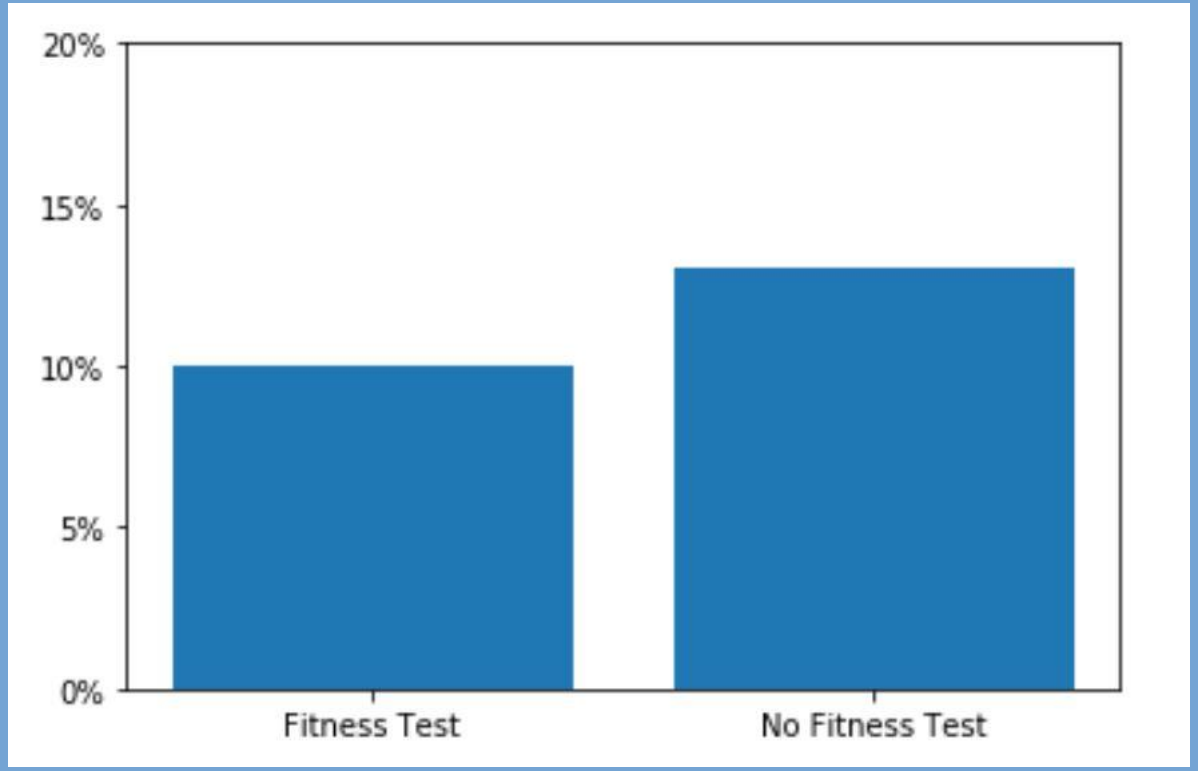
Since one table is easier to deal with than four, a series of left joins were made to consolidate the tables. Pandas was imported in order to do the analysis. Then columns were added to a dataframe for the two control groups. A groupby was done to verify that the groups were correctly divided in half. A pie chart was added to show the percentage of visitors within the test. To examine how many visitors filled out the application another dataframe was made and a groupby was done. In order to calculate the percentage of visitors who filled out an application, a pivot table was created. Chi2 contingency was used to verify the data that was found throughout the analysis. The same process was followed to calculate how many visitors became members. Bar graphs were used to show all of the findings.

MuscleHub Dataset Summary

After the left joins were completed the total number of rows were 5004. An investigation of the two test groups revealed that there were four more visitors in test group A than B which proved to be an .8% difference between them. After more analysis the creation of a lambda function, and the counting of the number of visitors who filled out an application, it appeared to be an astonishing find. It showed that 75 more visitors within test group B filled out an application than in A. But statistically it turned out to be insignificant. The next bit of information that was shown was how many visitors purchased a membership. In order to calculate the number of visitors who purchased a membership a pivot table was created which showed that 25 more test subjects became members in test group B than A with a .3% difference. It look as though people that took the fitness test were more likely to purchase a membership and statistically the data looks significant. But when all visitors were considered in the analysis it proved that Janet`s hypothesis was correct and statistically it showed the same significance.

Visitors who Applied to MuscleHub

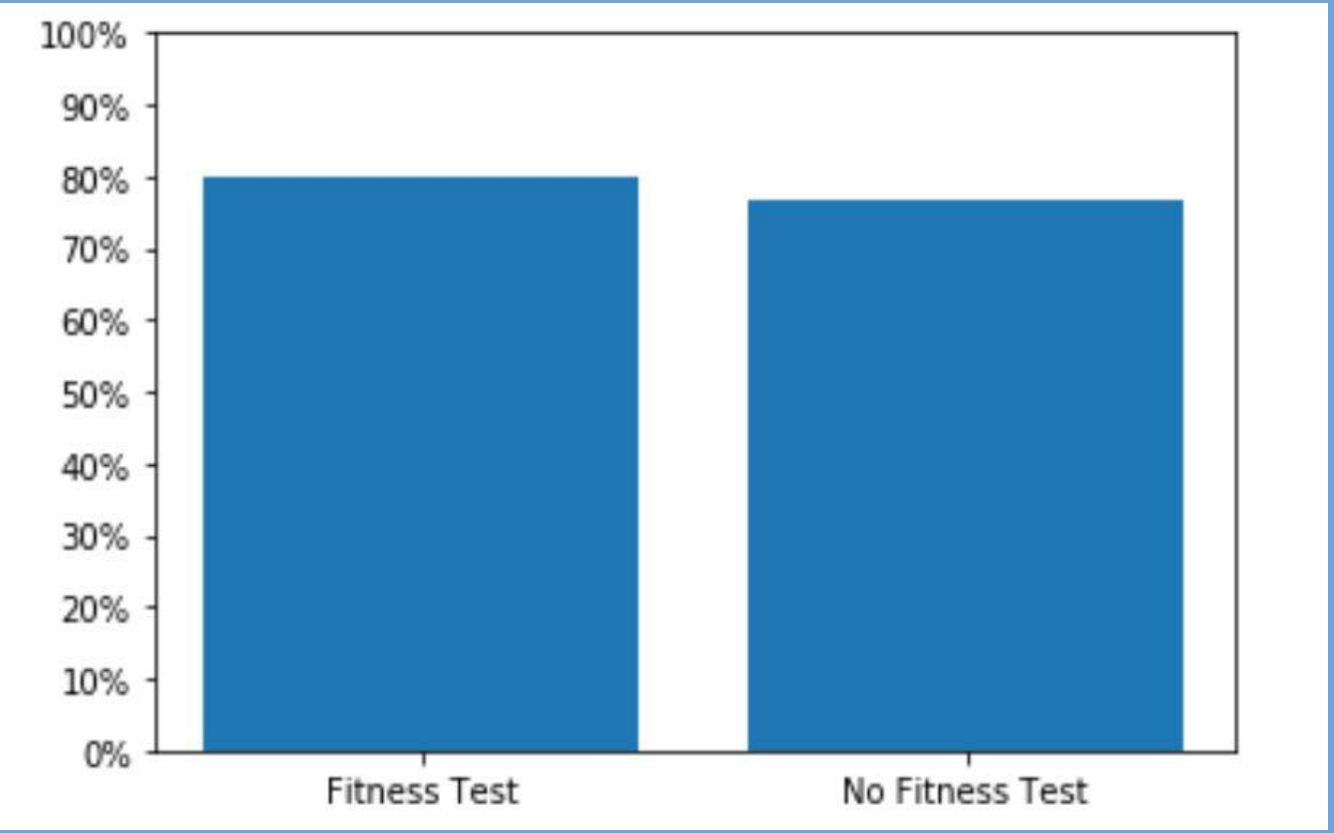
Within this analysis Group B had a slight 3% edge on Group A which seems to prove Janet right.



Chi2 Contingency was used because MuscleHub used an A/B Hypothesis test. The Chi2 Contingency table is used for two or more categorical dataset comparisons.

Applicants who Became Members

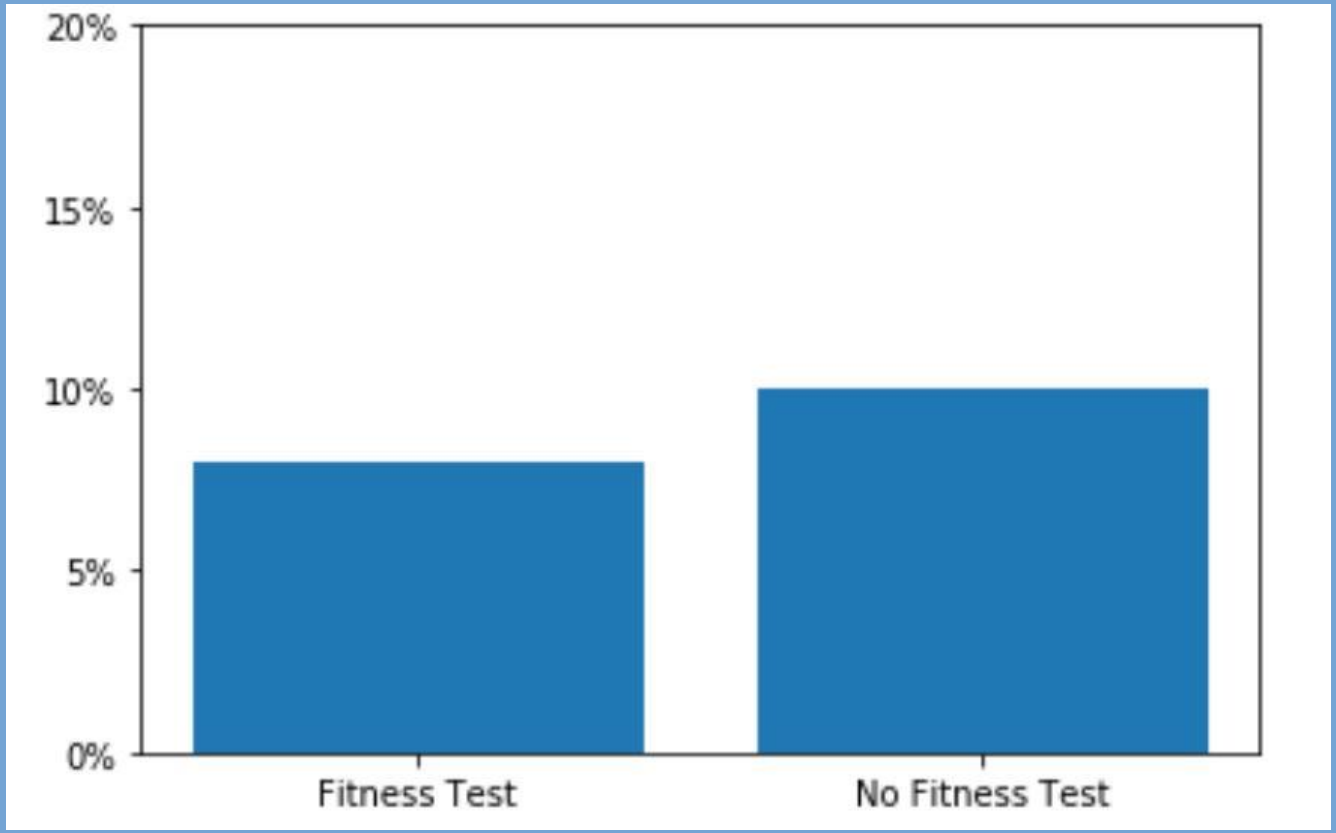
Within this analysis Group A had a slight 3% edge on Group B which appears to prove Janet`s hypothesis wrong.



Chi2 Contingency was used because MuscleHub used an A/B Hypothesis test. The Chi2 Contingency table is used for two or more categorical dataset comparisons.

Visitors who Became Members

The final analysis shows that Group A had a 3% difference between group B which proves Janet was right. More visitors who didn't take the test purchased a membership.



Chi2 Contingency was used because MuscleHub used an A/B Hypothesis test. The Chi2 Contingency table is used for two or more categorical dataset comparisons.

Qualitative Data Summary

Four visitors of MuscleHub were interviewed about their experience. Two of the interviewers provided positive feedback and two provided negative feedback. The fitness test was a turn off for one of the interviewers which may mean the test may be too hard, and an unclean gym for another. Between the other two interviewers one enjoyed the fitness test, the other had a great customer service experience. These interviews provide a look into the minds of the visitors and shows how they feel about the gym. And gives a different look at the fitness test and what the effect has on the visitors.