

SEARCH



RESOURCES

CONCEPTS

- ✓ 28. Notebook + Quiz: Drawing Con...
- ✓ 29. Solution Notebook: Drawing C...
- ✓ 30. Other Things to Consider - Im...
- 31. Other Things to Consider - Wh...
- 32. Other Things to Consider - Ho...
- 33. Notebook + Quiz: Impact of Sa...
- 34. Solution Notebook: Impact of ...
- 35. Notebook + Quiz: Multiple Tests
- 36. Solution Notebook: Multiple te...
- 37. Hypothesis Testing Conclusion
- 38. Quiz + Text: Recap



Mentor Help

Ask a mentor on our Q&A platform



Peer Chat 2

Chat with peers and alumni



One of the most important aspects of interpreting any statistical results (and or overlooked) is assuring that your sample is truly representative of your population.

Particularly in the way that data is collected today in the age of computers, **representativeness** is an important concept to keep in mind. In the 2016 U.S election, polls conducted by many news organizations showed a staggering difference from the reality of poll results. You can read about how a role [here](#).

Hypothesis Testing vs. Machine Learning

With large sample sizes, hypothesis testing leads to even the smallest of findings being **statistically significant**. However, these findings might not be practically significant at all.

For example, imagine you find that **statistically** more people prefer beverage 1 in a study of more than one million people. Based on this you decide to open a shop selling beverage 1. You then find out that beverage 1 is only more popular than beverage 2 by 0.001 (a statistically significant amount with your large sample size). Practically, maybe you should have sold both.

Hypothesis testing takes an aggregate approach towards the conclusions made from these tests as they are aimed at understanding population parameters (which are aggregate values).

Alternatively, machine learning techniques take an individual approach towards data as they attempt to predict an outcome for each specific data point.

In the final lessons of this class, you will learn about two of the most fundamental approaches used in practice: **linear** and **logistic** regression.