**Assignment**

Now it's time to flex your critical evaluation skills. Read the following descriptions of an experiment and its analysis, identify the flaws in each, and describe what you would do to correct them.

1. The Sith Lords are concerned that their recruiting slogan, "Give In to Your Anger," isn't very effective. Darth Vader develops an alternative slogan, "Together We Can Rule the Galaxy." They compare the slogans on two groups of 50 captured droids each.

In one group, Emperor Palpatine delivers the "Anger" slogan. In the other, Darth Vader presents the "Together" slogan. 20 droids convert to the Dark Side after hearing Palpatine's slogan, while only 5 droids convert after hearing Vader's. The Sith's data scientist concludes that "Anger" is a more effective slogan and should continue to be used.’’

FLAWS:

Ceteris paribus fails to hold. All other things are not equal.

We have possible selection bias, observer bias and contextual bias.

FIX

A/A testing would be helpful. As would having the person deliver the slogan

1. In the past, the Jedi have had difficulty with public relations. They send two envoys, Jar Jar Binks and Mace Windu, to four friendly and four unfriendly planets respectively, with the goal of promoting favorable feelings toward the Jedi.

Upon their return, the envoys learn that Jar Jar was much more effective than Windu: Over 75% of the people surveyed said their attitudes had become more favorable after speaking with Jar Jar, while only 65% said their attitudes had become more favorable after speaking with Windu. This makes Windu angry, because he is sure that he had a better success rate than Jar Jar on every planet. The Jedi choose Jar Jar to be their representative in the future.

FLAWS:

Ceteris paribus fails again. Likely selection bias, contextual bias and observer bias.

For starters we don’t even of they visited the same planets. And even if they did, we know nothing the nature of their audience. So very unlikely the people surveyed were anywhere near homogenous.

FIX

Redo the tours and ensure they both visit the same planets and see similar audiences although this will not negate contextual bias.

1. A company with work sites in five different countries has sent you data on employee satisfaction rates for workers in Human Resources and workers in Information Technology. Most HR workers are concentrated in three of the countries, while IT workers are equally distributed across worksites. The company requests a report on satisfaction for each job type. You calculate average job satisfaction for HR and for IT and present the report.

FLAW

Simpson’s paradox

FIX

From the lesson:

The lesson of Simpson's Paradox is this: pay attention to how groups differ from one another before averaging across them.

1. When people install the Happy Days Fitness Tracker app, they are asked to "opt in" to a data collection scheme where their level of physical activity data is automatically sent to the company for product research purposes. During your interview with the company, they tell you that the app is very effective because after installing the app, the data show that people's activity levels rise steadily.

FLAWS

Clearly an issue of self-selection. These people were not chosen at random. Also do not have a control group and a treatment group

FIX

Choose people at random. Have two separate samples – control and treatment. - of equal size and homogeneity among groups.

1. To prevent cheating, a teacher writes three versions of a test. She stacks the three versions together, first all copies of Version A, then all copies of Version B, then all copies of Version C. As students arrive for the exam, each student takes a test. When grading the test, the teacher finds that students who took Version B scored higher than students who took either Version A or Version C. She concludes from this that Version B is easier, and discards it.

FLAW

Many. The students are not chosen at random. They do not take the same test. They are not taking the test at the same time – contextual bias.

FIX

Select three groups at random from her students.

A/A testing