Martin Russell

Lesson 4 Challenge

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.

    a. **Faults:** There are two different Sith Lords, presenting two different slogans. We don't know if Palpatine is giving the better slogan, or if he's just a more convincing speaker. And captured droids, is that really a random sample of droids, or a random sample of the galaxy's inhabitants?

    b. **The Fix:** Have Palpatine present the "Together" slogan to group A, and then present the "Anger" slogan to group B. Also we need a much larger sample of droids than just the machines we captured at the last battle. And if we are truly concerned about recruiting throughout the galaxy, we'll need to sample from various planets throughout the galaxy.

2. 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.

    a. **Faults:** The major fault here is that our samples are coming from different populations, so the test of who is the best speaker is invalid from the start, some populations may be biased against Windu and some biased in favor of Binks. These are not random samples.

    b. **The Fix**: Draw samples from each world, mix them together, then divide them into group A, and group B. Then allow Binks to give a prepared speech to group A, then have Windu give the same speech to group B. The speeches need to be done on the same day, and in the same environment.

3. 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.

    a. **Faults:** You cannot get accurate statistics for either job type by combining the samples from each country. There is too many cultural, economic, and environmental varables in play to

aggregate the country samples. What might make workers satisfied in Vietnam could be completely different from what makes workers happy in Jamaica.

b. **The Fix:** This really needs to be 10 different studies, an IT study for each country, and an HR study for each country. Also, the questions asked to determine satisfaction will probably need to differ from one country to another due to cultural norms.

4. 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.

    a. **Faults**: Have we really established what “effective” means? Also, who exactly is wearing these fitness trackers, what are they tracking (heart rate, steps, calories burned). How do we know people aren’t cheating when using the app/fitness trackers?

    b. **The Fix**: We need to come up with a metric for “effective” and a way to keep people honest in there reporting/wearing of the devices. Also, there needs to be a baseline for what people’s fitness was before they had the tracker/app, perhaps this was a trend that was already occurring before they had the tracker/app.

5. 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.

    a. **Faults**: Are the number of each test taken the same (sample sizes)? Also, since smart students associate with each other, and stupid students associate with each other, when they came in the room together the group of stupid students may have chosen all of their tests from the same stack, the same goes for the smart students. So the samples are hardly random.

    b. **The Fix**: The tests should have been distributed evenly in one stack (a, b, c, a, b, c…), to make this a truly random sample, and the teacher should have passed the tests out herself instead of the students choosing the test on their own.