Unit 1: Lesson 4: Project 1: A/B Tests Getting Testy

1. Does a new supplement help people sleep better?

Problem:

Women in perimenopause can experience increasingly poorer quality of sleep as the body's hormone production levels change.

Restate the question:

Does a new supplement help perimenopausal women living in Pacific Northwest of US spend more time in deeper sleep stages?

Sample:

Randomly assign women to supplement vs placebo 50/50 into 4 groups. Each group will take the supplement for 3 month trial beginning in different quarters during the year: Jan, Apr, Jul, Oct.

Hypothesis:

If perimenopausal women spend more time in deep sleep, symptoms related to changing physiology will be improved. Supplement X will help perimenopausal women living in Pacific Northwest of US spend significantly more time in deeper sleep stages than a placebo.

Primary Outcome of interest:

Length of time in deeper sleep stages as measured by wearable technology that measures activity during sleep and determines stage of sleep by algorithmic estimation. Supplementation will result in significantly longer length of time in deeper sleep.

Secondary Outcome of interest:

Cortisol, estrogen, progesterone, and testosterone levels at wakeup as measured by saliva test. Supplementation will be associated with significant baseline change in cortisol levels (increase or decrease depending on symptoms). Supplementation will be associated with more stable range of estrogen levels, and an increase in progesterone and testosterone levels.

Sample population:

Women who are perimenopausal and live in a similar geographic region will ensure that a/b samples are similar. Randomizing the intervention will minimize bias.

Note:

Potential factors that could bias a/b test would be: socio-economic status, race, history of childbirth, work status, exercise levels, mental health status, actual time in bed, etc.

2. Will new uniforms help a gym's business?

Problem:

A new gym owner would like to focus on a different demographic market that can afford to pay for premium training/workout services: men 35 - 50 years old who are mid to upper level career professionals who want to optimize their fitness levels.

For the last year, sign-ups within this demographic have not converted at desired levels even though initial free consultations with this demographic have been steady and follow up surveys indicate good-to-high satisfaction. Follow-ups with these individuals indicate perception of non-professional presentation ("messy", "not professional", etc).

The new gym owner believes new gym sign-ups in this market would increase significantly if staff/trainers appeared more professional in their presentation. Currently, trainers/staff wear no standard attire.

Restate the question:

Will requiring staff/trainers to wear uniforms increase gym membership signups for new premium training/workout services within the 35-50 yr old male demographic that receives an initial free consultation?

Sample:

Men 35-50 yrs old who receive an initial free consultation

Hypothesis:

Mid-to-upper level career men in this demographic will be significantly more likely to sign up for premium training/workout services if they perceive the staff/trainers as having a "professional" appearance.

Primary Outcome of interest:

Conversion from initial free consultation to paid monthly/yearly membership.

Secondary Outcomes of interest:

Time to convert: payment at end of free consultation vs following 72 hours.

Sample population:

Given the intervention of new uniforms, it's not possible to expose a sample of the test population to staff/trainers with new uniforms and the other sample of the test population to "as-usual" no uniforms. This is an "all on" or "all off" example. However, given that the gym has tracked signups within this demographic over the last year, there is already a baseline comparison "a" group. All new free initial consultations with the proposed demographic, could then be assigned to the "b" group.

Given that seasonal changes can affect gym membership, comparisons could be made for sign up conversions based on quarterly periods or seasonal categories of time. Also, given that weekday vs weekend, am vs pm initial free consultation times could affect likelihood of signup, "b" group could also be compared with the "a" group based on these classifications.

Note:

There are so many factors that could confound the results of this a/b test that it would be advisable to investigate them prior to implementing this uniform change.

Individual trainer conversion rates may indicate which factors might lead to significantly higher vs. lower conversion rates. Other potential factors that could affect conversion rates: trainer's education or certification levels or years of experience, actual time spent with potential customer, goodness of fit between expectations and actual experience of initial consult, initial consult process, trainer's ability to communicate deliverables/past successes, gym organization, time of day consultation occurs and , whether or not the potential client "sees" himself as belonging/wanting to belong to the group of people who are working out at the gym, etc. Also, sample size is an issue - it would be difficult to determine impact of intervention with smaller sample sizes.

3. Will a new homepage improve my online exotic pet rental business?

Problem:

My website does not generate online reservations for rentals of exotic pets. I think it's because my homepage is not performing due to no "snazzy" photos of my Mr. Snuffleupagus, gaggle of Hippogriffs, pair of Kneazles, or my singular Crups.

Restate the question:

Will "snazzy" photos of my exotic pets on the homepage of my website significantly improve rental rates?

Sample:

Website visitors will be randomly assigned in equal numbers to "a": existing homepage design with no "snazzy" photos of my exotic pets above the fold; or "b": new homepage design that differs in that "snazzy" photos of my exotic pets are visible above the fold.

Hypothesis:

A homepage design with images that show visitors the exotic pets will be more effective in driving online rentals than a homepage design with no visuals.

Primary Outcome of interest:

Online rental rate of exotic pets driven by homepage

Secondary Outcome of interest:

Rental reservation rate

Sample population:

Online visitors to my exotic pet rental website homepage split randomly to see "a" condition: current home page design or "b" condition: new homepage design that features "snazzy" photos of my exotic pets for rent above the fold.

Note:

Similar to the previous prompt, there are so many potential factors contributing to online rental rate of the product that it would be important to explore those potential factors prior to implementing this A/B test. Assumptions are made as to what causes the poor online rental rate that may be inaccurate.

Are people who are interested in exotic pet rentals even arriving at the homepage after their online search? Once they get there, how long do the spend on the page? Where do they focus their attention on the page? Are they able to get to the online reservation process successfully? Do they start the online reservation process, but get stuck in the reservation process?

The problem is much too vague to be able to set up an A/B test that can really provide meaningful insight. The sample sizes may also be an issue. Since it's such a niche business, even a "successful" online exotic pet rental business may not receive the number of visitors to its site that are needed to have a big enough sample to derive any meaningful insight with confidence.

4. If I put 'please read' in the email subject will more people read my emails?

Problem:

Recipients of my newsletter are not opening the emails I send, so they're not checking out the amazing wizdoodles I make for them that will send their productivity into overdrive.

Restate the question:

Can I improve my click-through-rates to my wizdoodles from this email campaign by integrating a simple call-to-action in my email subject line?

Sample:

Email newsletter subscribers will be randomly assigned to condition A or B with gender evenly split among the groups (gender status captured at signup). Both groups will be subscribers within the United States (based on signup IP address). Both groups will receive the same email with the same subject line, except condition B email subject line will contain: "Please Read:" before the subject line text starts.

Hypothesis:

The call to action "Please Read" will significantly improve the click-through-rate for this email campaign. If people read the email, they will want to check out my whizdoodles.

Primary Outcome of interest:

Click-through-rate from links within my email

Secondary Outcome of interest:

Open rate of call-to-action email vs. no call-to-action email

Tertiary Outcome of interest:

Length of time call-to-action email open vs. no call-to-action email open

Sample population:

Email newsletter subscribers in the US who signed up for my email newsletter. Since gender and IP are captured at signup, use these two variables to increase chances of similarity in population across conditions. Eliminate potentially large regional differences influencing likelihood of opening email. Also eliminate chance that call-to-action is less/more effective due to difference in gender.

Note:

While the initial problem described wanting to get email campaign readers to "read" the email, it seems like a poor primary outcome. What do I want my readers to do as a result of reading my email?

Open rates seem to be a clearer metric to know that an email recipient at least saw the email content. While opening an email is a precursor to reading it and length of time the email is open can be a helpful indicator of interest, it does not indicate action toward a goal I have for my readers.

From the perspective of the sender, it would seem to be more important to measure how strongly the reader is influenced to take a particular kind of action.

The process could be broken down into two specific actions: 1) getting to the site; 2) doing something specified on the site (buying/donating/volunteering, etc).

Because the ultimate goal would be for a reader to engage with the site in some meaningful way, checking out the whizdoodles that I think can improve my readers' lives seems to be a reasonable metric to measure that addresses the first of two desired options.

Once rates of getting email campaign subscribers to the site is optimized, then follow up A/B testing could focus on optimizing the email campaign for sales/donations/volunteering hours, etc.