1. Does a new supplement help people sleep better?

This could be set up as an A/B test as follows: We would need to select a random sample of people from the population and then split the sample up into two groups. It’s important to make sure that the groups are split up randomly and compare the demographics of the two groups to ensure that there aren’t any substantial differences between them. For example, having a group of just males and another of just females to test the supplement would provide an ineffective design, because differences that may be observed may be due to gender instead of the supplement.

The hypothesis would be that the new supplement does help people sleep better.

After randomization, one group would get the supplement and the other would not – or perhaps, even better, they would get a placebo supplement. Sleep quality would be measured in the two groups with both scientific tests (like REM monitoring, for example) and also self reported sleep quality assessment of the subjects. These would be the measured variables and even the outcome of interest, which could be composite definition of ‘better’ sleep quality.

1. Will new uniforms help a gym’s business?

This would be set up as follows:

There is no way to really randomize, because people would come into the gym on different days and there is no way to make sure that only some of them would see one uniform vs another one. Here the trial would have to run multiple times to capture everyone customer’s exposure to the uniforms and to ensure that things like different day of the week did not throw the results.

The hypothesis could be that there is no difference in how the uniform affects business and could be measured in either the revenue generated during a certain interval of time or maybe the number of new memberships during a certain interval of time.

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

Here, metrics could be collected that are associated with the ‘old’ home page – things like web traffic and number of pet sales. Then the new web page could be launched and the metrics could be remeasured.

Here the hypothesis could be that there is no difference in the number of sales or web traffic between the two different web pages. There is also no real way to randomize which customer seas which web page.

1. Will putting ‘please read’ make people read e-mail more often?

Randomize the list of contacts into two groups. The groups have to be balanced demographically to ensure that nay measured difference is due to the factor being measured instead of some other difference between the groups.

The hypothesis could be that putting ‘please read’ has no effect on how many emails get read.

The emails would get sent out to the different contacts at the same time, to ensure a temporal effect doesn’t get introduced. The result could be measured based on how many of each different email type got opened.