**Original Flawed Research Proposal:**

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.

**Why was it flawed:**

1) Many people install apps, use them once, but then stop using them and do not delete them. Thus, this test does nothing to determine who is an active user and who is not.

2) Cause and Effect Question: There is no control group. Perhaps everyone who installs the app has created a resolution to be more active. And it's not the app that is helping them be more active.

3) Seasonality. It's possible the test started when the weather started to warm up causing people to become more active.

**My modified Research Proposal:**

**Problem**: Determine if Happy Days Fitness Tracker app is effective in helping people increase activity levels.

**Solution / Hypothesis**: People who use the app experience higher activity levels within a month of using the app.

**Experiment**

Compare the change in activity levels for people who don't use the app vs people who do use the app. (across various subgroups such as geography, age, gender)

Design

Change the data collection method:

Before registering for the app, users are forced to agree to terms of use which indicates that data may be aggregated for research purposes and to improve the app over time.

Data is automatically collected for all users who install the app.

Two Groups will be analyzed:

1) Control group: Those who install the app, but do not meet the active user criteria.

2) Test group: Those who install the app and meet the active user criteria

To be considered an active user, you must interact with the app at least once per week over the first 4 weeks after installation.

Once the 2 groups are established, we can divide them into subgroups such as location, age, and gender.

Analysis

Change in activity level will be analyzed for both groups (and all subgroups) for 2 months after installation.

Benchmarks

For the hypothesis to be true, the test group must have experienced an increase in activity from Month 1 to Month 2 that is at least 5% higher than the control group. (With 95% confidence)