Case Study

Let's imagine we have three potential assets that we can show users at the end of an article:

Asset A is a quiz designed to show users the best medications to manage their gut health. If a user goes on to buy a recommended product listed at the end of the quiz, we will get **$5** in revenue per conversion.

- Conversion = user took the quiz and bought a product
- Nets Health.inc $5 of revenue

Asset B is a sponsored newsletter that users can sign up for to get support for their certain gastrointestinal chronic conditions. We can assume that one newsletter signup (i.e. conversion) nets us **$7** in revenue.

- Conversion = user signed up for the newsletter
- Nets Healthl.inc $7 of revenue

Asset C is an advertisement for a probiotic that users can buy through Amazon and nets us **$2.50** per conversion.

- Conversion = user clicked on amazon link and purchased the probiotic
- Nets Health.inc $2.50 per conversion

In this example, let's assume that our assets were being randomly shown to users in the past, but we want to start using the data we have to understand what variables are most important in eliciting conversion. We hope that with the right information, we will be able to show the appropriate assets for each user and page visit to maximize **conversion rate** and **revenue**.

The business team has asked you to help make sense of their data and has provided one month's worth of page views from a subset of page categories. They want to know if they can increase **revenue per page view** by matching certain user and page information to specific assets, assuming the ability to do real-time-decisioning to match users to assets at the end of an article.

There are three broad areas that they would like you to report on:

1. Tell us about the variables.
   Things to consider: Which variables are the most helpful in understanding if a user visiting a page is going to buy a product on one of the three assets or not? Are there certain variables that are not useful to collect?
2. Tell us about the different assets.
   Things to consider: What is each asset's average conversion rate? How does the revenue per conversion differ for the same asset across different dimensions? Are there any clear differences between the three different assets on what type of users and pages lead to conversion?
3. Which user/page combinations should we should match to which asset?
   Things to consider: Are there types of users or pages that would be a better fit for a different asset? How much incremental revenue could we have made if we changed certain groups of users more appropriately?

Data Dictionary:

Pageview ID (string) - unique ID to represent the a each unique page view

User ID (string) - unique ID to represent the user who viewed the webpage

Known Diagnosis (string) - known condition of a user

Page Category (string) - categorical grouping of the page being viewed

Device Type (string) - device the user utilized to reached to the site

Page Topic Description (string) - more detailed for a page category

Session Start Time (timestamp) - the time at which the user's Health.inc session started

Asset Loaded Time (timestamp) - the time at which the asset was loaded on the page and displayed to the user

Time of Day (int) - hour of the day during which the user visited the page

Return Visitor (bool) - an indicator if a user has visited the site before

Asset Shown (string) - indicates which of the 3 assets was shown

Conversion (bool) - an indicator of whether the user converted and generated revenue for Health.inc