Below we have a series of questions for you to translate into a technical plan. For each question, describe how you would make it testable and translate it from a general question into something statistically rigorous. Write your answers down in a shareable document and submit the link below.

1.       You work at an e-commerce company that sells three goods: widgets, doodads, and fizzbangs. The head of advertising asks you which they should feature in their new advertising campaign. You have data on individual visitors' sessions ([activity on a website](https://en.wikipedia.org/wiki/Session_(web_analytics)), [pageviews](https://en.wikipedia.org/wiki/Page_view" \t "_blank), and purchases, as well as if those users [converted](https://en.wikipedia.org/wiki/Conversion_marketing) from an advertisement for that session. You also have the cost and price information for the goods.

We want to determine which good produces the highest expected profit outcome based on prior visitor data. The expected profit for each good = (Price – Cost) \* Number of sessions \* Conversion rate per session

Other items to consider:

Are increased page-views is correlated with increased conversion rates? (Do we need to target customers with more page views and/or longer sessions?)

What types of products had the highest conversion rates historically? Which of our three products is the most similar in functionality?

2.       You work at a web design company that offers to build websites for clients. Signups have slowed, and you are tasked with finding out why. The [onboarding funnel](https://en.wikipedia.org/wiki/Funnel_analysis) has three steps: email and password signup, plan choice, and payment. On a user level you have information on what steps they have completed as well as timestamps for all of those events for the past 3 years. You also have information on [marketing spend](https://en.wikipedia.org/wiki/Marketing_spending) on a weekly level.

I would start by looking at total page-views, if I have access to that information. It could be the case that overall visitor volumes have fallen, which is driving lower signups. Compare this to the weekly marketing spend – maybe we have invested less money in advertising which has reduced our exposure. (We can test different marketing spend levels to see what impact it has on visitor counts.)

Next, look at the percentage of visitors remaining after each step in the ‘funnel’ over time. Group the data by week and look at the trends over time. Are there more people who don’t even sign up with an email and password? Are there more people stopping at the plan choice? Perhaps the plans are too complicated, or perhaps the plans are not as comprehensive as they need to be because the market has changed. (We can test adding new plans, modifying existing plans, removing plans, or rearranging the existing plans to see if that has any effect on follow-through.) Are more people stopping at the payment step than they were historically? Perhaps the plans are too expensive now – have we changed the price recently? Perhaps we need to make it easier to pay by accepting more forms of payment? (We can test different pricing levels for each plan to see what the impact of price is on follow-through.)

Looking at all of this data over time can help us solve the problem.

3.       You work at a hotel website and currently the website ranks search results by price. For simplicity's sake, let's say it's a website for one city with 100 hotels. You are tasked with proposing a better ranking system. You have session information, price information for the hotels, and whether each hotel is currently available.

We can potentially enhance the rankings by taking attributes of the unavailable hotels and giving them higher priority in a new ranking algorithm (such as location). Hotels that are frequently unavailable (assuming we have some historical data) will also be given a higher rank due to demonstrated outperformance. We should also remove the hotels that are currently unavailable from the ranking list entirely, or at least move them to the bottom of the rankings.

We can also improve the ranking system by focusing on the following session information:

1.       Hotels with the highest number of page-views (more views should increase the hotel rank).

2.       Hotels with short sessions that were not booked should be penalized in the rankings, but should rank above hotels that were not looked at.

3.       Hotels with long sessions that were booked should be increased in the rankings.

Over time, you can see how the bookings match up with the rankings, and adjust the weights of each of the above factors accordingly to get the rankings to better align with bookings.

4. You work at a social network, and the management is worried about [churn](https://en.wikipedia.org/wiki/Churn_rate) (users stopping using the product). You are tasked with finding out if their churn is atypical. You have three years of data for users with an entry for every time they've logged in, including the timestamp and length of session.

Check to see whether average session lengths have changed over time – this could indicate a gradual decrease in interest, or it could be sudden. Check to see how often the customers are logging in – again, looking at this trend over time. Perhaps we changed a feature that guests did not like, and hence they stopped logging in altogether. Check to see whether the average session length or the login frequency changed, and determine whether it was gradual or sudden. Compare these changes over time to feature changes/introductions/removals from the site, and you can begin to make inferences about what could be causing the churn.