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.

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**, **pageviews**, and purchases), as well as whether or not those users **converted** from an advertisement for that session. You also have the cost and price information for the goods.

We can use the session data to see what percent of people ultimately buy which products. I would track how many bought their first viewed item without further activity, went to a second item and bought that or went to a third item or back to original and purchased. Giving us how

many bought item X vs how many bought item X after Y page views needed. I wouldn do the same with the current advertisement to see if people are buying the advertised item or viewing/ purchasing others.

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** 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** on a weekly level.

On the on-boarding side I would see if more people are stopping before plan choice or payment. Whichever is more of a problem we could try reaching out to users on if they think the plans aren't robust enough or pricing is too high. How is the marketing team using resources. Are they spending 80% of the budget on newspaper ads

when only 10% of users come from those ads? Data from user polls would probably help significantly with slowed signups.

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.

The rankings could have an optiion for availablity. We can also break down how people travel the site using session information and create other ranks based on that.

4. You work at a social network, and the management is worried about **churn** (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.

I would break down the information into 2 groups. Those who have never signed back on and those that do. With the group that signs back on how long in between do they take. Are more people above or below the average. You could compare statistics on length of sessions on people who use it periodically vs people who use it habitually.