As we outlined in the previous section, a lot of data science work can be classified as translation: taking a question and translating it into mathematical or statistical tests or taking statistical results and translating them into something everyone can understand.

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, 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.
   1. I would try to understand which items are bringing a high level of profit to the company as these items might warrant more attention in advertising
   2. I would see which items are being viewed the most and purchased the most to get an idea of what customers want, what is trending, and has potential to grow.
   3. I would especially focus on users converted from an advertisement as these are likely new customers and not repeat customers who already know the brand
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.
   1. I would compare market spending to the slowing of signups. Are these correlated? Is this a bigger trend in the market or something unique to the company?
   2. I would look at the whole sign up process to see if there is a bottleneck somewhere – is it that customers are particularly burdened or turned off at some point in this process? As it always been the case for the last 3 years?
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.
   1. I would look at all the hotels ever chosen by customers in the session information and find how it correlates with other variables that are not price. I would show the many ways that customers in fact predictably follow other criteria in choosing their hotels, for example hotel quality and location. I would propose we rank hotels by location and/or quality in addition to price.
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.
   1. I would look at the data to see the structure of the churn, that is to say, if it happened suddenly, or if its been constant, or if it is periodic. I would see if the churn of the company correlates with bigger market trends in the city or market to capture if this is a larger trend or not.

Discuss your answers to each of these questions with your mentor during your next session.