Ten questions an audience would ask with answers:

1. Do you think that people answer truthfully on these types of questionnaires?
   a. This is a good question, because if the participants are not honest then the results would be useless. However, because of several more recent events in the United States and globally, attitudes toward mental health issues have shifted and now it is regarded as important to get help and address a potential mental health issue. This is a fantastic change from the past. Because of this, I do think people will answer more truthfully in these surveys.
2. Did you think about doing a different feature reduction method? Why did you pick chi-square?
   a. I researched this topic and for the type of data that I have (26 out of 27 variables are categorical) using a chi-square method was the best choice to reduce the number of features. The criterion of < 0.2 is a standard cut-off.
3. Do you think you could draw more conclusions about the data?
   a. I think the data pointed out that people may not be aware of benefits for mental health services from their employer and may not know how to use them. Being that "care options" and "benefits" were important factors as well as the EDA figure for "how easy is it to take medical leave" where the finding is really in the lower ends, when those on treatment have a higher rate than those not on treatment point to a trend that employees need more assistance in navigating mental health benefits.
4. What other aspects of the data did you think would be worthy of more investigation?
   a. I think it would be important to do more EDA and see what other relationships may pop out in the data. I covered this but it could have been done more extensively.
5. Do you think that the main reason that employers don't offer these benefits is cost?
   a. At this stage, I think having comprehensive benefits can be a challenge for smaller businesses. The government should find a way to offer services to such small businesses.

Also, since Covid, there is a huge uptick on seeing a mental health profession online – though it would be important to understand if this treatment is as effective as it is in person. Perhaps online could be less expensive too.

6. What about exploring these types of data with more demographic characteristics like race or type of employment – is that level of analysis possible?
    a. This is a great question. For this dataset, this was not possible as these variables were not part of the dataset. However, it would be ideal to do further research and include such characteristics.
7. Why did you decide to impute the value of 'interfered with work' the way you did? Is that statistically sound?
    a. I imputed the value of 'interfered with work' by treatment because there were about 258 missing for Treatment = No and 4 missing for Treatment = Yes. The most frequent score for Treatment is Sometimes, however that is driven by Treatment = Yes. Thus, I think it would introduce a bias to treat those who are not treated the same as those who were.
8. Don't you worry that this type of research puts the burden of medical and mental healthcare on the employer? What if the employer cannot afford to have good benefits?
    a. I think that the United States is set up in a certain way where the burden of many things lies on the employer. We do not have a national health care program and because of this, it falls to the employer. The same is true for mental health care. As I mentioned earlier, there are different alternatives to supporting mental health benefits for employers and these should be investigated prior to assuming it is not possible to provide.
9. Did your reduced feature logistic regression model outperform your initial model?
    a. Yes, it did. The base model performed well, and the reduced feature model had a slightly higher accuracy score.
10. Why didn't you analyze the data for the United States separately?

a. This was mostly due to time – there are more analyses to run with this dataset as well as getting additional years of data to develop a more comprehensive model.