Hi Jenna,

Thank you for sharing your project with me, I think it is a pretty interesting idea. It's unfortunate they don't have better demographic data to help you out. See my feedback below and please let me know if you have any questions.

Best,

Becca

1. Strengths of the project (things you particularly like about it):
    - The project has merit because, like many of the resource sharing businesses, Airbnb has a pretty new business model. The behavior of their clientele is likely to be different from typical boarding websites.
    - It is reasonable and expected that the demographics of the person booking an Airbnb would be indicative of which country they book in.
2. Comments about things you think could be improved:
    - The Timestamp feature could be improved by breaking out the time of booking. Time of day could be an interesting and potentially predictive feature.
    - You're currently using KNN as a prediction model. Do you plan to use any other type of prediction model? Could you use decision trees or random forests to improve your accuracy?
3. Comments about their code:
    - I looked into your prediction model and, as you know, it doesn't change much based on the features added. I tried a couple of different things:
        - By changing the KNN value, I was able to change the accuracy significantly (made it much worse). It seems like it's working.
        - I used a number of different features for the model and (with KNN at 30) I could change the accuracy, but it did stay very close to 70%.
        - Using age and language I found that the ideal KNN was 15. With that KNN value I was able to push the accuracy up to 71.2%...not much better but improved!
    - Comment conclusions: I think you were concerned there was something wrong with the workflow of your model, but I didn't find anything. Maybe check your work with a different model? It's possible that the features just aren't very predictive. The null accuracy is already pretty high because so many people are booking in the US. Since the demographics aren't very telling, it could be difficult to improve over the null accuracy of 70%.

4. Suggestions for next steps:
    - **Date/Time stamp**: pull out time to use as a feature
    - **Model**: try something besides KNN
    - **Hypothesis: Change your hypothesis** - Instead of predicting country based on your features, change your response variable to a binary predictor for USA vs NOT USA. I bet your timestamp and language feature will give you a model with more predictability than the null with that binary response variable.