Jared Yu
Module 9

I think that your discussion is interesting. I had a similar conclusion, where I think we both are referencing the same section from the textbook that directly discusses this topic. In my approach, I tried to connect the problem equation to the generalized form seen within the text, and then to make my argument from there. I see however that you include insight into the concept where in the b_1 only model, the model is becoming biased towards b_1, where I suspect that in reality there is additional impact coming from b_2 which is being neglected. This makes sense to me, and I like that you also included the part about how the math shows this fact, where we must also consider the possibility that $b_2 = 0$. You also mention that when looking only at the subset, we are not accounting as much for the variance that would come from b_2 being included in a full model. I didn't actually fully understand the textbook argument, since some of the math didn't connect for me. However, I think from having studied statistics and machine learning that there is inevitably a bias-variance tradeoff. If we increase the bias towards b_1 in the subset model, then we are probably reducing the variance at the same time. I am not sure entirely how this can be logically connected, except maybe for that the subset model is less accurate but at the same time more stable.