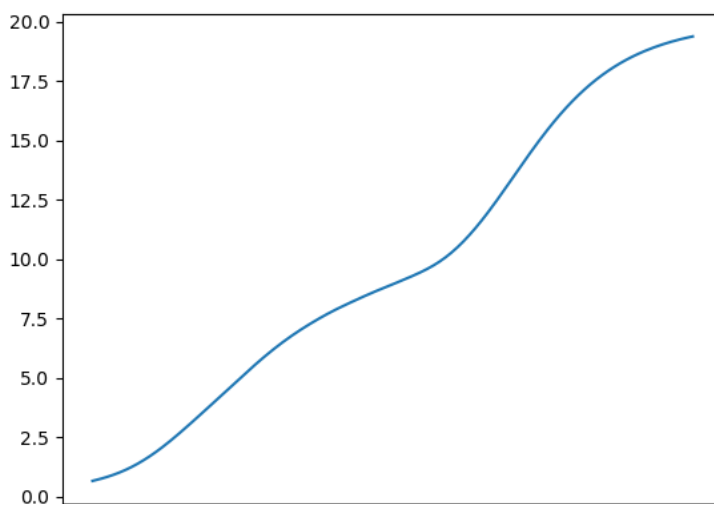


Question 2 c)

As τ increases, the value of a decreases, which means bigger τ leads to less sensitivity to close neighbors. Thus, I would expect for the training error to increase as τ increases, because the model is giving more significance to the data points further away that likely have different properties than the test point. When τ is small, it's overfitted because a large portion of the weight is contributed by the few training points close to it that are very similar in nature, and so I expect the training loss to be small. For testing error, the situation is a bit more complex as very small τ below a certain limit can lead to overfitting (discussed above) and an increase in testing error. I expect the error to drop as τ increase.

The training error is exactly what I expected. The testing error curve is a bit unexpected as it started to increase again beyond a certain limit and plateaued, which I did not expect. Although that would make sense because the model will likely perform worse by giving too much significance to the far away data points that likely behave very differently than neighboring data points, and results in underfitting. The plateau at the end likely means the model is starting to perform more like a normal linear regression where local weights don't have much more significance than "far-away" weights, and the contribution of weights by different data points become more evenly. The first part of the validation graph was as expected that beyond a below τ value the loss started to increase due to overfitting.

Graph 1: Training Loss (y-axis) vs Tau (x-axis)



Graph 2: Validation Loss (y-axis) vs Tau (x-axis)

