Problem 3

1. The Gauss-Markov theorem states that under certain assumptions, the least squares estimator in linear regression is ultimately the best estimator among any other unbiased estimator as it has the lowest possible variance. That is important because we know that we have a reliable estimator that is not only unbiased but also has the least variance possible which effectively fits the data well (by reducing the reducible error to a minimum leaving only the irreducible error) provided the assumptions hold.
2. The Gauss-Markov error term assumptions:
   a. The errors are uncorrelated: the error for a given variable should be independent from that of another variable.

      $Cor(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j) = 0 : i \neq j;\ \forall i, j$

   b. The errors have equal variances (Homoscedasticity): the error across variables is constant and not a function of X.

      $Var(\boldsymbol{\varepsilon}_i) = c,\ \forall i$

   c. The errors should have an expected value of 0: the mean of each of the errors is 0.

      $E[\boldsymbol{\varepsilon}_i] = 0,\ \forall i$
3. Theoretically, having no bias and the least variance possible is promising and should mean that we have the best test error compared to other models including

   (1) those with less variance but are biased which would underfit.

   (2) those with higher variance but are unbiased which would overfit.