1. **Nursing Home Utilization** data for the following questions. this exercise involves data filename "WiscNorsingHome" Frees page 59.

   You will need to read a .csv file into R by read.csv(file,header=TRUE) for this questions.

   You decide to examine the relationship between total patient years (LOGTPY) and the number of beds (LOGNUMBED), both in natural logarithmic units, using cost-report year 2001 data.

   You are asked to perform below in **both excel workbook and verify with R**

   (a) Estimation of coefficients, β_0, β_1

      1. Calculate the 2×2 matrix, $\mathbf{x}^T\mathbf{x}$, $\left(\mathbf{x}^T\mathbf{x}\right)^{-1}$ and $\mathbf{x}^T\mathbf{y}$

      2. Calculate the 2×1 estimate $\hat{\beta}$

   (b) Calculate the diagonal element of the hat matrix $\mathbf{H}$, h_{ii}.

      1. Calculate the inverse of 2×2 matrix, $\mathbf{x}^T\mathbf{x}$

      2. Calculate $\mathbf{x}_i^T \left(\mathbf{x}^T\mathbf{x}\right)^{-1} \mathbf{x}_i$, where $\mathbf{x}_i^T$ is ith row of $\mathbf{x}$, $(1, x_i)$.

   (c) Calculate the fitted value $\hat{\mathbf{y}}$

   (d) Calculate R^2, adjusted R^2, $F-$stat, p-value, and the mean squared error (MSE) s^2.

   (e) Calculate the standard residual vector $\mathbf{r}$

      1. Calculate the residuals e_i

      2. Calculate the standard residual $r_i = \frac{e_i}{s\sqrt{1-h_{ii}}}$.

   (f) You are asking to accessing the effectiveness of predictor, LOGNUMBED. Also compute the corresponding p-value.

      1. Hypothesis testing: $\mathcal{H}_0 : \beta_1 = 0$ *versus* $\mathcal{H}_1 : \beta_1 \neq 0$ at the 5% levels of significance using a t-statistic.

      2. Compute the p-value. what is your assessment of the estimate $\hat{\beta}_1$?

      3. Provide a 95% confidence internal (CI) corresponding to the point estimate for β_1.

      4. Provide a 99% CI corresponding to the point estimate for β_1.

   (g) At a specified number of beds estimate $x_* = 100$, do these things:

      1. Find the predicted value of LOGTPY.

      2. Obtain the 95% prediction interval.

   (h) (**Perform in R**) Fit the basic linear model using LOGTPY as response variable and LOGNUMBED as explanatory variable. Compare results with what you calculate above.