**Name: Taylor Ward**

**BIOS 6102**

9. A project evaluated whether the four factors (InCostAid, Admit, GradRate, and OutCostAid) are predictors for debt in dollars at graduation (AveDebt) using the dataset of “**BESTVAL.csv**.”

|  |  |
| --- | --- |
| Variable name | Contents |
| AveDebt | debt in dollars at graduation |
| InCostAid | in-state cost per year after need-based aid |
| Admit | The admittance rate |
| GradRate | Four-year graduation rate |
| OutCostAid | Out-of-state cost after aid |

1. Provide descriptive statistics and histogram for all variables of interest (outcome and predictors)
2. Provide pairwise Pearson correlations for the four predictors and check multicollinearity. Comment on these correlations using the criteria provided in class and whether it is appropriate to consider all 4 predictors for building a multivariable model.
3. Create scatter plots for each predictor with AveDebt. Comment on the plots.
4. Perform univariate linear regression for each predictor and complete the following table.

|  |  |  |
| --- | --- | --- |
|  | Beta (SE) | p-value |
| InCostAid |  |  |
| Admit |  |  |
| GradRate |  |  |
| OutCostAid |  |  |

1. Conduct a multivariable linear regression using both (1) univariate model approach and (2) automatic variable selection. Comment what your final model is and your reason for choosing it as the final model. <10 points>
2. Obtain the Jackknife residuals, leverage, and Cook’s distance from your final multivariable model and comment on the problematic points.