**Download the file Golf.jmp. Use JMP to develop a multiple linear regression model to predict the Earnings/Event using the data found in Golf.jmp. Consider the four independent variables listed in the table below. Find the best model and check assumptions.**

|  |  |  |
| --- | --- | --- |
|  | EARNINGS | Average Earnings per Event |
|  | SCORE | Average Score |
|  | DRIVE\_D | Average Drive Distance |
|  | DRIVE\_A | Average Drive Accuracy |
|  | PUTTS | Average Putts per Round |

**[1] Create a correlation matrix for the variables EARNINGS, SCORE, DRIVE\_D, DRIVE\_A, and PUTTS using JMP.**

|  |
| --- |
|  |

**[2] What is the correlation coefficient for EARNINGS and SCORE? Interpret the linear relationship between the two variables.**

|  |  |
| --- | --- |
| Correlation  Coefficient | Interpret the linear relationship |
|  |  |

**[3] Does the correlation matrix indicate a potential multicollinearity problem? If so, which pair(s) of independent variables are a concern? If your answer is no, state why.**

**Use JMP the fit the full model (all of the independent variables). Include the confidence interval and VIF for each of the regression parameters. Paste JMP output below. Include Summary of Fit, Analysis of Variance, and Parameter Estimates**

**[4] Summary of Fit, Analysis of Variance, and Parameter Estimates**

|  |
| --- |
|  |

**[5] State the null and alternative hypothesis statements for testing a multiple regression model. (F-test)**

|  |  |
| --- | --- |
|  |  |
|  |  |

**[6] What is your conclusion about the overall model?**

**[7] State the null and alternative hypothesis statements for testing the regression coefficient of an independent variable. (t-test)**

|  |  |
| --- | --- |
|  |  |
|  |  |

**[8] What are your conclusions for each of the independent variables in the model?**

**(Is the variable significant or not)**

|  |  |  |
| --- | --- | --- |
| Variable |  | Significant at the 0.05 level? |
| SCORE |  |  |
| DRIVE\_D |  |  |
| DRIVE\_A |  |  |
| PUTTS |  |  |

**[9] Is there a variable that should be removed from the multiple regression model? If so, which variable and why should it be removed?**

**[10] Fit a new multiple regression model with the remaining independent variables.**

**Summary of Fit, Analysis of Variance, and Parameter Estimates**

|  |
| --- |
|  |

**[11] What are your conclusions for each of the independent variables in the model?**

**(Is the variable significant or not)**

|  |  |  |
| --- | --- | --- |
| Variable |  | Significant at the 0.05 level? |
|  |  |  |
|  |  |  |
|  |  |  |

**[12] Is there a variable that should be removed from the multiple regression model? If so, which variable and why should it be removed?**

**[13] Fit a new multiple regression model with the remaining independent variables.**

**Summary of Fit, Analysis of Variance, and Parameter Estimates**

|  |
| --- |
|  |

**[14] What are your conclusions for each of the independent variables in the model?**

**(Is the variable significant or not)**

|  |  |  |
| --- | --- | --- |
| Variable |  | Significant at the 0.05 level? |
|  |  |  |
|  |  |  |

**[15] Do the values of Variable Inflation Factors indicate any potential problems for this multiple regression? Why or why not?**

**[16] State and interpret the Confidence Interval for the estimate of the regression coefficient of the independent variable SCORE.**

**[17] What proportion of the variability of EARNINGS is explained by the multiple regression model. Interpret this value in terms of the problem.**

**[18] Copy and paste the Residual by Predicted Plot**

|  |
| --- |
|  |

**[19] Describe the shape of the distribution. Does the residual by predicted plot support the assumption of constant variance? (Is there an extreme change in the variance?)**

**[20] Use the multiple regression model to estimate the earnings per event for a golfer with an average score of 70 and an average of 32 putts per round. Compute a point estimate and construct a 95% Confidence Interval for the Mean. Interpret the confidence interval in terms of the problem.**

|  |  |
| --- | --- |
| **Point estimate** |  |
| **Confidence Interval for the Mean** |  |
| **Interpretation of the**  **Confidence Interval for the Mean** |  |