Name\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Finance 4355: Exam 1

Predictive Analytics

*Instructions: Do the exam yourself, and sign a statement stating that you did not receive help from other persons besides the instructor. You may use a calculator/excel but must show handwritten work where it is requested. Upload a word processed document on Blackboard by Sunday, October 3rd 11:59 pm. If you cheat, you will receive a score of zero on the exam and a grade of F in the class. Show your work. And, Remember, the Force will be with you. Always.*

Data

|  |  |  |  |
| --- | --- | --- | --- |
| **Obs.** | **X** | **Y** | **Y hat** |
| 1 | 5 | 23 |  |
| 2 | 5 | 16 |  |
| 3 | 6 | 25 |  |
| 4 | 8 | 28 |  |

Part I: Regression Basics

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT A |  | |  | |  |  |
|  |  | |  | |  |  |
| *Regression Statistics* | | |  | |  |  |
| Multiple R | 0.97688814 | |  | |  |  |
| R Square |  |  | |
| Adjusted R Square | 0.939080583 | |  | |  |  |
| Standard Error | 3.277208147 | |  | |  |  |
| Observations | 5 | |  | |  |  |
|  |  | |  | |  |  |
| ANOVA |  | |  | |  |  |
|  | *df* | | *SS* | |  |  |
| Regression | 1 | | 672.9797 | |  |  |
| Residual | 3 | | 32.22028 | |  |  |
| Total | 4 | |  | |  |  |
|  |  | |  | |  |  |
|  | Coefficients | | Standard Error | | t Stat | P-value |
| Intercept | 2.531468531 | | 3.604616 | | 0.702285 | 0.533087 |
| X | 3.43006993 | | 0.433317 | | 7.915839 | 0.004203 |

Model A:

1. Use the Excel Output from Model A to answer the following questions.
   1. Is X a significant predictor of Y? Interpret the coefficient of X. (10 points)
   2. Find the Total Sum of Square. Find Mean Squared Regression (MSR) and Mean Squared Error (MSE) (10 points)
   3. What is the F-statistic? Is the F statistic significant? (assume critical value is 3.5) (10 points)
   4. Calculate Y hat. Generate a new column of predictions based on the output from Model A (FYI: You are using your model to forecast outcomes for a set of points that were not used to build your model). Feel free to copy and fill in the empty column on previous page. (10 points)
   5. Find the R square. Is this a good model? Explain why or why not. (10 points)
2. What does the scatterplot below tell us, if anything? (20 points)

Model B: where Y = Earnings, X = Experience, and X Square dis Experience Squared

1. Suppose now that we’ve squared X and added that to the model.
   1. Are the predictors significant in this model B? Explain. (10 points)
   2. Is this a good model? Compare this model B to model A. Which model should you use for predictive analytics purposes? Why? (10 points)
   3. Interpret the coefficients from Model B. (10 points)
   4. Can we interpret the coefficients from Model B as a causal parameters? (20 points).

|  |  |
| --- | --- |
| *SUMMARY OUTPUT B* | |
| *Regression Statistics* | |
| Multiple R | 0.978084 |
| R Square | 0.956648 |
| Adjusted R Square | 0.913296 |
| Standard Error | 3.909722 |
| Observations | 5 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
| Intercept | 8.670659 | 19.1831 | 0.451994748 | 0.695563 |
| Experience | 1.886228 | 4.729599 | 0.39881345 | 0.728582 |
| Experience squared | 0.080838 | 0.246166 | 0.328388911 | 0.773812 |

Part II: More Regression

1. Carefully state and explain the L.I.N.E assumptions in your own words. (20 points)
2. Regression Line (40 points):

Use the table below to find the regression coefficients for and . Explain what and represent. (See back for formula). Show **all of your handwritten** work for calculations.

**Calculation of Regression Coefficients (40 points)**

|  |  |
| --- | --- |
| **Beer (X)** | **Cigarettes (Y)** |
| 8 | 16 |
| 6 | 13 |
| 2 | 4 |
| 4 | 9 |

Formulas