Question 1

The office of the Registrar took a random sample of 427 students and obtained their grade point averages in college (COLGPA), high school (HSGPA), verbal scores in the SAT (VSAT) and math scores in SAT (MSAT). The following model was estimated (subscript $t$ is omitted for simplicity).

$$\text{COLGPA} = \beta_1 + \beta_2 \text{HSGPA} + \beta_3 \text{VSAT} + \beta_4 \text{MSAT} + u$$

a. The unadjusted $R^2$ was 0.22. Because this is very low, we might suspect that the model is inadequate. Test the model for the overall goodness of fit (using a 1 percent level of significance). Be sure to state the null and alternative hypotheses, the test statistic and its distribution, and the criterion for acceptance or rejection. What is your conclusion?

ANSWER:

First we test the hypothesis:

$$H_0: \beta_2 = \beta_3 = \beta_4 = 0$$
$$H_1: H_0 \text{ is not true and not all slope coefficients are zero.}$$

Next we test the level of significance where $\alpha = 0.01$ which is a given value.

Degree of Freedom numerator

\[= k-1\]
\[= 4-1\]
\[= 3\]

Degree of Freedom denominator

\[= n-k\]
\[= 437-4\]
\[= 423\]

With these values we then test the statistics using the equation as follows.

\[F = \frac{R^2}{\frac{k-1}{(1-R^2)} \frac{n-k}{(n-k)}}\]

It should be noted that a specific decision rule will be used. We reject $H_0$ if the calculated $F$ value is greater than the tabulated $F$ value. Therefore, if calculated $F$ is greater than 1.7816 we will reject $H_0$.

Calculating using above equation:
Based on the results obtained, we will reject the H0 because our F calculated value is greater than the F tabulated value (39.8 > 3.7816). Based on the results, the null hypothesis is rejected and it is concluded that at minimum one of the regression coefficients is non zero.
b. Test each regression coefficient for significance at the 1 percent level against the alternative that the coefficient is positive. Are any of them insignificant? In the table fill in the test statistic and indicate whether the coefficient is significant or insignificant. State their distribution and degrees of freedom.

**ANSWER:**

For the particular questions we are testing individual regression coefficients. Hence we will use the T test for testing the hypothesis about individual partial regression coefficient.

\[ H_0: \beta_1 \leq 0, \ H_1: \beta_1 > 0 \]
\[ H_0: \beta_2 \leq 0, \ H_1: \beta_2 > 0 \]
\[ H_0: \beta_3 \leq 0, \ H_1: \beta_3 > 0 \]
\[ H_0: \beta_4 \leq 0, \ H_1: \beta_4 > 0 \]

Using these we will test for significance with \( \alpha = 0.01 \), one tailed t test.

The following formulas are used:

\[
\hat{\beta}_1 = t = \frac{\hat{\beta}_1 - \beta_1}{Se (\hat{\beta}_1)}
\]
\[
\hat{\beta}_2 = t = \frac{\hat{\beta}_2 - \beta_2}{Se (\hat{\beta}_2)}
\]
\[
\hat{\beta}_3 = t = \frac{\hat{\beta}_3 - \beta_3}{Se (\hat{\beta}_3)}
\]
\[
\hat{\beta}_4 = t = \frac{\hat{\beta}_4 - \beta_4}{Se (\hat{\beta}_4)}
\]

The decision rule used for this is as follows, reject the null hypothesis (H0) if the calculated t is greater than the tabulated t or t-critical. In order to calculate this the t-critical value at 0.01 is 2.33. The test statistics used for this calculation is t-distribution test.

**For \( \hat{\beta}_1 \):**

\[
\hat{\beta}_1 = t = \frac{\hat{\beta}_1 - \beta_1}{Se (\hat{\beta}_1)}
\]

\[
t = \frac{0.423 - 0}{0.220}
\]

\[
t = 1.92
\]
For $\hat{\beta}_2$:

$$\hat{\beta}_2 = t = \frac{\beta_2 - \beta_2}{Se (\hat{\beta}_2)}$$

$$t = \frac{0.398}{0.061}$$

$$t = 6.52$$

For $\hat{\beta}_3$:

$$\hat{\beta}_3 = t = \frac{\beta_3 - \beta_3}{Se (\hat{\beta}_3)}$$

$$t = \frac{7.375 \times 10^{-4}}{2.807 \times 10^{-4}}$$

$$t = 2.63$$

For $\hat{\beta}_4$:

$$\hat{\beta}_4 = t = \frac{\beta_4 - \beta_4}{Se (\hat{\beta}_4)}$$

$$t = \frac{0.001015}{2.936 \times 10^{-4}}$$

$$t = 3.46$$

Completed Table:

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard error</th>
<th>Test Statistic</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.423</td>
<td>0.220</td>
<td>t-test/distribution</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.398</td>
<td>0.061</td>
<td>t-test/distribution</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>7.375e-04</td>
<td>2.807e-04</td>
<td>t-test/distribution</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>0.001015</td>
<td>2.936e-04</td>
<td>t-test/distribution</td>
</tr>
</tbody>
</table>

Based on the results as presented in the table above, values of $\beta_2$, $\beta_3$, $\beta_4$ are greater than 2.33 (critical value). This is not true for the constant $\beta_1$ which is 1.92, below the critical value. It is concluded that the coefficients of $\beta_2$, $\beta_3$, $\beta_4$ are all significant at the at the one percent level ($\alpha=0.01$), except for the constant of $\beta_1$. 
c. Based on your results, write down the names of variables that are candidates for omission from the model.

d. Suppose a student took a special course to improve her SAT scores and increased the verbal and math scores by 100 points each. On average, how much of an increase in college GPA could she expect?

**ANSWER:**

\[ \Delta \text{COLGPA} = 0.0007375 \Delta \text{VSAT} + 0.001015 \Delta \text{MSAT} \]
\[ = 0.07375 + 0.1015 \]
\[ = 0.17525. \]

Thus, the expected average increase in COLGPA is 0.175.

e. Suppose you want to test the hypothesis that the regression coefficients for VSAT and MSAT are equal. How will you test this hypothesis? [Suppose that \( \text{cov} (\beta_3, \beta_4) = -0.25 \).]

**ANSWER:**

The general unrestricted model is

\[ \text{COLGPA} = \beta_1 + \beta_2 \text{HSGPA} + \beta_3 \text{VSAT} + \beta_4 \text{MSAT} + u \]

The marginal effect of VSAT is \( \beta_3 \) and the marginal effect of MSAT is \( \beta_4 \). The test is therefore \( \beta_3 = \beta_4 \). The alternative is that these two coefficients are unequal. Hence our hypotheses for the test are as follows,

\[ H_0: \beta_3 = \beta_4 \text{ or } \beta_3 - \beta_4 = 0 \]
\[ H_1: \beta_3 \neq \beta_4 \text{ or } \beta_3 - \beta_4 \neq 0 \]

The variance of the estimated differences of \( \hat{\beta}_3 - \beta_4 \) is given by the following:

\[ \text{Var} (\hat{\beta}_3) + \text{Var} (\hat{\beta}_4) - 2 \text{Cov} (\hat{\beta}_3, \hat{\beta}_4) \]

Using the direct two tailed t-test method the calculated t-statistic is therefore as follows,

\[ t_c = \frac{\hat{\beta}_3 - \hat{\beta}_4}{\sqrt{\text{Var} (\hat{\beta}_3) + \text{Var} (\hat{\beta}_4) - 2 \text{Cov} (\hat{\beta}_3, \hat{\beta}_4)}} \]

For this test, if the numerical value of \( t_c \) is greater than \( t^*_{n-k} \) (significance level of 2), then our \( H_0 \) hypothesis is rejected.

The following are the calculations;
\[
t_c = \frac{(\hat{\beta}_3 - \hat{\beta}_4) - (\beta_3 - \beta_4)}{Se(\hat{\beta}_3 - \hat{\beta}_4)}
\]

**Equation 0-1**

\[
t_c = \frac{\beta_3 - \beta_4}{\sqrt{Var(\hat{\beta}_3) + Var(\hat{\beta}_4) - 2Cov(\hat{\beta}_3, \hat{\beta}_4)}}
\]

**Equation 0-2**

\[
Se(\hat{\beta}_3 - \hat{\beta}_4) = \sqrt{Var(\hat{\beta}_3) + Var(\hat{\beta}_4) - 2Cov(\hat{\beta}_3, \hat{\beta}_4)}
\]

\[
Se(\beta_3 - \beta_4) = \sqrt{7.879 \times 10^{-8} + 8.62 \times 10^{-5} - 2(0.25)}
\]

\[
= \sqrt{.50}
\]

\[
Se(\hat{\beta}_3 - \hat{\beta}_4) = 0.707
\]

Now we place the value obtained from equation 2 into equation 1.

\[
t_c = \frac{(\hat{\beta}_3 - \hat{\beta}_4) - (\beta_3 - \beta_4)}{Se(\hat{\beta}_3 - \hat{\beta}_4)}
\]

\[
= \frac{(0.0007375 - 0.001015) - (0)}{0.707}
\]

\[
t_c = -0.0003925
\]

Based on our results, the t-tabulated value is greater than the t-calculated value. Therefore, we must accept the H\textsubscript{0} hypothesis at the level of one percent significance concluding that H\textsubscript{0}: \( \beta_3 = \beta_4 \) or \( \beta_3 - \beta_4 = 0 \).
Question 2
The following is known as the **transcendental production function** (TPF), a generalization of the well-known Cobb–Douglas production function:

\[ Y_i = \beta_1 L_i^{\beta_2} K_i^{\beta_3} e^{\beta_4 L_i + \beta_5 K_i} \]

where \( Y = \) output, \( L = \) labor input, and \( K = \) capital input.

After taking logarithms and adding the stochastic disturbance term, we obtain the stochastic TPF as

\[ \ln Y_i = \beta_0 + \beta_2 \ln L_i + \beta_3 \ln K_i + \beta_4 L_i + \beta_5 K_i + u_i \]

where \( \beta_0 = \ln \beta_1 \).

a. For the TPF to reduce to the Cobb–Douglas production function, what must be the values of \( \beta_4 \) and \( \beta_5 \)?

**ANSWER:**

The function allows the marginal products of labour and capital to increase before they experience a downfall that is bound to happen. In case of a standard Cobb-Douglas production function, the marginal products decrease from the beginning. Hence, this function allows for the experience of variable elasticity of substitution, then compared to the commonly used Cobb-Douglas model. Hence, the values of \( \beta_4 \) and \( \beta_5 \) must be zero for the TPF to reduce the Cobb-Douglass production function, in this case the value of \( e^{\beta_0} \) is equal to 1 making it the standard model.

b. If you had the data, how would you go about finding out whether the TPF reduces to the Cobb–Douglas production function? What testing procedure would you use?

**ANSWER:**

If data were available, the testing procedure that will be used to see if TPF reduces to the Cobb-Douglas production function is through the use of the F test for restricted least squares. The following is a general outline of the procedure that would be used with the availability of data.

When conducting the multiple regression F test the following hypotheses will be used in which.

\[ H_0 = \beta_4 = 0, \beta_5 = 0 \]
\[ H_1 = \text{rejection of } H_0 \text{ as it is untrue} \]

After establishing this an un-restricted regression is run on

\[ \ln Y_i = \beta_0 + \beta_2 \ln L_i + \beta_3 \ln K_i + \beta_4 L_i + \beta_5 K_i \]

Afterwards, RSS\text{UR} (unrestricted) is calculated before proceeding to the calculations of restricted regression by including \( \beta_4 \) and \( \beta_5 \) into the equation as follows,
\[ \ln Y_i = \beta_0 + \beta_2 \ln L_i + \beta_3 \ln K_i \]

Then its RSS_{R} (unrestricted) is calculated using the formula:

\[ F = \frac{(RSS_{R} - RSS_{UR})}{RSS_{UR}} \frac{1}{J(n-k)} \]

For this equation J is equals the number of restrictions, n is the total of number of observations, and k is the number of parameters.

**Question 3**

Marc Nerlove has estimated the following cost function for electricity generation:

\[ Y = AX^{\beta} P^{\alpha_1} P^{\alpha_2} P^{\alpha_3} u \] (1)

where \( Y \) = total cost of production  
\( X \) = output in kilowatt hours  
\( P_1 \) = price of labor input  
\( P_2 \) = price of capital input  
\( P_3 \) = price of fuel  
\( u \) = disturbance term

Theoretically, the sum of the price elasticities is expected to be unity, i.e., \( \alpha_1 + \alpha_2 + \alpha_3 = 1 \). By imposing this restriction, the preceding cost function can be written as

\[ \left( \frac{Y}{P_3} \right) = AX^{\beta} (P_1/P_3)^{\alpha_1} (P_2/P_3)^{\alpha_2} u \] (2)

In other words, (1) is an unrestricted and (2) is the restricted cost function. On the basis of a sample of 29 medium-sized firms, and after logarithmic transformation, Nerlove obtained the following regression results

\[ \ln \hat{Y}_i = -4.93 + 0.94 \ln X_i + 0.31 \ln P_1 \]

\[ \text{se} = (1.96) \quad (0.11) \quad (0.23) \]

\[ -0.26 \ln P_2 + 0.44 \ln P_3 \]

\[ (0.29) \quad (0.07) \]

\[ \text{RSS} = 0.336 \]

\[ \ln \left( \frac{Y}{P_3} \right) = -6.55 + 0.91 \ln X + 0.51 \ln \left( \frac{P_1}{P_3} \right) + 0.09 \ln \left( \frac{P_2}{P_3} \right) \]

\[ \text{se} = (0.16) \quad (0.11) \quad (0.19) \quad (0.16) \quad \text{RSS} = 0.364 \]

a. Show the transition from equation (1) to equation (2).
ANSWER:

We begin by taking the log of both sides in equation 1 to derive equation 2.

\[ Y = AX^\beta P^{\alpha_1} P^{\alpha_2} P^{\alpha_3} u \]  \hspace{1cm} \text{(equation 1)}

\[ \ln Y = AX^\beta (\ln P_1^{\alpha_1})(\ln P_2^{\alpha_2})(\ln P_3^{\alpha_3}) u \]

\[ \ln Y = AX^\beta (\alpha_1 \ln P_1)(\alpha_2 \ln P_2)(\alpha_3 \ln P_3) u \]  \hspace{1cm} \text{(equation 2)}

Next a restriction is imposed on \( \alpha_3 \) resulting in,

\[ \alpha_1 + \alpha_2 + \alpha_3 = 1 \]
\[ \alpha_3 = 1 - \alpha_1 - \alpha_2 \]

Now we substitute the value obtained for \( \alpha_3 \) into equation 2.

\[ \ln Y = AX^\beta (\alpha_1 \ln P_1)(\alpha_2 \ln P_2)(\ln P_3 - \alpha_1 \ln P_3 - \alpha_2 \ln P_3) u \]

We then take the common of \( \alpha_1 \) & \( \alpha_2 \) in order to obtain the following result.

\[ \ln Y = AX^\beta \alpha_1 (\ln P_1 - \ln P_3) \alpha_2 (\ln P_2 - \ln P_3) u \]

\[ \frac{\ln Y}{\ln P_3} = AX^\beta (\frac{P_1}{P_3})^{\alpha_1} (\frac{P_1}{P_3})^{\alpha_2} u \]

b. Interpret Eqs. (3) and (4).

ANSWER:

Both the models discussed above are log-linear, that is taking into consideration the estimated slope coefficients that represent the partial or elasticity of the dependent variable with respect to the regressor that is being taken into consideration. For example, the coefficient of 0.94 in equation 3 basically tells us that if the output in kwh were to increase by 1%, then on average the total cost of production will also increase by 0.94%. When looking at equation 4, the trend described is about the same. If the price of labour comparative to the price of fuel goes up by about 1% then once again on average the relative cost of production also increases by 0.51%.

c. How would you find out if the restriction \( (\alpha_1 + \alpha_2 + \alpha_3) = 1 \) is valid? Show your calculations.

ANSWER:

In order to find out if the restrictions \( (\alpha_1 + \alpha_2 + \alpha_3) = 1 \) is valid, the F statistics will be used under the formula, this formula is because the dependent variables in equations 3 and 4 are not the same;

\[ F = \frac{(RSS_R - RSS_{UR})}{\frac{N}{n-k}} \]
Inputs of the formula includes $n = 29$, $k = 5$, $\alpha = 0.05$ with the $df$ for numerator being 1, and $df$ for denominator being 24. Our decision rule for this equation is that if calculated $F$ is greater than the critical $F$ value then our $H_0$ is to be rejected.

$$F = \frac{(RSS_R - RSS_{UR})}{\frac{NR}{n - k}}$$

$$F = \frac{(0.364 - 0.336)}{\frac{1}{(1 - 0.336)}}$$

$$F = 1.012$$

The tabulated $F$ value is 4.26. Based on the results the $F$ obtained is concluded to be insignificant because the 5% critical value of $F$ for 1, and 24 numerator and denominator $df$ is 4.26. Therefore, we cannot reject the null hypothesis that the sum of the price of elasticities is 1.

**Question 4**

Suppose theory suggested that annual income ($Y$) depended on gender ($G$), degree received ($D$), and years of experience ($X$): $Y = f(G, D, X)$. We define four dummy variables:

- $D_1 = 1$ if non-graduate; 0 otherwise
- $D_2 = 1$ if graduate; 0 otherwise
- $G_1 = 1$ if male; 0 otherwise
- $G_2 = 1$ if female; 0 otherwise

First consider the following results obtained from regressing gender dummy with $Y$:

$$\hat{Y} = $ 8000 - 1500 G_2$$

The sample size is 10.

a. Why is $G_1$ not included in the equation?

**ANSWER:**

$G_1$ is not included in the above equation because it allows us to avoid the trap of a dummy variable or the perfect multi-collinearity. It is possible to include $G_1$ into the equation but only if we are to drop the intercept or constant.

b. What is the mean income for male and female in the sample?
ANSWER:

Using the equation the mean income for males and females can be found.

Mean Income for males:

\[
\hat{Y} = $8000 - 1500 G_2 \\
\hat{Y} = $8000 - 1500 (0) \\
\hat{Y} = $8000
\]

Mean Income for females

\[
\hat{Y} = $8000 - 1500 G_2 \\
\hat{Y} = $8000 - 1500 (1) \\
\hat{Y} = $8000 - 1500 G_2 \\
\hat{Y} = $6500
\]

c. Do females earn significantly less than males? (α = 0.05)

ANSWER:

To being we formulate hypotheses for testing;

\[H_0: \text{Females earn significantly less than males} \]
\[H_1: \text{Females earn significantly more than males} \]

Our level of significance is at α=0.05, it is a one tail t-test so \(df = n - 1\); therefore, 10-1=9, \(n = 9\).

The following formula is used;

\[
t = \frac{\hat{\beta}_2 - \beta_2}{Se(\beta_2)}
\]

The decision rule for this equation will be that if \(t_{\text{calculated}} > t_{\text{critical}}\) then our null hypothesis is rejected.

\[
t = \frac{\hat{\beta}_2 - \beta_2}{Se(\beta_2)}
\]

\[
t = \frac{-1500 - (0)}{600}
\]
\[ t = -2.5 \]

\[ t_{\text{critical}}(0.05, 9) = 1.833. \]

Based on the results we accept the null hypothesis as \( t_{\text{calculated}} < t_{\text{critical}} \). It is therefore concluded that females earn less than males.

d. What would \( b_1 \) and \( b_2 \) be if the same sample was used to estimate

\[ Y = b_1 + b_2 G_1 + \hat{u} \]

To control for educational attainment and years of experience as well as gender, the following estimates were obtained from the same sample used to estimate equation (1):

\[ Y_i = 7000 - 2000 G_{2i} - 6000 D_{2i} + 500 X_i + 1600 G_{2i} D_{2i} + \hat{u}_i \]

**ANSWER:**

If we substitute \( G_2 \) for \( G_1 \) in the equation using the same sample the equation will be as follows

\[ \hat{Y} = 6500 + 1500 G_1 \]

Where \( b_1 = 6500 \) and \( b_2 = 1500 \)

e. What is the mean income for a male who is a non-graduate and has no work experience?

**ANSWER:**

Using the equation

\[ Y_i = 7000 - 2000 G_{2i} - 6000 D_{2i} + 500 X_i + 1600 G_{2i} D_{2i} + \hat{u}_i \]

The mean income for a male who is a non-graduate and has no work experience is as follows with variables defined.

For male \( G_2 = 0 \)
For non-graduate \( D_2 = 0 \)
For no work experience \( X = 0 \)

Plugging in the values into the equation we get.

\[ Y_i = 7000 - 0 - 0 + 0 + 0 \]

\[ Y_i = 7000 \]

Hence, mean income for males of this category is $7000.

f. What is the mean income for a female graduate with 10 years of experience?
ANSWER:

Using the equation,

\[ Y_i = $7000 - 2000 G_{2i} - 6000 D_{2i} + 500 X_i + 1600 G_{2i} D_{2i} + u_i \]

And the following variables,

For female G2 = 1
For graduate D2 = 1
For work experience X = 10.

Plugging in,

\[ Y_i = $7000 - 2000 (1) - 6000 (1) + 500 (10) + 1600 (1)(1) \]
\[ Y_i = $7000 - 2000 - 6000 + 5000 + 1600 \]
\[ Y_i = $5600 \]

Mean income for female graduates who have ten years of experience is $5600.
Question 5

Table 1.1 gives data on the GDP (gross domestic product) deflator for domestic goods and the GDP deflator for imports for country XYZ for the period 1968-1982. The GDP deflator is often used as an indicator of inflation in place of CPI.

a. If country XYZ is a closed economy, its inflation level can be gauged by GDP deflator for domestic good (Y). How can we estimate the growth rate of GDP deflator (Y) from 1968-82? Estimate the instantaneous growth rate and the growth rate over time [i.e. \( r \) in \( Y_t = Y_0 (1+r)^t \)].

**ANSWER:**

We begin by taking ln of the equation: \( Y_t = Y_0 (1+r)^t \) to estimate the growth rate of GDP deflator Y.

\[
\begin{align*}
Y_t &= Y_0 (1+r)^t \\
\ln Y_t &= \ln Y_0 + t \ln(1+r) \\
\ln Y_t &= \ln Y_0 + \beta_2 t \\
\ln Y_t &= \beta_1 + \beta_2 t \\
Y &= $16.0898 + 0.5340 \\
Se &= (40.5031) (0.0217) \\
r^2 &= 0.97889672
\end{align*}
\]

To find instantaneous growth rate;
\[
\beta_2 = \ln(1 + r) \\
r = \ln \beta_2 - 1 \\
r = 1.706 - 1 \\
r = 0.706 \\
r = 70.6\
\]

To find growth rate over time;
\[
r = \beta_2 \times 100 \\
= 0.534 \times 100 \\
= 53.4\% \\
\]

The rate of growth for durable goods expenditure was about 53.4% but the estimated coefficients are concluded to be indirectly statistically significant. The P values of the set are extremely low meaning it would be inadequate to run a double log model for the circumstance.

\[
\begin{align*}
\ln \text{Exp} &= \beta_1 + \beta_2 \ln \text{time} + u_2 \\
\ln Y_t &= \ln Y_0 + t \ln(1+r) \\
&= \ln $16.0898 + 1$ (1 + r)
\end{align*}
\]
b. If country XYZ is a small open economy heavily dependent on foreign trade for its survival. To study the relationship between domestic and world prices, you are given the following linear model:

\[ Y_t = \beta_1 + \beta_2 X_t + u_t \]

i. What are your prior expectations about the relationship between Y and X?

**ANSWER:**

**SUMMARY OUTPUT**

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
<th>Lower 95.0%</th>
<th>Upper 95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>40.5631118</td>
<td>12.72313</td>
<td>1.03265E-08</td>
<td>428.4585552</td>
<td>603.7211059</td>
<td>603.7211059</td>
</tr>
<tr>
<td>X Variable 1</td>
<td>0.021744585</td>
<td>24.55642</td>
<td>2.81131E-12</td>
<td>0.486992938</td>
<td>0.580945579</td>
<td>0.580945579</td>
</tr>
</tbody>
</table>
The following table represents the calculations conducted to establish the values of $\beta_1$ and $\beta_2$.

Do your results comply with your prior expectations? What is the economic meaning of the parameter estimates?

<table>
<thead>
<tr>
<th>Year</th>
<th>Y</th>
<th>X</th>
<th>Yhat = $\beta_0 + \beta_1 X + \beta_2 X^2$</th>
<th>$Yhat - Y$</th>
<th>$Yhat - xbar$</th>
<th>$Yhat - ybar$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1969</td>
<td>1023</td>
<td>422.3</td>
<td>-19.38</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1968</td>
<td>1053</td>
<td>455.7</td>
<td>-24.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>1079</td>
<td>498.7</td>
<td>-29.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>1105</td>
<td>541.7</td>
<td>-34.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>1131</td>
<td>585.7</td>
<td>-39.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1964</td>
<td>1157</td>
<td>628.7</td>
<td>-44.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1963</td>
<td>1183</td>
<td>672.7</td>
<td>-49.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1962</td>
<td>1209</td>
<td>716.7</td>
<td>-54.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td>1961</td>
<td>1235</td>
<td>761.7</td>
<td>-59.51</td>
<td>20.01</td>
<td>8.82</td>
<td></td>
</tr>
</tbody>
</table>

**ii.** Estimate $\beta_1$ and $\beta_2$. Do your results comply with your prior expectations? What is the economic meaning of the parameter estimates?
<table>
<thead>
<tr>
<th>Year</th>
<th>1980</th>
<th>1981</th>
<th>1982</th>
<th>SUMS</th>
<th>MEANS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1841</td>
<td>1959</td>
<td>2033</td>
<td>21836</td>
<td>1455.7</td>
</tr>
<tr>
<td></td>
<td>2621</td>
<td>2777</td>
<td>2735</td>
<td>26396</td>
<td>1759.7</td>
</tr>
<tr>
<td></td>
<td>385.26</td>
<td>503.26</td>
<td>577.26</td>
<td>57477.9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>861.26</td>
<td>1017.2</td>
<td>975.26</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>Year</td>
<td>2033</td>
<td>1959</td>
<td>1982</td>
<td>SUMS</td>
<td>MEANS</td>
</tr>
<tr>
<td>1980</td>
<td>74178</td>
<td>10348</td>
<td>95114</td>
<td>574779</td>
<td>33</td>
</tr>
<tr>
<td>1981</td>
<td>331817.3</td>
<td>511956.4</td>
<td>562988.9</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td>1982</td>
<td>378</td>
<td>0.3</td>
<td>378</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>74178</td>
<td>10348</td>
<td>95114</td>
<td>574779</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>331817.3</td>
<td>511956.4</td>
<td>562988.9</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>378</td>
<td>0.3</td>
<td>378</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>74178</td>
<td>10348</td>
<td>95114</td>
<td>574779</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>331817.3</td>
<td>511956.4</td>
<td>562988.9</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>378</td>
<td>0.3</td>
<td>378</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>74178</td>
<td>10348</td>
<td>95114</td>
<td>574779</td>
<td>33</td>
</tr>
<tr>
<td>1980</td>
<td>1915.6</td>
<td>1998.9</td>
<td>1976.4</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td>1981</td>
<td>233</td>
<td>225</td>
<td>958</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td>1982</td>
<td>-</td>
<td>39.922</td>
<td>25</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>74.623</td>
<td>-</td>
<td>0.4</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>14295</td>
<td>-</td>
<td>-</td>
<td>3069147.933</td>
<td>33</td>
</tr>
<tr>
<td>1980</td>
<td>14843</td>
<td>19920</td>
<td>19859</td>
<td>21836</td>
<td>33</td>
</tr>
<tr>
<td>1981</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>21836</td>
<td>33</td>
</tr>
<tr>
<td>1982</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>21836</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>5568.6</td>
<td>1593.8</td>
<td>3192.7</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>1841</td>
<td>1959</td>
<td>2033</td>
<td>21836</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>5568.6</td>
<td>1593.8</td>
<td>3192.7</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>1841</td>
<td>1959</td>
<td>2033</td>
<td>21836</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>5568.6</td>
<td>1593.8</td>
<td>3192.7</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>1841</td>
<td>1959</td>
<td>2033</td>
<td>21836</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>5568.6</td>
<td>1593.8</td>
<td>3192.7</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>1841</td>
<td>1959</td>
<td>2033</td>
<td>21836</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>5568.6</td>
<td>1593.8</td>
<td>3192.7</td>
<td>35330.29</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>35330.29</td>
<td>33</td>
</tr>
</tbody>
</table>

Finding $\beta_2$:

$$\beta_2 = \frac{x_iy_i}{x_i^2}$$

$$\beta_2 = \frac{3069147.933}{5747799}$$

$$\beta_2 = 0.533969$$

Finding $\beta_1$:

$$\beta_1 = \text{Mean of } Y - \beta_2 \times \text{Mean of } X$$

$$\beta_1 = 1455.733 - 0.533969 \times 1759.733$$

$$\beta_1 = 516.0898$$
iii. Also estimate the model using the double log functional form. Provide the economic meaning of the parameter estimates?

**ANSWER:**

<table>
<thead>
<tr>
<th>$\ln Y$</th>
<th>$\ln X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3.009875634</td>
<td>3.01786772</td>
</tr>
<tr>
<td>3.017033339</td>
<td>3.03822264</td>
</tr>
<tr>
<td>3.036229544</td>
<td>3.04336228</td>
</tr>
<tr>
<td>3.0539184618</td>
<td>3.04532298</td>
</tr>
<tr>
<td>3.108903128</td>
<td>3.0993528</td>
</tr>
<tr>
<td>3.1717264544</td>
<td>3.24278981</td>
</tr>
<tr>
<td>3.182129214</td>
<td>3.24797327</td>
</tr>
<tr>
<td>3.188365926</td>
<td>3.27623196</td>
</tr>
<tr>
<td>3.195068996</td>
<td>3.29534715</td>
</tr>
<tr>
<td>3.201943063</td>
<td>3.30427505</td>
</tr>
<tr>
<td>3.2340108184</td>
<td>3.35410844</td>
</tr>
<tr>
<td>3.265053789</td>
<td>3.41846702</td>
</tr>
<tr>
<td>3.292044436</td>
<td>3.44357588</td>
</tr>
<tr>
<td>3.308137379</td>
<td>3.43695733</td>
</tr>
</tbody>
</table>

**SUMMARY OUTPUT**

```
Regression Statistics
Multiple R  0.990685216
R Square    0.981457197
Adjusted R Square    0.980030828
Standard Error  0.014909774
Observations    15
```
## ANOVA

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>0.152961262</td>
<td>0.152961</td>
<td>688.0806</td>
<td>1.21139E-12</td>
</tr>
<tr>
<td>Residual</td>
<td>13</td>
<td>0.002889918</td>
<td>0.000222</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>0.155851118</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Coefficients and Standard Errors

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
<th>Lower 95.0%</th>
<th>Upper 95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.08295234</td>
<td>0.078944813</td>
<td>13.71784</td>
<td>4.14E-09</td>
<td>0.912402441</td>
<td>1.25350224</td>
<td>0.912402441</td>
<td>1.25350224</td>
</tr>
<tr>
<td>X Variable 1</td>
<td>0.642829358</td>
<td>0.024506203</td>
<td>26.23129</td>
<td>1.21E-12</td>
<td>0.589886924</td>
<td>0.695771792</td>
<td>0.589886924</td>
<td>0.695771792</td>
</tr>
</tbody>
</table>

### SUMMARY OUTPUT

Regression Statistics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.977612883</td>
</tr>
</tbody>
</table>
R Square 0.955726949
Adjusted R Square 0.95232133
Standard Error 0.023038441
Observations 15

ANOVA

<table>
<thead>
<tr>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>0.148951173</td>
<td>0.148951</td>
<td>280.6324</td>
</tr>
<tr>
<td>Residual</td>
<td>13</td>
<td>0.006900007</td>
<td>0.000531</td>
<td>0.000531</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>0.15585118</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Coefficients

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
<th>Lower 95.0%</th>
<th>Upper 95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.868031702</td>
<td>0.017925961</td>
<td>159.9932</td>
<td>8.37E-23</td>
<td>2.829305017</td>
<td>2.906758386</td>
<td>2.829305017</td>
</tr>
<tr>
<td>X Variable 1</td>
<td>0.00016098</td>
<td>9.60953E-06</td>
<td>16.75208</td>
<td>3.51E-10</td>
<td>0.00014022</td>
<td>0.00018174</td>
<td>0.00014022</td>
</tr>
</tbody>
</table>