Unit 2 Assignment

1. The professor of a mechanical engineering course would like to determine whether a student's midterm score can be used to predict his or her final exam score. The midterm score (out of 50) and the final exam score (out of 100) for a sample of eight students:

Student	1	2	3	4	5	6	7	8
Midterm Score (/50)	43	28	36	30	24	48	29	40
Final Exam Score (/100)	92	47	80	65	40	94	55	71

For these data, it can be calculated that $\overline{x} = 34.75$, $s_x = 8.36$, $\overline{y} = 68.00$ and $s_y = 20.03$.

(a) Create a scatterplot of the data and comment on the apparent relationship between Midterm Score and Final Exam Score.
(b) Calculate the value of the correlation between Midterm Score and Final Exam Score.
(c) Find the equation of the least squares regression line and draw it on the scatterplot.
(d) Interpret the meaning of the least squares regression slope.
(e) Interpret the value of r^2.
(f) Find the predicted final exam xcore for a students with midterm scores of 15 and 32. Is either of these predictions more reliable than the other? Explain.
(g) Find the residual value for the Student # 2. What does the sign of the residual tell you?
(h) Suppose that Midterm Score had instead been calculated as a percentage. What would be the value of the correlation between Midterm Score and Final Exam Score? What would be the equation of the least squares regression line? (Do not convert the data into percentages. Use the definitions of the correlation, intercept and slope to answer the question).

2. We would like to study the impact of Temperature X on the Viscosity Y of toluene-tetralin blends. The following table gives the data from an experiment, which are for blends with a 0.4 molar fraction of toluene:

Observation	1	2	3	4	5	6	7	8
Temperature (°C)	24.9	35.0	44.9	55.1	65.2	75.2	85.2	95.3
Viscosity (mPa·s)	1.12	0.99	0.85	0.82	0.68	0.64	0.53	0.45

For these data, it can be calculated that $\bar{x} = 60.10$, $s_x = 24.635$, $\bar{y} = 0.76$ and $s_y = 0.228$.

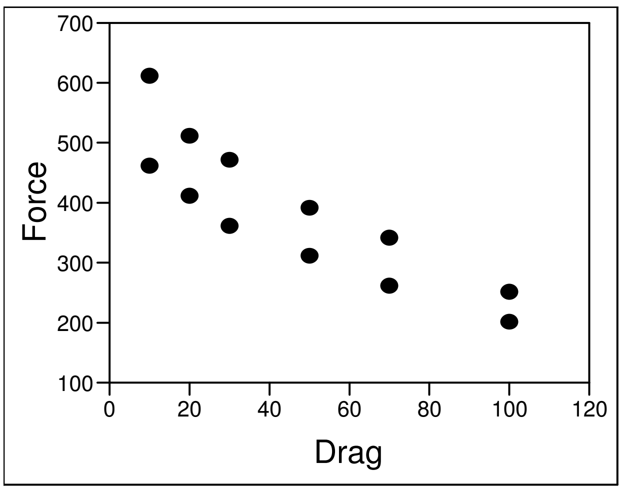
(a) Create a scatterplot of the data and comment on the apparent relationship between Temperature and Viscosity.
(b) Calculate the value of the correlation between Temperature and Viscosity.
(c) Find the equation of the least squares regression line of Viscosity on Temperature.
(d) Interpret the meaning of the least squares regression slope.
(e) Interpret the value of r^2.
(f) Find the predicted Viscosity for temperatures of 15°C and 60°C. Is either of these predictions more reliable than the other? Explain.
(g) Find the residual value for the 3rd observation. What does the sign of the residual tell you?
(h) Suppose that Temperature had instead been measured in degrees Fahrenheit. What would be the value of the correlation between Temperature and Viscosity? What would be the equation of the least squares regression line? (Do not convert the data into °F. Use the definitions of the correlation, intercept and slope to answer the question).

3. (a) Show that the least squares regression line must always pass through the point $(\bar{x}, \bar{y})$.
(b) Show that the sum of residuals in least squares regression must always be zero.

4. The least squares regression methods we have used are appropriate only in the case of a linear relationship. However, for other types of relationships, it may still be possible to use the least squares methods by transforming one or both of the variables in such a way that the transformed data are linearly related. Consider the following data set from the testing of automobile tires. A tire under study is mounted on a test trailer and pulled at a standard velocity. Using a braking mechanism, a standard amount of drag (measured in %) is applied to the tire and the force (in pounds) with which it grips the road is measured. The following data are from tests on 12 different tires of the same design made under the same set of road conditions. Note that x = 0% indicates no braking and x = 100% indicates the brake is locked.

Drag X	10	10	20	20	30	30	50	50	70	70	100	100
Force Y	460	610	510	410	470	360	390	310	260	340	250	200

From the scatterplot of these data shown below, we see that the relationship between X and Y appears to be nonlinear:



In fact, physical theory can be called upon to predict that instead of being linear, the relationship between X and Y is of the form $y = \beta_0 e^{\beta_1 x}$. Note that if natural logarithms are taken on both sides of this expression, we get $\ln(y) = \ln(\beta_0) + \beta_1 x$. We can rewrite this as $y^* = \beta_0^* + \beta_1 x$, where $y^* = \ln(y)$ and $\beta_0^* = \ln(\beta_0)$. In other words, the relationship between $y^* = \ln(y)$ and x is linear.

(a) Make a scatterplot of $y^* = \ln(y)$ vs. x. Does this relationship appear more linear than the one shown in the scatterplot above?

(b) Find the equation of the least squares regression line of y* on x. It can be calculated that $\bar{x} = 46.667$, $s_x = 32.287$, $\bar{y}^* = 5.896$, $s_{y^*} = 0.325$ and $r^* = -0.913$, where r* is the correlation between x and y*. Draw the regression line on the scatterplot.

(c) What is the predicted grip force for a tire on which 60% drag is applied?