Problem Set 06: User-Defined Functions

New Learning Objectives under Evaluation

12.00 Perform linear regression

Learning Objective	Evidence
12.01 State the purpose of regression	Being able to take given input argument values to one UDF and manually track the value of all computed and passed output arguments and input arguments through a series of UDFs linked through function calls
12.02 Compute and present in equation form the coefficients of a best-fit linear model using visual approximation and the two-point method	Given a plot of raw x-y data, draw a best-fit linear model Use two points on the drawn line to compute the linear model coefficients: slope (a) and intercept (b) Presentation of the slope (a) and intercept (b) in the linear model equation
12.03 Manually compute the SSE	Use the dependent variable (y) data and the predicted values of the dependent variable to compute the sum or squares of error
12.04 Manually compute the SST	Use the dependent variable (y) data and the mean of the dependent variable data to compute the sum or squares of deviation
12.05 Manually compute the r-squared value from SSE and SST	$r^2 = 1 - SSE/SST$ Recognize that the result falls between 0 and 1 inclusive
12.06 Add a trendline to a scatter plot of raw x-y data (Excel)	Create a scatter plot of raw bivariate data in Excel Use Excel chart tools to add a linear trendline to the plot
12.07 Display the equation and r-squared value of a trendline added to a scatter plot (Excel)	Create a scatter plot of raw bivariate data in Excel Use Excel chart tools to add a linear trendline to the plot Use Excel chart tools to display the equation on the plot Revise the equation so that the x and y variables are descriptive Use Excel chart tools to display the r^2 value on the plot
12.08 Manually compute and present in equation form the coefficients of a best-fit linear model using least-squares method	Correct calculation of the summations needed for the two least squares equations Correct placement of the summations in the two least squares equations Simultaneous solution of the two least squares equations for the correct linear model coefficients: slope (a) and intercept (b) Presentation of the slope (a) and intercept (b) in the linear model equation

Problem Set 06: User-Defined Functions

New Learning Objectives under Evaluation

<table>
<tr><td>12.09 Compute the coefficients of a best-fit linear model using least-squares method (MATLAB)</td><td>Correct syntax for <code>polyfit</code>
<code>output1 = polyfit(independent_ vector,dependent_vector,order_of_polynomial)</code>
<code>polyfit</code> is the MATLAB built-in function
Three input arguments separated by commas: vector1, vector2, and a scalar
One output variable for the coefficients
For linear models, the order of the polynomial is 1
Correct identification of the independent variable (x) and dependent variable (y) used in <code>polyfit</code></td></tr>
<tr><td>12.10 Compute predicted values using the best-fit linear model (MATLAB)</td><td>Correct syntax for <code>polyval</code>
<code>dependent_predicted_array = polyval(model_coeff_vector,independent_array)</code>
<code>polyval</code> is the MATLAB built-in function
Two input arguments separated by commas
One output argument variable assignment
Correct order of input arguments to the <code>polyval</code>
Use appropriate linear model coefficients vector that contains slope and intercept (in that order) in call to <code>polyval</code>
Correctly identified independent variable for the input argument
Appropriate scalar value (for single value predictions) or range of values (e.g., for subsequent plotting) for the independent variable input argument
Correctly identified dependent variable for the output argument</td></tr>
<tr><td>12.11 Plot the best-fit linear regression line on a plot of raw x-y data (MATLAB)</td><td>Use the <code>plot</code> command to plot the raw data
Correctly identify the independent <u>and</u> dependent variables when plotting the raw data
Use the <code>plot</code> command to plot the linear model
Correctly identify the independent <u>and</u> dependent variables when plotting the linear model
Use an appropriate method (e.g., <code>hold</code>) to place both the raw data and linear model on the same plot</td></tr>
</table>

Problem Set 06: User-Defined Functions

New Learning Objectives under Evaluation

12.12 Display the results of linear regression (MATLAB)	Display the linear model equation using `fprintf` The linear model equation has variable names appropriate to the context of the problem (e.g., does not use variable names x, y) Correctly identify which `polyfit` output is the linear model slope Correctly identify which `polyfit` output is the linear model intercept Display the SSE value using `fprintf` Display the SST value using `fprintf` Display the r-squared value using `fprintf` Manage decimal places appropriately
12.13 Define, explain the use of, and relate SSE, SST, r-squared	SSE is a measure of the squared difference between the dependent raw data and predicted values (fit to the data) SST is a measure of the squared difference between the dependent raw data values and the mean of these values (variability in the data) r^2 is a measure of the extent to which a model explains the variation that exists in the data Explain how changes in the data set impact SSE, SST, and r^2
12.14 Interpret the slope and intercept of a best-fit linear model	The slope is the rate of change and has units (dependent variable units/independent variable units) The intercept is the value of the dependent variable when the independent variable equals zero and has units of the dependent variable Recognize that the intercept may have no meaning with regards to the context of the data
12.15 Interpret the r-squared value	r^2 is a measure of the extent to which a model explains the variation that exists in the data An r^2 closer to 1 means that the model does a good job of explaining the variation that exists in the data An r^2 closer to 0 means that the model does a poor job of explaining the variation that exists in the data Depending on the context of the data, a low r^2 might indicate a linear model does a good job of explaining the variation that exists in the data

Problem Set 06: User-Defined Functions

New Learning Objectives under Evaluation

12.16 Compare data sets based on their best fit linear models and r-squared values	Comparison based on slope Comparison based on intercept Comparison based on the extent to which a linear model explains the variation that exists in the data (r^2)
12.17 Use the best-fit linear model to make predictions only when appropriate	Independent variable values within the range of the original data set (domain of the model) can be used to make predictions Independent variable values outside the range of the original data set (domain of the model) must be acknowledged or justified when making predictions Predicted numerical values must be consistent with the linear model used to make the prediction Presentation of numerical predictions with appropriate units Management of the decimal places of numerical predictions