Code	**RIDGE_LASSO_REGRESSION.PY**
Author	Nathaniel Heatwole, PhD (heatwolen@gmail.com) (GitHub) (LinkedIn)
Summary	Uses ridge, lasso, and elastic net regression to predict body mass index (BMI)[1] using weight, height, and age.[2] Selects hyperparameters using built-in cross-validation.

Methods/ process	Ridge / lasso / elastic net regression:
	- Extends concepts of linear regression to the case where there are significant correlations among the predictor variables (multicollinearity).
	- Avoids tendency of simple linear regression to overfit when multicollinearity exists, improving predictions for out-of-sample test cases.
	- Adds an additional term (penalty) to the typical least-squares regression:
	    - *Ridge*: sum of *squares* of the regression coefficients.
	    - *Lasso*: sum of *absolute values* of the regression coefficients.
	    - *Hyperparameter* (α): multiplied onto the term in either case, controlling its influence in the optimization ($\alpha = 0$ is simple linear regression).
	- *Elastic net*: combines ridge and lasso, using a weighted average, introducing an additional hyperparameter, controlling the lasso weight (versus ridge).
	- Ridge shrinks all coefficients proportionally, whereas lasso does not. As such, lasso can perform variable selection (i.e., by zeroing out some coefficients).
	- Continuous features (both X and Y) must be transformed (e.g., z-scores) to prevent some variables from have undue influence in the regression.
	- Intercept term generally unneeded, because of what these regressions optimize.
	- Model fit typically assessed the same as in ordinary linear regression.[3]
	
	Steps:
	1. Import and clean training data (including z-scores).
	2. Apply all three regression types (ridge, lasso, elastic net).

Training data	BMI data – empirical health-related data for 741 persons (from Kaggle).
Output	Plot - Training data scatter matrix (univariate distributions, 2D relationships) Summary - Training data stats - Regression coefficients - Regression summary
Result	All regressions are highly accurate (adj. R^2 = 0.97). The coefficients signs for weight and height are as expected (i.e., BMI increasing in weight and decreasing in height). However, the hyperparameters for lasso and elastic net are minimal ($\alpha = 0.01$).

[1] BMI is a height-normalized measure of weight. It is defined as: weight / height^2.
[2] These predictors are expected to be correlated, and so ridge/lasso regression is appropriate.
[3] For example: R^2, adjusted R^2, AIC, BIC, log-likelihood, MSE.