Code	**LOGISTIC_REGRESSION.PY**
Author	Nathaniel Heatwole, PhD (heatwolen@gmail.com) (GitHub) (LinkedIn)
Summary	Uses logistic regression to predict probability of passing an exam using number of hours studied (using both sklearn and statsmodels)
Methods/ Process	Logistic regression - Extends concepts of linear regression (OLS) to model a binary (0/1) rather than continuous quantity - Basis is the logistic function, which takes the form of an s-curve, with dual asymptotes at zero and one - Therefore, functions as binary classifier, outputting the probability that Y = 1 - Argument (X) of the logistic function is a linear equation containing one or multiple variables (continuous or categorical) - Accordingly, the logistic function maps an unbounded quantity (X) onto the bounded probability space (0-1) - Coefficients for linear component (X) often estimated using maximum likelihood[1] - Model fit typically assessed by examining rates of true/false positives - Extensions of logistic regression exist for non-binary outcome variables (Y), or those that can take on more than two possible levels[2]
Training Data	Exam data – synthetic data on whether 20 students passed an exam and number of hours studied (from *Wikipedia*)
Results	Predictions from the two logistic regression modules (sklearn and statsmodels) are similar, although the statsmodels fit is slightly better (based on log-likelihood)

[1] Informally, this means choosing the model parameters so as to maximize the likelihood that the particular training data would be observed. More rigorously, it means selecting the s-curve parameters (location, steepness) so as to minimize the aggregate distance between the s-curve and the true outcomes (0 or 1).

[2] If there is a "natural" order (monotonicity) to the groups (e.g., small/medium/large), ordinal logistic regression is useful. Otherwise, if no such natural order exists (e.g., favorite color), multinomial logistic regression relaxes the monotonicity assumption, and allows the regression to select the order of the groups.