DIAML_HW4_Python Resources
Prepared by the TAs

Linear Regression

1. Option 1: SciPy:

https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.linregress.html

2. Option 2: scikit-learn:

Perform linear regression with one and multiple explanatory variables (fitting and prediction):
https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

3. Option 3: statsmodels:

https://www.statsmodels.org/dev/generated/statsmodels.regression.linear_model.OLS.html

Additional helpful resource for univariate and multivariate linear regression:
https://realpython.com/linear-regression-in-python/
https://scikit-learn.org/stable/auto_examples/linear_model/plot_ols.html
AIC/BIC in regression:
https://machinelearningmastery.com/probabilistic-model-selection-measures/

Correlation

This can be easily computed using either Numpy, Pandas or SciPy.

See more information here:
https://realpython.com/numpy-scipy-pandas-correlation-python/#:~:text=Correlation%20coefficients%20quantify%20the%20association,comprehensive%2C%20and%20well%2Ddocumented.

Stepwise Linear regression

You can install and use the package below:
https://pypi.org/project/stepwise-regression/

You may want to modify the forward regression function to return the model summary statistics as well as. To do this:

DIAML_HW4_Python Resources
Prepared by the TAs

1. You can go to the github repo for this package: https://github.com/AakkashVijayakumar/stepwise-regression/blob/master/stepwise_regression/step_reg.py
2. Copy the forward_regression() function to a cell on jupyter notebooks
3. Modify the function to return the model statistics
4. Use the function as you would any other python function.

Note: If you do this, you will not have to install the package. To do the step-wise regression, just use the forward_regression function as any other python function.