Linear Regression homework assignment with hands-on activities

The following problems will help you understand how to perform linear regression analysis. You will learn how to process training data as well make a prediction for a linear relationship. You will learn how to interpret the output of the fit and quantify regression accuracy using metrics such as variance. Before starting with the assignment make sure you go over the accompanying lecture and hands-on tutorial. For the assignments below, you will work with the *Machine Learning for Materials Science: Part 1 – Using a linear regression model to predict material properties* notebook in the following tool: https://nanohub.org/tools/mseml.

Problem 1. Definitions. Research and provide for definition of the mean square error and variance score in terms of the linear model and data. You can find more information in the following links: https://scikit-learn.org/stable/modules/model_evaluation.html.

Problem 2. Correlations. What correlations do you find between the three pairs of properties studied in the Jupyter notebook? Explain the physical origin of these correlations.

Problem 3. Correlations. Compute the linear correlation between a pair of properties not explored in the notebook. Include the resulting graph and explain the correlation, if any, that you find.

Problem 4. Errors. Note the mean squared error in the various models. What error would you make, in average, if you used the melting temperature to predict Young's modulus?