1). Produce some numerical and graphical summaries of the data. Do there appear to be any patterns?

2). Use the full data set to perform multivariate linear regression with sensors 3 and 4 as the response and the rest columns as the predictors. Do any of the predictors appear to be statistically significant? If so, which ones?

3). Do feature selection, dimensionality reduction, normalizing and selection of loss function have any effect on the results? Find the set that seems to perform well on this data set, and justify your answer. Make sure that you are evaluating model performance using validation set error, cross-validation, or some other reasonable alternative, as opposed to using training error. Separate test set shall also be provided later for evaluation.

4). Does your chosen model involve all of the features in the data set? Why or why not? Does regularization have any effect on model performance?

NaN or 0 values represent invalid or missing data respectively.
Please provide graphs and illustrations depicting your results with justification.
You can use Python or MATLAB for experiments. Python is preferred.
Tutorials and detailed reading material will be shared soon for understanding.