Name HW5

Description

Rubric Detail

Criteria	Levels of Achievement		
	Novice	**Competent**	**Proficient**
Q1-Preparing the log time-log sulfate plot **Weight** **5.00%**	**0 %** Not Done	**80 %** Taking only log of time or log of concentration for model construction.	**100 %** The data points fit the a negative-slope line almost well, but the axis labels show that the plot is in log-log coordinates.
Q1-Preparing the time - sulfate plot **Weight** **5.00%**	**0 %** Not Done	**80 %** Taking only log of time or log of concentration for model construction.	**100 %** The data points still lie on the fitted model, however, the predictive model is an exponentially decaying curve
Q1- Plotting the residual against fitted values in log-log coordinates **Weight** **5.00%**	**0 %** Not Done	**80 %** Different behavior than the reference figure (Please note that there is no randomness in the method, so the results ideally should look the same).	**100 %** An almost sinusoidal shape is supposed to be seen.
Q1- Plotting the residual against fitted values in original coordinates **Weight** **5.00%**	**0 %** Not Done	**80 %** Partially correct plot.	**100 %** An almost sinusoidal shape is supposed to be seen, but the scale of the residuals is substantially different.

Criteria	Levels of Achievement		
	Novice	**Competent**	**Proficient**
Q1-Reasoning about the model fit **Weight** **5.00%**	**0 %** Not Given.	**80 %** Partially correct reasoning.	**100 %** Any correct explanation is accepted with no regard to just the yes/no answer. One correct answer could be that "There is a non random pattern to the residuals which indicates that a linear regression in log log space may not capture the entire relationship."
Q2- Building a regression model predicting the body mass from diameters, and plotting the residuals against the fitted values **Weight** **7.50%**	**0 %** Not Done	**80 %** Partially correct plot/fitting.	**100 %** Basically three lines of R code can do the trick. The random residual pattern must be fine.
Q2-Regressing the cube root of mass against diameters and plotting the residual against the fitted values in both cubic and original coordinates **Weight** **7.50%**	**0 %** Not Done	**80 %** Partially correct plot/fitting.	**100 %** Results being similar to the reference ones. Note that there is no randomness in the methodology.
Q2-Explanation **Weight** **10.00%**	**0 %** Not Given	**80 %** Partially correct explanation.	**100 %** Both methodologies and plots seem effective, but the cubic root regression seems more physically understandable. However, the box-cox transformation might indicate otherwise... Any correct explanation is acceptable.
Q3-Fitting Linear Regression of age wrt to every feature ignoring gender, and plotting the residuals vs fitted values **Weight** **7.50%**	**0 %** Not Done	**80 %** Partially correct plot/fitting.	**100 %** The range of residuals should be acceptable.

Criteria	Levels of Achievement		
	Novice	Competent	Proficient
Q3-Fitting Linear Regression of age wrt to every feature including gender, and plotting the residuals vs fitted values **Weight** **7.50%**	**0 %** Not Done	**90 %** The gender was neither encoded as one-hot, nor treated as a factor. Partially correct plot/fitting.	**100 %** The gender values should be regarded as factor, or encoded as one-hot. The range of residuals should be acceptable.
Q3-Fitting Linear Regression of log of age wrt to every feature ignoring gender, and plotting the residuals vs fitted values **Weight** **7.50%**	**0 %** Not Done	**80 %** Partially correct plot/fitting.	**100 %** The range of residuals should be acceptable.
Q3-Fitting Linear Regression of log of age wrt to every feature including gender, and plotting the residuals vs fitted values **Weight** **7.50%**	**0 %** Not Done	**90 %** The gender was neither encoded as one-hot, nor treated as a factor. Partially correct plot/fitting.	**100 %** The gender values should be regarded as factor, or encoded as one-hot. The range of residuals should be acceptable.
Q3- Choice of Model Among the last ones **Weight** **10.00%**	**0 %** Not Given	**90 %** Partially correct explanation.	**100 %** Students can use box-cox transformations or r-squared or p-value values in order to justify their choice. Any correct explanation with acceptable evidence is enough.
Q3-Doing Cross Validation and regularization **Weight** **10.00%**	**0 %** Not Done	**90 %** Partially correct explanation.	**100 %** Including a plot of cv-error vs lambda could be very helpful. However, the general answer is that the cv mean squared error at its best value, is still worse than no regularization, and larger lambdas yield larger cv error rates.

View Associated Items

Print Close Window