

Goal: project.txt includes 305 simulated observations for a response y and a predictor x . Use the first 300 observations to construct an appropriate polynomial model. Then use your model to predict the last 5 observations and report the prediction error.

Please follow these steps:

1. Explore the data.
2. Fit the initial model $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. Report b_0 and b_1 .
3. Perform diagnostics on the initial model.
4. Fit a more appropriate polynomial model. Use this model for the rest of the project.
5. Perform diagnostics on the polynomial model.
6. Perform an F test for lack of fit at $\alpha = 0.05$. Your model needs to “pass” this test.
7. Report both point and interval estimates for your regression coefficients.
8. Use your model and the last 5 x observations to predict the corresponding y 's.
9. Suppose we define prediction error as

$$\frac{1}{5} \sum_{i=301}^{305} (y_i - \hat{y}_i)^2,$$

where y_i is the i th (response) observation and \hat{y}_i is the i th prediction. Report the prediction error for your model.

10. Write a *concise* conclusion.

The steps above should make up the skeleton of your report. Where appropriate, you need to show all your work, provide plots, or provide *concise*, yet complete and correct explanations.

Some of the steps are vague on purpose. Please do not ask Dmitriy or Christine if you have done “enough;” there is no length requirement. You are done with this project when you feel you have done a good job.

Treat this project as a take-home exam. Do not work with anyone outside of your 3-4 person group. Please do not ask questions that you would not ask during an in-class exam, such as how to do specific steps. If you have specific questions about the instructions, please feel free to ask Christine.

Format:

- Turn in a typed report in which you correctly and *concisely* explain every step of your analysis. Your submission must read as a report; i.e. no numbering system or bullet points.
- You must turn in exactly one report per 3-4 person group.
- You must include a title page with the names of every person in your group.
- Do not include R code in the body of your report; any and all R code must be neatly appended to the back of the report.

Due 3pm on Monday, March 14th. Submit your printed report to Christine Cai’s mailbox in Mathematical Sciences Building 4118.

Good luck!