Assignment 1 (Linear regression):

The data for this question is in the file restaurant.xls. The data is from the Swiggy restaurant guide. There are 114 observations and each observation correspond to a restaurant. There are 4 variables:

- price: the price of a typical meal

- food: the Swiggy rating for the quality of food.

- service: the Swiggy rating for the quality of service.

- decor: the Swiggy rating for the quality of the decor.

We want to see how the price of a meal relates the quality characteristics of the restaurant experience as measured by the variables food, service, and decor. Pl answer all the questions on a randomly selected set of 50 instances.

(a) Plot price vs. each of the three independent variables. Does it seem like y (price) is related to the x's (food, service, and decor)?

(b) Suppose a restaurant has food = 125, service=148, and decor=265. Run the regression of price on food, decor, and service and give the 95% predictive interval for the price of a meal.

(c) What is the interpretation of the coefficient estimate for the explanatory variable food in the multiple regression from part (b)?

(d) Suppose you were to regress price on the one variable food in a simple linear regression? What would be the interpretation of the slope? Plot food vs. service. Is there a relationship? Does it make sense? What is your prediction for how the estimated coefficient for the variable food in the regression of price on food will compare to the estimated coefficient for food in the regression of price on food, service, and decor? Run the simple linear regression of price on food and see if you are right! Why are the coefficients different in the two regressions?

(e) Suppose I asked you to use the multiple regression results to predict the

price of a meal at a restaurant with food = 20, service = 3, and decor =17. How would you feel about it?

** You need to submit a doc file for the answers, a .py file for the codes (properly commented), a doc file summarizing a demo run. Pl mention the honour code in each of the file.