MTH 325, Quiz #4, Spring 2023 Name _____

Instructions: Answer each question as thoroughly as possible. Round answers to 4 decimal places as needed. Exact answers are best when possible. Be sure to answer all parts of each question.

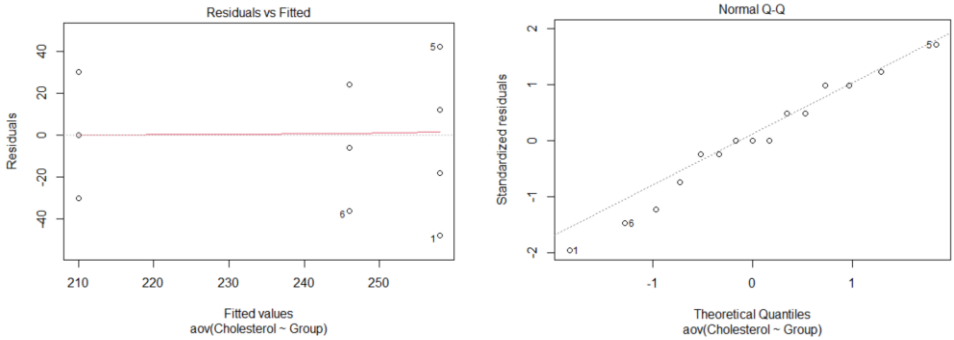
1. A pharmaceutical company conducts an experiment to test the effect of a new cholesterol medication. The company selects 15 subjects randomly from a larger population. Each subject is randomly assigned to one of three treatment groups. Within each treatment group, subjects receive a different dose of the new medication. In Group 1, subjects receive 0 mg/day; in Group 2, 50 mg/day; and in Group 3, 100 mg/day. After 30 days, doctors measure the cholesterol level of each subject. The results for all 15 subjects appear in the table below and in the file **325quiz4data.xlsx**.

Dosage		
Group 1, 0 mg	**Group 2, 50 mg**	**Group 3, 100 mg**
210	210	180
240	240	210
270	240	210
270	270	210
300	270	240

   a. Conduct a one-way ANOVA test to see if there are effects to the medication and dosage level (groups, as a factor), under the assumption that no other levels of the medication are of interest. Clearly state your hypothesis, check your normality assumptions and state your conclusion in the context of the problem.

$$H_0: \mu_i = \mu_j, \forall i \neq j$$
$$H_a: \mu_i \neq \mu_j, for\ some\ i \neq j$$

```
            Df Sum Sq Mean Sq F value Pr(>F)
Group        2   6240    3120    4.16 0.0424 *
Residuals   12   9000     750
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```



Since the p-value is less than 0.05, we can reject the null hypothesis, so we know that at least one mean is different from the others.

   b. State the general linear model from your ANOVA analysis.

```
(Intercept)      Group2      Group3
        258         -12         -48
```

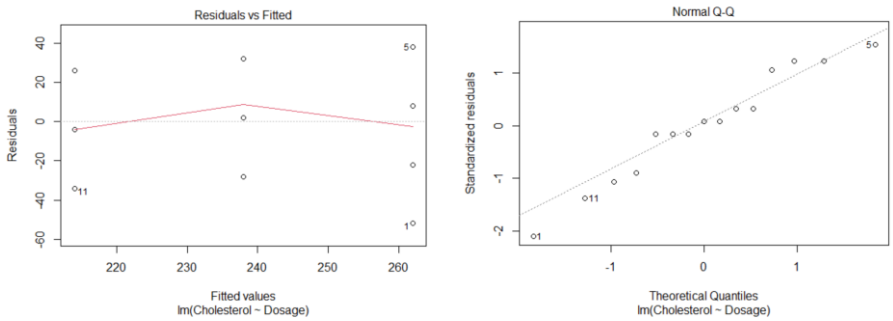
$$\hat{y} = 258 - 12(Group2) - 48(Group3)$$

c. Redo the problem as a simple linear regression using the dosages as the explanatory variable (x) and cholesterol as the response variable (y). Conduct a model test. Clearly state your assumptions and conclusion in context.

```
Call:
lm(formula = Cholesterol ~ Dosage, data = data1)

Residuals:
   Min     1Q Median     3Q    Max
   -52    -13      2     17     38

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 262.0000    11.0244   23.77 4.27e-12 ***
Dosage       -0.4800     0.1708   -2.81   0.0147 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 27 on 13 degrees of freedom
Multiple R-squared:  0.378,  Adjusted R-squared:  0.3301
F-statistic: 7.899 on 1 and 13 DF,  p-value: 0.01473
```

$$H_0: \beta_1 = 0, H_a: \beta_1 \neq 0$$

Assumes normality of residuals and linearity of relationships, steps between dosage levels are constant.



Seems mostly okay.

The p-value is low enough to reject the null. There is a relationship between dosage and cholesterol levels.

d. What is the regression model for this data that you found? How does the regression model compare to the (linear) ANOVA model?

$$\hat{y} = 262 - 0.48Dosage$$

Slightly higher level for 0mg, but predicts the same decrease for 100mg. Does not predict the same for 50mg.

e. Create an appropriate residual plot.
See above

f. If you had to predict the effects of a dosage of 150 mg dosage, which model would be more appropriate and why? Would using either model be appropriate? If not, why not?

While extrapolation is dicey generally speaking, the linear model would be more appropriate to make such an extrapolation. The ANOVA doesn't relate dosage levels, only the one given, so any interpolation or extrapolation is impossible.

g. In general, which model do you think is best for this data and why? Explain.

If the only dosage levels in question are available (maybe due to pill size), then the ANOVA does slightly better in that it predicts the 50mg dose level independent of the 100mg dose level. However, if you need to account for dosages other than those given, the linear model is better and easier to calculate.