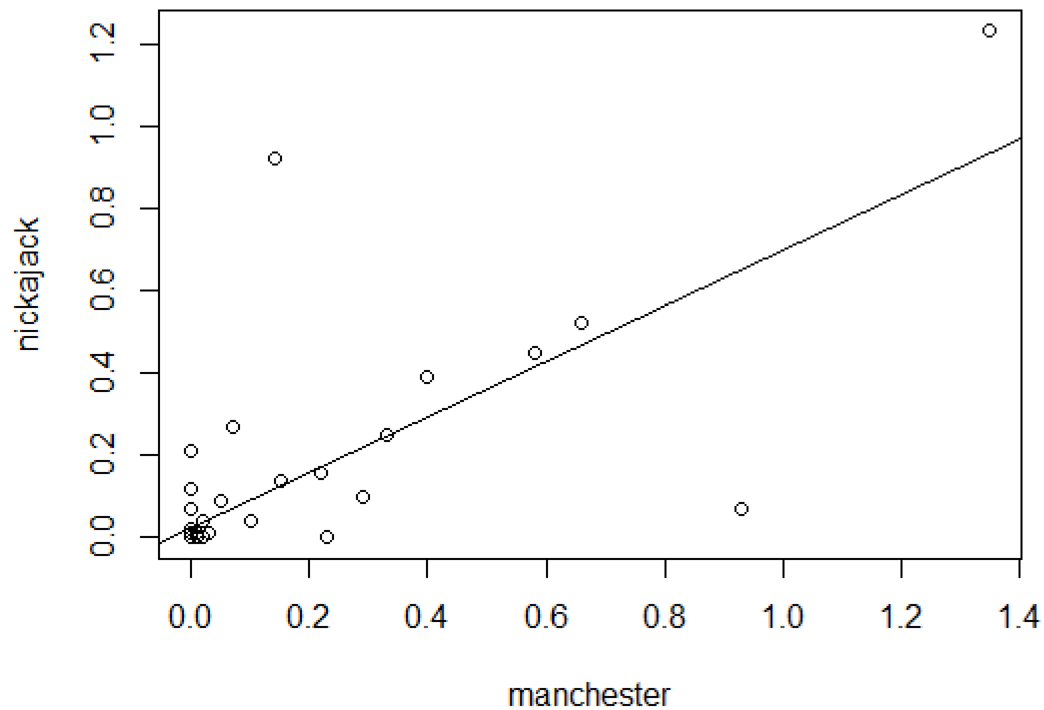
Math 216 Spring 2012
Problem Set 5 Answer Key

Part One

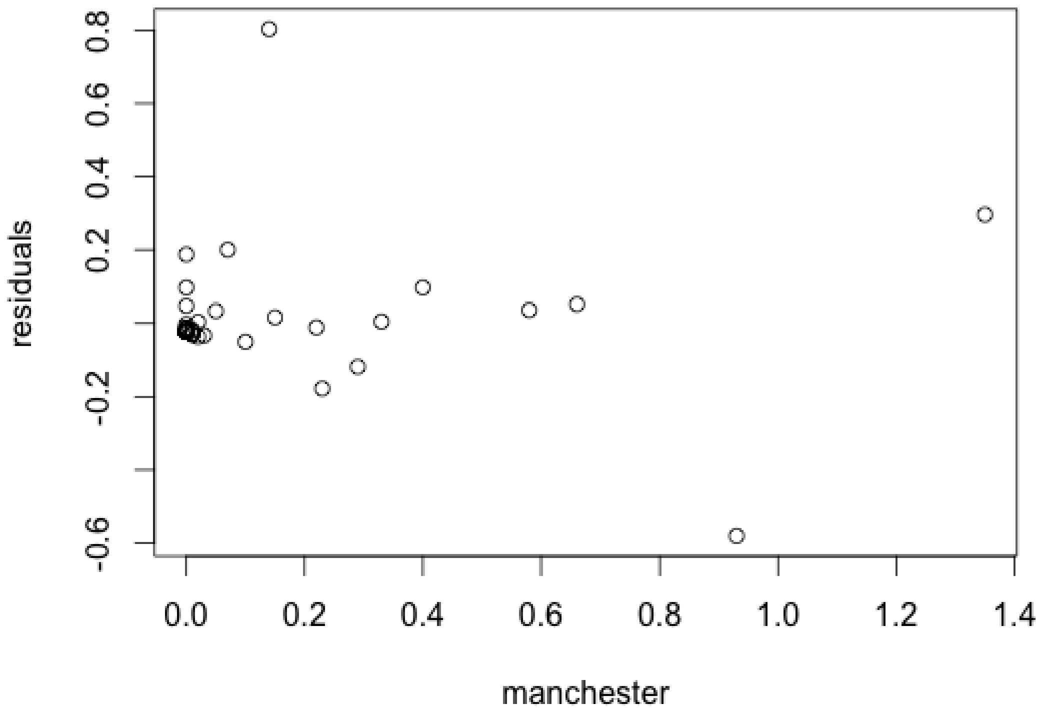
1. $y = 0.02266 + 0.67498x$ [**6pts**]

2. Scatterplot [**6pts**]



3. $r = 0.7473656$ This indicates a moderately strong, positive linear correlation between rainfall at Nickajack and Manchester dams. [**6pts**]

4. Residual Plot [**6pts**]



5. $r^2 = 0.5586$, so about 56% of the variance in rainfall at Nickajack dam can be explained by the linear relationship between rainfall in Manchester and Nickajack dams. [**6pts**]

6. $H_0 : \beta_1 = 0$
   $H_A : \beta_1 > 0$ [**2pts for hypotheses**]
   p-value $= P(\beta_1 \geq 0.67498 \mid \beta_1 = 0) = \mathbf{6.93 \times 10^{-12}}$
   We reject the null hypothesis and conclude that there is strong evidence of a positive linear relationship between rainfall at Manchester and Nickajack dams. [**7pts total**]

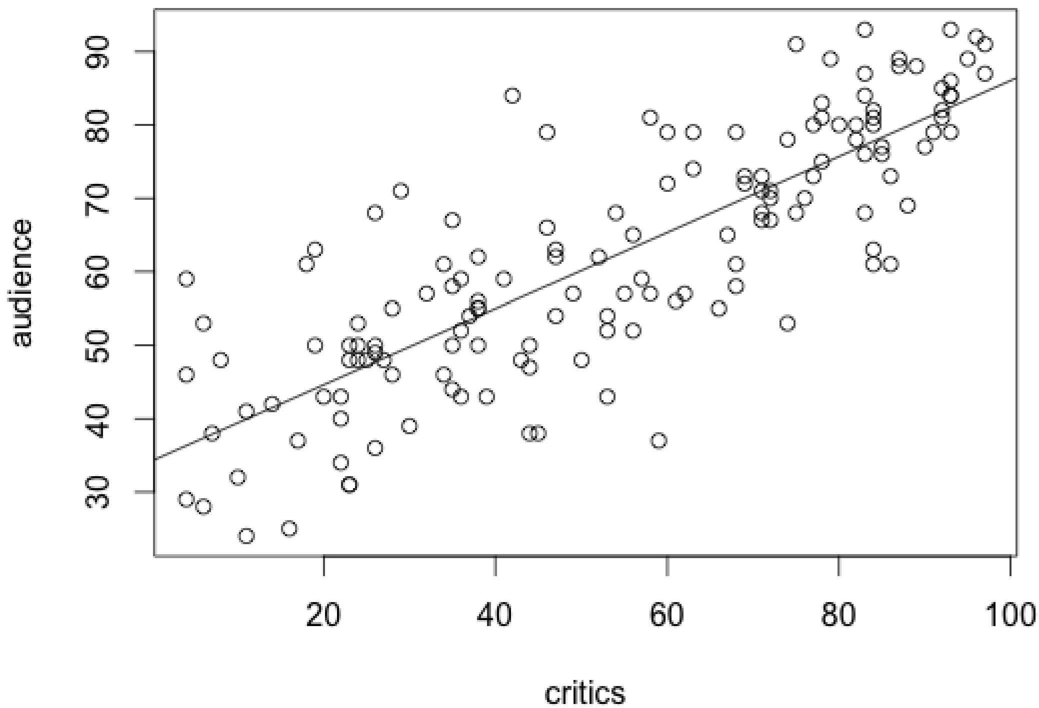
7. Based on the linear regression model, if Manchester receives an addition inch of rain beyond what is forecasted in a given day, Nickajack will receive **0.67498** additional inches of rain. [**6pts**]

8. Residuals are not exactly nearly normal. There are a few major outliers in the residual plot. More importantly, the residuals don't have constant variability. (As x increases, the residuals increase in absolute value). Therefore the least squares line may not be the best fit for the data. [**6pts : 4pts for saying "either not normal" or "lack of constant variability" but not both**]

Part Two

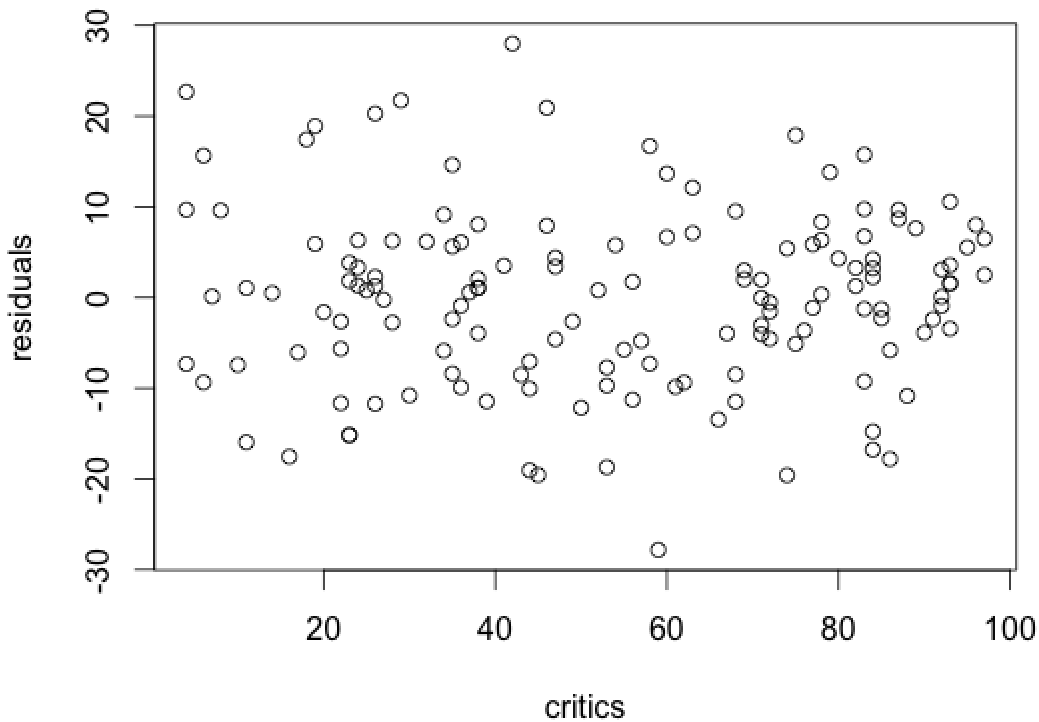
1. $y = 34.25562 + 0.51612x$ [**6pts**]

2. Scatterplot [**6pts**]



3. $r = 0.8200528$ This indicates a moderately strong, positive linear correlation between the audience scores and the critics scores. [**6pts**]

4. Residual plot [**6pts**]



Residual plot appears nearly normal, and we have what appears to be constant variability with no obvious outliers, so least squares appears to be a good fit.

5. $r^2 = 0.67$, so about 67% of the variance in average audience scores can be explained by a linear relationship between average audience scores and average critics scores. [**6pts**]

6. $H_0 : \beta_1 = 0$
   $H_A : \beta_1 > 0$
   p-value $= P\left(\beta_1 \geq 0.51612 \mid \beta_1 = 0\right) = \mathbf{2.2 \times 10^{-16}}$
   We reject the null hypothesis and conclude that there is strong evidence of a positive linear relationship between average audience scores and average critics scores. [**7pts: 2 for hypotheses**]

7. $H_0 : \beta_1 = 0.5$
   $H_A : \beta_1 > 0.5$
   p-value $= P\left(\beta_1 \geq 0.51612 \mid \beta_1 = 0\right)$
   $$= P\left(z \geq \frac{0.51612 - 0.5}{0.03013}\right) = P\left(z \geq 0.6014\right) = \mathbf{0.2743}$$
   We fail to reject to the null hypothesis. There is not strong evidence that the linear relationship between average critic scores and average audience scores has a slope

greater than 0.5. [**8pts: 2 for hypotheses, 4 for computation of z-score (2 for SE), 2 for p-value**]

8. $20 = 34.25562 + 0.51812x$ Solve: $\mathbf{x = -27.51}$ The prediction is problematic because we cannot have negative scores. [**6pts: 2pts for explanation**]