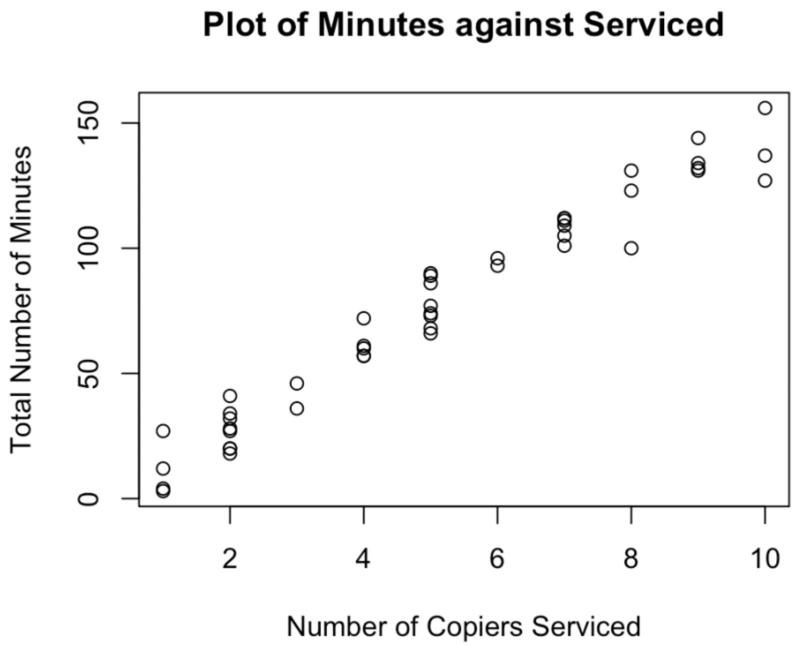
Yunlu Li (yl4df)
STAT 5120
Homework 2

1. (a)The linear regression model appears appropriate.



(b) $\hat{y} = -0.5802 + 15.0352x$

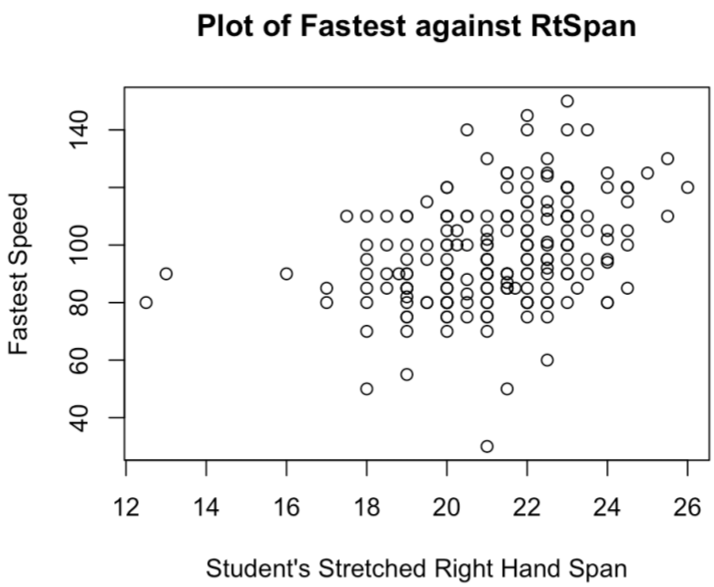
(c) The change in the estimated service time is 15.0352 when the number of copiers serviced increases by one. The 95% confidence interval is (14.061010, 16.009486). We have 95% confidence that this interval contains the true change in the estimated service time when the number of copiers serviced increases by one.

(d) $H_0: \beta_1 = 0, H_1: \beta_1 \neq 0$. The p-value is $2 * 10^{-16}$. It is less than significance level, so we reject null hypothesis. We conclude that there is a linear association.

(e) The result from (c) is consistent with the result from (d). The lower bound of 95% confidence interval is much greater than 0 and the p-value is very small. Both indicate that there is a linear association.

(f) The mean service time when 5 copiers are serviced is 74.59608. The 95% confidence interval is (71.91422, 77.27794). We have 95% confidence that this interval contains the true mean service time when 5 copiers are serviced.

2. (a) $\hat{y} = 30.5199 + 3.1441x$.
The p-value is $7.3 * 10^{-7}$. The p-value indicate that there is a linear association.



(b) The p-value for male subset is 0.43071, so there is no linear relationship between two variables for male. The p-value for female subset is 0.788, so there is no linear relationship between two variables for female.

(c) As we can see, both subsets do not show linear relationship. However, the combined data shows a linear relationship. That is why the results differ.