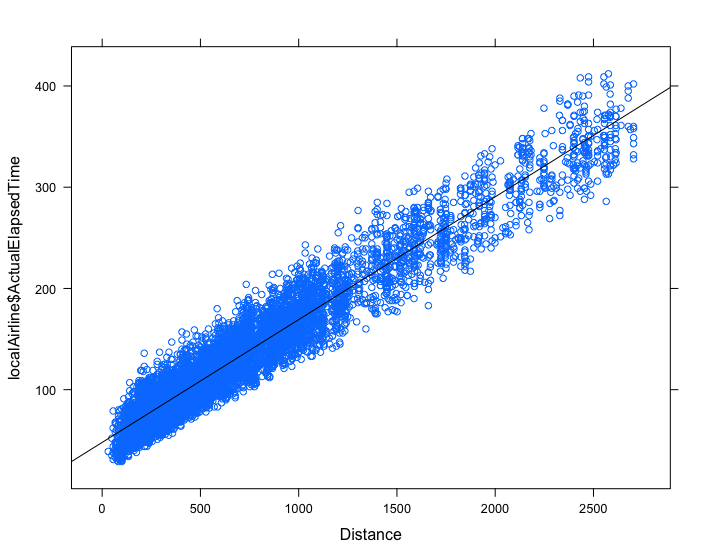
Lab 9

Tian Qiu

2.



3.

This scatter plot shows that distance and actual elapsed time are positive and strong linear relationship. I cannot see any obvious outliers. Yes it is linear.

4.

[1] 0.960428

The correlation coefficient between distance and actual elapsed time is 0.960428. This looks like there is strong association between distance and actual elapsed time. Therefore, the strength is tremendous and it is the same as problem 3.

5.

Yes, the correlation is a good numerical summary of the graphical display in the scatterplot. Because correlation is used to describe the linear relationship between two continuous variables. In general, correlation tends to be used when there is no identified response variable.

6.

1. (5 pts) Look at the scatterplot for these data that you made in part (2). Is the correlation a good numerical summary of the graphical display in the scatterplot? Please explain by discussing the reasons why correlation can or cannot be used.
2. (6 pts) Obtain the equation of the least-squares regression line for predicting the actual elapsed time from the distance. What is r2 for these data?
3. (5 pts) Predict the actual elapsed time when the distance is 297, and calculate the residual. This part may be done by hand.  Bonus: 5 pts.: Determine the residual. Use the data point on 11/7 from MSP to MKE with a departure time of 15:29.
4. (5 pts) Obtain the residuals and plot them versus the distance. There is no need to have a listing of the residuals. Is there anything unusual to report? If so, explain. Are the conclusions from the residual plot the same as from the scatterplot (parts 3 and 5)? If they are different, provide a possible explanation for the difference.
5. (5 pts) Do the residuals appear to be approximately Normal? Explain your answer. Be sure to include the appropriate graph(s) in your answer.
6. 

STAT 350 (Spring 2016) Lab 9 2

1. (5 pts) Based on your answers to parts, (2), (8), and (9), do the assumptions for the linear regression analysis appear reasonable? Explain your answer.
2. (12 pts) Construct and interpret a 99% confidence interval for the slope and the intercept. What is the significance of the result for the slope? Is the inference on the intercept of interest in this problem? Why or why not?
3. (10 pts) Is there significant evidence that distance is associated with actual elapsed time at a 0.01 significance level? Please perform the 4-step process (identify the variables, state hypotheses, give a test statistic and P-value, and state your conclusion).
4. (6 pts.) How are the results from parts 11 and 12 similar? How are they different?
5. (11 pts.) Write a short paragraph in complete English sentences summarizing the results which is understandable to non-statisticians. The summary should contain the following parts: a) is the model appropriate to use, b) What is the relationship between the distance and the actual elapsed time? c) Is this situation good for prediction? d) Is there any causality in this situation? e) Can you generalize this situation to November of 2015? f) In addition, provide a justification for not including distances over 2704 miles which does not include anything concerning making the assumptions valid.