1. **How would you characterize the deviation of the red points from the points on the blue line in figures on slide 3 (write the equation of al line and use the equation as a basis for your explanation)?**

TV: y = 0.0475x + 7.033

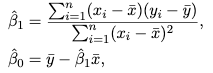
Radio: y = 0.2025x + 9.312

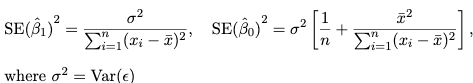
Newspaper: y = 0.0546x + 12.35

The TV model is the best-fitted line which indicates there’s a direct correlation between sales and TV advertisements.

The Newspaper model is the least-fitted line which indicates that newspaper advertisements don’t have that much of an impact on sales

1. **What is the definition of RSS and why is the preferred metric for an optimization algorithm?**  
   RSS is the residual sum of squares and shows the difference between actual data and predicted data. It is the preferred metric for an optimization algorithm because one can use it to analyze their algorithm. If the value of RSS is small, then the data predicted from the algorithm will be very close to actual data, which means the algorithm is accurate and is a good model.
2. **What the key elements of an optimization algorithm?**Objective functions, predictor values, and target values are the key elements of an optimization algorithm. In a model, the objective function will use a set of predictor values to compute the target value.
3. **Explain what is a 95% confidence interval? If you were to determine the 95% interval for the 3 figures on slide 3 what would be the process? (Explain your answer qualitatively)**A 95% confidence interval is defined as a range of values such that with 95% probability, the range will contain the true unknown value of the parameter.  
     
   There is a simple process/formula to determine the 95% confidence interval.
4. Take the linear regression models calculated earlier in part A
5. Compute the least squares coefficient estimates (formula shown below)



1. Compute the RSS (RSS = e12 + e22 + · · · + en2)
2. Compute the standard errors associated with the coefficient estimates (formula shown below)  
   
3. Compute the confidence interval with the standard error from step 4 (β1±2·SE(β1))