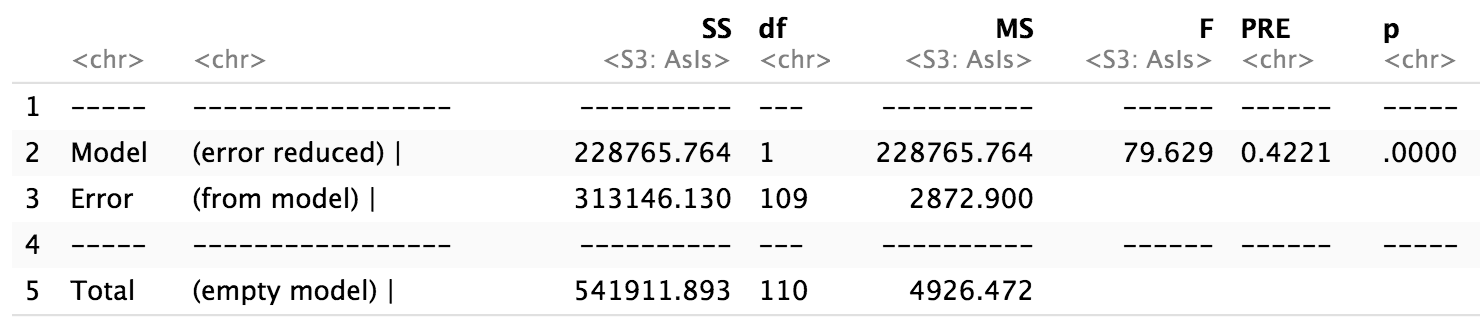
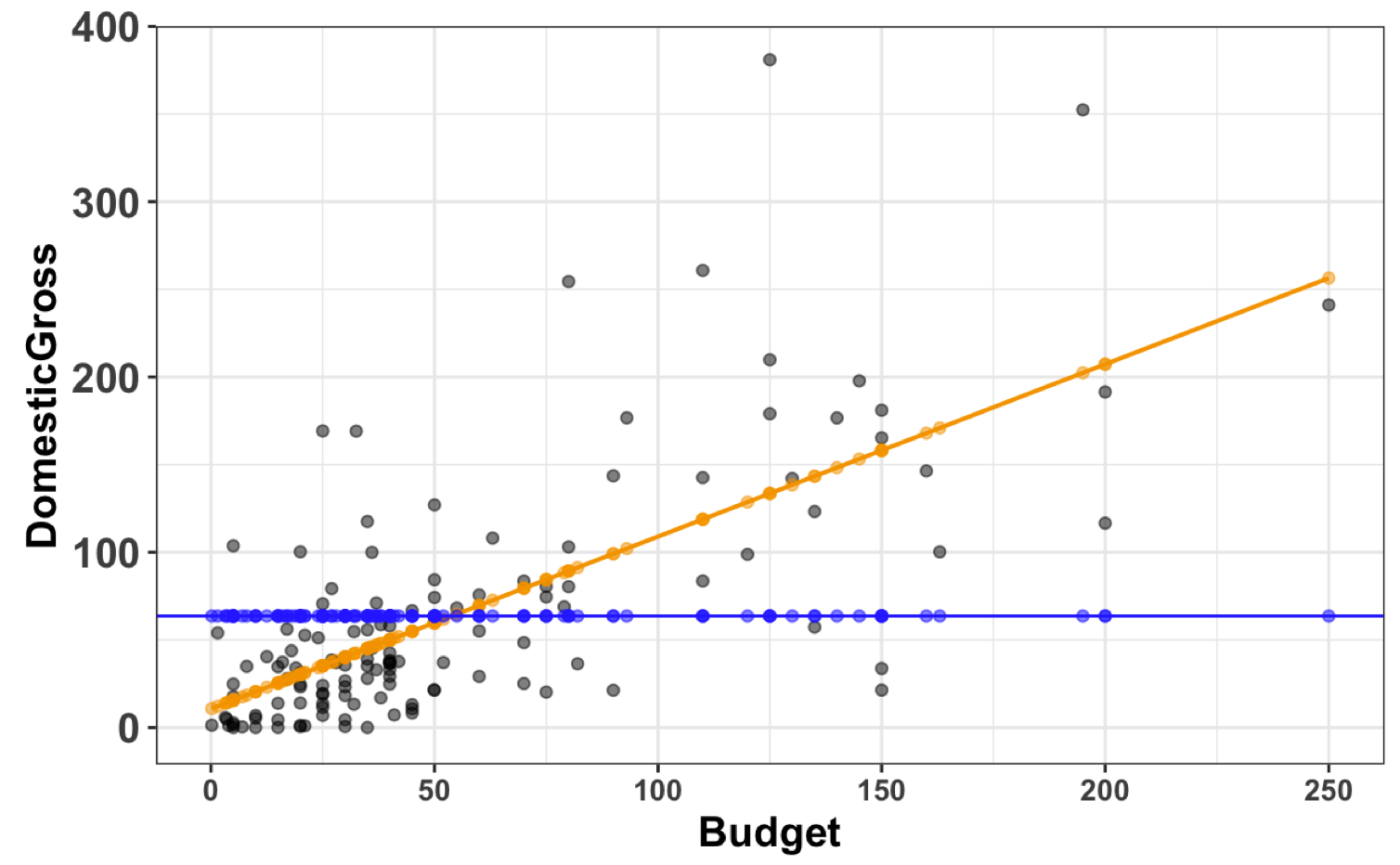
# Name:

# **Classwork 15: How much error has been explained by the Budget model?**

Here is the supernova table for the Budget model (note we are using the data frame **Movies** from the previous classwork). Let’s delve into these different measures of how much error has been explained a little more deeply.



1. What is the relationship between MS from the empty model and standard deviation?
2. Draw the average residual from the empty model. Draw the average residual from the budget model. Which one is smaller? What does that mean?



1. Interpret the PRE (0.42) from the supernova table. What does it mean?
2. Wow, almost half the variation in **DomestricGross** is explained by **Budget**! Do these data and the analysis we did demonstrate that spending more money on a movie causes it to make more money in the domestic market? Why or why not?
3. What is the correlation coefficient between **DomesticGross** and **Budget**?
4. What is the relationship between the correlation coefficient (often notated as *Pearson’s r* – not to be confused with the R programming language!) and PRE?
5. **DomesticGross** is measured in millions of dollars. Create a new variable called **DGdollars** that is the **DomesticGross** measured in dollars.
6. If you knew **DGdollars**, do you think you could do a better job of predicting **DomesticGross**? How good of a job would you be able to do?
7. If you created a model predicting **DomesticGross** with **DGdollars**, what would be its SS Error? What would be the PRE? (Make a prediction here – and then try it.)
8. Why is the SS Error so low? Why is the SS Total the same as when we ran supernova with the Budget model?

