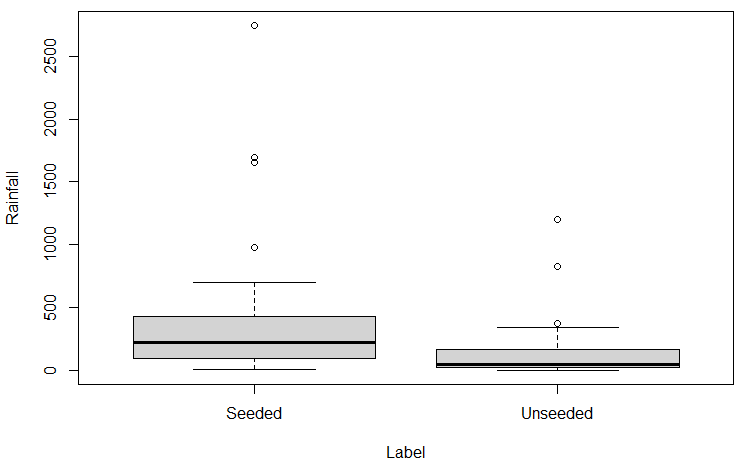
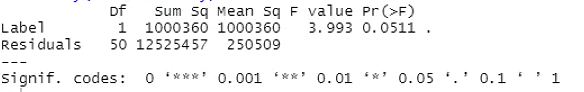
**1.1 Plot two box plots side-by-side of data from the two groups. Discribe the distributions.**

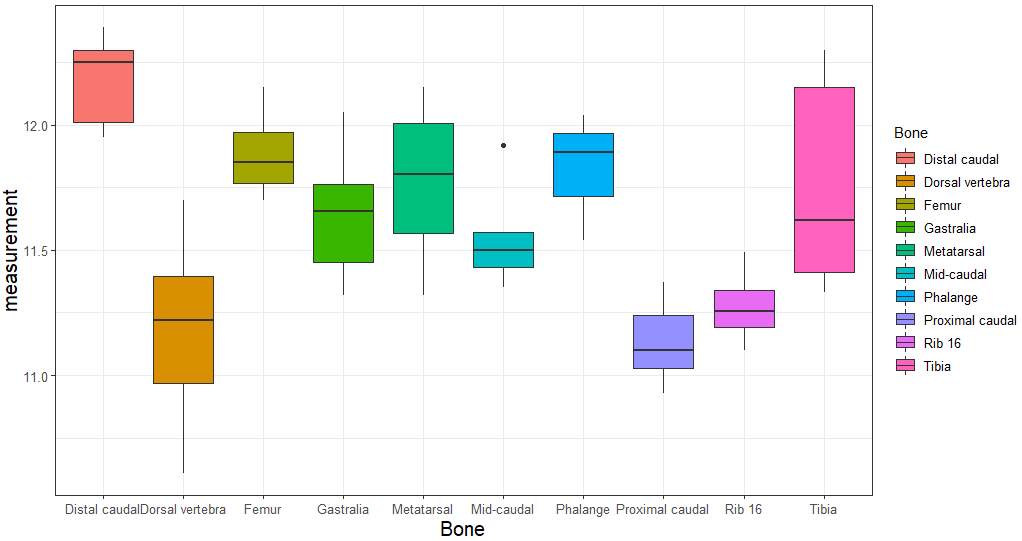


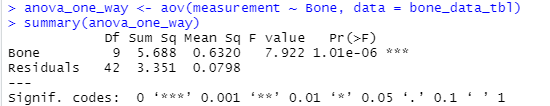
**1.2 Did cloud seeding have an effect on rainfall in this experiment? If so, how much?**



从以上结果看出，cloud seeding对降雨没有明显作用

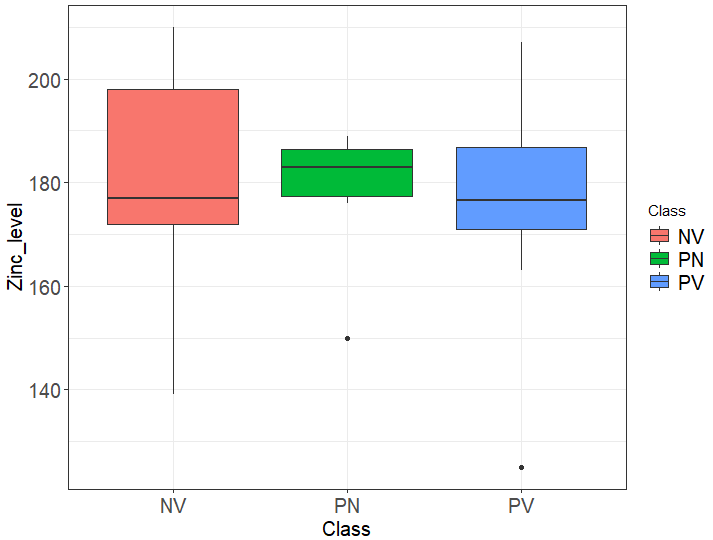
**2. Is there evidence that the means are different for the different bones? Does the dataset support Tyrannosaurus Rex is warm-blooded or not?**

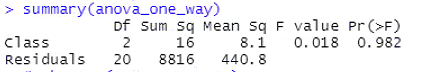




由箱线图和anova\_one\_way的结果可以看出，不同骨骼的温度有非常显著的差异，因此Tyrannosaurus Rex不是warm-blooded。

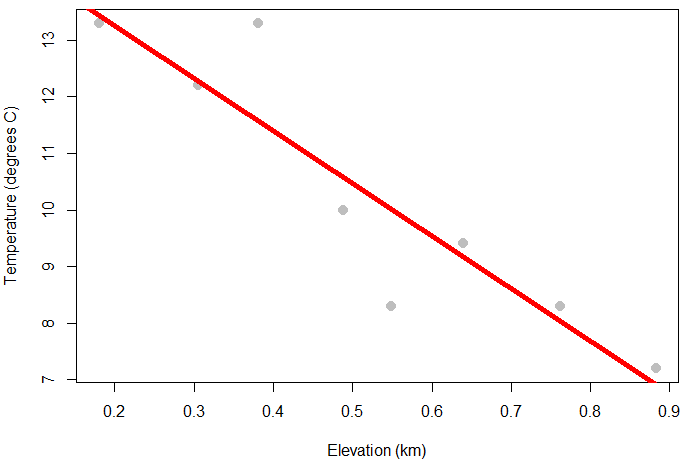
**3. What evidence is there that pregnant vegetarians tend to have lower zinc levels than pregnant nonvegetarians?**

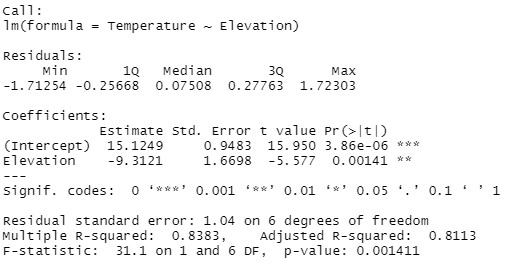




由箱线图和anova\_one\_way结果可以看出，参与调查的对象中PV和PN以及NV的锌水平并没有明显差异。

**4. Draw a scatter plot with regression line, and investigate if the lapse rate is 9.8 degrees C km-1.**

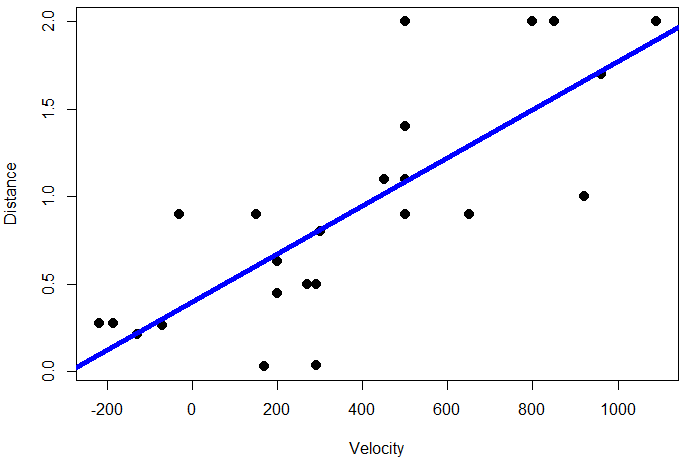




可以看出，斜率并不是9.8 degrees C km-1

**5.1 Make a scatter plot with distance as the Y-axis and recession velocity as the X-axis. Describe what you see.**

**5.2 Add a simple linear regression line to the above scatter plot.**



由散点图可以看出，虽然整体上随着Velocity的增大，Distance的值也增大，但是这些点都很分散。

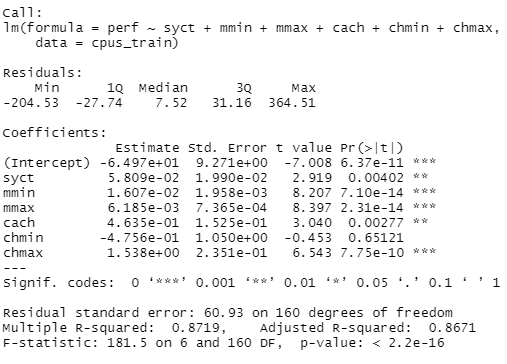
**5.3 If Hubble’s Big Bang Theory is correct, explain why the following two assumptions about the regression line you made in 5.2 need to be true:**

* **The intercept should be zero**
* **And the slope is the age of the universe**

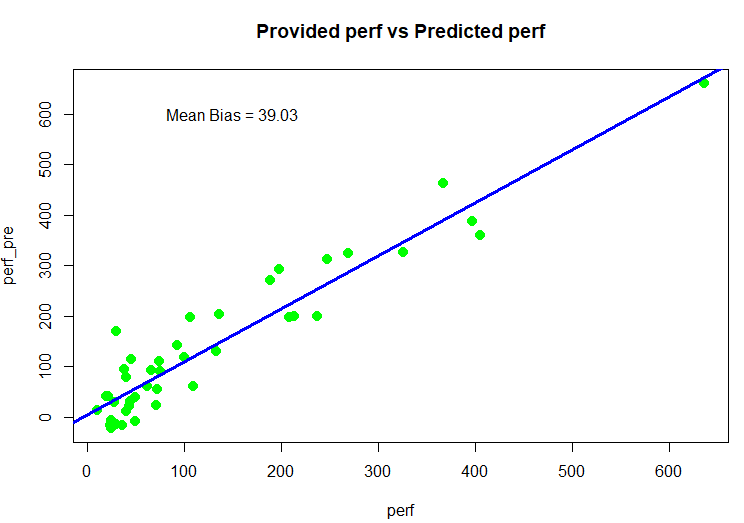
**Address the first assumption with your regression results; and estimate the age of the universe.**

**5.4 Explain why improved measurement of distance would lead to more precise estimates of the regression coefficients.**

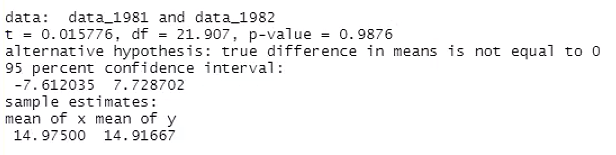
**6.1 For the train set, fit the best subset regression between predictor variable perf and response variables including syct, mmin, mmax, cach, chmin, and chmax.**



**6.2 Apply the best regression model to the test set, and compare your predicted perf values with the actual values that provided in the test set. Quantify the mean bias between predicted perf values and provided perf values.**

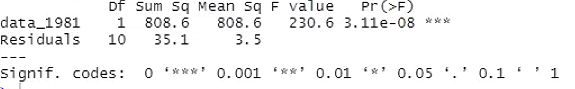


**7.1 Define a simple research question that can be tested with the t-test. Test your question with R, and describe your findings.**



该结果比较了1981和1982年某站点的年平均温度有无明显差异，可以发现没有明显差异。

**7.2** **Define a simple research question that can be tested with the ANOVA. Test your question with R, and describe your findings.**



**相同的数据，此时表现为明显差异，可能是两种方法侧重的点不同。**

**7.3 Define a simple research question that can be tested with a simple linear regression model. Test your question with R, and describe your findings.**