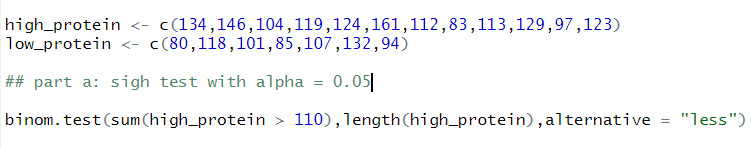
Assignment3 for MAST90058

Name: Mu Tong

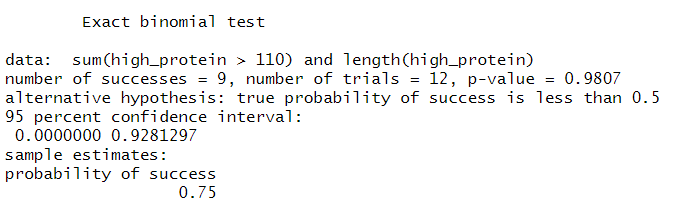
Student Number: 1004452

Question 1：

1. R script:



Output:

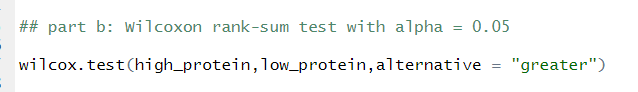


Null hypothesis: the median weight gain in the first group is 110 (m = 110)

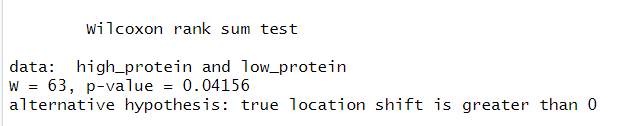
Alternative hypothesis: the median weight gain in the first group is smaller than 110 (m < 110)

We cannot reject null hypothesis because p-value (0.98) is bigger than 0.05.

1. R script:



R output:

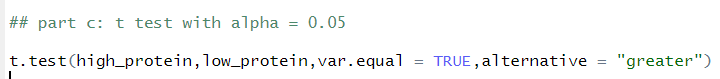


Null hypothesis: the median weight gain in the first group is the same as that in the second group.

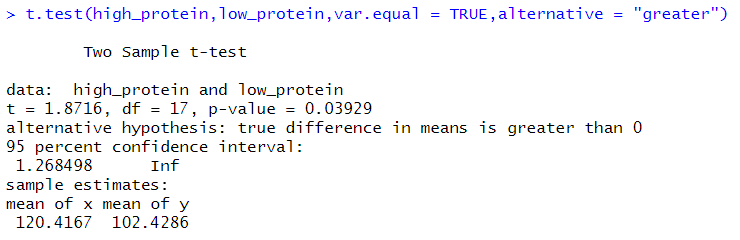
Alternative hypothesis: the median weight gain in the first group is larger than that in the second group.

We can reject null hypothesis because the p-value (0.04156) is smaller than 0.05.

1. R script:



Output:

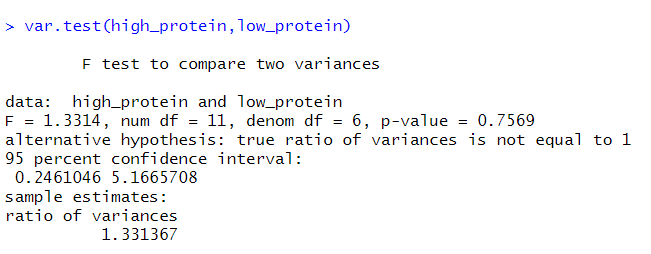


Null hypothesis: the mean weight gain in the first group is the same as that in the second group.

Alternative hypothesis: the mean weight gain in the first group is larger than that in the second group.

We can reject null hypothesis because the p-value (0.039) is smaller than 0.05.

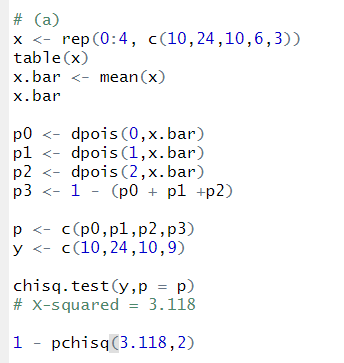
To check the assumption of the same variance.



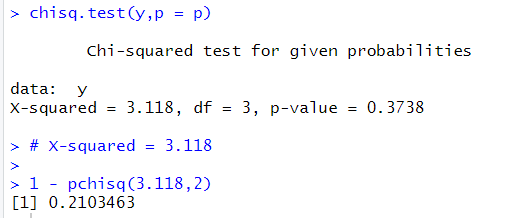
We can get the high p-value is 0.7569, so we cannot reject null hypothesis of the same variance.

Question 2

1. R script:



Output:

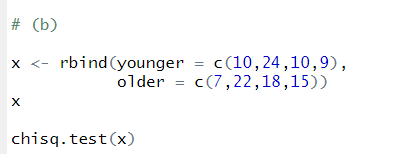


Null hypothesis: the hours of exercise for the younger age group follows a Poisson distribution.

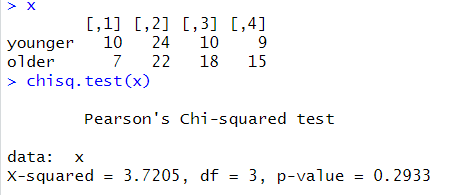
Alternative hypothesis: the hours of exercise for the younger age group doesn’t follow a Poisson distribution.

We used goodness of fit test to test, and we find the p-value is greater than 0.05, therefore, we fail to reject null hypothesis, which means that the data follows a Poisson distribution.

1. R script:



Output:



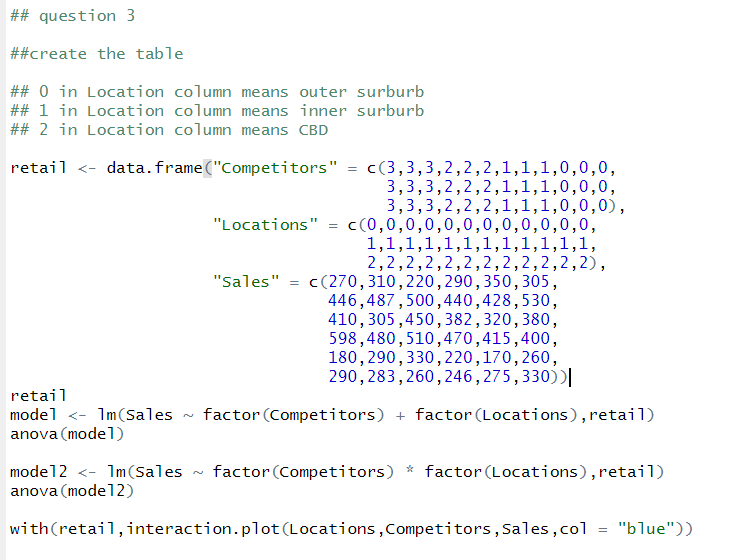
Null hypothesis: age is independent of hours of exercise.

Alternative hypothesis: age is not independent of hours of exercise.

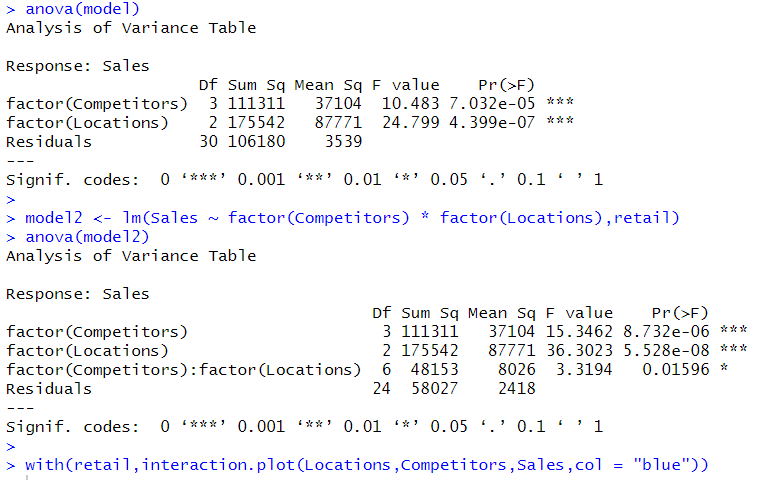
We use contingency tables to test the independence, we find the p-value is 0.2933 which is much greater than 0.05, therefore, we fail to reject null hypothesis at 5% significance level, which means that age is independent of hours of exercise.

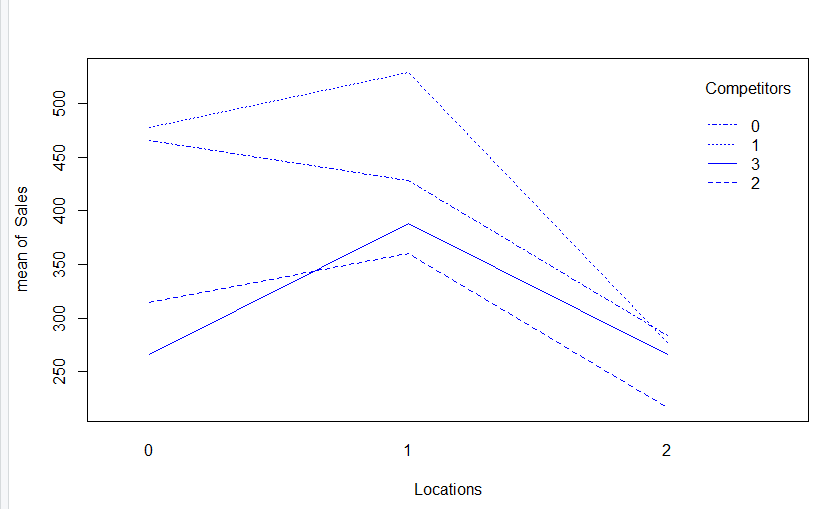
Question 5:

R script:



Output:





In the first model, we make some assumptions. The first one is that they both follow normal distribution, and the second one is that factors are independent, and the last one is that all populations have the same variance.

we find that the F-values for number of competitors and locations factors are 10.483 and 24.799 respectively, and the p-value for these two factors are 0.00007032 and 0.0000004399 respectively, which are much less than 0.05. Therefore, we get the conclusion about there is a clear difference in retail sales between the number of competitors (We reject null hypothesis of factor competitor) and also between the locations (we reject null hypothesis of factor location).

It is possible to test for interaction, and the interaction plot is showed above. From the interaction plot, we can find that the interaction p-value is 0.01596 which is smaller than 0.05, therefore, we can say that there is no interaction between these two factors.