The problem with part 1 of the analysis is that we have to look at 8 different displays to see what is happening! One thing we could do is to put 8 displays side by side. Another thing we could do is to combine the categories to get an overall score. For the first seven questions a 5 is a good score and a 1 is a bad score, so if we add up the scores for questions 1-7 we get a scale that goes from 7 to 35, with 35 being best and 7 being worst.

Total of Questions 1-7

```
          Means with the same letter are not significantly different.

              Tukey Grouping           Mean      Lab      Avg Score

                     A               27.136      11        3.88
                     A
                 B   A               26.150      10        3.74
                 B   A
                 B   A    C          24.348       5        3.48
                 B   A    C
                 B   A    C          23.182      16        3.31
                 B        C
                 B   D    C          22.444       7        3.21
                     D    C
                 E   D    C          21.056       2        3.01
                 E   D
                 E   D               18.636       6        2.66
                 E
                 E                   18.095       9        2.59
```

Lab 11 had the highest observed average, but we can be 95% sure that one of 11, 10, 5, or 16 will have the highest quality.
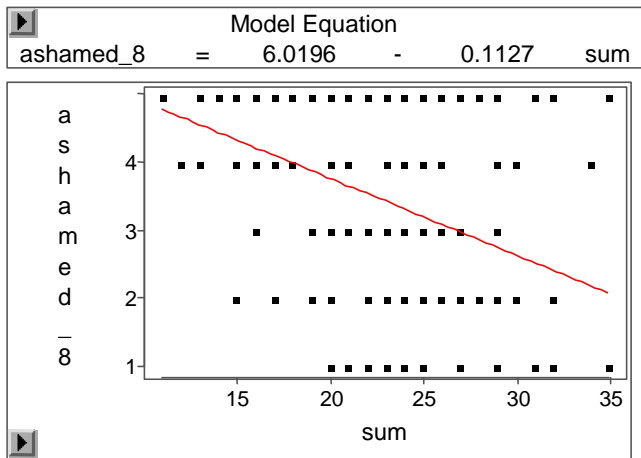
Lab 9 had the lowest observed average, and we can be sure that one of 9, 6, or 2 actually has the lowest overall quality based on the student reviews.

Note we don't have a large enough sample size to say anything conclusive about many of the videos. For example we can't tell if lab 2 is any better than lab 9 or lab 2 is any worse than lab 5.

Say we wanted to add question 8 into the total too. The problem with that is that a high number is bad. (It means the prof should be very ashamed for showing it!)   We could reverse the scores by adding in 6 minus the score on 8. This would change a 1 to a 5 and a 5 to a 1.

Question 8 has another problem though. It looks like many people didn't read the question very well and never realized that a high score was bad on this question!

Look at the following regression for predicting question 8's response from the total of the other 7 questions:

Model Equation
ashamed_8   =   6.0196   -   0.1127   sum

Notice that the regression equation starts right... the higher the score the less ashamed you should be.

But notice that someone gave a video a perfect score of 35 and still said the instructor should be very ashamed (gave it a 5 on question 8 too!). Someone else gave a video a total score of only 15 on the 7 questions but gave it a 2 for question 8 (don't need to be very ashamed).

Clearly this question has problems!