Cracking Professor Moody's Grading Scale the Old Fashioned Way

BY FAUZAN AMJAD

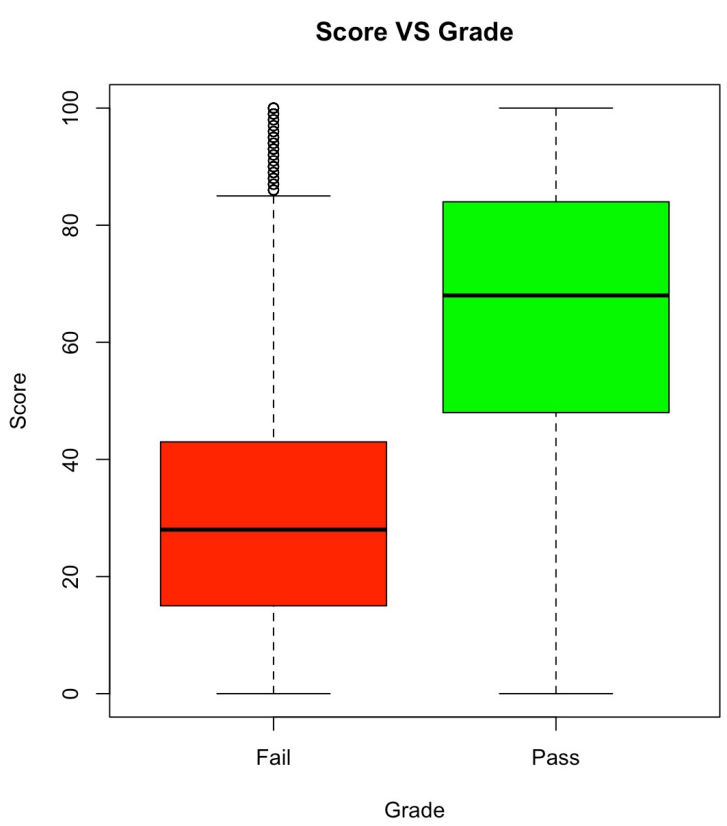
The Mission

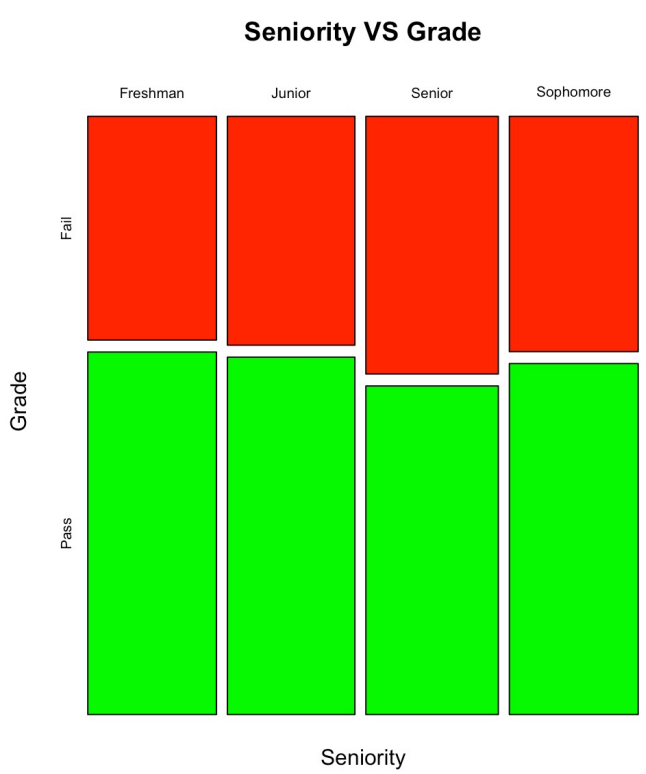
Professor Moody's class is run on a Pass/Fail grading, scale.

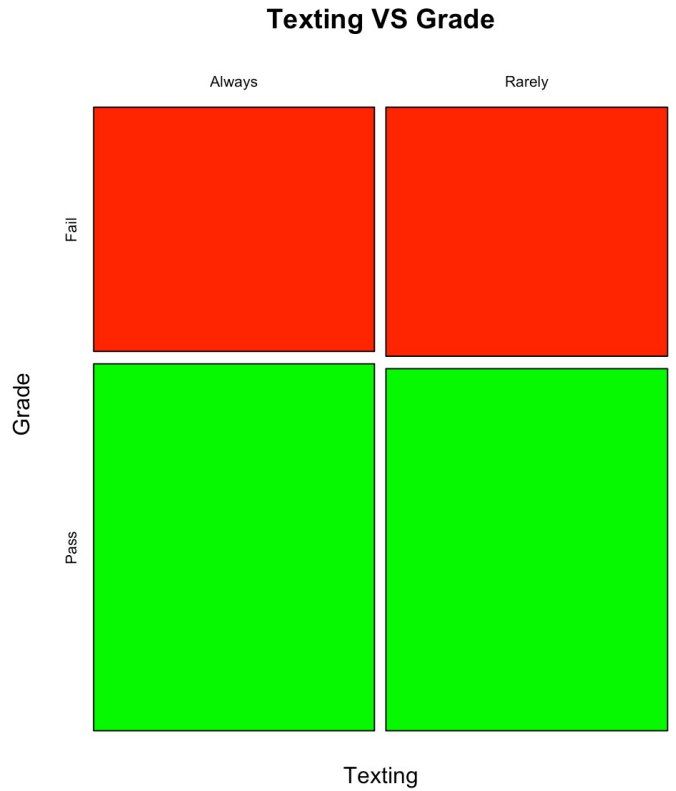
The only issue is that the numerical score is not the only determinant of this categorical grade.

We have to analyze to values such as attendance, major, frequency of asking questions, and frequency of texting to make our interpretation.

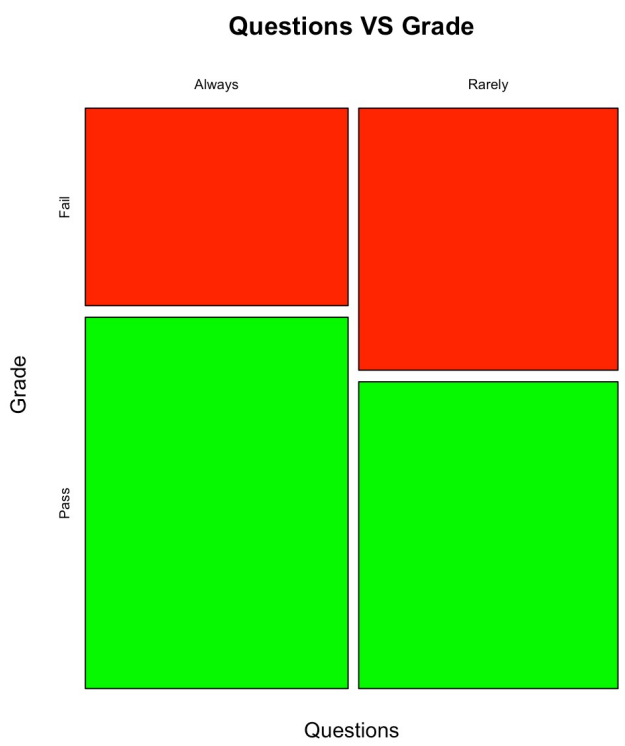
Making the Basic Plots

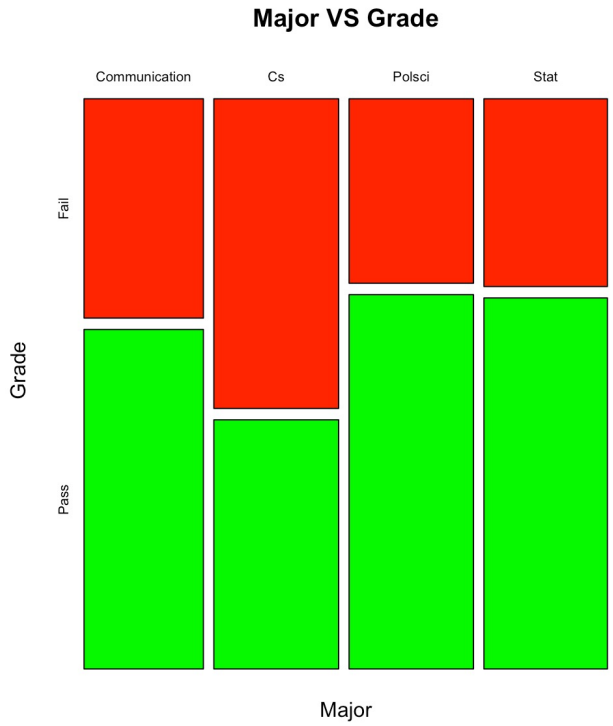






Making the Basic Plots (cont.)





Analysis of Basic Plots

From a broad perspective, plot every variable against the grade in the class reveals very little.

However, we can see how there is a huge gap disparity between the median score of failed students (roughly 30) and the median score of passing students (roughly 70) via the Score VS Grade boxplot. This means that score plays a major role in shaping grade.

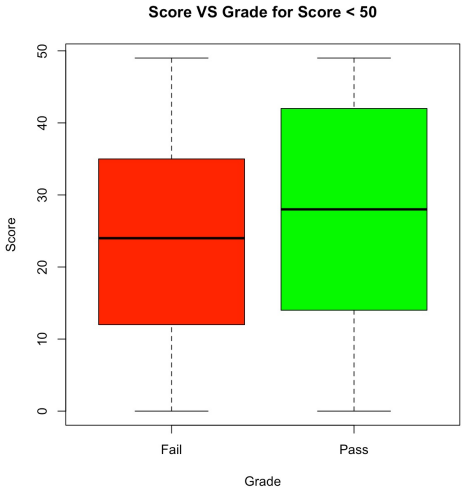
Another thing to note is that the percentage of computer science students failing sticks out the most within the mosaic plot. This will be a pattern I'll analyze in depth later.

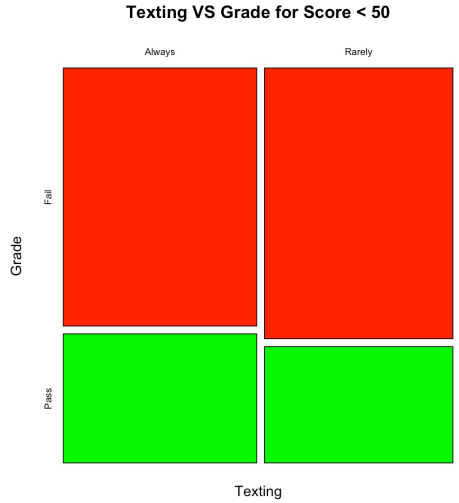
Since this is the most statistically revealing plot, I decided to subset my data by splitting it into two groups: students who had a score less than 50 and students who have a score greater than 50.

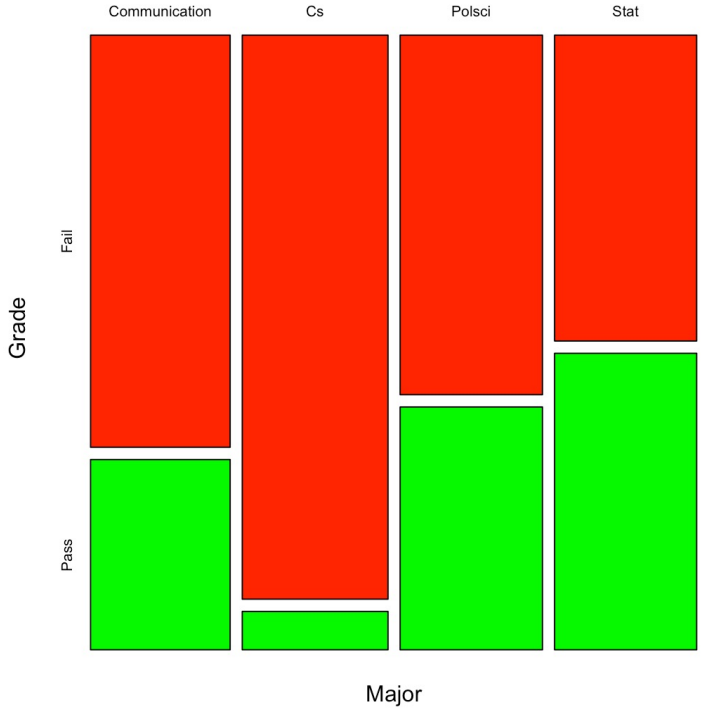
- Reasoning -> I simply added the two median and divided it by two -> (30+70)/2

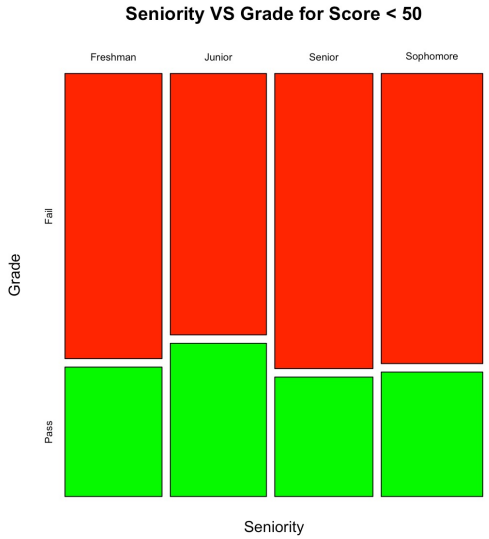
Analyzing Students Who Have a Score Less than 50 to find Certain Trends

Plots for Score < 50

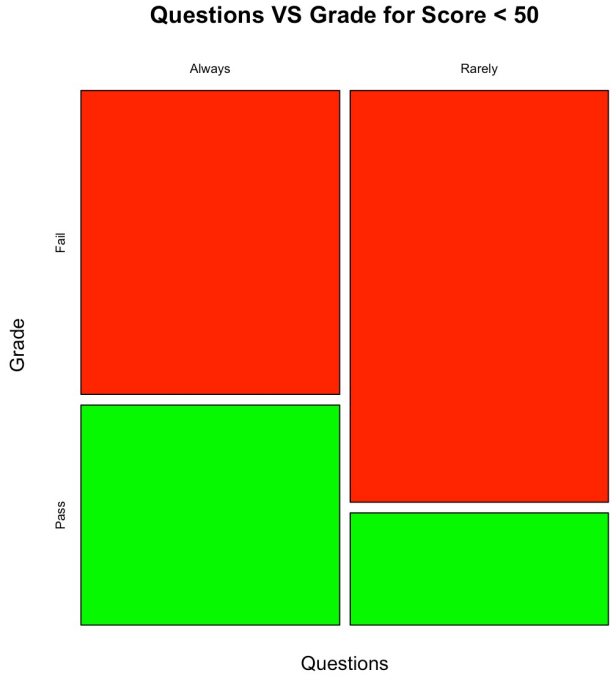








The bigger the plot, the more I want to emphasize it.



Analyzing the Score < 50 Subset

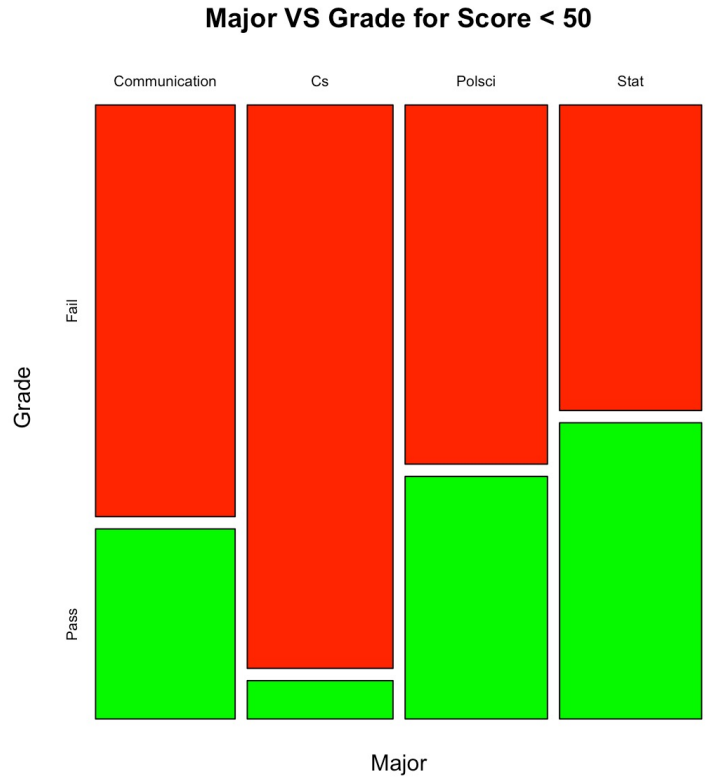
The most significant plot that stuck out to me from the previous slide was the mosaic plot of Grade VS Major when we subsetted the data to include student that have attained a grade less than 50.

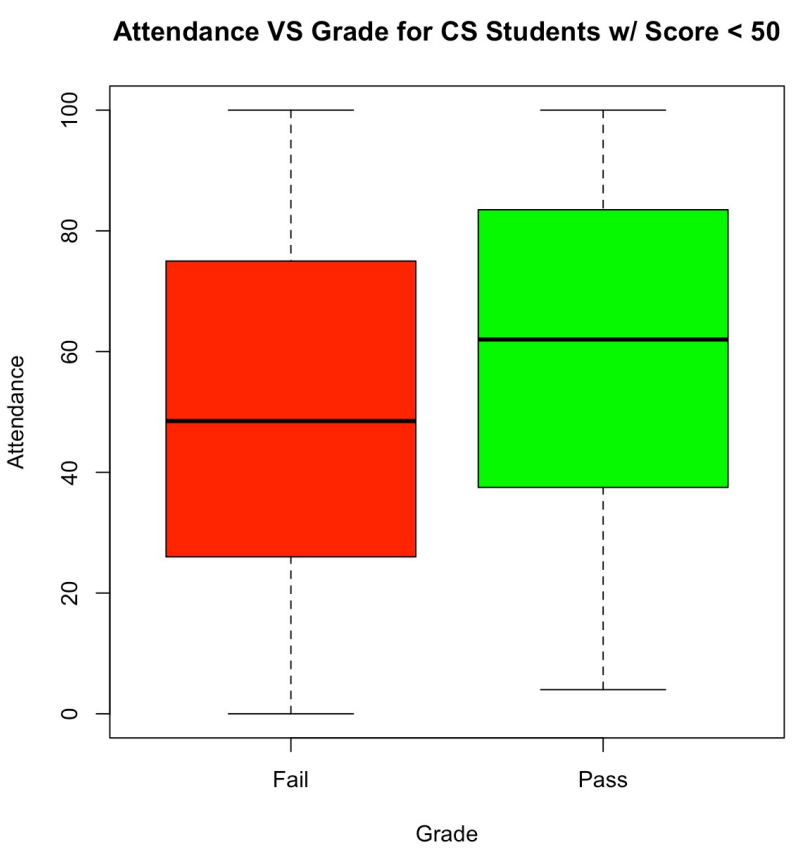
An overwhelming majority of computer science students that have gotten a score less than 50 ended up failing.

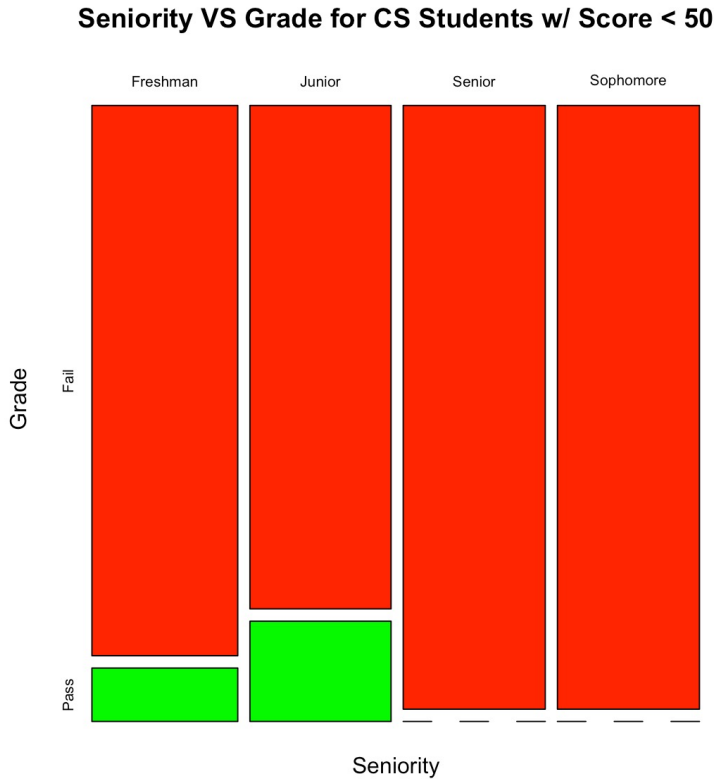
- We will have to investigate why some portion still end up passing.

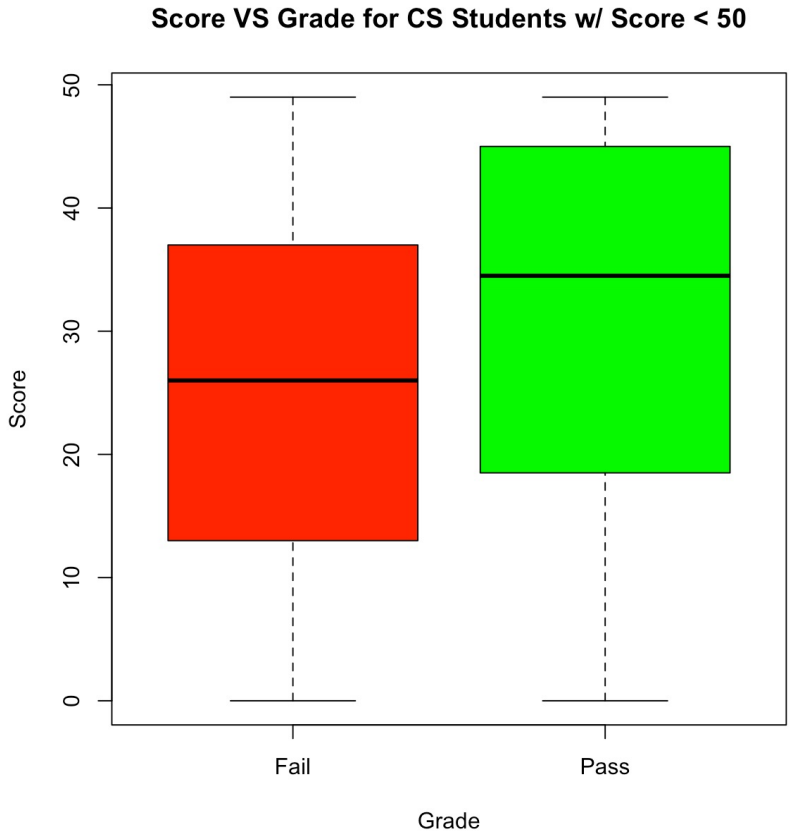
To further explore this data, I separated this data into two parts: computer science and the rest of majors.

- It's clear that computer science is it's own unique scenario.









Plots for CS Students w/ Score < 50 (Most Notable Plots)

Analyzing CS Students w/ Score < 50

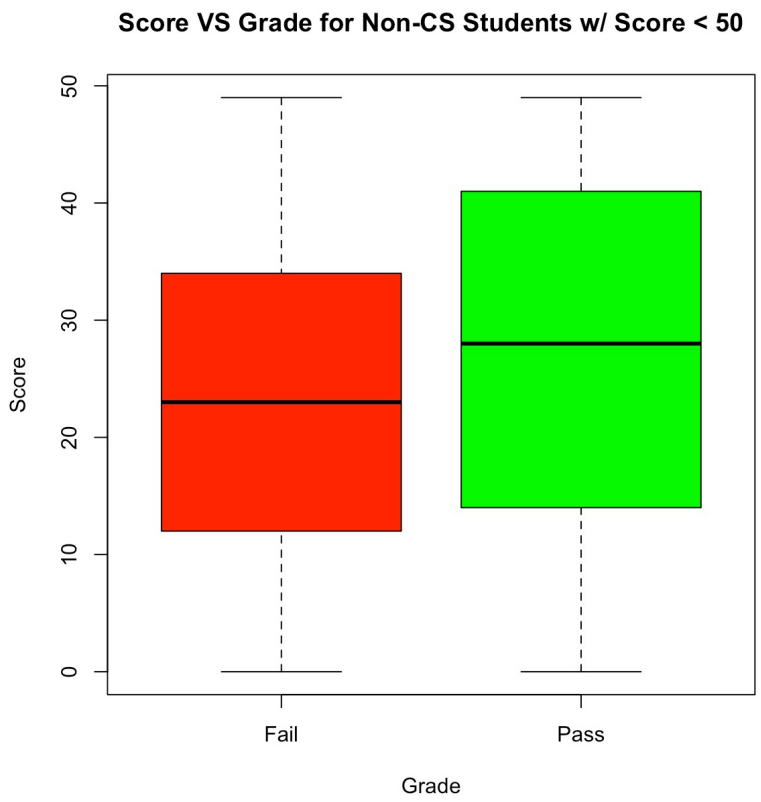
All Seniors and Sophomores that fit this category failed. Although this isn't the trend we'll exhibit all the time, this could potentially imply that seniors and sophomores may have to work harder on the other predictor values.

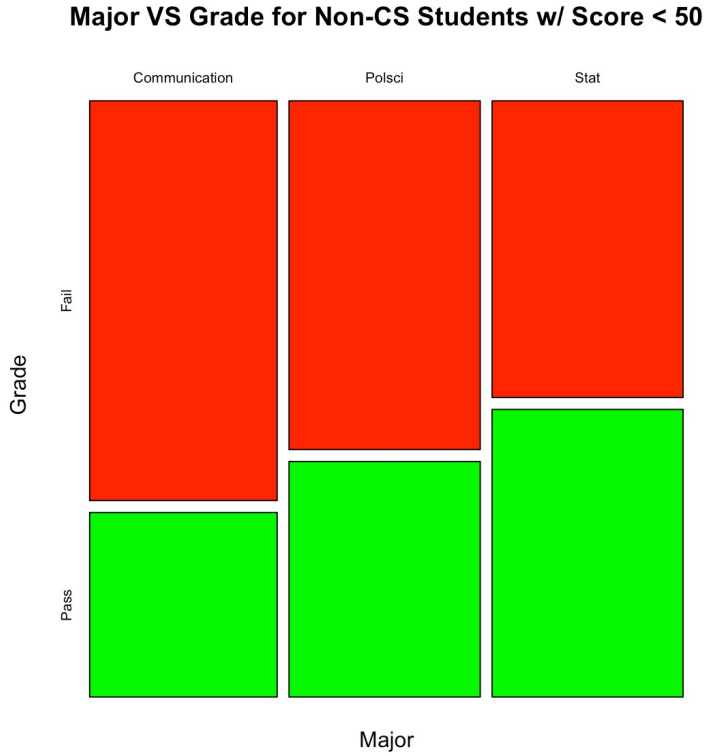
- Hopefully, we can use this trend with other majors down the line. We'll analyze that later if that's the case.

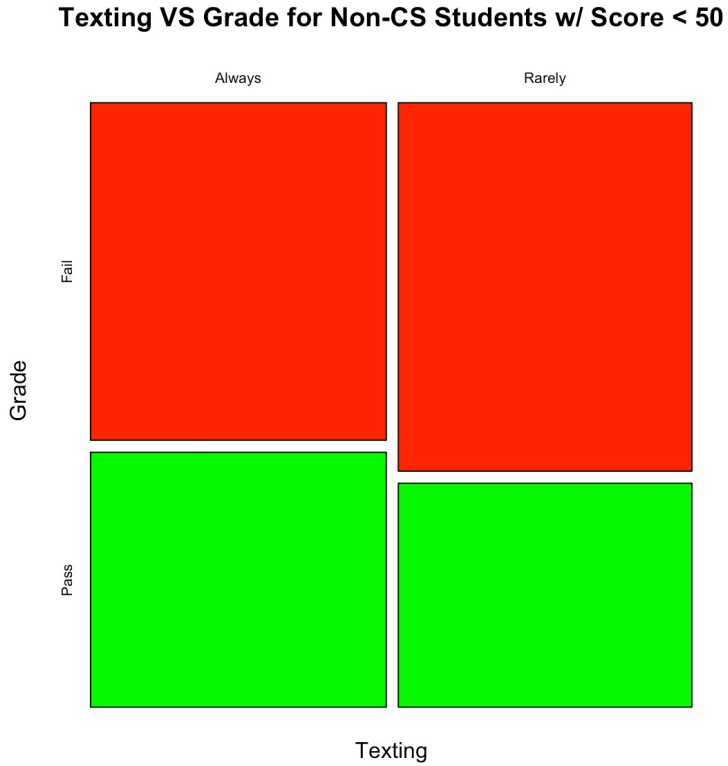
CS students that passed has higher numerical scores (median approximately 38), as the trend suggests, and also had a higher attendance score.

- Hopefully, we can use the trend of higher attendance (>60 as the boxplot median shows) leads to a higher likelihood of passing on other majors and other subsets.

Overall, to make up for the fact that you have a low grade as a CS student, try attending class a lot more, and if you're lucky, be a junior or freshman.







Plots for Non-CS Students w/ Score < 50 (Most Notable Plots)

Analyzing Non-CS Students w/ Score < 50

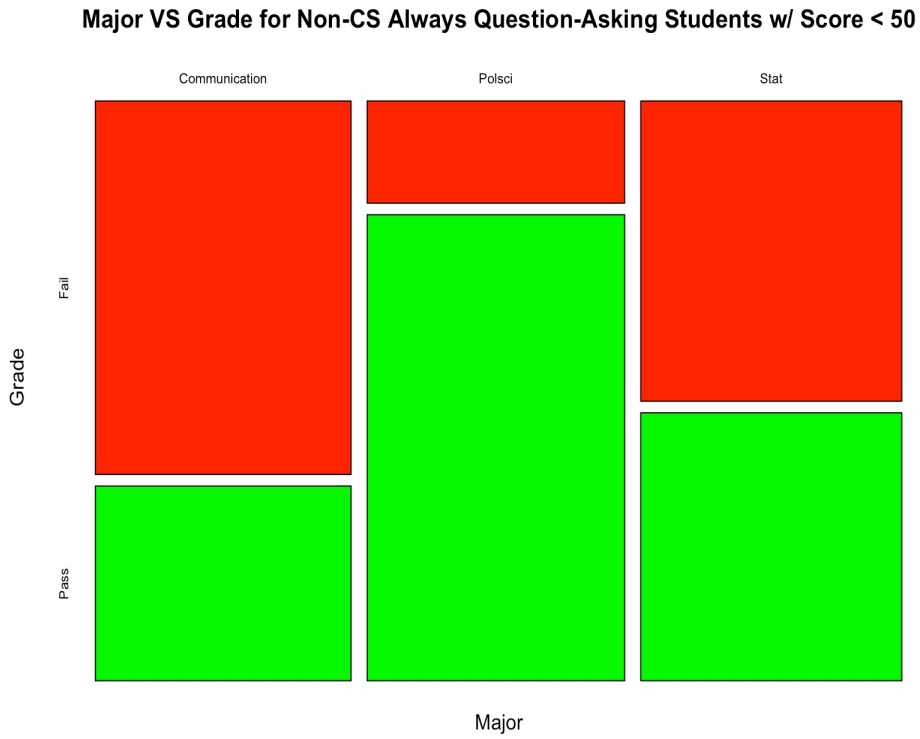
There was very little to interpret from these plots; however, the the reason why many people were passing and many people were failing wasn't being illustrated.

Degree looked like it played a role given how different the percentages looked visually within the mosaic plot.

The numerical score, like previous examples, shows the the median of passing students was higher than the median of failing of students; however, the gap isn't as so clear as previous examples.

I made an assumption that maybe the predictor values could be different in the context of of another potential predictor factor's value.

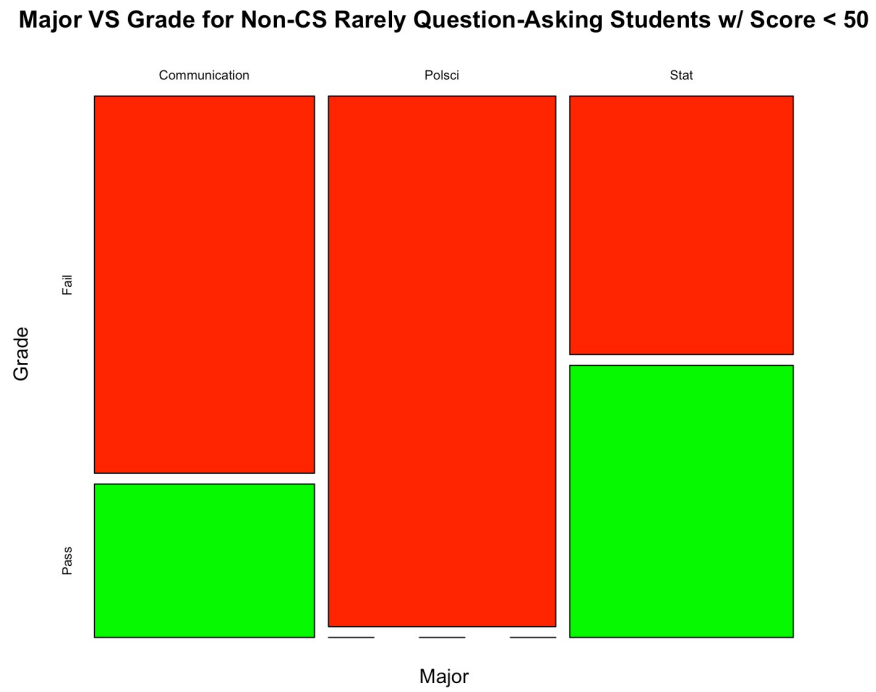
- Although the categorical factor of questions wasn't yielding any huge disparity, maybe the other values were more significantly different when separating from each other.



Political Science has a very high pass rate and likely the major itself will exhibit the same trend. Upon closer inspection of the data, it is likely that the people who failed in the other majors, such as communications and statistics, had different numerical scores required to pass the class.

PLOT(S) FOR NON-CS STUDENTS ALWAYS QUESTION-ASKING W/ SCORE < 50 (MOST NOTABLE PLOTS)

Major VS Grade for Non-CS Rarely Question-Asking Students w/ Score < 50



All Political Science Majors in this category failed.

- This would most likely have more to do with their major than it does with their numerical scores.

Statistics Majors have roughly the same amount of people passing and failing in this category.

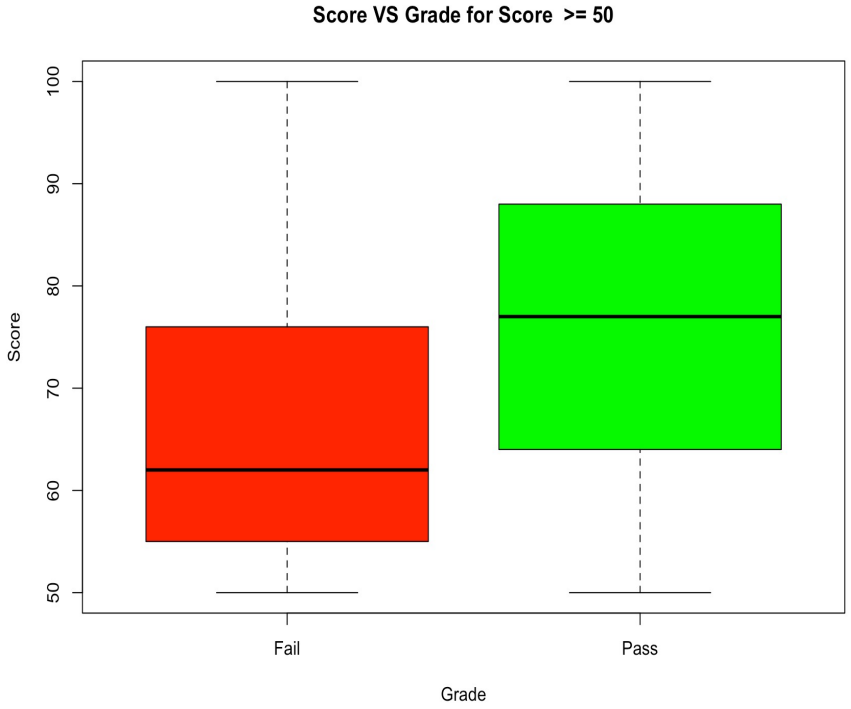
- We will test out medians and different numbers of numerical scores within this scenario to see what is the threshold for numerical scores.

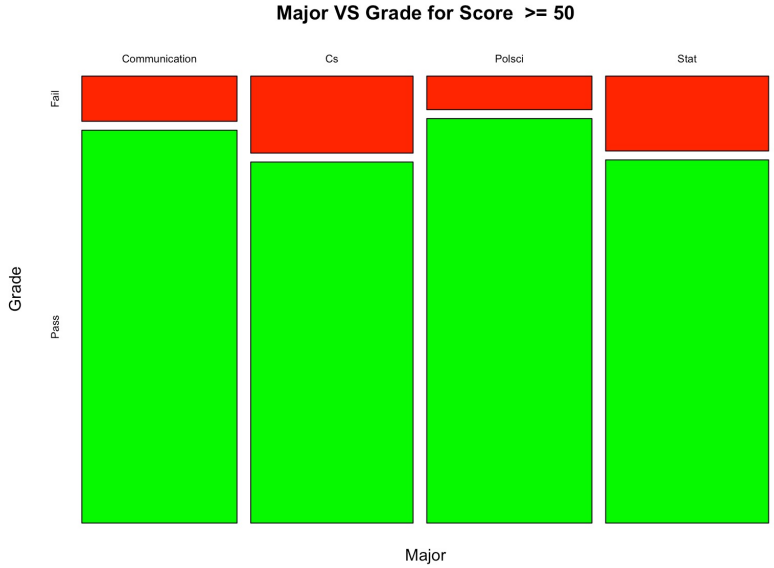
Communication Majors have more people failing than passing, but they have people passing nonetheless.

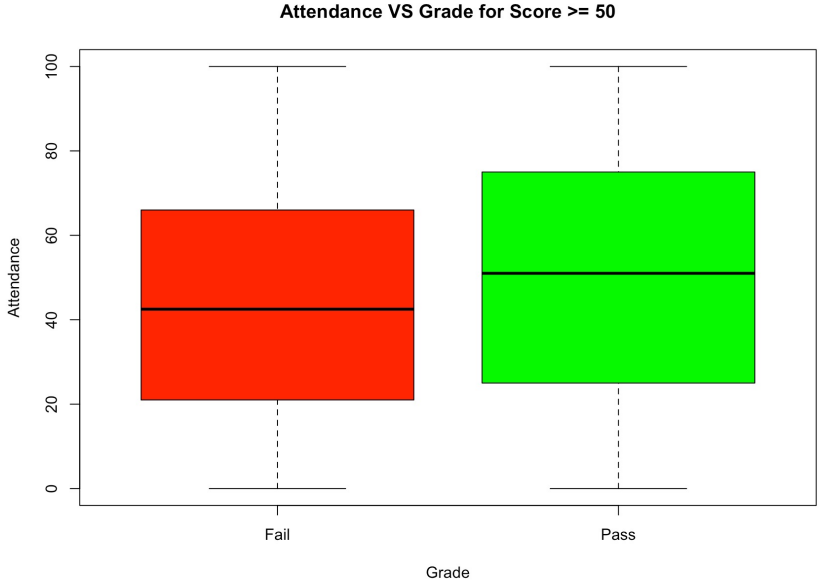
- Again, we will test out different numerical numbers to see what numerical value would be the best threshold for the decision.

Analyzing Students Who Have a Score Greater than or Equal to 50

Plots for Students w/ Score >= 50







Analysis of Students w/ Score >= 50

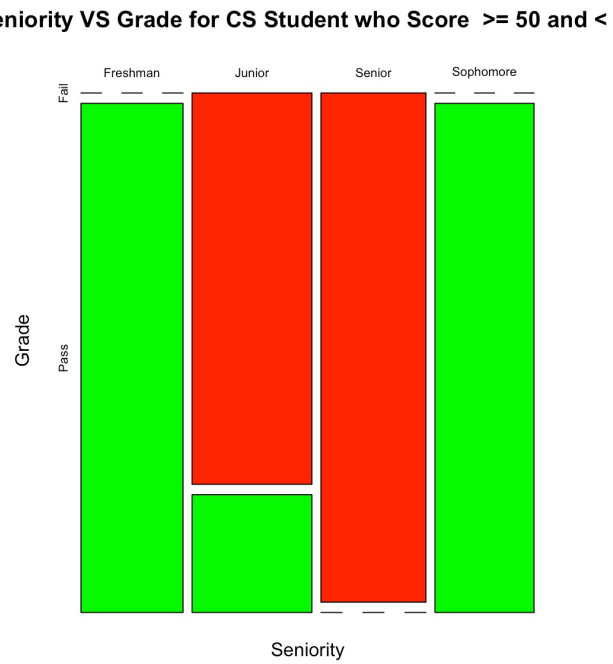
Attendance for students was higher for students that were passing; however, it wasn't high enough to warrant a Boolean expression.

Cs and Stats students had a higher failing rate than communication and polsci. Therefore, major could come into play

There was a huge score disparity between passing students and failing students. It was consistent with the other subsets we've seen throughout the data – the higher your score, the more chance you have of passing the class.

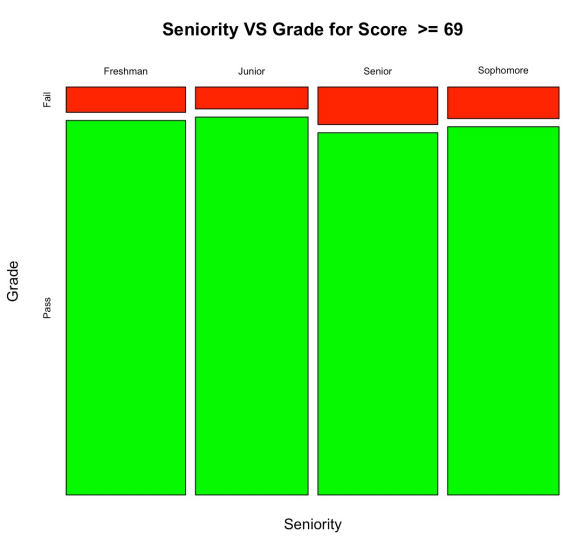
- Since this disparity gives us the most insight, I used split the data using score of 69 as the line. I used the approximate medians of the two boxplots from looking at it and averaged them to get an approximate.
  - (63+75)/2 = 69

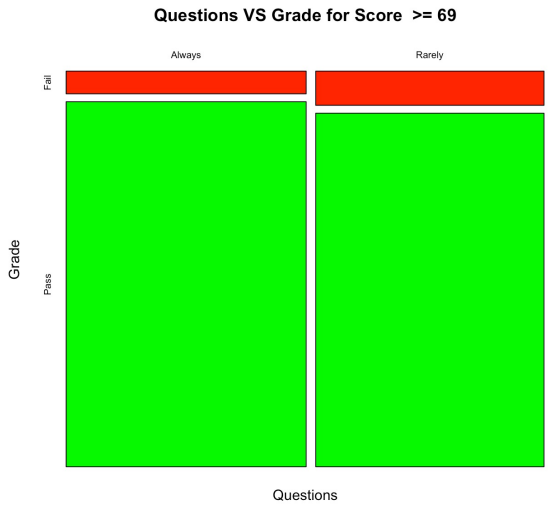
CS Students who Score >= 50 and < 69

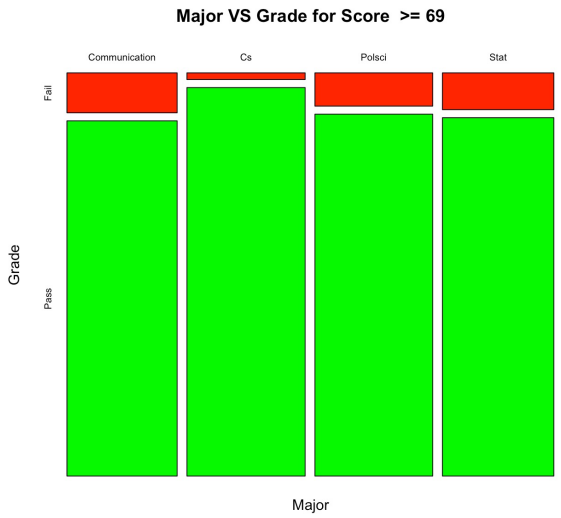


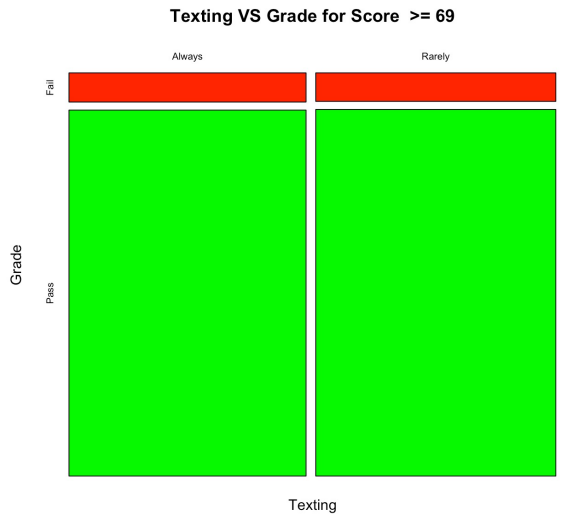
Freshman and Sophomores pass 100%. Seniors fail 100%. Most juniors fail. General trend in this case would be that seniors and juniors fail.

Plots of Students Who Score >= 69









The general trend here seems to be that if you have a score higher than 69, you have a really good chance that you are going to Pass