**Project Presentation Sample Answers/Discussion topics**

-Q1-Elaboration

**Why is the question of cohabitation of the students’ parents in the home important?**

The marital status of the various students' households may have a major effect on how the students’ performance in their classwork. Marital status, whether single or married, can be considered a negative or a positive influence on the student’s overall performance, along with their mental and emotional health. Since marital status can have such a strong negative or positive effect on the students’ grades and health (mentally/emotionally) it wise to take a look on the actual findings of the data and what that story can tell, whether it is confirming common assumptions about each living situation (i.e. 2 parent household= positive effect, single parent=negative effect), or explaining a trend that may also have relation to a separate grade influencing factor i.e. - 1 parent household=less income=no access to internet - may lead to poor grades.

What data was used to answer the question?

-Data needed was the total amount of students classified by their household parental status, and of those two categories (Apart and Together) the number of prior failed classes for those students \*(not sure if broke that down correctly)\*\*\*\* A great spot for Donterrius or Corban to speak on how they were able to make the charts.

-Using Pandas to merge datasets for Pstatus value counts and failures value counts

-Used to create pie charts to portray the percentage of failures per household classification

**Explaining what the data shows/Potential story being told**

Many would assume a single parent household may contribute to student failure due to potential hardships that may come with that cohabitation status, however the data reflects that this does not support this theory based on observation. Although there is a relatively small difference in percentage of passes and fails for the two groups compared in the visualization, the data shows that cohabitation of parents has a limited effect on student class failure.

There is a 1.13% larger rate of passing all of your classes as a student with parents who are apart vs students with parents who are together. This may not be enough of a significant difference to use in a comparison of the two groups, however it would not follow the commonly expected stereotype that single household students would be at a lesser advantage to do as good or even better than students in a 2 parent household.

Another interesting find from this visualization is that students from a single parent household have a 3.66 lower percentage of failing 1 class(8.26%) vs their counterparts from a 2 parent household (11.92%).

17.67% failing 1 or more classes (Household Together) vs. 16.53% failing 1 or more classes (Household Apart)

Very small difference in percentage in the two groups that could be considered insignificant upon observation.

-Q2-Elaboration

**Why/How would internet access have such a major effect on student success?**

In the age of fast paced technological advances, school systems are at the forefront of technological innovation. A student attending a school that requires some sort internet use at home to complete homework assignments,take exams, etc., may be left in a compromised situation that, most would assume, inevitably leads to poor grades due their inability to complete required assignments using an internet connection.

**What data was used to create this visualization and what did the data display from the visualization?**

From the data frame we were able to determine what information could be used to draw a true comparison of grade results for students with and without internet access from home. Using an average of the 3 grades reported for each student in the dataset, and matching the corresponding letter grade to each student’s average we were able to get the total values for each letter grade ranging from A-F. From there combining the total counts of those with/without internet and itemizing each letter grade to its total internet counts to create a dictionary that helped read the values into a double bar graph. The double bar graph was ultimately chosen to adequately and clearly display the percentage of those with or without internet for each letter grade. \*\*possibly use this point to refer back to when describing any issues with the data frame and how tricky it was to choose/create the proper display for this data\*\*\*\*

**-What trends/effects can be determined about student success with or without internet access ?**

The effect of internet access in regard to passing or failing is most likely not the driving factor for failure, the interpreted data visualization shows that despite having internet access, many students are still failing. Although the data shows that of the students who failed, a majority did not have internet access, there was still a very high number of those who failed with adequate internet access. Therefore internet access most likely does play a part in the lower grades; it is not the driving factor of failure. Based on the data there must be other factors playing a role in student failure.

\*\*Not too sure what else to elaborate on this point(did not take a lot of notes down when this analysis was being completed), feel free to add some additional information or thoughts that I know I’m missing!!\*\*\*\*

-Q3- Elaboration

**Does the time spent studying have a serious effect on the grades a student will receive?**

This question was important to determine if the actual length of a student’s study time will have a major effect on the grades received, or if like the previous questions there is no definite trend resulting from this factor. Usually most would assume that the more time spent studying should ultimately lead to a higher letter grade, however after creating a visualization of the data frame there may be a conclusion that is not as straightforward.

**What data was used to create the visualization for study time vs grade?**

-Used dictionary to itemize letter grades and number of hours spent studying \*\*Please help add to this explanation!\*\*

Similarly to the internet access visualization, a bar graph was used to best display a percentage distribution per x-axis category. In this visualization the distribution of study time ranging from <2 hours - >10 hours was shown for each letter grade (A-F).

To get a more in depth visual comparison of students who received a high grade “A” or a failing grade “F” vs the number of hours studied for each letter grade, a pie chart made for the two groups was a better fit. Ultimately this data set used 2 different visualization methods to adequately compare grades vs study time not just for a range of grades, but also for a specific comparison of study time for students with a high passing grade and those with a failing grade.

**Is there a noticeable trend or effect on grades received due to length of study time?**

When analyzing the data shown in the bar graph using the entire range of the letter grades (A-F) it is difficult to reach a real conclusion on how studytime affects letter grade, due to the unusual trend shown in the 2-5 hour study time mark. Approximately 50% of students in the C-F category were students who studied for 2-5 hours. Less than 40% of the C-F categories were made up of students who studied for less than 2 hours. On the contrary 30% of those who received an A studied for less than 2 hours, and only 17.65% of those in the A category studied for 2-5 hours. Most would assume that those who make a very high letter grade would have very lengthy study times, however in this data set that is not the exact case. A conclusion that could be drawn from this specific data set could potentially be that students with higher letter grades may be getting extra help for the class in ways that would not technically be considered independent study time such as taking additional courses in relation to that subject.

**Cleanup, Exploration, and Analysis**

\*\*\*I believe those who created and provided the graphs would have the best points for this section!\*\*\*

**Discussion**

Comparing each groups’ results: study time vs grades, home life (parents being together), Effect of internet access on grades

Comparing the visualizations presented to determine which factor may have had the most influence on the students’ grades.

1. While each data set stands by themselves in their own respective manner, using the 3 factors that we looked at, the factor with the least noticeable trend would be the effect of family life on the student as it displays that most students will pass all of their classes regardless of a single or 2 parent household. There is not enough of a significant percentage difference between students excelling in their school work in a single or 2 parent household.
   1. Family life did provide another point of view in showing that students failing 2 classes had a greater percentage coming from a single parent household. However more students failed >2 classes in a 2 parent household vs a single parent household, not following the expected trend yet again.
   2. This confirms that contrary to a common assumption a 2 parent household does not guarantee the student will always excel greater than a student coming from a single parent household.
2. \*\*Possible comparison/connection between Internet access and home life?

\*\*\*Feel free to correct anything or add more comparisons\*\*\*\*\*

From the analysis of our findings it is safe to assume that the trends shown could be considered unusual/unexpected. For instance students with internet access received failing grades at a rate that was comparable to those without any internet access, or the observation that almost half of the students who received a failing grade were studying for multiple hours, unlike the A category of students. Most of the data would not fit the commonly expected trend for the situations being analyzed, but had a slightly different and interesting outcome. The inferences made from these data sets can definitely lead to further exploration and comparisons of the topics presented from the data set.

\*\*\*These are just quick comments that I thought of please add more if necessary\*\*\*

**Any difficulties that arose?**

-Picking a visualization for the data being analyzed

-Choosing the best data sets to analyze/compare

-\*\*\*\*I know there were more so please add more here!

**CLASS QUESTIONS/END**

\*\*\*I hope that this is helpful, I know my interpretations may not be spot on, but I wanted to get some ideas down to refer to or definitely build off of! You guys are literally experts(in my opinion) so please feel free to correct or add to anything!