My plan is to go through weekly ipnyb assignments and apply what I learned to the project. I have a couple of bookmarked examples if needed. Item 15 is what my project is.

For my project, I plan on using the following dataset - <https://www.kaggle.com/janiobachmann/math-students>. This dataset is called student-mat.csv and contains the final math scores of students in secondary school. This dataset was interesting to me because it has a lot of different variables that can be tested to see whether it is apparent that they are impacting the final grade.

Some of the variables that I am most interested in are sex, family size, alcohol consumption, whether they are in a relationship, family support. Some of my primary questions will be:

1. How does the increase in weekday alcohol consumption impact grades or does it?
2. Does the level of parental education impact grades?
3. Does having a romantic relationship impact grades?
4. Is there a difference in impact between weekday and weekend alcohol consumption?
5. Does family size have an impact on grades at all?
6. Does gender play a factor in grades?
7. How does the number of school absences impact grades?

My initial assumptions are that alcohol use has a large impact on grades. I also think the number of absences will have a large impact. I also think parental education will have a significant impact on grades as well. I am less confident that family size or romantic relationship will have much of an impact, but I am open to surprises.

I feel like alcohols impact on grades could be a project in its own but would probably require additional datasets and I really like the number of variables for this one so I think exploring some of these impacts will give a good picture of what external factors play a part in determining math grade performance.

Reference

P. Cortez and A. Silva. Using Data Mining to Predict Secondary School Student Performance. In A. Brito and J. Teixeira Eds., Proceedings of 5th FUture BUsiness TEChnology Conference (FUBUTEC 2008) pp. 5-12, Porto, Portugal, April, 2008, EUROSIS, ISBN 978-9077381-39-7.