Statistics 110 and 201 Midterm Review

The midterm exam will cover Chapters 1, 2, 3, 5, 6 and Sections 7.1 to 7.3, excluding sections explicitly omitted on the daily list on the class webpage. (Sections covered and omitted are listed with the homework assignments.) There are certain topics in the covered sections that we skipped, and you are not responsible for those. The exam is open book and notes and you should bring a basic calculator.

Here is what you should know how to do and/or interpret:

1. Write the population model in algebraic and matrix forms, including assumptions.
2. Interpret each term in the population model (coefficients, error, population standard deviation). Be able to do this for multiple regression as well as simple linear regression, and for models that include indicator variables.
3. Write the sample equation in notation and numerically (once you have computer output).
4. Use the sample equation to find a predicted value and a residual for an individual in the sample.
5. Interpret everything in the computer output given by Stata following the regress command.
6. Test individual coefficients (using t*), get confidence intervals for them, and interpret the results.
7. Test the overall regression equation (using F*) and know what is being tested.
8. Understand when it can be concluded that a change in X *causes* a change in Y, and when it can only be concluded that there is a relationship (experiments versus observational studies; also see page 8 of the textbook).
9. Construct an ANOVA table, including how the pieces fit together (e.g. SSR + SSE = SSTO, and so on).
10. Understand how to estimate and predict for a new case with values $\mathbf{X_h}$, including how to interpret $\hat{Y}_h$, a confidence interval for $E\{Y_h\}$ and a prediction interval for $Y_{h(new)}$.
11. Find and interpret R^2 and know why we would sometimes use R^2-adjusted instead.
12. Understand how to check for outliers, lack of normality and non-constant variance, including what plots should be examined.
13. Understand what to do about outliers.
14. Understand how transformations can be used to correct deficiencies in the assumptions, and what to do after using a transformation to get results in the original units.
15. Understand how to use the general linear test for full and reduced models, including the concept of extra sums of squares, and how to construct the test from the extra sums of squares. Also be able to state the null and alternative hypotheses being tested in any such test, and give the appropriate numerator and denominator degrees of freedom.
16. Understand the idea behind the F test for "lack of fit."
17. Work with basic matrix manipulation.
18. Work with random vectors and matrices. For example, know that $\mathbf{\sigma^2\{AY\} = A\sigma^2\{Y\}A'}$, know that the variance-covariance matrix of a random vector has the variances on the diagonal and the covariances as the other elements, and so on.
19. Understand the concept of least squares estimation.
20. Understand why the Bonferroni method is used, and how to use it.
21. Understand why you shouldn't extrapolate after getting a regression equation.
22. Understand the circumstances in which regression is useful – for description, control and prediction. See page 8 of the textbook.