**C105 final EXAM Study Guide**

That test can't drop me. They say it's an IQ test. I came to the combine for football.   
I looked at the test, and wasn't any questions about football. I didn't see no   
point in the test. I'm not in school anymore. I didn't complete it.   
I only finished 15 or 18 questions.



*Morris Claiborne   
Cowboys Cornerback on the NFL’s Wunderlic Test, April 27, 2012*

### final exam details

Date: Thursday, December 18  
Time: 5-7pm  
Location: Ballantine Hall 228  
Format: 50 multiple choice questions (please bring a pencil)  
Content: The final exam is *not* cumulative. It will only include topics outlined below.

### Week 9: Human decision Making

Readings: Blink\_TheSecretLifeOfSnapDecisions / Scorecasting\_GoForIt

* What are “normative” and “descriptive” statements? Does human behavior vary from the normative account in systematic or random ways?
* How does classical decision theory (or more specifically, “expected utility theory”) suggest we evaluate decision making alternatives? How is this different from the heuristics that we usually employ?
* How does the way risk is “framed” affect decision making? Do people make decisions on the basis of \*actual\* risk or their \*perceptions\* of risk?
* Understand the reasons that Kevin Kelley’s high school football team never punts. –and if it’s so much better to always go for it on 4th down, why don’t more coaches go for it? Can you explain the irrational reliance on punting in the context of the confirmation bias? -- the representativeness heuristic?

### WeekS 10 & 11: Correlation and Regression

No readings—review the PowerPoints for days 14, 15, and 16

* You should understand how correlation can be used to describe relationships between variables, but not causal connections between variables.
* What information is in a correlation coefficient (namely, “r”)? If shown scatterplots of data, be able to identify an approximate & appropriate value for r. What is the coefficient of determination, and how is it different from r? What does it tell you about the relationship between two variables?
* What is measured by a p-value? How small should the p-value be for most people to accept that a result (in this case, a correlation) is reliably different from 0?
* What is regression analysis for, and how is it different from correlation?
* Understand the equation for a regression line, including the intercept and slope/coefficient. Be able to identify the correct line given an equation, and given a line, be able to identify the equation that correctly defines the line.
* What does it mean for the best-fitting regression line to minimize squared error?
* If we do a linear regression analysis with multiple predictor variables, how would this change the equation for the regression line? Be able to compare coefficients and p-values between predictor variables.

### Week 12: Factor Analysis

Readings: DiscoveringStatsUsingSPSS\_FactorAnalysis

* What the heck is factor analysis? What is a latent factor?
* What is a correlation matrix? If shown a correlation matrix, be able to make reasonable guesses about what might be found after running factor analysis.
* Understand, and be able to distinguish between the following: specific variance, error variance, and common variance. If a measure has low common variance (low communality), will it be useful in a factor analysis? Why not?
* What’s an eigenvalue? What’s a scree plot? How can you use these to determine the number of factors to include in your factor analysis?
* Be familiar with factor loadings, and how these can help you interpret the results of a factor analysis.

### Week 14: Cluster Analysis

Reading: ClusterAnalysis\_IntroToClassificationAndClustering

* What’s the difference between cluster analysis and factor analysis?
* How does k-means clustering work? What does it require as input, how does it initialize, and what are the specific steps that it takes to assign observations into clusters? What is measured by the “distance to cluster center”?
* Understand some situations when cluster analysis might not work well on a dataset.

### Week 15: Simulation Methods

Readings: HowDoesAccuscoreWork / MonteCarloMethod (Wikipedia page)

* In Week 15 we created a [more-or-less] optimal lineup for an auction league. How’d we do that? -- or better yet, how did *Excel* maximize value while keeping cost beneath the salary cap?
* How (generally speaking) does Accuscore work? --and how is this simulation approach different from the regression models that we created in the first half of the course? Why wouldn’t you always get the same basic result with simulation as with modeling? (hint: simulation works best when inputs are correlated with each other)
* The output of a simulation analysis is a frequency distribution; be able to interpret one of these. What are the horizontal and vertical axes?

### Week 16: Skill and Gambling

Readings: FantasyFootballGamblingOrSkill

* The author of this week’s reading writes “In general, the rule for determining whether an activity is gambling rests on the answers to two questions…” What are these questions? Understand the reasons that “betting” on your team’s fantasy success is legal.