Exam 1 list of concepts

Below you will find a skeleton outline of what we have covered so far. Exam 1 will finish with the lecture from 10/24. For the exam you will not need to manually calculate critical values, but need to understand the underlying logic behind why they vary with sample size. For the exam you can have either a half sheet of paper (large notecard) front and back **or** a single sided full sheet of paper. I advise bringing a calculator to the exam, but you will not be required to produce any R code (you might be asked to interpret it or tell me what went wrong in the code).

1. Unit 1- What is a model
   1. Model overview
      1. What I mean when I say “all models are wrong, but some are better than others”
      2. Role of uncertainty in modeling
   2. Model basics
      1. Independent and dependent variable
      2. Role of error in our model
         1. Residuals and their role in making a regression line
      3. Using regression equation to predict values of DV
   3. Statistics vs parameters
      1. Inference
   4. Data
      1. Unit of analysis, structure of data like rows and columns
      2. Types of variables
         1. Quantitative, categorical (nominal v ordinal) and binary
2. Simple linear regression
   1. Understand what each part of following formula means: y=+X+ε
   2. Logic of least squared regression
      1. Sum of squared errors (what it represents, why we square it, importance in regression, etc).
   3. Understand intercept and slope
      1. What they represent in a line, how they are interpreted.
3. Conditions for linear model (know conceptually what they mean and how to identify)
   1. Understand how to assess conditions (where applicable)
      1. Linearity
      2. Normally distributed residuals with zero mean
      3. Independence
      4. Uniform spread/constant variance/homoskedasticity
      5. Randomness
   2. Standard error of the regression line
      1. Differences between SE of regression line and coefficient SE
   3. Data transformation
      1. Two classes – interpretability vs transformation due to non-linearity
         1. Interpretability
            1. Centered and standardized

Know how to interpret/explain both transformation and model results

* + - 1. Log transformation
         1. Understand when to use
         2. Understand how to interpret using methods we learned in class (when IV/DV, or both transformed)
  1. Outliers and why we care about them
     1. How to address them in regression
  2. Dangers of extrapolation

1. Inference with regression
   1. Logic of central limit theorem
      1. How it relates to setting up null and alternative hypotheses
      2. Role of t distribution
   2. 3 considerations for inference
      1. Test for statistical significance
      2. Test for intervals
      3. Effect size (statistical vs practical/substantive significance)
   3. T-test for slope coefficient
      1. Critical value
      2. How to convert to confidence interval
      3. Hypothesis testing with and without p-value
      4. P-value definition and interpretation
   4. F-statistic
      1. Note- just need to understand what the null/alternative hypothesis is and what a p<.05 means
   5. R-squared and adjusted r-squared
      1. How to interpret and what adjustment means
   6. Prediction interval vs confidence interval
2. Multiple regression
   1. Simpsons paradox
   2. Interpretation of parallel slopes model
   3. Interpretation of interaction model
   4. Polynomial regression (interacting variable with itself)
3. Regression output
   1. Understand what every piece of regression output refers too
      1. Residuals
      2. Intercept and slope coefficient, SE, t-value, and p-value
      3. Se of regression line
      4. R-squared and adjusted R-squared
         1. Calculation of R squared
      5. F-test
      6. Using critical values for confidence intervals and hypothesis testing