# Final Review Sheet

## Linear Regression

- Four assumptions of linear regression model
- Model a linear relationship, not a causal effect
- Exploratory analysis with scatter plot (visual) or Pearson/ Spearman correlation (quantitative)
- Multicollinearity issue
    - Detect multicollinearity among predictors through scatter plot matrix or VIF
    - What happens if highly correlated predictors are in the model? Check simulation studies in Week10 materials
- Model selection
    - Automatic selection (forward, backward or stepwise selection)
    - Best subset approach
        - ✓ Balance between good prediction and simple model
          e.g., Adjusted $R^2$, AIC, BIC, Mallow's Cp
- Interpretation of R output
- Interpretation of $R^2$
- Determine significance of parameters and interpret the model
- Specify estimated regression line based on estimated parameters
    - When all predictors are continuous
    - When model includes categorical predictor
    - When interaction between categorical and continuous predictors is included
- Assess and remedy diagnostic issues. Guess which assumption is violated based on diagnostics
    - residual plots
    - histogram and QQ-plot
    - Cook's distance
- Transformation of Y


## Logistic Regression (binary logit)

- Response variable is binary
- Important to set an event of interest (decide what is our event of interest)
- Goal of logistic regression
- Odds and Odd Ratio
- Choose the best model (model selection)
- Specify estimated logistic regression line based on estimated parameters
- Assess diagnostic issues (e.g. residual plots, cook's-d measure)
- Determine significance of parameters and interpret the model
- Interpretation of estimated coefficients with **Odds Ratio** (specifically)
- Guess what kind of observation would have **the highest chance to have an event**
- Howser-Lemeshow test (Goodness of fit test)
- Classification tool based on estimated probabilities
    - How we can classify each observation based on given cut-off
    - Example of cut-off's: 0.5, sample proportion..