Stat 340 - Potential Quiz 2 Topics

Some things I might ask you to do on a quiz about regression:

Interpret linear model coefficients (main effects and interaction effects, categorical and quantitative variables)

- Main effects for quantitative variables (assuming no interactions or quadratic terms involving that variable in the model): Increasing x by one unit while holding all other explanatory variables fixed results in an increase/decrease in predicted y of (coefficient value) units.
- Main effects for categorical variables (assuming no interactions involving that variable in the model): the difference in predicted value for an observational unit in the baseline category and the category corresponding to this coefficient is (coefficient value) units, holding all other explanatory variables in the model fixed.
- Interactions between categorical and quantitative variables: The difference in slopes in the quantitative variable between the baseline category and the category corresponding to this coefficient is (coefficient value).
- Be able to write down a full estimated model equation, special cases obtained by fixing the categorical variable value (e.g., MLAA rating category G or MPAA rating category PG-13), and if there's only one quantitative variable, draw the two corresponding lines.

Diagnose model issues from a residuals plot, and recommend possible strategies for addressing these problems

- Identify non-linear trends in plots of residuals vs. fitted or residuals vs. explanatory variables: consider adding a polynomial term
- Identify a lack of normality in a density plot or histogram of residuals: consider a transformation, starting with the response
- Identify heteroskedasticity/unequal variance of residuals in plots of residuals vs. fitted or residuals vs. explanatory variables: consider a transformaiton, starting with the response

Conduct appropriate t and F tests using R output from `summary` and `anova`

- t tests for a single coefficient (from `summary` output)
- F tests for multiple coefficients - either for all coefficients in the model (from `summary` output) or for a subset of coefficients in the model (from `anova` output)