NHANES in-class assignment

Using the plotting and modeling skills you have learned over the past few weeks, you will explore the `NHANES` data set. You can answer these questions on this sheet or in your code, which I would like for you to email to me at `slauer@schoolph.umass.edu`. Start by loading and checking out the data:

```
install.packages("NHANES")
library(NHANES)
data("NHANES")
?NHANES
View(NHANES)
```

1. Choose a health-related outcome (e.g. `Diabetes`, `SleepTrouble`, `DaysMentHlthBad`, etc.). You will either try to find an association with another covariate or try to predict this outcome directly. Think about what covariates you need to account for. If you are looking for an association, what covariates could confound that association? If you are doing prediction, which predictors make sense?

2. Clean up your data as necessary, removing NA values. Plot your data using `ggplot2`. Do the data look like you thought it would?

3. Split your data into training and testing sets (see last week's code to see how). Fit a few models of varying complexity (i.e. a null model, a single term model, and a full model) on your training data. How well do each of these models describe the data? Remember to use the correct model type (logistic for binary data, normal for continuous data).

4. Choose a couple of models to predict the testing data. Is there an association between your covariate of interest and your outcome? How well does your best model predict the data? How confident would you be in this model going forward? How would you improve this model if you had more, different data (i.e. what other covariates do you think would help you predict your outcome)?