This dataset is a part of the longitudinal dataset collected by the General Hospital of the city of St. John’s, Canada for six years from 1985 to 1990.  The outcome is the count variable: Visit, which is the number of visits to a physician by each individual during a given year. Covariates include

* Year (1-6)
* Gender: (1=male, 2 =female)
* Chronic:  Chronic disease status (0=no chronic disease, 1=one chronic disease and so on), treated as continuous
* Education: Education level (1 = less than high school, 2 = high school, 3 =university graduate, 4 = post graduate), treated as nominal
* Age\_base: Age at 1985

In addition, Family ID is family identification and ID is the family member identification. The data has been provided on the blackboard. Refer to Lecture 12 and answer the following questions.

**Questions:**

1. Draw a histogram for the count data of Visit. Does the plot appear to be right-skewed?
2. Use the model building strategies to build a model for these data. In the intercept-only model at the first step, if the variance at the highest level is not significant, then remove the random statement for the highest level so there is zero variance at the highest level. In the meantime, the random statement for lower levels must consider the nested relationship between units at lower levels and the highest level.  In addition, if there is overdispersion, a scale parameter must be added to the model by using the random \_residual\_ statement.
3. From the final model, use the type 3 analysis to find significant effects and interpret them.