**grade:** The grade in school of the student (most 15-year-olds in America are in 10th grade)

**male:** Whether the student is male (1/0)

**raceeth:** The race/ethnicity composite of the student

**preschool:** Whether the student attended preschool (1/0)

**expectBachelors:** Whether the student expects to obtain a bachelor's degree (1/0)

**motherHS:** Whether the student's mother completed high school (1/0)

**motherBachelors:** Whether the student's mother obtained a bachelor's degree (1/0)

**motherWork:** Whether the student's mother has part-time or full-time work (1/0)

**fatherHS:** Whether the student's father completed high school (1/0)

**fatherBachelors:** Whether the student's father obtained a bachelor's degree (1/0)

**fatherWork:** Whether the student's father has part-time or full-time work (1/0)

**selfBornUS:** Whether the student was born in the United States of America (1/0)

**motherBornUS:** Whether the student's mother was born in the United States of America (1/0)

**fatherBornUS:** Whether the student's father was born in the United States of America (1/0)

**englishAtHome:** Whether the student speaks English at home (1/0)

**computerForSchoolwork:** Whether the student has access to a computer for schoolwork (1/0)

**read30MinsADay:** Whether the student reads for pleasure for 30 minutes/day (1/0)

**minutesPerWeekEnglish:** The number of minutes per week the student spend in English class

**studentsInEnglish:** The number of students in this student's English class at school

**schoolHasLibrary:** Whether this student's school has a library (1/0)

**publicSchool:** Whether this student attends a public school (1/0)

**urban:** Whether this student's school is in an urban area (1/0)

**schoolSize:** The number of students in this student's school

**readingScore:** The student's reading score, on a 1000-point scale

The goal is to build a linear regression to predict readingScore based on one or more of the other features. Train the model using the training data and observe test accuracy on the test data.

1. Explore the relationships between different variables and readingScore. Use appropriate plots to do this.

2. Build one or multiple regression models to predict readingScore. You might build multiple models that uses one independent variable to predict the response variable or use combinations of variables to create a multivariate regression.

Write a report with your findings about the relationships between different independent variables and the response variable. Also report the regression results and how you picked the best model (and variables) to predict readingScore.