1. Prices of diamonds are determined by what is known as the 4 Cs: cut, clarity, color, and carat weight. The prices of diamonds go up as the carat weight increases, but the increase is not smooth. For example, the difference between the size of a 0.99 carat diamond and a 1 carat diamond is undetectable to the naked human eye, but the price of a 1 carat diamond tends to be much higher than the price of a 0.99 diamond. In this question we use two random samples of diamonds, 0.99 carats and 1 carat, each sample of size 23, and compare the average prices (standardized by weight) of the diamonds. The data are provided in the coding practice associated with today's class. The variable `carats_99` represents the sampled prices of the 0.99 carat diamonds, and the variable `carats_1` represents the sampled prices of the 1 carat diamonds.

   (a) (Coding practice) Conduct a hypothesis test via a mathematical model to evaluate if there is a difference between the average standardized prices of 0.99 and 1 carat diamonds. Make sure to state your hypotheses, check conditions (coding practice), and interpret your results in context of the data.

   (b) Describe in words how you would obtain the null distribution via simulation for your hypotheses from (a). Then write code to conduct this hypothesis test (can use the same coding practice `.Rmd`). How does your result here compare to (a)?

2. A market researcher wants to evaluate average car insurance savings at a competing company. Based on past studies he is assuming that the standard deviation of savings is \$100. He wants to collect data such that he can get a margin of error of no more than \$10 at a 96% confidence level. How large of a sample should he collect?

3. In July 2008 the US National Institutes of Health announced that it was stopping a clinical study early because of unexpected results. The study population consisted of HIV-infected women in sub-Saharan Africa who had been given single dose Nevaripine (a treatment for HIV) while giving birth, to prevent transmission of HIV to the infant. The study was a randomized comparison of continued treatment of a woman (after successful childbirth) with Nevaripine vs Lopinavir, a second drug used to treat HIV. 240 women participated in the study; 120 were randomized to each of the two treatments. Twenty- four weeks after starting the study treatment, each woman was tested to determine if the HIV infection was becoming worse (an outcome called virologic failure). Twenty-six of the 120 women treated with Nevaripine experienced virologic failure, while 10 of the 120 women treated with the other drug experienced virologic failure.

   (a) Create a two-way table presenting the results of this study.

   (b) State appropriate hypotheses to test for difference in virologic failure rates between treatment groups.

   (c) Complete the hypothesis test at the $\alpha = 0.05$ level and state an appropriate conclusion. (Reminder: Verify any necessary conditions for the test.)

   (d) Obtain a 90% confidence interval for the difference in failure rates between treatment groups.