HW11, Your Name

Problem #1

This question deals with our beloved Airbnb data.

a. Conduct a t-test on finding out whether the average price of NYC listings significantly differs from 150$ per night. What population we are trying to infer about when conducting this test? Make sure to formulate the hypotheses (in parameter notation), report both the p-value and (Cohen's) effect size. Interpret the result. Does the confidence interval agree with hypothesis test?

b. Is the normality assumption satisfied? Are there are any extreme outliers? Given that, should we be concerned about the legitimacy of the conducted t-test?

c. (**Similar to what we did with** *TVhours* **data set in lecture code**). Piggy-backing off part (b), proceed to pretend as if your data set is the "entire population" of NYC Airbnb listings (of size 48,864). Repeat the following process 10,000 times:

- Randomly sample 30 listings from that "population".
- Conduct the t-test based on this sample of whether population mean price differs from $mean(listings\$price)$ (which is the true mean of our "population").
- Record the p-value.

In the end, calculate the % of times your p-value was less than 0.05. Is it what you expected with significance level $\alpha = 0.05$? Why/Why not?

Problem #2

This question deals with data on working hours per week.

a. Proceed to check the normality assumption for weekly working hours of 1) males and 2) females. Does it look satisfied? If not - do you believe that t-test results on comparing males and females with respect to weekly working hours might be suspect? Why/why not?

b. Conduct a two-sample t-test comparing males and females with respect to weekly working hours. Make sure to formulate the hypotheses (in parameter notation), report both the p-value and (Cohen's) effect size. Interpret the result (along with confidence interval). Does the confidence interval agree with hypothesis test? Explain.

Problem #3

9.56

10.14

10.24

10.49, 10.58 (please use *R* **here, along with** $t.test()$ **function to find confidence intervals)**