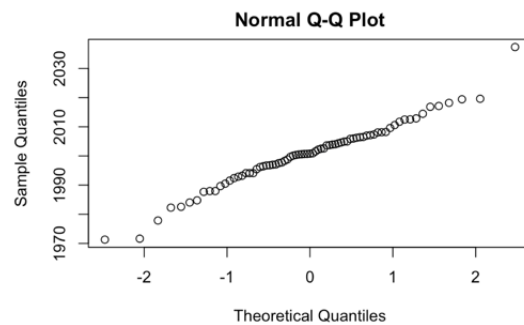
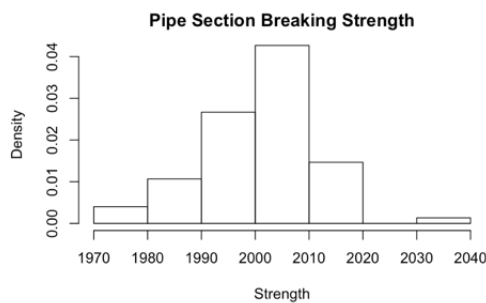


# Discussion 8 Soln

- Specifications for a water pipe call for a mean breaking strength  $\mu$  of more than 2000 lb per linear foot. Engineers will perform a test to decide whether or not to use a certain kind of pipe. A random sample of 1 ft sections of pipe is selected and their breaking strengths are measured. The pipe will not be used unless the engineers can conclude (statistically, not with certainty) that the mean breaking strength is greater than 2000.

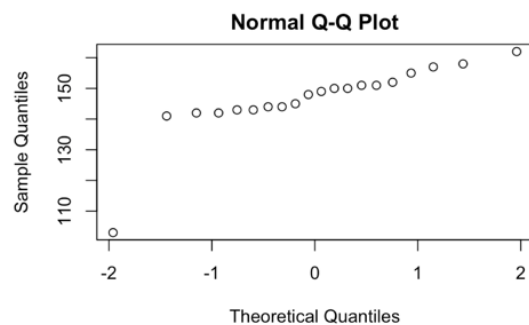
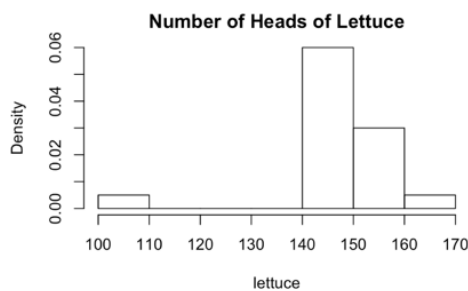
- Specify appropriate null and alternative hypotheses for this situation.
- Based on last week's analysis, the engineers chose to obtain a sample of 75 random 1 foot pipe sections. The qqplot of the sample data and summary statistics from the sample are given below. Perform a one sample t-hypothesis test at the 5% level after checking that the assumptions for testing are well met and interpret the results in context.



Pipe Strength Sample Mean: 2001.98; Pipe Strength Sample Standard Deviation: 11.281

- Another scientist in the lab suggests instead of a t test, a z test could be performed. Explain why either a t or z test will give nearly equivalent p values and conclusions in this case.
- A crop scientist evaluating lettuce yields plants 20 plots, treats them with a new fertilizer, lets the lettuce grow, and then measures yield in numbers of heads per plot, with these results:  
145, 142, 144, 141, 142, 155, 143, 157, 152, 143, 103, 151, 150, 148, 150, 162, 149, 158, 144, 151  
The old fertilizer led to an average yield of 145 heads per plot. Test whether the new fertilizer leads to an improved yield via the following steps.

- A histogram and QQ plot of the data is given below. Is the assumption of Normal population appropriate? Will the CLT save us? Why or Why not? Which test should we use?



- (b) Run a bootstrap test with `set.seed(1)`.
    - i. What assumptions are you making?
    - ii. Which hypotheses are you testing?
    - iii. What is your observed test statistic?
    - iv. Find a p-value with the help of 6000 bootstrap replicates.
    - v. Draw a conclusion at significance level  $\alpha = 0.05$
  - (c) Perform a t test after stating the assumptions you are using. Compare the p-value and conclusions you draw from the t test to those with the bootstrap.
  - (d) Test whether the lettuce data are compatible with a population median of 145, or rather are strong evidence of a median greater than 145 with a sign test. What assumptions are we making?
3. A state's Division of Motor Vehicles claims that 60% of teens pass their driving test on the first attempt. An investigative reporter examines an SRS of the DMV records for 125 teens; 86 of them passed on their first try.
- (a) Is there convincing evidence at the  $\alpha = 0.05$  significance level that the DMV's claim is incorrect? Do an appropriate hypothesis test.
  - (b) Construct and interpret a 95% confidence interval for the proportion of all teens in the state who passed their driving test on the first attempt. Explain what the interval tells you about the DMV's claim.
  - (c) Suppose instead, the reporter gets their information by interviewing a random sample of teens who had taken their driver's test that year. How, if at all, might the test and confidence interval differ from that calculated in a and b?