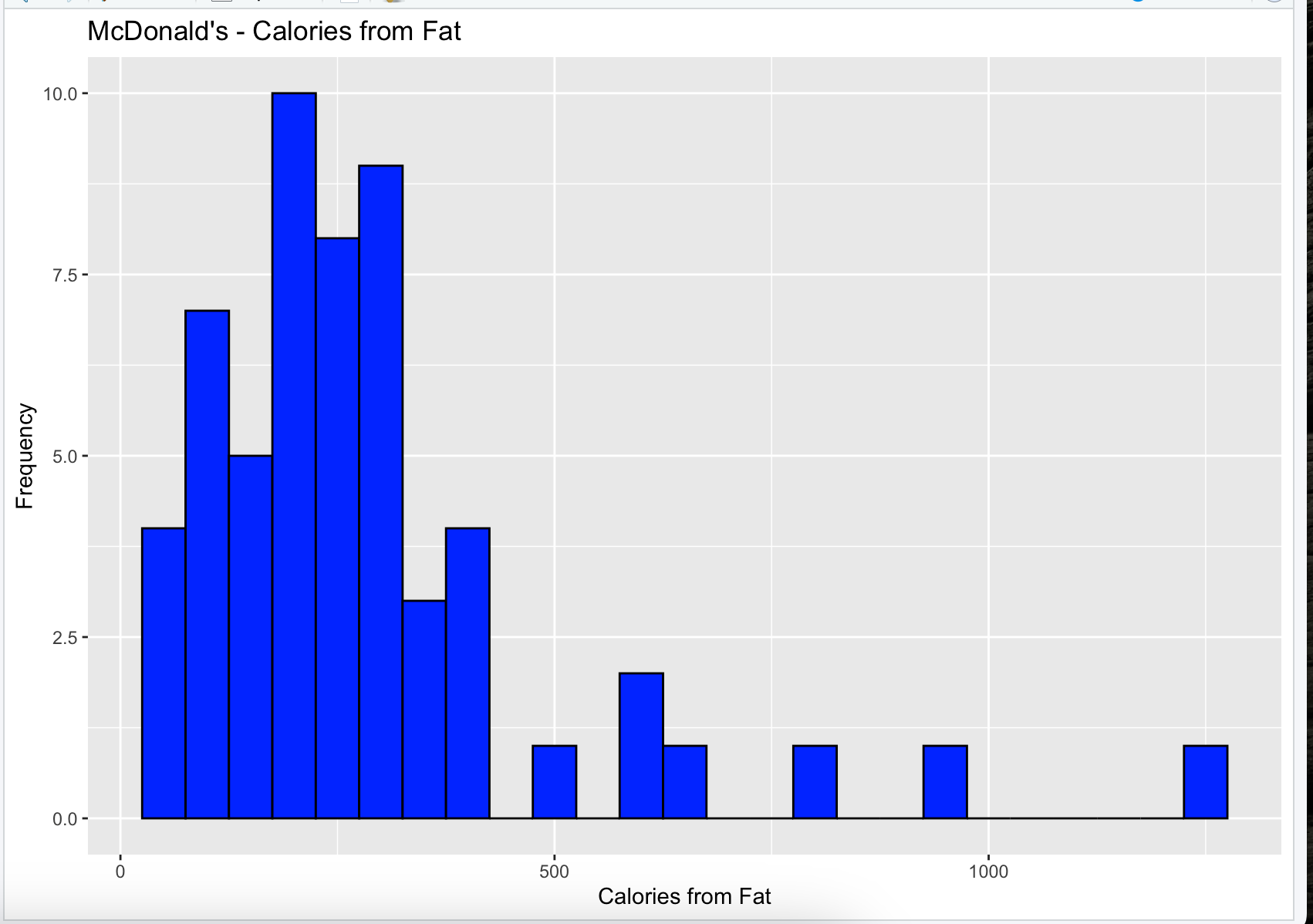
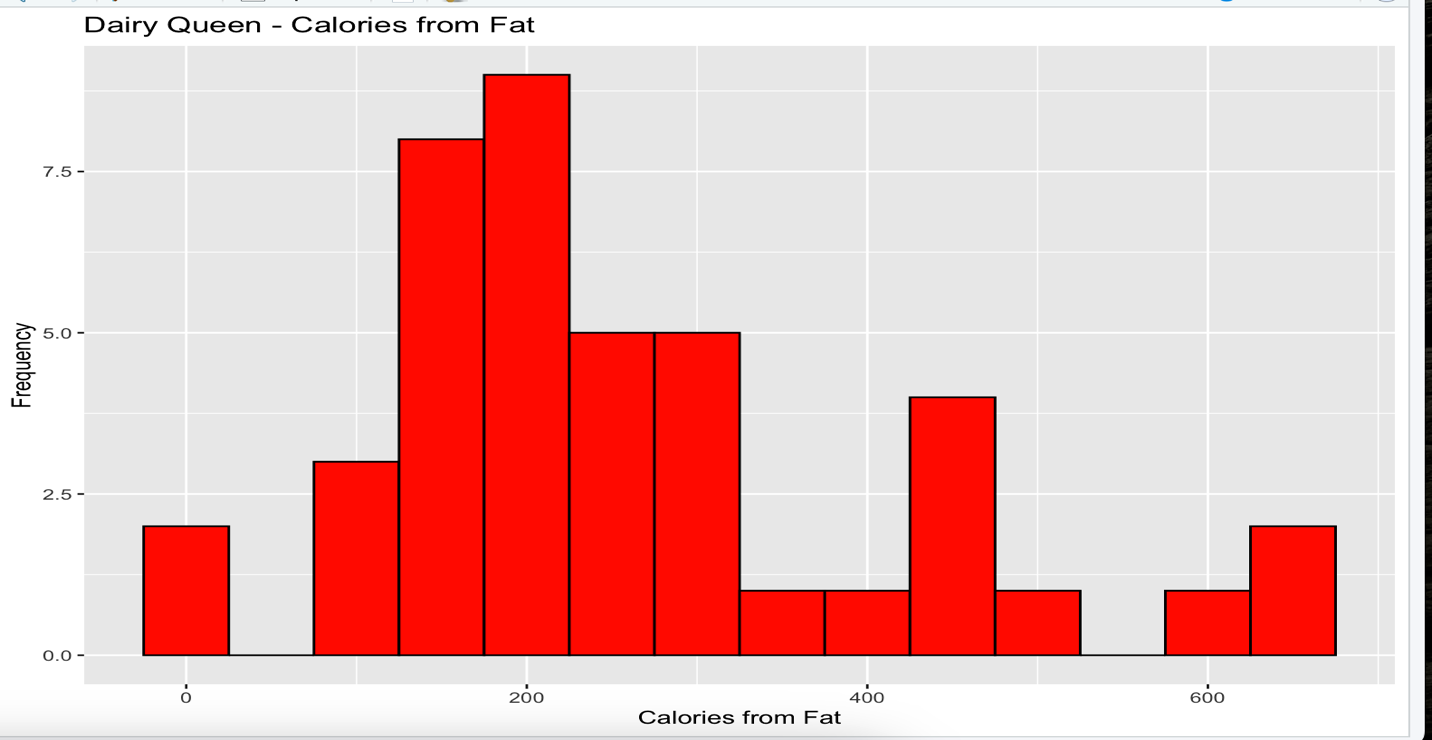
Lab 3: Distributions of random variables

Exercise 1

Make a histogram to visualize the distributions of the amount of calories from fat for each restaurant, Dairy Queen and McDonalds. Include labels in your plots. How do their centers, shapes, and spreads compare? *Include graphs and complete sentences in your write-up.*

* The McDonald's histogram shows the calories from fat are skewed to the right. Most of the data hangs around 200 calories, but the spread is pretty wide, going from almost zero to over 400.
* The Dairy Queen histogram also looks skewed to the right. The center seems to be about 300 calories. The spread is wider than what we saw with McDonald's, reaching over 600 calories from fat. Dairy Queen's distribution seems to have more items packed with higher fat compared to McDonald's.

Exercise 2 Based on this plot, does it appear that the data follow a nearly normal distribution? Include your plot and a complete sentence in your lab write-up.

* Based on the plot, the data does not perfectly follow a normal distribution. While the histogram has a unimodal shape, it appears to be skewed to the right. The normal curve does not perfectly align with the histogram, especially in the tails, indicating deviations from normality.A graph with a line going up

  AI-generated content may be incorrect.

Exercise 3: Make a normal probability (Q-Q) plot of sim\_norm. Do all of the points fall on the line? How does this plot compare to the Q-Q plot for the real data you made for Dairy Queen’s calories from fat at the start of this section?

* The Q-Q plot of the simulated normal data shows points closely following the line. In contrast, the Q-Q plot for Dairy Queen's calories from fat shows deviations from the line, especially in the upper tail. This suggests that the Dairy Queen data is not perfectly normally distributed. A graph of a line graph

  AI-generated content may be incorrect.

Exercise 4: Even better than comparing the original Q-Q plot to a single plot generated from a normal distribution is to compare it to multiple examples of plots generated from a normal distribution. Run the prior sim\_norm code 3 more times to create 3 more examples of normal Q-Q plots. Include the Q-Q plot for the calories from fat from Dairy Queen. Does the normal probability plot for Dairy Queen’s cal\_fat look similar to the plots created for the simulated data? That is, do the plots provide evidence that the calories from fat at Dairy Queen are nearly normal?

* The normal probability plot for Dairy Queen's cal\_fat does not look similar to the plots created for the simulated data. The simulated normal Q-Q plots are very linear. The Dairy Queen data has considerable deviation from the line, especially in the upper tail, indicating that the calories from fat at Dairy Queen are not close enough to a normal distribution.

A screenshot of a graph

AI-generated content may be incorrect.

Exercise 5: Using the same technique, determine whether or not the calories from the McDonalds data (calories, not cal\_fat this time) appear to come from a normal distribution. *Provide a histogram of the calorie distribution, with normal curve overlaid, and provide a Q-Q plot. Also provide an explanation.*

* The histogram of McDonald's calories shows a distribution that appears to be right-skewed, with the normal curve not perfectly aligning with the histogram, particularly in the tails. The Q-Q plot shows deviations from the line, especially in the upper tail, suggesting that the calorie distribution is not perfectly normal. Overall, the distribution of McDonald's calories is not close enough to normal.

A graph of a calories chart

AI-generated content may be incorrect.

Exercise 6: For each of the following probability questions, calculate the probability using both the theoretical normal distribution as well as the empirical distribution (two ways of computing each probability). Label them carefully.

1. Probability that a randomly selected McDonalds item has more than 1000 units of sodium.  
     
   McDonalds Sodium > 1000:

Theoretical Probability: 0.2214719

Empirical Probability: 0.2433036

1. Probability that a randomly select Dairy Queen item has between 20 and 30 units of protein  
     
   Dairy Queen Protein between 20 and 30:

Theoretical Probability: 0.2396108

Empirical Probability: 0.2222222

1. Probability that a randomly selected Dairy Queen item has less than 100 units of cholesterol.  
     
   Dairy Queen Cholesterol < 100:

Theoretical Probability: 0.8705785

Empirical Probability: 0.8888889