

Name _____

Investigation 14: How Long Do the Subway Doors Stay Open?

Worksheet 14.3 Analyze the Data

Statistical Question _____

Collection of Data

Over a period of 18 school days, a group of students in Brooklyn, New York recorded how long the doors on the F train at the subway stop near their school stayed open. They collected 65 lengths of time, to the nearest second. Each measurement was taken at approximately the same time each day, usually during 3 lunch periods.

Below are the data collected:

Length of time subway doors were open on the F train

31 17 19 20 33 29 25 25 26 17 18 29 22 24 24 26 30 27 21 23 28 25 24 21
20 31 28 27 21 24 28 29 30 21 24 27 25 24 23 22 30 26 26 25 25 24 29 24
25 27 27 24 26 22 23 27 22 24 26 28 23 24 25 23 27

Analysis of the Data

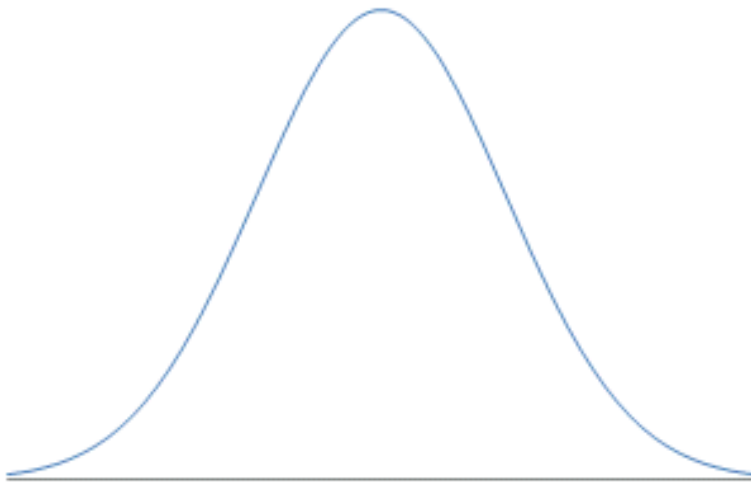
1. Construct a dot plot of the open-door times.
2. Describe the distribution of the subway door-open times. Include in your description the shape, an estimate for the mean, and an estimate for the standard deviation.
3. Interpret the mean and standard deviation in this context.
4. Using the distribution of subway door-open times, how would you answer the question, "What is the length of time the doors on the F train on the New York City subway route typically stay open?"

5. To help further understand the distribution of times, complete the frequency chart below.

Lengths of Time(sec)	Frequency		Lengths of Time (sec)	Frequency	
17			25		
18			26		
19			27		
20			28		
21			29		
22			30		
23			31		
24			32		
			33		
			Total		

6. What percent of door-open times were:
 - a. less than or equal to 24 seconds?
 - b. more than or equal to 30 seconds?
 - c. between 22 and 28, inclusive?
7. Add a column to the frequency table. Label the column Relative Frequency. Convert the frequencies to relative frequencies and record in the new column.
8. Use technology to construct a histogram (bin width of 1 sec.) with **Relative Frequency** as the y-axis, of the subway door-open times. Use technology and find the mean and standard deviation. Copy the histogram below.

9. What are the similarities and differences between the two graphs (dot plot and histogram)?
10. Do you think a Normal distribution is an appropriate model for the length of time the doors on the F train on the New York City subway route stay open?
11. On the Normal curve below, locate the mean and the times within one and two standard deviations of the mean by drawing a vertical line through these points.



12. Using the **Empirical Rule**, what proportion of the data is within one standard deviation of the mean? Show this on the Normal curve.
13. Using the **Empirical Rule**, what proportion of the data is within two standard deviations of the mean? Show this on the Normal curve.

14. Using the data table of the relative frequencies, approximately what percent of the open-door times are within one standard deviation of the mean?
15. Using the data table of the relative frequencies, approximately what percent of the open-door times are within two standard deviations of the mean?
16. How do the proportions that you found in questions 14 and 15 compare with the Empirical Rule percentages based on the Normal curve as a model for the times that the subway doors are open?

Interpret the Results in the Context of the Original Question

17. Do you think a Normal distribution is an appropriate model for the length of time the doors on the F train on the New York City subway route stay open?
18. Use the model (Normal curve) that you have created and decide if the urban myth that the subway doors typically stay open for 30 seconds is true.
19. Another method to determine if there are outliers is to use the standard deviation. If the distribution is mound shaped (approximately Normal) then any data point more than 2 standard deviations from the mean can be considered an outlier. Using this definition, are there any outliers in the subway door-open times? Is the urban myth of the doors staying open 30 seconds an outlier?