Activity: Statistical Analysis of Forest Growth in a Pellet Mill Region

Grade Level: 9th grade Algebra 1

Alignment to Algebra 1 Standards:

- **A1.S.ID.A.2** Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.

Learning Objectives:

- Use technology to find the values of mean, median and standard deviation for a data set
- Compare mean and median, and explain differences between them using extreme values in the data set.
- Use mean and median to compare two data sets.
- Use standard deviation to compare variability in two data sets.

Materials: Graphing calculator

Time Required: 10 – 25 minutes (depending on students' comfort with using the necessary functions in the calculator)

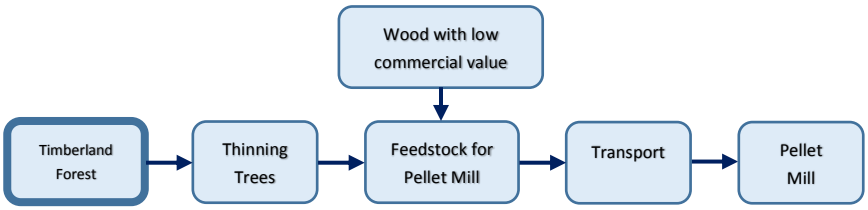
Photograph of wood pellets (www.ehrhartenergy.com)

Background Information: Bioenergy is energy derived from living matter on the surface of the earth. One biofuel increasingly used in Europe is wood pellets. Wood pellets (seen right) are derived from leftover wood from other commercial uses, tree cut to thin a forest, or trees that do not have other commercial value. Pellets are burned and the resulting energy is converted into electricity.

The data in this activity come from an extensive forest area known as the Chesapeake fuelshed (seen on map, right). This area is over 12 million hectare, where one hectare is 10,000 square meters. Since 2009, the wood pellet industry has increased their activity in this area to help meet the demand for wood pellets. Forest Inventory and Analysis (FIA) data collected by the USDA Forest Service are being monitored to see if the wood pellet industry has had a negative impact on the forest area. Data source:

Parish, ES, Dale VH, Tobin E, Kline KL (2017) Dataset of timberland variables used to assess forest conditions in two Southeastern United States' fuelsheds. Data in Brief 13C (2017) pp. 278-290. Available at http://www.sciencedirect.com/science/article/pii/S2352340917302391

The flow chart below shows the general process for timber to become wood pellets.



Name:_____ Date:_____ Period:_____

Biofuel Application: The table of data provides the area of live trees, in thousands of hectares, in the Chesapeake fuelshed by size of tree for select years from 2002 to 2014. Answer each of the following questions.

Year	Area of Live, Small Diameter Trees	Area of Live, Large Diameter Trees
2002	331	555
2003	244	556
2005	125	529
2006	129	427
2007	221	545
2009	222	559
2010	236	603
2011	214	602
2012	154	579
2013	164	617
2014	154	622

1) Find the mean and median area covered by small diameter trees. Round answers to the nearest tenth.

2) Which value, the mean or the median, is higher? What do you notice in the data that would explain your result in part b?

3) Find the mean and median area covered by large diameter trees. Compare these values to the values for small diameter trees and draw a conclusion about the two data sets.

4) What are the standard deviations (rounded to the nearest tenth) for each set of data? What conclusions can you make about the two data sets using these values?

Answer Key

Biofuel Application: The table of data provides the area of live trees, in thousands of hectares, in the Chesapeake fuelshed by size of tree for select years from 2002 to 2014. Answer each of the following questions.

Year	Area of Live, Small Diameter Trees	Area of Live, Large Diameter Trees
2002	331	555
2003	244	556
2005	125	529
2006	129	427
2007	221	545
2009	222	559
2010	236	603
2011	214	602
2012	154	579
2013	164	617
2014	154	622

1) Find the mean and median area covered by small diameter trees. Round answers to the nearest tenth.

$\bar{x}$ = 199.5 thousand hectares

med = 214 thousand hectares

2) Which value, the mean or the median, is higher? What do you notice in the data that would explain your result in part b?

The median is higher. Student answer should discuss the two low values (125 and 129) that are affecting the mean, but not the median.

3) Find the mean and median area covered by large diameter trees. Compare these values to the values for small diameter trees and draw a conclusion about the two data sets.

$\bar{x}$ = 563.1 thousand hectares

med = 559 thousand hectares

There is significantly more (about 2 and half times more) area that is occupied by large diameter trees as compared to small diameter trees.

4) What are the standard deviations (rounded to the nearest tenth) for each set of data? What conclusions can you make about the two data sets using these values?

small diameter trees: σ_x = 58.5 thousand hectares

large diameter trees: σ_x = 52.1 thousand hectares

The area occupied by small diameter trees has greater variability in this 12-year timeframe than area occupied by large diameter trees.