## Determine degree days to predict fuel usage

Program Task: Determine degree days to predict fuel usage

## Program Associated Vocabulary:

DEGREE DAYS, MEAN, MAXIMUM, MINIMUM, AVERAGE, LOW, HIGH, SWING

## Program Formulas and Procedures:

Most people do not track their fuel usage and depend on their oil supplier to somehow "know" when their tank is getting low.
Oil suppliers use a system known as the Degree Day system, which uses a simple averaging formula to predict how much fuel a customer has used. If it is known how much fuel was in the tank on a start date, and how much fuel has been used
 since then, it makes sense that the amount of fuel left can be estimated. It's not magic, it's math.

Heat is typically not needed until the outdoor temperature drops below $65^{\circ} \mathrm{F}$. Degree Days are a measure of how far below $65^{\circ} \mathrm{F}$ a given day's mean (average) temperature drops. The formula is :
Degree Days $=65-(($ High Temp + Low Temp $) / 2)$
Example: In Pottstown, PA, on January 20, 2008, the low temperature of the day was $14^{\circ} \mathrm{F}$ and the high temperature of the day was $32^{\circ} \mathrm{F}$.

$$
\begin{aligned}
& \text { Degree Days }=65-((32+14) / 2) \\
& \text { Degree Days }=65-23 \\
& \text { Degree Days }=42
\end{aligned}
$$

There are roughly 5,000 degrees day per year in southeastern Pennsylvania. Suppliers simply add degree days to the total as each day passes. At first, the supplier uses trial and error to estimate how much fuel a customer uses per degree day. Once a baseline is established, the supplier simply starts counting, adding up degree days for a given customer, restarting at zero each time the customer receives a fill up. If customer Jones needs a fill up every 1,200 degree days, the supplier simply keeps track and sends the truck to the Jones house when 1,200 degree days have passed since the last Jones delivery.
This system is remarkably effective and accurate when high and low temps do not differ drastically in one day (called a swing). On some odd days, a low temp may be $50^{\circ} \mathrm{F}$ and a high temp $80^{\circ} \mathrm{F}$, which calculates to 0 degree days (average is $65^{\circ} \mathrm{F}$ ). Does that mean no oil is used? Hardly. Consider how extreme outliers can spell failure to a system based on a steady average.

Summarize, represent, and interpret data on a single count or measurement variable
PA Core Standard: CC.2.4.HS.B. 1
Description: Summarize, represent, and interpret data on a
single count or measurement variable.

## Math Associated Vocabulary:

MEAN, MEDIAN, MODE, OUTLIER

## Formulas and Procedures:

| Outlier | An extreme value in a set of data which is much <br> higher or lower than the other numbers. |
| ---: | :--- |
| Mean <br> (Average) | The average of set of data that is calculated by <br> dividing the sum of the data by the number of <br> items in the set. |
| Median | The middle value when data are arranged in <br> numerical order or the average of the two middle <br> numbers when the set has an even number of <br> data items. |
| Mode | The value that occurs most frequently in a set of <br> data. |

Measures of central tendency are mean, median and mode. Outliers affect the mean value of the data but have little effect on the median or mode of a given set of data.

Example: A student receives a zero on a quiz and subsequently has the following scores:

$$
0,70,70,80,85,90,90,90,95,100
$$

Outlier: 0

## Mean:

$$
\frac{0+70+70+80+85+90+90+90+95+100}{10}=77
$$

Median: since the data set has 10 values, there are two middle numbers, so one must find the mean of these two values, 85 and 90.

$$
\frac{85+90}{2}=87.5
$$

Mode: The score 90 occurs more frequently than the other values (three times), so 90 is the mode.

Receiving a zero on a quiz significantly affects a student's mean, or average. Notice that the outlier had a small effect on the median and mode of the data.

It should be noted that because outliers affect the mean and have little effect on the median, the median is often used to describe "average" income. Often, one hears that the median income for a group is a certain value. Mean is not typically used because outliers, people who make significantly more or make no money at all, affect this measure.

## Instructor's Script - Comparing and Contrasting

HVAC technicians typically use mean for measuring the central tendency of data. Since outliers affect mean, HVAC technicians must compensate for outliers when using means to make predictions. Take notice in the three examples provided on page three that although all three days saw a temperature of 42 degrees, the mean was significantly affected by the high and low temperature. At some point during each of the days, it was 42 degrees, but the mean and therefore fuel usage was significantly affected by the highs and lows. The further apart the high and low are, the less reliable the predictive value of the mean when determining degree days.

## Common Mistakes Made By Students

Calculator error when finding the mean: Students often forget to use parenthesis when finding the mean of a data set. For instance, to find the average of 40 and 50 , parenthesis must be used for the sum before dividing by two. Students often enter $40+$ $50 / 2$, which yields an answer of 65 instead of entering $(40+50) / 2$ which yields the correct answer of 45 .

Changing the divisor: When determining how an outlier affects the mean of a data set, the student must find the mean with the outlier, then find the mean again once the outlier is removed. Removing the outlier decreases the number of data by one and therefore you must decrease the divisor. For instance, when you find the mean of $0,10,10,12,12$, you must divide the sum by 5 , but when you remove the outlier of 0 , you must then divide by 4 .

## When calculating the median, students must list the data need in numerical order.

Finding the median of an even set of data: Finding the median or middle number, of a set of data is simple when there is an odd number of data. When there is an even number, there are two middle numbers, and these numbers must be averaged to obtain the median. For instance, the median of $1,1,2,3,3$ is 2 because 2 is the middle number. If the data set is $1,2,3,3$, then 2 and 3 are the middle numbers and must be averaged to obtain the median of 2.5 .

## CTE Instructor's Extended Discussion

You can reinforce this math concept by maintaining a running Degree Day Chart in your lab. The table below serves as an example. Have the students fill in the cells daily, reinforcing the simple algebraic formula each day. Have students discuss the finer points of making predictions using data such as mean temperatures. And of course, you will want to discuss how outliers can affect the safety, comfort, and satisfaction of a customer.
DEGREE DAY CHART

| Date | Low Temp | High Temp | Mean Temp | Daily Deg Days | Running Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $10 / 30 / 08$ | 32 | 51 | 42 | 23 | 23 |
| $10 / 31 / 08$ | 30 | 63 | 46 | 19 | 42 |
| $11 / 1 / 08$ | 41 | 70 | 56 | 9 | 51 |
| $11 / 2 / 08$ | 39 | 52 | 45 | 20 | 71 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

For many years, the Degree Day system has worked well for many oil customers. It is no longer state of the art, however, as digital technology has once again changed the face of our trade. Solar powered, wireless tank monitors are available that physically monitor the level of fuel in a tank and have the ability to report back to the users and suppliers via email, text messages, or even a fax.

## Problems Occupational (Contextual) Math Concepts <br> Solutions

1. Determine the degree days when the high is $42^{\circ} \mathrm{F}$ and the low is $13^{\circ} \mathrm{F}$.
2. Determine the degree days when the high is $42^{\circ} \mathrm{F}$ and the low is $40^{\circ} \mathrm{F}$.
3. Determine the degree days when the high is $65^{\circ} \mathrm{F}$ and the low is $42^{\circ} \mathrm{F}$.
Even though all 3 days in these problems saw $42^{\circ} \mathrm{F}$ at some point, consider how distance from mean affected fuel usage for any given day.

## Problems

Related, Generic Math Concepts

## Solutions

4. Sally earned $60,65,65$, and 80 on 4 tests. How would scoring 100 on a fifth test affect the mean?
5. Tom recorded his daily caloric intake for 5 days. The results were as follows: 2500, 2600, 2600, 2400, and 3900. How would removing the outlier affect the mean, median and mode of the data?
6. Angela recorded the number of hours she spent watching TV for one week. The results were as follows: $6,2,2,1.5$, $3,2.5,2$. How would removing the outlier affect the mean, median, and mode of the data?

## Problems

PA Core Math Look
Solutions
7. Which of the following measures of central tendency does an outlier affect the most?
a) Mean
b) Median
c) Mode
8. Which measure of central tendency would best depict the following data: $10,200,200,300,325,350,400$ ?
a) Mean
b) Median
c) Mode
9. How would removing the outlier affect the mean of the following data: 1200, 2400, 2400, 2500, 9000 ?

| Problems Occupational (Con | extual) Math Concepts Solutions |
| :---: | :---: |
| 1. Determine the degree days when the high is $42^{\circ} \mathrm{F}$ and the low is $13^{\circ} \mathrm{F}$. | $\begin{aligned} & \text { Degree Days }=65-((42+13) / 2) \\ & \text { Degree Days }=65-28 \\ & \text { Degree Days }=37 \end{aligned}$ |
| 2. Determine the degree days when the high is $42^{\circ} \mathrm{F}$ and the low is $40^{\circ} \mathrm{F}$. | $\begin{aligned} & \text { Degree Days }=65-((42+40) / 2) \\ & \text { Degree Days }=65-41 \\ & \text { Degree Days }=24 \end{aligned}$ |
| 3. Determine the degree days when the high is $65^{\circ} \mathrm{F}$ and the low is $42^{\circ} \mathrm{F}$. <br> Even though all 3 days in these problems saw $42^{\circ} \mathrm{F}$ at some point, consider how distance from mean affected fuel usage for any given day. | $\begin{aligned} & \text { Degree Days }=65-((65+42) / 2) \\ & \text { Degree Days }=65-54 \\ & \text { Degree Days }=11 \end{aligned}$ |
| Problems Related, | neric Math Concepts Solutions |
| 4. Sally earned $60,65,65$, and 80 on 4 tests. How would scoring 100 on a fifth test affect the mean? | $\begin{aligned} & \text { Initial Mean }=\frac{60+65+65+80}{4}=67.5 \\ & \text { Mean with outlier }=\frac{60+65+65+80+100}{5}=74 \end{aligned}$ <br> The mean increased by 6.5 |
| 5. Tom recorded his daily caloric intake for 5 days. The results were as follows: $2500,2600,2600,2400$, and 3900. How would removing the outlier affect the mean, median and mode of the data? | The mean would decrease from 2800 to 2525 . The median would decrease from 2600 to 2550. The mode would remain constant at 2600. |
| 6. Angela recorded the number of hours she spent watching TV for one week. The results were as follows: $6,2,2,1.5$, $3,2.5,2$. How would removing the outlier affect the mean, median, and mode of the data? | The mean would decrease from 2.71 to 2.17 The median would remain constant at 2. The mode would remain constant at 2 . |
| Problems PA | ore Math Look Solutions |
| 7. Which of the following measures of central tendency does an outlier affect the most? <br> a) Mean <br> b) Median <br> c) Mode | a) Mean |
| 8. Which measure of central tendency would best depict the following data: $10,200,200,300,325,350,400$ ? <br> a) Mean <br> b) Median <br> c) Mode | Median, because the outlier of 10 would make the average lower, and the mode of 200 would represent the lower data. The median is 300 . |
| 9. How would removing the outlier affect the mean of the following data: $1200,2400,2400,2500,9000$ ? | The mean would decrease from 3500 to 2125. |

