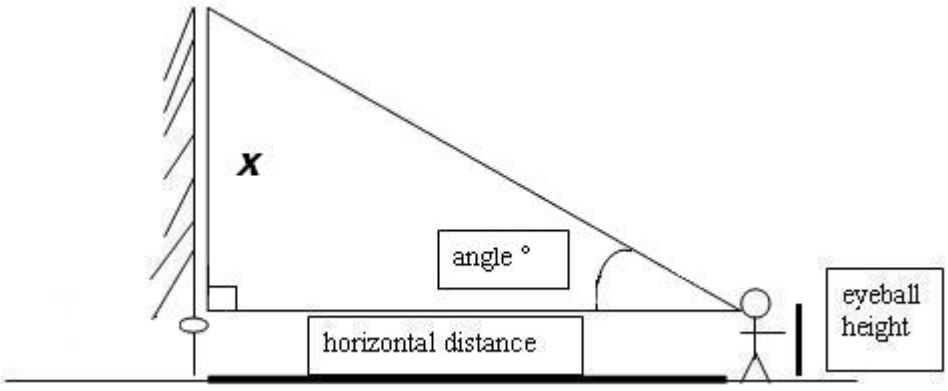
Exercise: Douglas Firs, Regression and Statistics

Name_____ Name _____

In this project we will focus on analyzing your field data, a sample of the Douglas fir population along the White River valley. **Attach your field trip data sheet to this project.**

1. Calculate tree heights.

The diagram below is similar to that found in your field trip information packet. The diagram shows the basic geometry of the tree measurement procedure. You measured the eyeball height, the angle of elevation from your eyeball to the top of the tree, and the horizontal distance.



To compute the height of the tree, you must find the distance labeled **x** and add that distance to the eyeball height. How do you find **x**? From trigonometry we know that the tangent of the angle of elevation is equal to the slope. But the slope of the inclined line in the diagram above is *x* divided by the horizontal distance (see below).

$$\tan(\textit{angle}) = \text{slope} = \frac{\text{rise}}{\text{run}} = \frac{x}{\text{hor. dist.}}$$

Solving the equation above for *x*, we get x = hor. dist. × tan(angle). Therefore the height of the tree is: **tree height = eye height + hor. dist. × tan(angle)**.

Important note: your calculator should be in **Degree Mode** when you do this work. Go to Mode, then go down two lines to select Degree. Press **ENTER**. See below.

Try this example: Suppose eye height is 1.7 meters, horizontal distance is 35 meters, and the angle is 55 degrees. Then the height of the tree is $1.7 + 35 \times \tan(55) \approx 51.7$ meters. Do you get this same result?

Compute the heights of all your trees and fill in the last column on your data sheets. Then enter the data for all trees in your calculator: **L1 = circumference**, **L2 = height**. Do not sort the data.

2. Your field data

Were there any factors/things/situations that greatly influenced the accuracy of the circumference measurements? If so, briefly list/describe.

Were there any factors/things/situations that greatly influenced the accuracy of the height measurements? If so, briefly list/describe.

We will compile all measurements of the circumference and height of Tree "B" at Stop #2 on the blackboard. Record all measurements below.

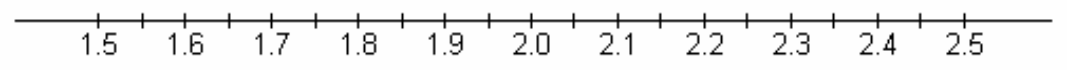
group initials	circum.	height	group initials	circum.	height

Enter the circumference and height data from the table above into lists **L3** and **L4**, respectively. Do not sort the data.

3. Statistics for tree circumference measurements

Calculate and give the 5–number summary for the circumference measurements only.

Draw a Box and Whisker plot of the circumference measurements (using the scale).

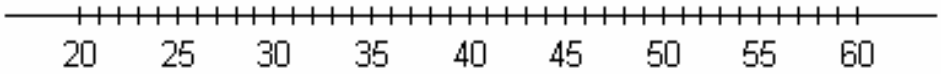


What do you think is the real value of circumference at chest height of Tree B? Explain.

4. Statistics for tree height measurements

Now lets examine all class measurements of the height of Tree "B" at Stop #2. Calculate and give the 5–number summary for these height measurements.

Draw a Box and Whisker plot of the height measurements (using the scale).



How does your tree height measurement compare to the class "average"? Explain.

5. Examining the spread in class measurements

What is the range in class circumference measurements for Tree B? What is the range in the height measurements?

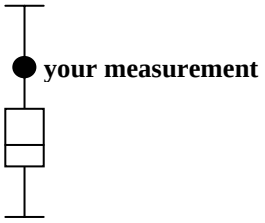
Express the **range as a percent of the median** for both height and circumference measurements of Tree B.

In your own words, explain what "range as a percent of the median" indicates about the data. Also explain why it is greater for height, rather than circumference measurements.

5. Modeling height vs. circumference for YOUR data

Rectangular graph paper is provided. Plot *height* vs. *circumference* (y vs. x) using a filled circle for each of your trees. Choose axes scales so that most of the graphpaper is used. Your height and circumference scales need not be the same. Label your axes.

At the point for Tree B on your graph, draw the Box and Whisker plot illustrating the class heights. See example below:



6. Finding the best-fitting model

There are many best-fitting regression models that are pre-programmed into your TI-83/84 under **STAT > CALC**. Go to this menu now and scroll down to take a look. Then compute the **linear**, **quadratic**, **exponential**, **power**, and **logarithmic** models that best fit your (L1, L2) tree data. *Note: the keystrokes on the TI-83/84 for all of these regressions are identical.* Record the equation and correlation coefficient results in the table below. Use H for height and C for circumference in each equation.

Graph the (circumference, height) data on your TI-83/84. Press **Zoom>9ZoomStat** to size the window. Then graph each regression model (one at a time) over the data. Look at the general shape of each regression graph and make a "thumbnail" sketch of each in the table below. Label axes in each diagram.

Model Type	**Equation**	**Model Predictive Power* (R^2)**	**sketch**
linear (LinReg)		$R^2 =$	
quadratic (QuadReg)		$R^2 =$	
exponential (ExpReg)		$R^2 =$	
power (PwrReg)		$R^2 =$	
logarithmic (LnReg) Helpful Hint: $a + b\ln(x) \approx$ $a + 2.3b\log(x)$		$R^2 =$	

* R^2 is the *Coefficient of Determination* and measures the predictive power of the function. The value of R^2 lies from 0 to 1. As with the correlation coefficient, the closer to 1 the better.

7. Model comparison

Do the **shapes** of the regression graphs match the **shape** of the data? Which models have the best shape? Which are not so good? Explain.

Which of the regressions seems to best fit the data? What final formula do you conclude is best for circumference to predict height? Explain your choice.

A member of our class says, *"We don't really expect a high correlation coefficient for any of the models because we don't have enough data points. If we measured more trees, we would know a lot more about how height can predict circumference for Douglas firs. In fact, if we had 1000 trees in our sample, the trees would fit one of the models perfectly. Then, for one of the regression models, we would get a perfect correlation coefficient of* $r = 1.0$ *!!!"* Do you think that your class member is accurate in his/her reasoning? Answer yes or no, and then argue your case.

8. The monster Douglas fir

Greg and Joe measured the height and circumference of the very old Douglas fir at the Dalles campground; it was 9.45 meters in circumference and 84.1 meters high. Which of your best fit equations is a better predictor of height given a circumference of 9.45 meters? Describe how you know; include work.

9. Sampling and the Douglas firs

Was your sample of the population of Douglas firs along the White River valley *random* or not? Explain briefly, using the statistical definition of "random" in your response.

Compute $\bar{x}$, the mean height of all trees in your sample:_____ How well does your value of $\bar{x}$ estimate the mean height of all Douglas firs along the White River Valley (the population)?

Suggest 2 ways in which your sampling method could be improved so that $\bar{x}$ would better approximate the population mean.

10. Forest succession

Sketch a **qualitative** graph for each forest relationship described below. A qualitative graph has labels but no scales/no numbers. The graphs should be fairly accurate with respect to shape, slope, concavity, intercept position, etc.

a) height versus age (y vs. x) of Douglas fir trees.

b) number of trees per acre versus age.

c) light illumination of forest floor versus age

d) biodiversity of understory versus biomass of understory

Remember to attach your field trip data sheet to this project.