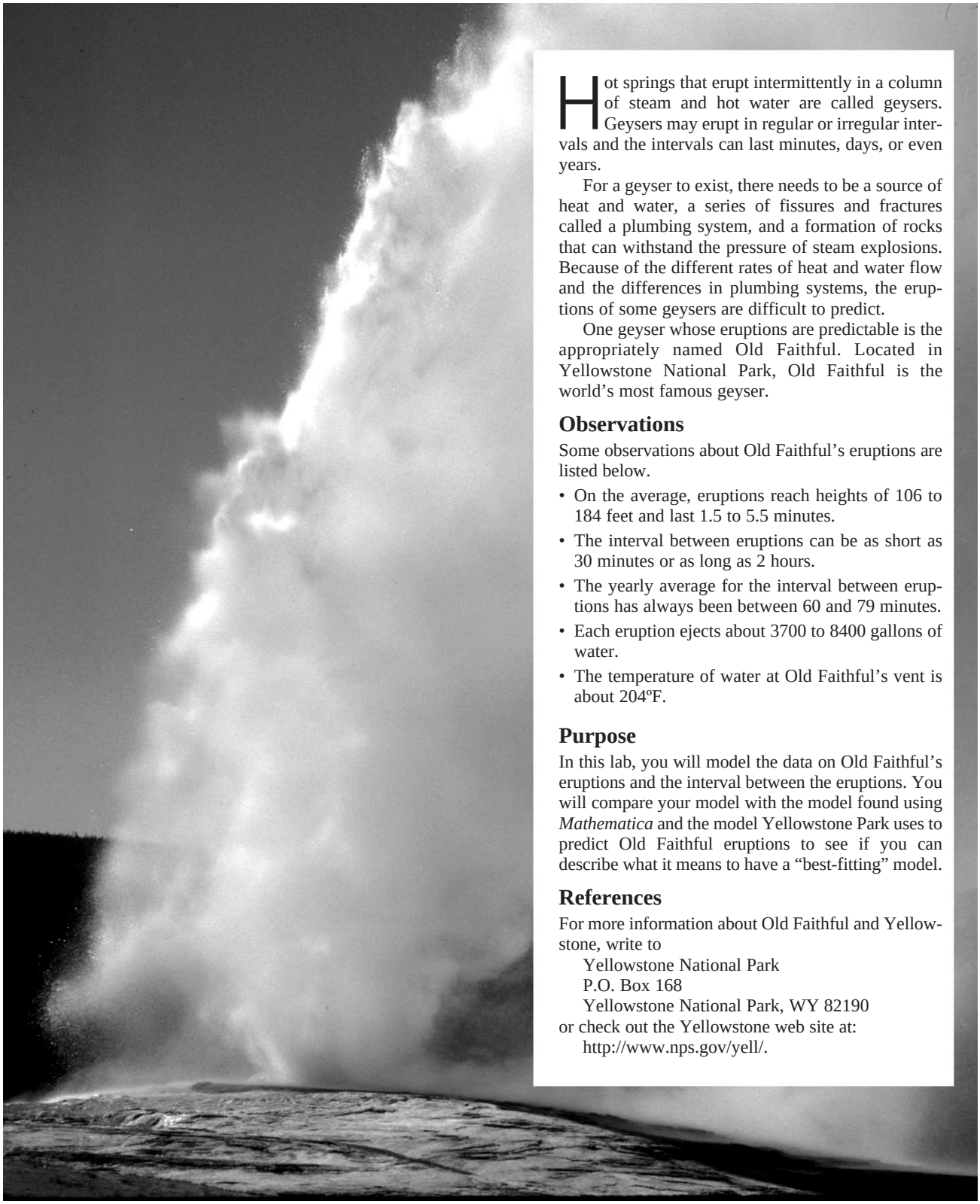
LAB 1

Modeling Old Faithful's Eruptions

Modeling Data



Hot springs that erupt intermittently in a column of steam and hot water are called geysers. Geysers may erupt in regular or irregular intervals and the intervals can last minutes, days, or even years.

For a geyser to exist, there needs to be a source of heat and water, a series of fissures and fractures called a plumbing system, and a formation of rocks that can withstand the pressure of steam explosions. Because of the different rates of heat and water flow and the differences in plumbing systems, the eruptions of some geysers are difficult to predict.

One geyser whose eruptions are predictable is the appropriately named Old Faithful. Located in Yellowstone National Park, Old Faithful is the world's most famous geyser.

Observations

Some observations about Old Faithful's eruptions are listed below.

- On the average, eruptions reach heights of 106 to 184 feet and last 1.5 to 5.5 minutes.
- The interval between eruptions can be as short as 30 minutes or as long as 2 hours.
- The yearly average for the interval between eruptions has always been between 60 and 79 minutes.
- Each eruption ejects about 3700 to 8400 gallons of water.
- The temperature of water at Old Faithful's vent is about 204ºF.

Purpose

In this lab, you will model the data on Old Faithful's eruptions and the interval between the eruptions. You will compare your model with the model found using *Mathematica* and the model Yellowstone Park uses to predict Old Faithful eruptions to see if you can describe what it means to have a "best-fitting" model.

References

For more information about Old Faithful and Yellowstone, write to

Yellowstone National Park
P.O. Box 168
Yellowstone National Park, WY 82190

or check out the Yellowstone web site at:

http://www.nps.gov/yell/.

© Houghton Mifflin Company. All rights reserved.

Data

The duration of Old Faithful's eruptions and the interval between the eruptions is given in the table below. Let x represent the duration of a geyser eruption in minutes. Let y represent the interval between geyser eruptions in minutes.

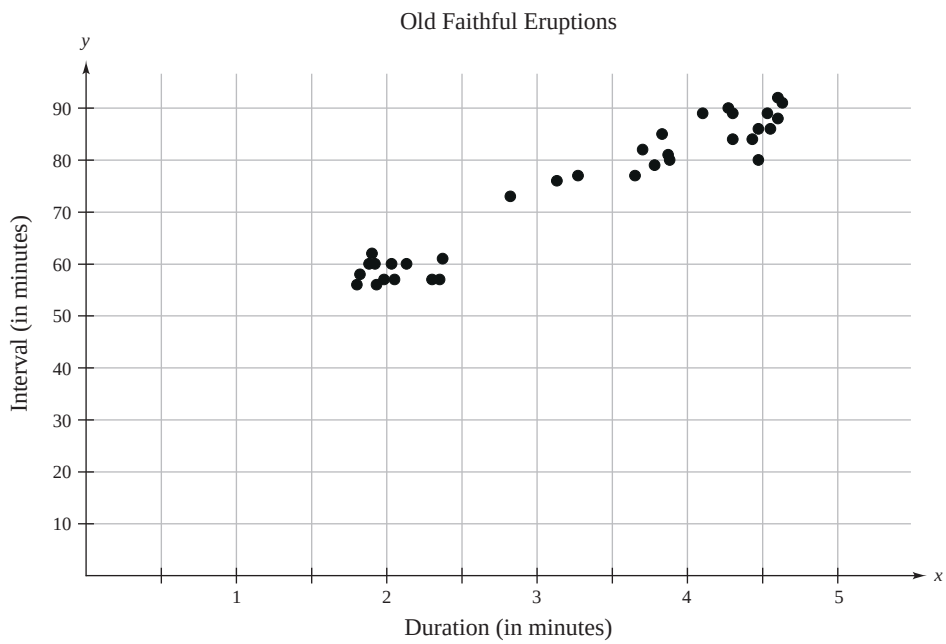
Duration, x	1.80	1.82	1.88	1.90	1.92	1.93	1.98	2.03	2.05	2.13
Interval, y	56	58	60	62	60	56	57	60	57	60

Duration, x	2.30	2.35	2.37	2.82	3.13	3.27	3.65	3.70	3.78	3.83
Interval, y	57	57	61	73	76	77	77	82	79	85

Duration, x	3.87	3.88	4.10	4.27	4.30	4.30	4.43	4.43	4.47	4.47
Interval, y	81	80	89	90	84	89	84	89	86	80

Duration, x	4.53	4.55	4.60	4.60	4.63
Interval, y	89	86	88	92	91

A scatter plot of the data is given below.



The data in the table and the scatter plot are stored in the *Mathematica* file called LAB01.NB.

© Houghton Mifflin Company. All rights reserved.

Exercises

Name _____

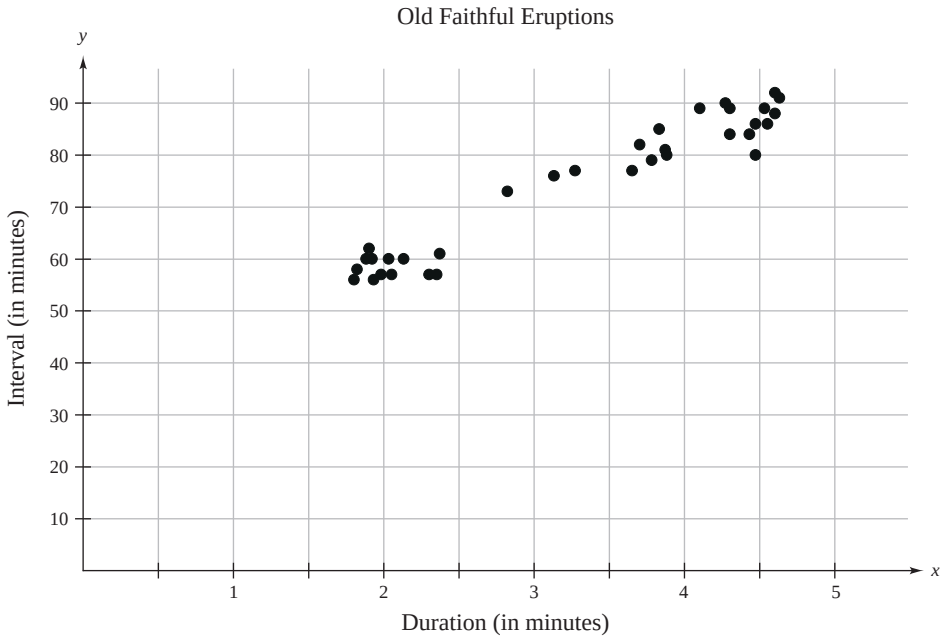
Date _____ Class _____

Instructor _____

1. ***Describe the Relationship.*** Describe in words the numerical relationship between x and y as given in this lab's Data. Also suggest a possible explanation of why x and y are related in this manner.

© Houghton Mifflin Company. All rights reserved.

2. ***Drawing a Best-Fitting Line.*** Draw the line that seems to best fit the data points and find its equation (by hand).



3. ***Is Your Model the Best?*** Use the table given in this lab's Data and *Mathematica* to compare the actual *y*-values and your equation's *y*-values. How good is your model? Do you think it is the best possible linear model? Explain your reasoning.

4. ***Smaller or Larger Sum?*** In this lab's *Mathematica* file the sum of the squares of the differences of the actual *y*-values and your equation's *y*-values is given. If you could find a better fitting linear model, would its sum be smaller or larger than the sum that is given? Explain your reasoning.

© Houghton Mifflin Company. All rights reserved.

5. ***"Least Squares" Method.*** In statistics the linear model that best fits the data is found using what is called a "least squares" method. Based on your answer to Exercise 4, what do you think this means?

6. ***Comparing Equations.*** The equation used by Yellowstone Park to predict the intervals between Old Faithful's eruptions is $y = 14x + 30$. Compare this equation with the one you found in Exercise 2. Which equation seems to be the "better" model? Explain your reasoning.

7. ***Choosing the Best Model.*** In this lab's *Mathematica* file, the least squares method is used to find the "best" linear model for the Old Faithful eruption data. Use the table to compare the actual y-values and the least squares model's y-values. Of these two linear models and your own, which is the best one? Why?

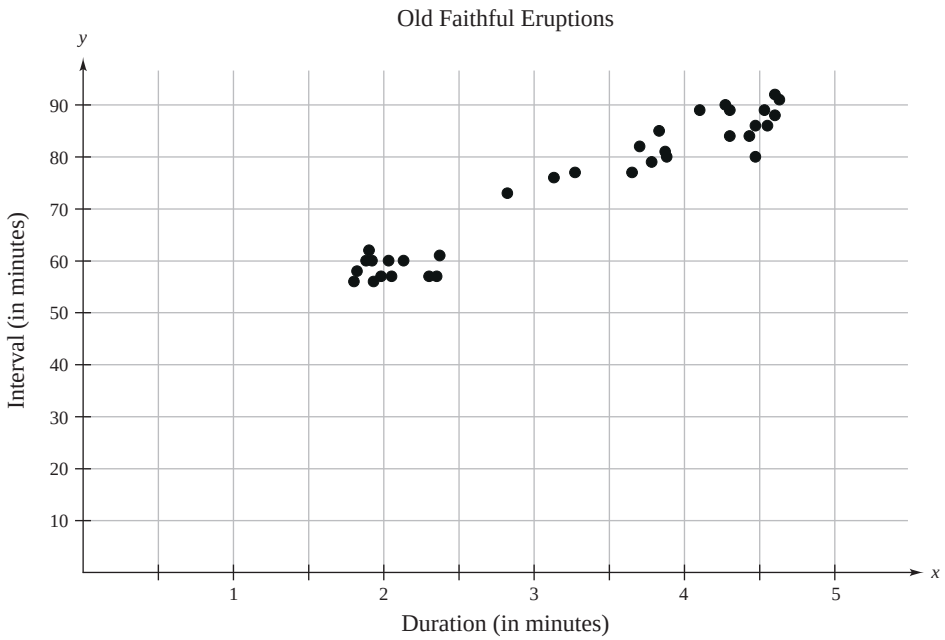
© Houghton Mifflin Company. All rights reserved.

8. ***Lab Summary.*** Summarize your results and sketch the graph of each.

Model from Exercise 2: _____

Yellowstone Park Model: _____

Least Squares Model: _____



9. ***A Better Type of Model?*** Do you think a quadratic model would fit the data better than a linear model in this case? Why or why not?

_____

_____

_____

_____

_____

_____

_____

_____

© Houghton Mifflin Company. All rights reserved.