IM2 Problem Set 4.4 - Data, Scatter Plots and Functions

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set when we collect information on two variables? • How do I describe and analyze bivariate data? • What functions can I use to model bivariate data sets? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?

PART 1 - Functions we will use to model data - Opening Investigation

Steelco has been implementing a CO_2 emissions policy since 2010. Here is the data set that records the levels of CO_2 emissions over the past few years for this steel factory in Hamilton, Ontario.

Year	2011	2012	2013	2014	2015	2016
Years since 2010	1	2	3	4	5	6
CO_2 (tonnes)	165	145	130	117	107	100

The company must reduce emissions to 70 tonnes of CO_2 per year before 2018. Is this an achievable goal for the factory?

1. Graph the data using DESMOS. Let the *x*-axis represent the number of years since 2010.

LINEAR MODELS

2. Does the scatter plot appear "linear"?
3. Start by modeling the data with a linear model. Determine the equation of the line of best fit using DESMOS.
4. Record the equation of the linear model for the data set as well as the correlation coefficient, the ***r*** value.
5. Using this linear model, can Steelco meet their targeted CO_2 emissions levels by 2018?
6. What does the slope of the linear model represent? (NOTE: Let's add the slope to each *y* value and see what happens.) Hence, does the trend in the data seem linear? Why/why not?

EXPONENTIAL MODELS

7. Does the scatter plot appear "exponential"?
8. We can also model data using exponential functions. The simplest of exponential functions has a general equation of $y = ab^x$, where b is the base of the exponential function. Determine the equation of the exponential curve of best fit using DESMOS.
9. Record the equation of the exponential model for the data set as well as the correlation coefficient, the ***r*** value.
10. Using this exponential model, can Steelco meet their targeted CO_2 emissions levels by 2018?
11. What does the base, b, of the exponential model represent? (NOTE: Let's multiply each y value by the base and see what happens.) Hence, does the trend in the data seem exponential? Why/why not?

QUADRATIC MODELS

12. Does the scatter plot appear "quadratic"?
13. We can also model data using quadratic functions. The simplest of quadratic functions has a general equation of $y = a(x - h)^2 + k$ where a, h and k are parameters that control different aspects of the appearance of quadratic functions. Determine the equation of the quadratic curve of best fit using DESMOS.
14. Record the equation of the quadratic model for the data set as well as the correlation coefficient, the ***r*** value.
15. Using this quadratic model, can Steelco meet their targeted CO_2 emissions levels by 2018? Why/why not?
16. Using DESMOS, label the lowest point on the quadratic model. Compare the point to the equation generated by DESMOS. What do h and k represent?
17. If the trend in the data is quadratic, what should happen to the levels of CO_2 emissions in the future?

18. Which model seems best for this data set? Explain your reasoning.

PART 2 - Applications - DESMOS, TI-84 and Lines (Curves) of Best Fit

We will now use DESMOS and regression functions to compare and contrast models of growth curves of bacterial populations. So Sripriya is now a world famous microbiologist and in two of her experiments, she grows bacteria in the lab and produces the following data sets. Her data tables show the population (in thousands of bacteria) as a function of time (how long they have been growing)

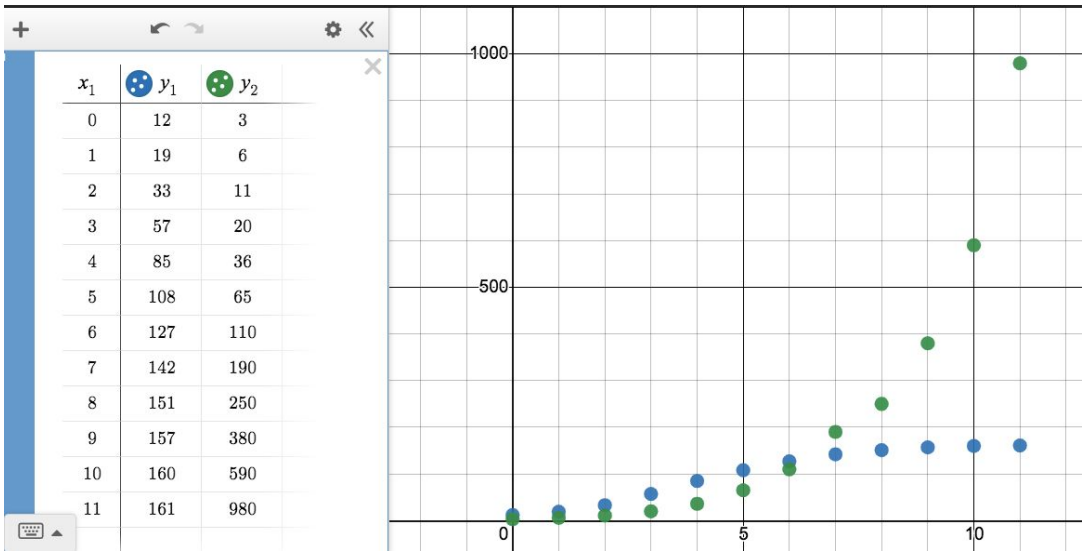
Bacteria Population #1

Time (h)	0	1	2	3	4	5	6	7	8	9	10	11
Population (thousands)	3	6	11	20	36	65	110	190	250	380	590	980

Bacteria Population #2

Time (h)	0	1	2	3	4	5	6	7	8	9	10	11
Population (thousands)	12	19	33	57	85	108	127	142	151	157	160	161

1. Prepare scatterplots of both data sets using DESMOS



2. Does either data set produce a "linear trend"? An "exponential trend"?
3. Since exponential models are modeled as $y = ab^x$, determine exponential models for both data sets. Record your equations and both ***r*** values.
4. How reasonable are the exponential models for each of the data sets? (HINT: multiply the y-values by the value of the base, b, in each case)

You should find that the exponential model doesn't really fit well for Population #2, as the data seems to suggest that the population levels off.

5. With Population #2, try the following exponential "adjustment" ⇒ $y_1 \sim ab^{x_1} + 165$. Describe what happens and why the changes make sense. What does this model imply should happen in the "long run"?

6. With Population #2, try a quadratic regression ⇒ $y_1 \sim a(x_1 - h)^2 + k$. Describe what happens and why the changes make sense. What does this model imply should happen in the "long run"?

PART 3 - Homework from Nelson 9 Chap 6.3

3. A basketball is dropped from a height of 200 cm. The table shows how high it bounces on each bounce.

Maximum Height (cm)	200	120	72	44	26	16	10	6	4
Bounce Number	0	1	2	3	4	5	6	7	8

a) Plot the data on a scatter plot.
b) Sketch a curve of best fit. Should you use a solid curve or a dashed curve? Explain.
c) Does it make sense to use the curve of best fit for interpolation? Explain.
d) Does it make sense to use the curve of best fit for extrapolation? Explain.

4. In order to obtain a medical image of a patient's thyroid gland, a **C** chemical is injected into the patient's bloodstream. The chemical's concentration in the blood gradually decreases with time.

Concentration (mg/L)	29.0	15.0	7.7	3.9	2.1	1.3	0.7	0.5	0.4
Time (h)	0	1	2	3	4	5	6	7	8

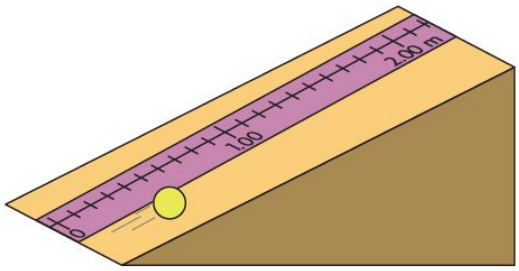
a) Plot the data on a scatter plot.
b) Sketch a curve of best fit. Did you use a solid curve or a dashed curve? Explain.
c) Describe the relationship between the variables.
d) Use your curve of best fit to estimate when the concentration of the chemical will be 6.1 mg/L.
e) Use your curve of best fit to estimate the concentration of the chemical after 12 h.

5. The ages and resting heart rates for some people are listed in the table.

Age (years)	21	24	26	29	31	35	39
Resting Heart Rate (beats per minute)	60	61	63	65	68	73	78

a) Plot the data on a scatter plot.
b) Sketch a curve of best fit. Did you use a solid curve or a dashed curve? Explain why.
c) Describe the relationship between the variables.
d) Does it make sense to use the curve of best fit for interpolation? Explain.
e) Does it make sense to use the curve of best fit to estimate the resting heart rate for an 85-year-old person? Explain.

6. In his experiments to study the Earth's gravity, Galileo rolled objects
A on inclined planes. In one such experiment, a ball is rolled up a plane, and then, rolls back down. The data are in the following table.



Time (s)	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5
Position (m)	0	1.13	1.50	1.88	2.00	1.88	1.50	1.13

a) Use a graph to estimate the position of the ball after 0.3 s.
b) Use a graph to estimate when the ball will return to the bottom of the inclined plane.

8. A herd of caribou is moved to a small, remote island where they have no predators. Data on the population of the herd were collected for 6 years.

Time (years)	0	1	2	3	4	5	6
Population	24	35	51	74	104	151	225

a) Sketch the data on a scatter plot.
b) Draw a line or curve of best fit through the plotted points. Explain which is more appropriate.
c) Describe the growth of the herd.
d) Predict the population of the herd after seven years.

9. In the Kingdom of Petrodalla, natural gas is the primary resource. The table shows the amount of natural gas produced each year.

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016
Natural Gas Produced (millions of m^3)	1.6	2.1	3.0	4.1	4.3	4.4	3.6	2.1	0.5

a) Sketch the data on a scatter plot.
b) Draw a line or curve of best fit through the plotted points. Explain which is more appropriate.
c) Describe how the production of natural gas changes over time.
d) Predict when natural gas production will decrease to zero.