(3.4) Analyzing Data Using Desmos

Questions:
DRAWING A SCATTER PLOT
1. **a)** What does the scatter plot suggest about how the number of smog advisories is changing over time?

    b) Describe any correlation between the variables (ie. Positive, Negative, Weak or Strong).

    c) Do you think it would be appropriate to model the relationship with a line of best fit? Justify your answer.

CREATING A LINE OF BEST FIT
2. **a)** The screen shows that the equation of the line is $y= mx+ b$ and gives the values of m and b. What are these values? (Round m to one decimal place and b to the nearest whole number).

    What does m tell you about the line of best fit?

    b) What is the correlation coefficient, r?

    What does r tell you about the relationship?

3. Describe the line of best fit.

Is it what you expected? Explain why or why not.

EXTRAPOLATING WITH THE LINE OF BEST FIT
4. **a)** Predict the number of smog advisories for 2007.

b) As of October 22, 2007, there were 13 smog advisories in Ontario. Is the total number for the year likely to be much greater than this? ***Explain.***

c) Was the prediction from the line of best fit close to the actual number (as of October)?

5. Predict the number of advisories for 2010. Explain what you did.

6. a) Which point appears to be an outlier?

b) The summer of 2005 was one of the hottest and most humid summers in Ontario. Toronto recorded 41 days with temperatures greater than 30°C. How might this have resulted in the unusual data for 2005?

REMOVING AN OUTLIER
7. Repeat the process from *Creating a line of best fit*. **a)** What are these values of *m* and *b*? (Round *m* to one decimal place and *b* to the nearest whole number)

What does *m* tell you about the line of best fit?

b) What is the correlation coefficient, *r*? What does it tell you about the relationship?

c) How does the correlation coefficient compare to the coefficient from question 2? What does this suggest about the new model?

8. Repeat the process from *Extrapolating with the line of best fit*.
a) What are the new predicted number of smog advisories for 2007 and 2010?

b) Which line of best fit gave a more accurate prediction for 2007?

c) Would you say that including outliers always weakens a model? Justify your answer.

9. The Ministry of the Environment says air quality in the province has been improving since 1988, but it expects the number of smog advisories to increase because it is doing a better job of monitoring air quality.
What does this show about drawing conclusions about cause and effect based only on numerical data?