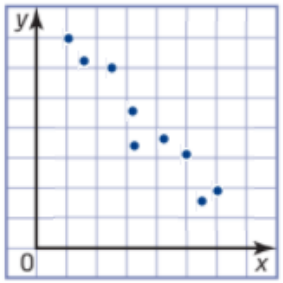
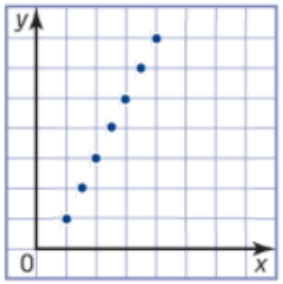
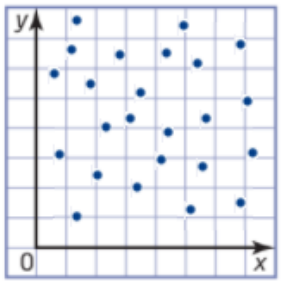
3.4 Homework Handout – Correlation in linear and non-linear data

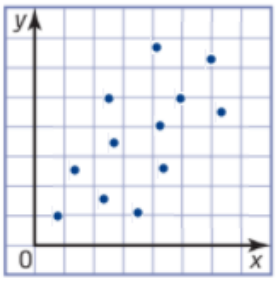
PART A

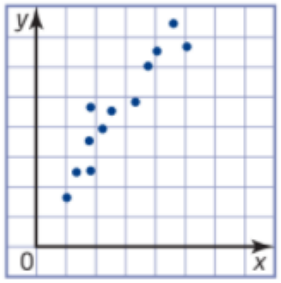
1. For the following scatter plots, estimate the correlation coefficient (r).

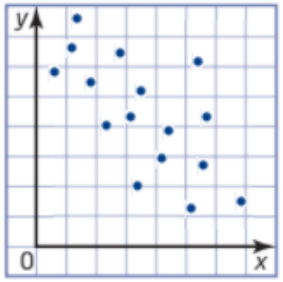
a) 

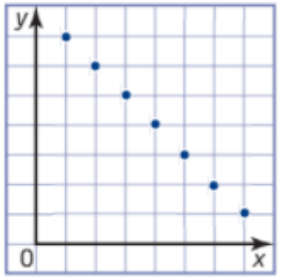
b) 

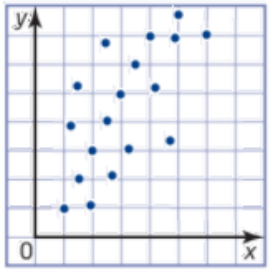
c) 

d) 

e) 

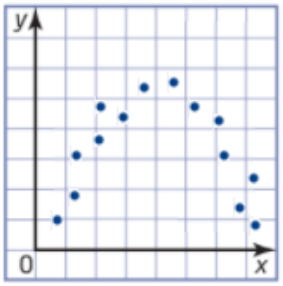
f) 

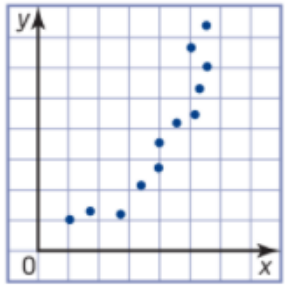
g) 

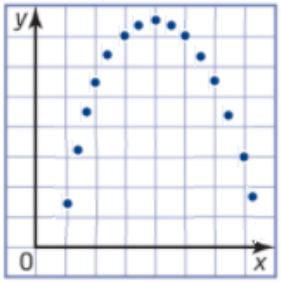
h) 

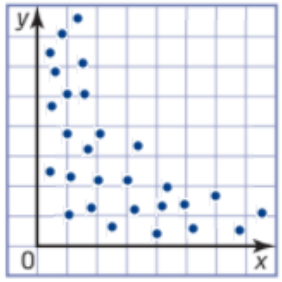
2. Sketch the general appearance of a scatter plot and describe the correlation associated with each value of the correlation coefficient (r). Include a line of best fit.
   a) $r = 0.8$ b) $r = -0.3$ c) $r = 0$ d) $r = 1$ e) $r = 0.4$

3. Describe the relationships below and estimate a value for their coefficient of determination (r^2). Include a curve of best fit.

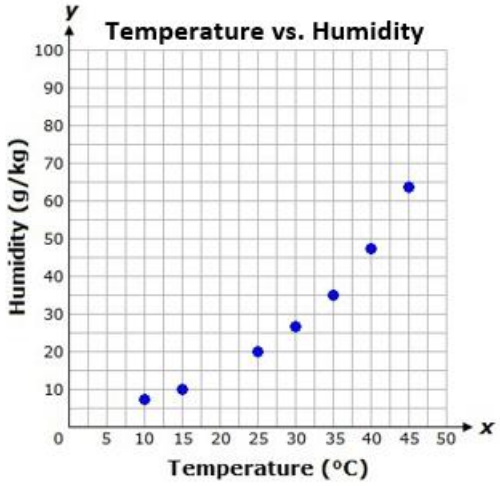
a) 

b) 

c) 

d) 

4. Phoebe collected their own data on humidity and temperature once a month for 7 months and presented their data in the scatter plot to the right.
   a) Describe the relationship in the data.
   b) Estimate a value for the coefficient of determination (r^2)
   c) Draw a curve of best fit for this data.
   d) What do you expect the temperature to be if the humidity is 15g/kg? Is this **interpolation** or **extrapolation**?
   e) What do you expect the humidity to be if the temperature is $50°C$? Is this **interpolation** or **extrapolation**?



PART B

5. The table below shows the mean air pressure at various altitudes.
   a) Make a scatter plot of the data. Put altitude on the x-axis and air pressure on the y-axis.
   b) Describe the relationship in the data.
   c) Do you think a line of best fit is appropriate for this data? Why or why not?
   d) Estimate a value for the coefficient of determination for this relationship.
   e) Estimate the air pressure at an altitude of 18km. Show traces of your work on your graph!

Altitude (km)	0	5	10	15	20	25	30
Air Pressure (kPa)	101.3	54.0	35.0	12.0	5.5	3.5	3.2

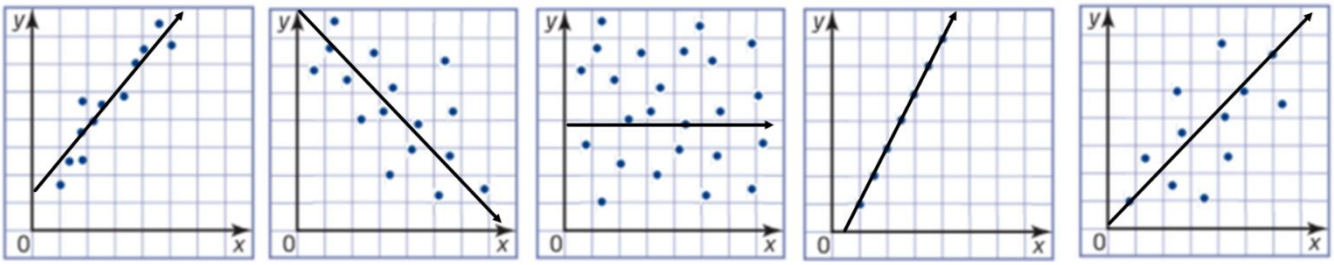
6. The table below shows a series of measurements of water temperatures at various depths below a research ship.
   a) Make a scatter plot of the data.
   b) Draw a line or curve of best fit.
   c) Describe the relationship between the variables.
   d) Estimate the water temperature at a depth of 700m.
   e) Estimate the water temperature at a depth of 1600m.
   f) Which of your two estimates is likely to be more accurate? Explain your reasoning.

Depth (m)	100	200	300	400	500	800	1000	1200
Temperature $(°C)$	19.4	19.0	18.1	17.5	16.0	9.7	6.2	6.0

ANSWERS

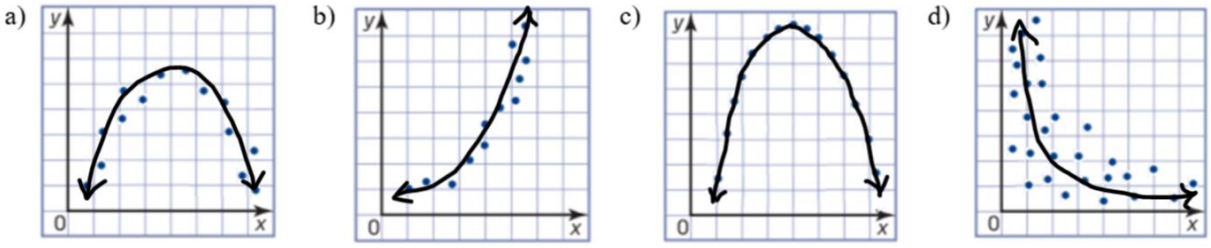
1. Answers may vary! As long as they're in the correct range.
(± 0 = null, $\pm 0.1 - \pm 0.5$ = weak, $\pm 0.5 - \pm 0.9$ = strong, ± 1 = perfect)
a) $r = -0.9$ b) $r = 1$ c) $r = 0$ d) $r = 0.3$ e) $r = 0.8$ f) $r = -0.4$ g) $r = -1$ h) $r = 0.4$

2. a) $r = 0.8$ b) $r = -0.3$ c) $r = 0$ d) $r = 1$ e) $r = 0.4$

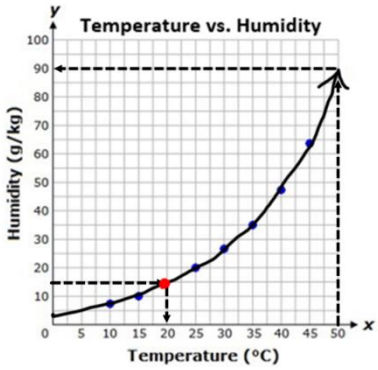


3. Answers may vary! As long as they're in the correct range. (0 = null, 0.1 – 0.5 = weak, 0.5 –0.9 = strong, 1 = perfect)
a) $r^2 = 0.8$ b) $r^2 = 0.9$ c) $r^2 = 1$ d) $r^2 = 0.3$

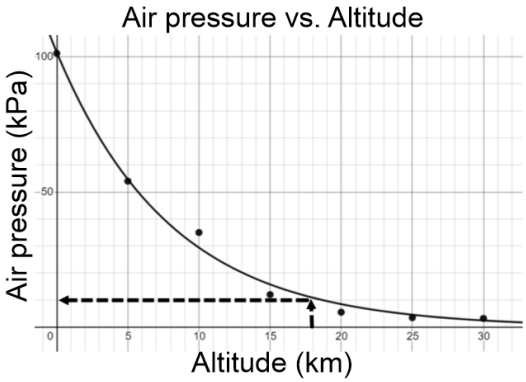
a) b) c) d)



4. a) Perfect, non-linear relationship. As the temperature increases, so does the humidity (faster and faster). b) $r^2 = 1$ (or very close to it!).
c) See graph. d) Approximately 20°C. e) Approximately 90 g/kg.



5. a) See graph. b) Strong, non-linear relationship. As the altitude increases, the air pressure decreases (faster and faster). c) No. The data seems to follow a non-linear trend. d) $r^2 = 0.95$ (answers may vary). e) There is an air pressure of approximately 10kPa at an altitude of 18km.



6. a) See graph. b) See graph. c) Strong, non-linear relationship. As the water depth increases, the temperature decreases faster and faster before levelling just below 5°C. d) Approximately 12°C. e) Approximately 5.5°C.
f) The first estimate should be more accurate because we are interpolating. In the second estimate we are extrapolating which is less accurate because there is no data to guide the curve.