1. The errors

$$\begin{aligned} e_1 &= 2 - 0.75(1) - 1.25 = 0, \\ e_2 &= 3 - 0.75(3) - 1.25 = -0.5, \\ e_3 &= 6 - 0.75(5) - 1.25 = 1, \\ e_4 &= 7 - 0.75(8) - 1.25 = -0.25. \end{aligned}$$

Sum of squares of the errors is $0^2 + 0.5^2 + 1^2 + 0.25^2 = 1.3125$.

2. a. The least squares best fit formula gives the mean as $\bar{x} = \frac{1+2+5+6}{4} = 3.5$. The a coefficient for the line is given by:

$$a = \frac{(1-3.5)2 + (2-3.5)3 + (5-3.5)5 + (6-3.5)7}{(1-3.5)^2 + (2-3.5)^2 + (5-3.5)^2 + (6-3.5)^2} = \frac{15.5}{17} = 0.91176$$

The mean of the y is $\bar{y} = \frac{2+3+5+7}{4} = 4.25$. The b coefficient satisfies:

$$b = \bar{y} - a\bar{x} = 4.25 - (0.91176)(3.5) = 1.0588$$

b. The errors

$$\begin{aligned} e_1 &= 2 - 0.91176(1) - 1.0588 = 0.029412, \\ e_2 &= 3 - 0.91176(2) - 1.0588 = 0.11765, \\ e_3 &= 5 - 0.91176(5) - 1.0588 = -0.61765, \\ e_4 &= 7 - 0.91176(6) - 1.0588 = 0.47059. \end{aligned}$$

Sum of squares of the errors is $e_1^2 + e_2^2 + e_3^2 + e_4^2 = 0.61765$.

3. The linear least squares formula gives the mean as $\bar{x} = \frac{6.7+6.9+7.1+7.3+7.5}{5} = 7.1$. The a coefficient satisfies:

$$\begin{aligned} a &= \frac{(6.7-7.1)0.9 + (6.9-7.1)1.2 + (7.1-7.1)1.4 + (7.3-7.1)1.2 + (7.5-7.1)1.5}{(6.7-7.1)^2 + (6.9-7.1)^2 + (7.1-7.1)^2 + (7.3-7.1)^2 + (7.5-7.1)^2} \\ &= 0.6 \end{aligned}$$

Since $\bar{y} = \frac{0.9+1.2+1.4+1.2+1.5}{5} = 1.24$, so the b coefficient satisfies:

$$b = \bar{y} - a\bar{x} = 1.24 - (0.6)(7.1) = -3.02.$$

b. The sum of squares of the errors is

$$\begin{aligned} J(0.6, -3.02) &= |0.9 - 0.6(6.7) + 3.02|^2 + |1.2 - 0.6(6.9) + 3.02|^2 + |1.4 - 0.6(7.1) + 3.02|^2 \\ &\quad + |1.2 - 0.6(7.3) + 3.02|^2 + |1.5 - 0.6(7.5) + 3.02|^2 \\ &= 0.068. \end{aligned}$$

4. a. Model A shows an increasing relationship, while Model B shows a decreasing relationship.

b. The sum of the squares of the errors for Model A is

$$\begin{aligned} J(0.4, 2.7) &= |4.2 - 0.4(1) - 2.7|^2 + |2.9 - 0.4(3) - 2.7|^2 + |5.7 - 0.4(5) - 2.7|^2 + |5.1 - 0.4(8) - 2.7|^2 \\ &= (1.1)^2 + (-1)^2 + (1)^2 + (-0.8)^2 = 3.85. \end{aligned}$$

The sum of the squares of the errors for Model B is

$$\begin{aligned} J(-0.4, 6.1) &= |4.2 + 0.4(1) - 6.1|^2 + |2.90.4(3) - 6.1|^2 + |5.7 + 0.4(5) - 6.1|^2 + |5.1 + 0.4(8) - 6.1|^2 \\ &= (-0.22)^2 + (-2.04)^2 + (1.88)^2 + (1.04)^2 = 8.826. \end{aligned}$$

Thus, Model A is better.

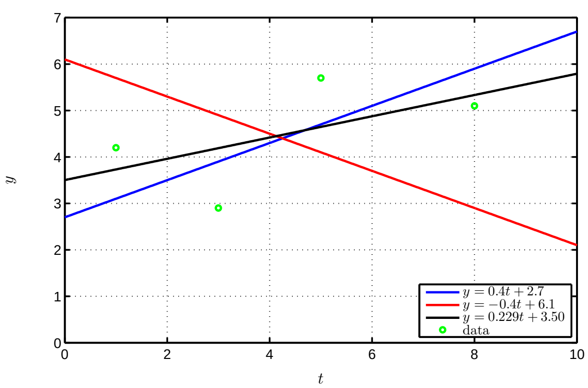
c. The linear least squares formula gives the mean as $\bar{x} = \frac{1+3+5+8}{4} = 4.25$. The a coefficient satisfies:

$$\begin{aligned} a &= \frac{(1-4.25)4.2 + (3-4.25)2.9 + (5-4.25)5.7 + (8-4.25)5.1}{(1-4.25)^2 + (3-4.25)^2 + (5-4.25)^2 + (8-4.25)^2} \\ &= \frac{6.125}{26.75} = 0.22897 \end{aligned}$$

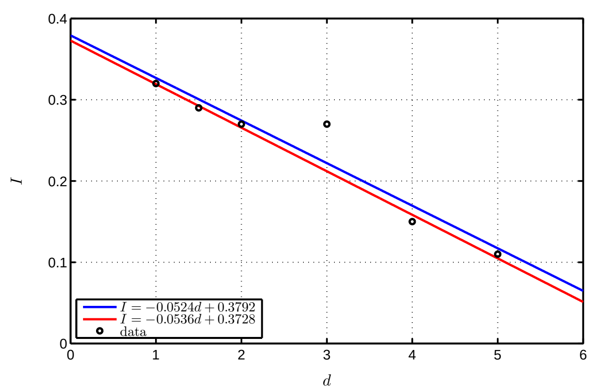
Since $\bar{y} = \frac{4.2+2.9+5.7+5.1}{4} = 4.475$, so the b coefficient satisfies:

$$b = \bar{y} - a\bar{x} = 4.475 - (0.22897)(4.25) = 3.5019.$$

which gives the least squares best fit line $y = 0.22897\,t + 3.5019$. The graph is shown below to the left. Researcher A had the better intuition by this measure, but clearly insufficient data were collected for a good analysis.



Problem 4



Problem 5

5. a. The sum of the squares of the errors is

$$\begin{aligned} J(-0.0524, 0.3792) &= |0.32 + 0.0524(1) - 0.3792|^2 + |0.29 + 0.0524(1.5) - 0.3792|^2 \\ &\quad + |0.27 + 0.0524(2) - 0.3792|^2 + |0.27 + 0.0524(3) - 0.3792|^2 \\ &\quad + |0.15 + 0.0524(4) - 0.3792|^2 + |0.11 + 0.0524(5) - 0.3792|^2 \\ &= 0.0068^2 + 0.0106^2 + 0.0044^2 + 0.048^2 + 0.0196^2 + 0.00728^2 = 0.00291796. \end{aligned}$$

b. The fourth data point is most likely in error, so $b = 3$ The graph is shown above to the right.

The sum of the square errors without the fourth point is

$$\begin{aligned} J(-0.0524, 0.3792) &= |0.32 + 0.0536(1) - 0.3728|^2 + |0.29 + 0.0536(1.5) - 0.3728|^2 \\ &\quad + |0.27 + 0.0536(2) - 0.3728|^2 + |0.15 + 0.0536(4) - 0.3728|^2 \\ &\quad + |0.11 + 0.0536(5) - 0.3728|^2 \\ &= 0.00012336. \end{aligned}$$

Percent error $= \frac{-0.0524+0.0536}{-0.0536} \times 100 = -2.238806\%$

6. The linear least squares formula gives the mean as $\bar{x} = \frac{0+10+20+30+40+50}{6} = 25$. The a coefficient satisfies:

$$\begin{aligned} a &= \frac{(0-25)75.7 + (10-25)91.9 + (20-25)105.9 + (30-25)122.4 + (40-25)131.5 + (50-25)151.4}{(0-25)^2 + (10-25)^2 + (20-25)^2 + (30-25)^2 + (40-25)^2 + (50-25)^2} \\ &= \frac{2569}{1750} = 1.468 \end{aligned}$$

Since $\bar{y} = \frac{75.7+91.9+105.9+122.4+131.5+151.4}{6} = 113.13$, so the b coefficient satisfies:

$$b = \bar{y} - a\bar{x} = 113.13 - (1.468)(25) = 76.433.$$

which is the least squares best fit line $P(t) = 1.468t + 76.433$

The sum of the squares of the errors is

$$\begin{aligned} J(1.468, 76.433) &= |75.7 - 1.468(0) - 76.4333|^2 + |91.9 - 1.468(10) - 76.433|^2 \\ &\quad + |105.9 - 1.468(20) - 76.433|^2 + |122.4 - 1.468(30) - 76.433|^2 \\ &\quad + |131.5 - 1.468(40) - 76.433|^2 + |151.4 - 1.468(50) - 76.433|^2 \\ &= 20.68133. \end{aligned}$$

b. The population at $t = 40$ according to the model is

$$P_e(40) = 1.468(40) + 76.433 = 135.153.$$

This gives:

Absolute error $= |P_e(40) - P_t(40)| = |135.153 - 131.5| = 3.653$

Relative error $= \frac{P_e(40)-P_t(40)}{P_t(40)} = \frac{135.153-131.5}{131.5} = 0.0278$

Percent error $= \frac{P_e(40)-P_t(40)}{P_t(40)} 100\% = \frac{135.153-131.5}{131.5} 100\% = 2.78\%$

c. The population at $t = 60$ according to the model is

$$P_e(60) = 1.468(60) + 76.433 = 164.513$$

Absolute error $= |164.513 - 179.3| = 14.787$

Relative error $= \frac{164.513-179.3}{179.3} = -0.08247$

Percent error $= \frac{164.513-179.3}{179.3} 100\% = -8.247\%$