For the Birds — Wildlife Management

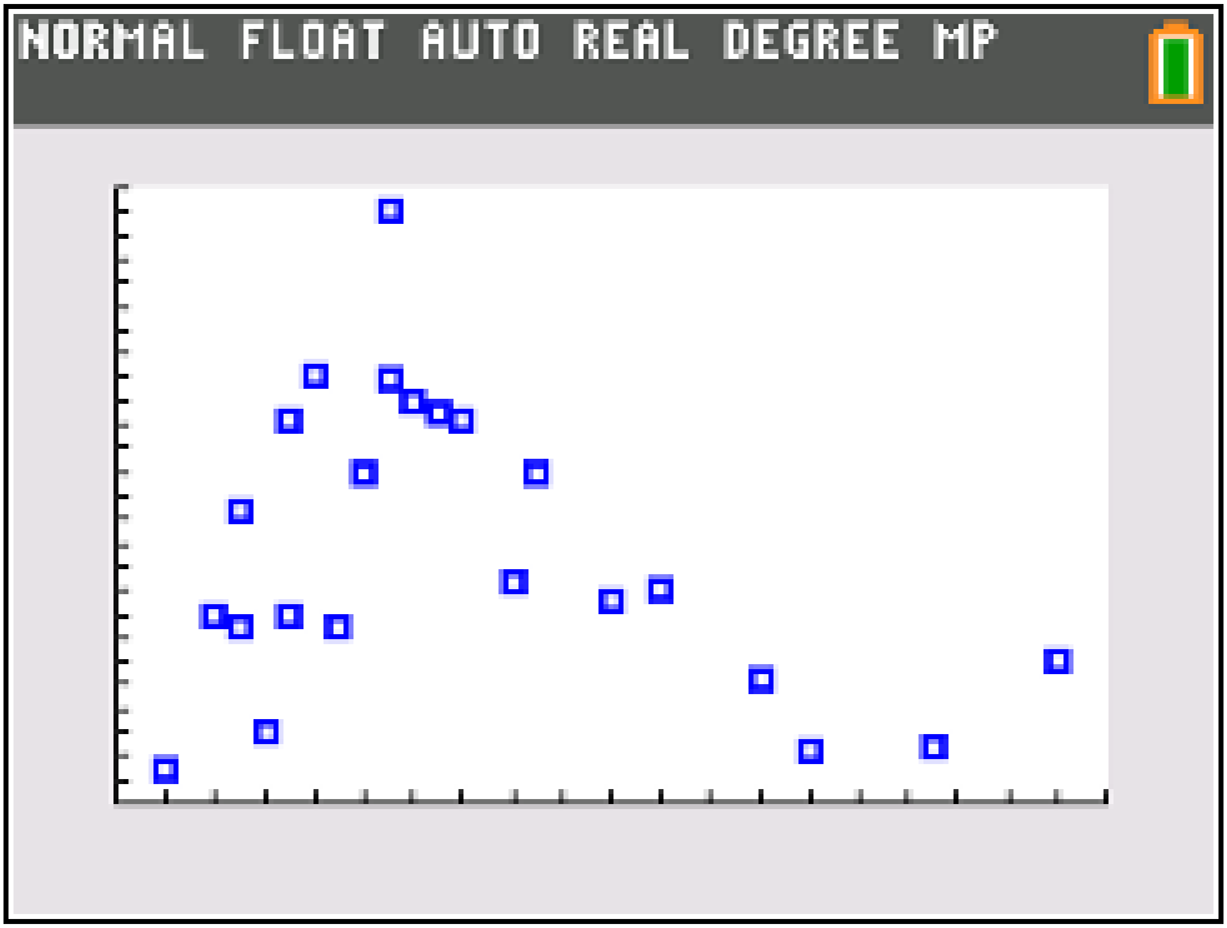
How does the presence of humans affect the population of sparrows in a park? Do more humans mean fewer sparrows? Or does the presence of humans increase the number of sparrows up to a point? Are there a minimum number of sparrows that can be found in a park, regardless of the number of humans? What can a mathematical model tell you?

In 1997, researchers set out to answer these questions. They observed the numbers of sparrows and pedestrians in several wooded parks. The approximate data are shown in the table (1 hectare = 107,639 ft^2).

Number of pedestrians per hectare, *x*	2	4	5	5	6	7	7	8	9	10	11
Number of sparrows per hectare, *y*	15	80	75	124	30	79	161	180	75	140	179

Number of pedestrians per hectare, *x*	11	12	13	14	16	17	20	22	26	28	33	38
Number of sparrows per hectare, *y*	250	169	165	162	94	140	86	90	53	22	24	60

1. Make a scatter plot of the data. Interpret the domain and range.



The domain of this relation is all the possible input, or *x*, values. The minimum number of pedestrians per hectare was 2 and the maximum was 38.

The domain is $2 \leq x \leq 38$.

The range of this relation is all the possible output, or *y*, values. The minimum number of sparrows observed per hectare was 15 and the maximum was 250.

The range is $15 \leq y \leq 250$.

2. Determine whether the data represent a function. Why or why not? What might explain the repeated x-values?

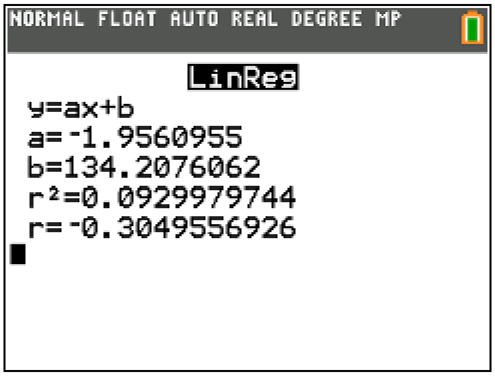
The data does not represent a function, as the same x-values have multiple y-values (x = 5, 7, 11). The repeated x-values can be explained by the fact that, on multiple occasions, the same number of pedestrians were sighted per hectare over different dates, times of observation, and locations.

3. What trends do you see in the data?

The data appears to be nonlinear and follows a cubic trend, with a local maximum at $x \cong 11$, and a local minimum at $x \cong 28$. The end behavior appears to be as $x \to \infty$, then $y \to \infty$. However, only the last 2 data points out of 23 data points show this upward trend, so the upward trend cannot be stated confidently. In addition, the point (11, 250) could be an outlier as it is farther away from the rest of the data points.

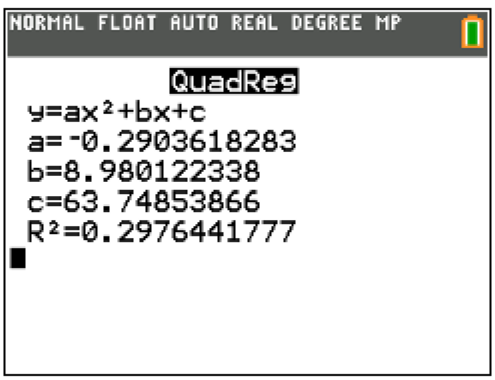
4. Use technology to perform regression to find three models for the data and the coefficients of determination.

a. a linear model



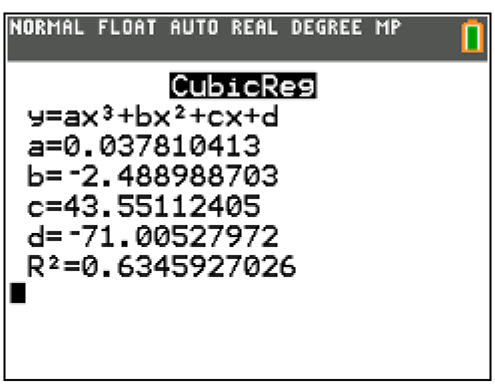
y = -1.9560955x + 134.2076062
Coefficient of determination: 0.0929979744

b. a quadratic model



$y = -0.2903618283x^2 + 8.980122338x$
$+ 63.74853866$
Coefficient of determination: 0.2976441777

c. a cubic model

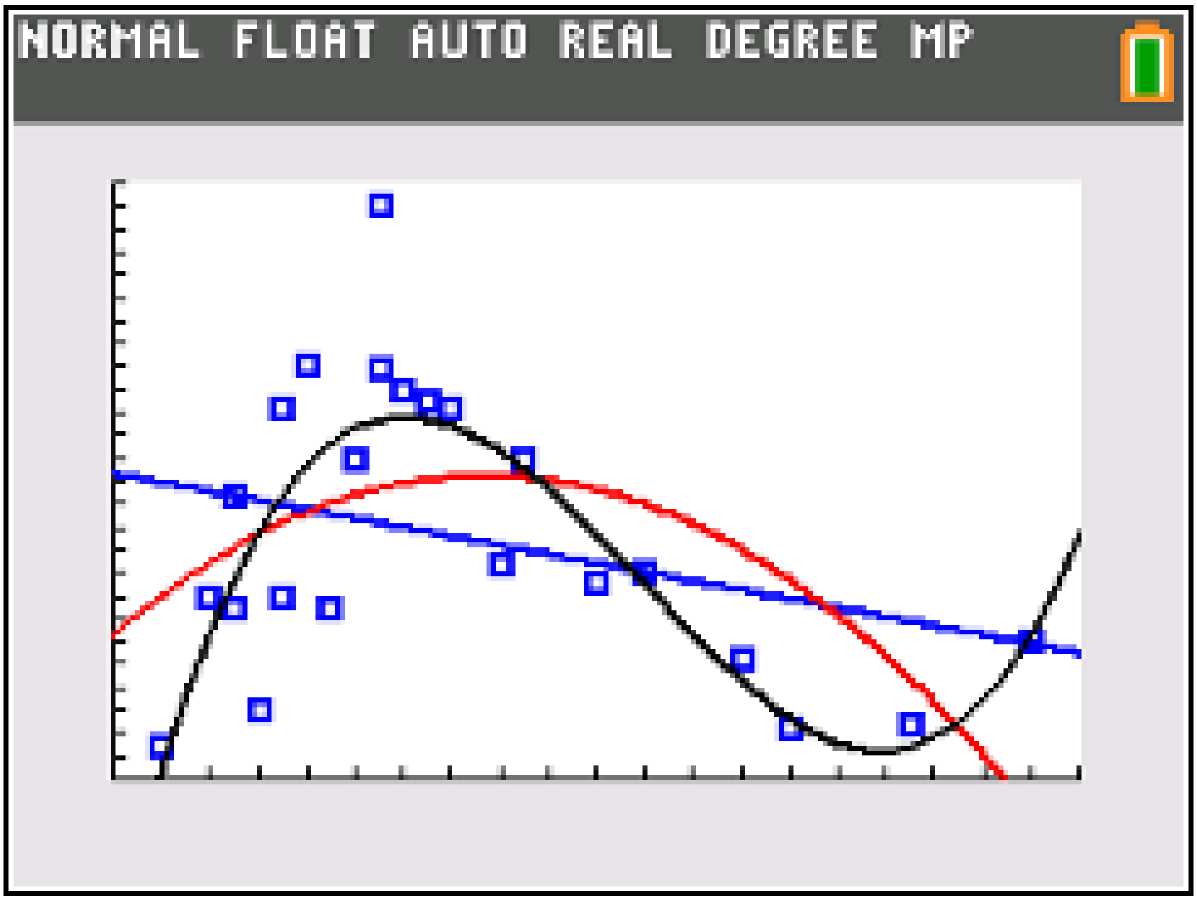


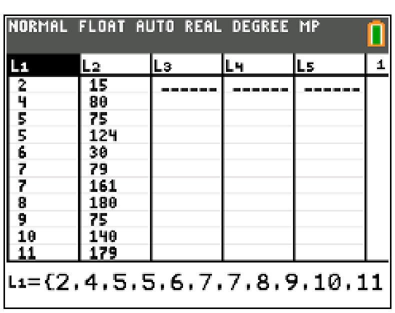
$y = 0.037810413x^3 - 2.488988703x^2$
$+ 43.55112405x - 71.00527972$
Coefficient of determination: 0.6345927026

5. According to the coefficients of determination, which model fits the data best? Why?

According to the coefficients of determination (R^2 values), the cubic model best fits the data, with a coefficient of determination of about 0.63, noticeably higher than the quadratic and linear models', 0.298 and 0.09 respectively. A higher R^2 indicates that the model is a better fit for the observed data set. Purely visually, the cubic model also appears to follow the general set of points best.

6. Graph the three regression equations with the scatter plot

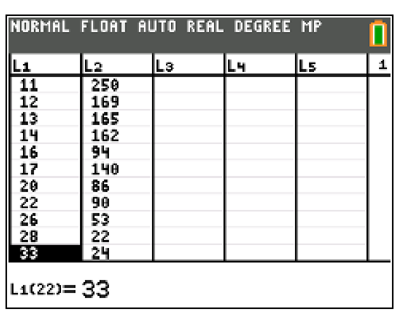




NORMAL FLOAT AUTO REAL DEGREE MP

L1	L2	L3	L4	L5	1
2	15	------	------	------	
4	80				
5	75				
5	124				
6	30				
7	79				
7	161				
8	180				
9	75				
10	140				
11	179				

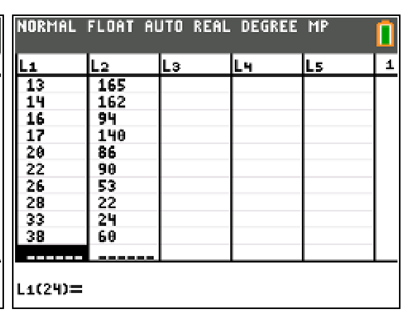
L1={2,4,5,5,6,7,7,8,9,10,11



NORMAL FLOAT AUTO REAL DEGREE MP

L1	L2	L3	L4	L5	1
11	250				
12	169				
13	165				
14	162				
16	94				
17	140				
20	86				
22	90				
26	53				
28	22				
33	24				

L1(22)= 33



NORMAL FLOAT AUTO REAL DEGREE MP

L1	L2	L3	L4	L5	1
13	165				
14	162				
16	94				
17	140				
20	86				
22	90				
26	53				
28	22				
33	24				
38	60				
------	------				

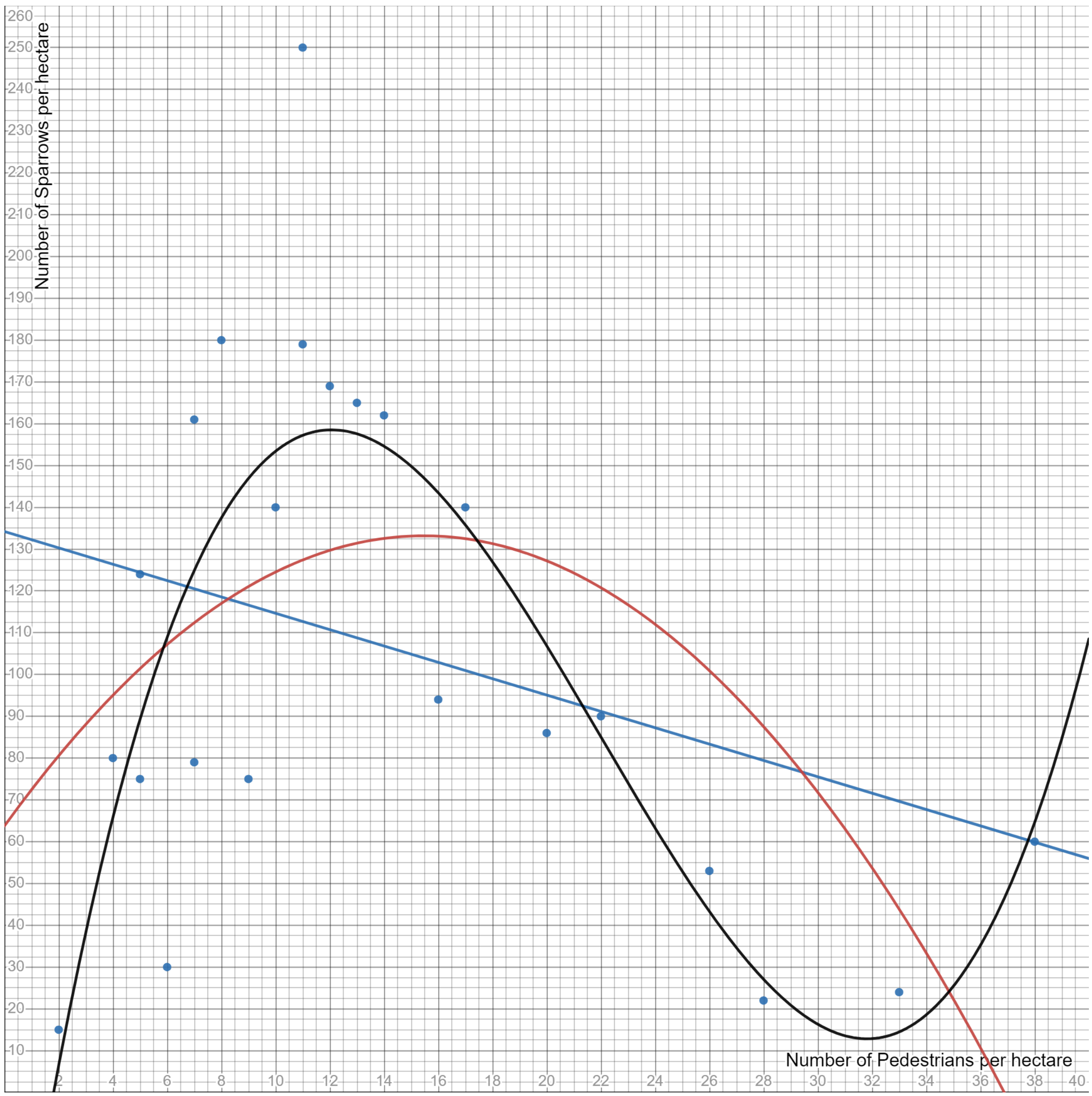
L1(24)=

Linear Model		
x	$y = -1.9560955x + 134.2076062$	y
0	$y = -1.9560955(0) + 134.2076062$	134.2076062
10	$y = -1.9560955(10) + 134.2076062$	114.64665
20	$y = -1.9560955(20) + 134.2076062$	95.085696
30	$y = -1.9560955(30) + 134.2076062$	75.524741
68.6099457	$y = -1.9560955(68.6099457) + 134.2076062$	0

Quadratic Model		
x	$y = -0.2903618283x^2 + 8.980122338x + 63.74853866$	y
0	$y = -0.2903618283(0)^2 + 8.980122338(0) + 63.74853866$	63.74853866
10	$y = -0.2903618283(10)^2 + 8.980122338(10) + 63.74853866$	114.64665
15.463677	$y = -0.2903618283(15.463677)^2 + 8.980122338(15.463677) + 63.74853866$	133.18139
20	$y = -0.2903618283(20)^2 + 8.980122338(20) + 63.74853866$	95.085696
30	$y = -0.2903618283(30)^2 + 8.980122338(30) + 63.74853866$	75.524741
36.880349	$y = -0.2903618283(36.880349)^2 + 8.980122338(36.880349) + 63.74853866$	0

Cubic Model		
x	$y = 0.037810413x^3 - 2.488988703x^2 + 43.55112405x - 71.00527972$	y
0	$y = 0.037810413(0)^3 - 2.488988703(0)^2 + 43.55112405(0) - 71.00527972$	-71.00527972
10	$y = 0.037810413(10)^3 - 2.488988703(10)^2 + 43.55112405(10) - 71.00527972$	153.4175
12.066499	$y = 0.037810413(12.066499)^3 - 2.488988703(12.066499)^2 + 43.55112405(12.066499) - 71.00527972$	158.53519
20	$y = 0.037810413(20)^3 - 2.488988703(20)^2 + 43.55112405(20) - 71.00527972$	106.90502
30	$y = 0.037810413(30)^3 - 2.488988703(30)^2 + 43.55112405(30) - 71.00527972$	16.31976
31.818911	$y = 0.037810413(31.818911)^3 - 2.488988703(31.818911)^2 + 43.55112405(31.818911) - 71.00527972$	12.840913
40	$y = 0.037810413(40)^3 - 2.488988703(40)^2 + 43.55112405(40) - 71.00527972$	108.52419

From Desmos (Higher Quality):



7. Consider the linear model:

a. Find and interpret the intercepts

$y = -1.9560955x + 134.2076062$

The y-intercept, 134.2076062, can be interpreted as, when there are 0 pedestrians per hectare present in the wooded park, there are around 134 sparrows per hectare.

The x-intercept is when the line intercepts the x-axis and $y = 0$:

$$0 = -1.9560955x + 134.2076062 \qquad \text{(Substitute } y = 0\text{)}$$

$$1.9560955x = 134.2076062 \qquad \text{(Add } 1.9560955x \text{ to both sides)}$$

$$x = 134.2076062 / 1.9560955 \qquad \text{(Divide both sides by } 1.9560955\text{)}$$

$$x = 68.6099457 \cong 69$$

Confirmed with a graphing calculator that the x-intercept is 68.6099457.

The x-intercept, 68.6099457, can be interpreted as, when there are about 69 pedestrians per hectare present, there will be no sparrows per hectare.

b. Interpret the slope

$y = -1.9560955x + 134.2076062$

The slope of the linear model, -1.9560955, is the predicted rate at which the number of sparrows per hectare changes for each additional pedestrian per hectare. The negative slope indicates that the number of sparrows per hectare decreases with an increase in the number of pedestrians per hectare. According to the model, approximately 2 fewer sparrows per hectare are observed for each new pedestrian per hectare.

c. What is realistic about this model? What is unrealistic?

The realistic part about the linear model is that it suggests that the number of sparrows per hectare will decrease proportionally for each additional pedestrian per hectare. This is perhaps due to the natural tendency of the sparrows to be frightened away by humans.

However, the model is unrealistic because of its lack of variability. It is possible that some pedestrians will feed the birds with bread crumbs or other food items, which would actually increase the number of sparrows per hectare as the number of pedestrians per hectare increases. Noticeably, the linear model does not follow the data provided well, with an R^2 of only about 0.093.

8. Consider the quadratic model.

a. Find and interpret the intercepts

$$y = -0.2903618283x^2 + 8.980122338x + 63.74853866$$

The y-intercept, 63.74853866, can be interpreted as, when there are no pedestrians per hectare, there are around 63 sparrows sighted per hectare.

The x-intercept is when the line intercepts the x-axis and $y = 0$. Using the graphing calculator, the x-intercepts are $x = -5.952997$ and $x = 36.880349$. However, the number of people per hectare cannot be negative, so $x = -5.952997$ is an extraneous solution and can be discarded.

The x-intercept, 36.880349, can be interpreted as, when there are approximately 37 pedestrians per hectare, there are no sparrows per hectare.

b. Find the maximum value of a function. What does it represent?

$\boldsymbol{y = -0.2903618283x^2 + 8.980122338x + 63.74853866}$

Determine the x-value of the vertex:

$x = -b / 2a$	(Axis of symmetry formula)
$x = -8.980122338 / 2\,(-0.2903618283)$	(Substitute values)
$x = -8.980122338 / -0.5807236566$	(Simplify the denominator)
$x = 15.46367577 \cong 15$	(Simplify)

Determine the y-value of the vertex:

$$y = -0.2903618283\,(-15.46367577)^2 + 8.980122338\,(-15.46367577) + 63.74853866$$

$$y = 133.18139 \cong 133$$

Confirmed with a graphing calculator that the vertex is (15.46368, 133.18139).

The maximum value of the quadratic model is around 133 sparrows per hectare. It represents a point where the number of sparrows observed per hectare is the maximum, probably because the sparrows are not yet scared by the approximately 15 pedestrians per hectare, but are no longer attracted by additional pedestrians with food.

c. What is realistic about this model? What is unrealistic?

The realistic part about the quadratic model is that it acknowledges that the sparrow count will not rise infinitely. Additionally, the model having a maximum number of sparrows per hectare is realistic as well, because the number of sparrows may not be the highest when there are 0 pedestrians. Most humans carry food items with them in a park, which attracts sparrows and would explain why the number of sparrows observed per hectare initially rises to a maximum as the sparrows are attracted to the food. However, once more pedestrians per hectare appear, the sparrows are possibly frightened by the noise or afraid of the number of people. Hence, the sparrows will begin to disperse until there are none left in the area because the number of pedestrians per hectare is too high.

However, the unrealistic part is how early the quadratic model suggests by the x-intercept that the sparrows per hectare will disappear completely, at around 37 pedestrians per hectare. There is already a point in the collected data at (38, 60), which means 60 sparrows per hectare were observed when 38 pedestrians per hectare were present, that counters this model.

9. Consider the cubic model.

a. Find and interpret the intercepts

$\boldsymbol{y = 0.037810413x^3 - 2.488988703x^2 + 43.55112405x - 71.00527972}$

The y-intercept is below 0, at -71.005, but it is impossible to have a negative number of sparrows per hectare. In the real world, this can be taken to mean that when there are no pedestrians per hectare, there are no sparrows per hectare.

The x-intercept is when the line intercepts the x-axis and $y = 0$. Using a graphing calculator, the only x-intercept is $x = 1.8130859 \cong 2$.

The x-intercept represents that, when there are around 2 pedestrians per hectare, there are no sparrows per hectare present.

b. Identify the points where the local maximum and local minimum occur. What do they represent?

The local maximum of the cubic model is at (12.066498, 158.53519), which represents the first turning point of the function when the point is higher than all nearby points. It shows that when there are about 12 pedestrians per hectare, the greatest number of sparrows per hectare of around 158 are present. This is possibly due to the fact that the pedestrians have food items, and are attempting to attract birds without scaring them away.

The local minimum of the cubic model is at (31.818913, 12.840913), which represents the second turning point of the function when the point is lower than all nearby points. It shows that when there are about 32 pedestrians per hectare, the least number of sparrows per hectare of around 13 are present. This is possibly due to the fact that the number of humans in the hectare is too great, and the birds are too frightened to enter the park. Somehow, beyond that point, the number of sparrows per hectare continues to rise with an increase of the number of pedestrians per hectare.

c. What is realistic about this model? What is unrealistic?

For one, it follows the general arrangement of the scatter plot, and has the greatest coefficient of determination, 0.63, showing that it fits the observed data the best. It also accounts for pedestrians trying to interact with the sparrows, as well as pedestrians trying to frighten the sparrows. Most humans carry food items with them in a park, which attracts sparrows and would explain why the number of sparrows observed per hectare initially rises to a local maximum, as the sparrows are attracted to the food. However, once more pedestrians per hectare appear, the sparrows are possibly frightened by the noise or afraid of the number of people. Hence, the sparrows will begin to disperse for each additional pedestrian up until the local minimum, which is also realistic, as it assumes there will still be some daring sparrows who remain in the area.

However, it is unrealistic because, beyond the local minimum, the model predicts that the number of sparrows per hectare will continue to rise, and there is no realistic reason to assume that. The rising end behavior of the cubic model is based on 2 data points only, which is not sufficient to confidently predict the end behavior. The model also assumes, without basis, that there will be no sparrows per hectare until more than 2 pedestrians per hectare appear, which seems doubtful.

10. The researchers chose a quadratic model for the data. Why do you think they chose a quadratic model even though the coefficient of determination is not the closest to 1 of all the models? Which model would you have chosen? Explain your reasoning.

The researchers most likely chose the quadratic model due to the fact that most models are made for explaining the collected data as well as for predicting. While the cubic model's coefficient of determination was the highest, it was incredibly unlikely for the sparrow count to rise beyond the local minimum indefinitely. While the quadratic model is not the one with the best fit for the data, it is more likely to give realistic predictions. However, most of these problems could be solved by taking more observations, especially of pedestrian per hectare counts greater than 38.

If I were allowed to restrict the domain to $2 \leq x \leq 32$, where x is the number of pedestrians per hectare, I would choose the cubic model to the local minimum of 32 pedestrians per hectare, as it is more likely to give a better prediction than the quadratic model.

However, if the domain were not limited, I would follow the scientists' example and use the quadratic model, as it provides the best fit considering real world explanations (for example, the number of sparrows per hectare cannot be expected to keep increasing with the number of pedestrians per hectare, as the cubic model suggests). Notably, none of the models can be used to predict values too far from the set of observed data points and more data would be needed to refine the model.