

Examples of Tasks from CCSS Edition Course 2, Unit 4

Getting Started

The tasks below are selected with the intent of presenting key ideas and skills. **Not every answer is complete**, so that teachers can still assign these questions and expect students to finish the tasks. If you are working with your student on homework, please use these solutions with the intention of increasing student understanding and independence. A list of questions to use as you work together, prepared in [English](#) and [Spanish](#), is available. Encourage students to refer to their class notes and Math Toolkit entries for assistance. Comments in red type are not part of the solution.

As you read these selected homework tasks and solutions, you will notice that some very sophisticated communication skills are expected. Students develop these over time. This is the standard for which to strive. See [Research on Communication](#).

The [Statistics and Probability](#) page might help you follow the conceptual development of the ideas you see in these examples.

Main Mathematical Goals for Unit 4

Upon completion of this unit, students should be able to:

- describe the shape of a cloud of points on a scatterplot and to describe the association between the two variables. (describir la forma de una nube de puntos en un diagrama de dispersión y describir la asociación entre las dos variables.)
- interpret the coefficients of the regression equation, learn some properties of the regression line, and understand that a regression line is an appropriate way to summarize the bivariate relationship only if the points form an elliptical cloud. (interpretar los coeficientes de la ecuación de regresión, aprender algunas propiedades de la línea de regresión, y entender que una línea de regresión es una forma adecuada para resumir la relación bivariado sólo si los puntos forman una nube elíptica.)
- compute and interpret Pearson’s correlation and to understand that a strong correlation does not imply that one variable causes the other. (calcular e interpretar la correlación de Pearson y entender que una correlación fuerte no implica que una variable causa el otro.)
- determine whether a point is influential on the correlation and on the equation of the least squares regression line. (determinar si un punto es influyente en la correlación y en la ecuación de la línea de regresión de los mínimos cuadrados.)

What Solutions are Available?

Lesson 1: Investigation 1—Applications Task 3 (p. 271), Connections Task 7 (p. 275),
Review Task 18 (p. 278)

Investigation 2—Applications Task 5 (p. 273), Connections Task 9 (p. 275),
Review Task 19 (p. 278)

Lesson 2: Investigation 1—Applications Task 1 (p. 305), Connections Task 11 (p. 313),
Reflections Task 18 (p. 316), Review Task 28 (p. 320)

Investigation 2—Applications Task 3 (p. 307), Review Task 30 (p. 321)

Investigation 3—Applications Task 6 (p. 310), Connections Task 15 (p. 315),
Reflections Task 20 (p. 316), Review Task 32 (p. 321)

Investigation 4—Applications Task 7 (p. 310), Review Task 34 (p. 321)

Selected Homework Tasks and Expected Solutions

(These solutions are for tasks in the CCSS Edition book.

For homework tasks in books with earlier copyright dates, see [Helping with Homework](#).)

Lesson 1, Investigation 1, Applications Task 3 (p. 271)

- a. Two possible observations are (dos posibles observaciones son las siguientes):

There is a moderate positive association; that is, if one country had a larger population than another country in 2010, it tends to be projected to have a larger population in 2050. (Existe una asociación positiva moderada, es decir, si un país tuvo una población mayor que otro país en el año 2010, tiende a ser proyectado a tener una población mayor en 2050.)

Both Brazil and Nigeria show the largest changes in rank, moving down 3 and up 3, respectively. (Brazil y Nigeria muestran los cambios más grandes en el rango, moviendo 3 para arriba y 3 para abajo respectivamente.)

- b. *Hint:* To answer this question, look at the table on page 261 in Problem 3. That problem walks you through how to find the sum of the differences squared. Then, use that answer and Spearman's rank correlation, r_s , formula on page 261 to calculate the answer. (*Pista:* Para responder a esta pregunta, mire la tabla en la página 261, Problema 3. Ese problema le guía a través de la forma de encontrar la suma de las diferencias cuadradas. Luego, utilice esa respuesta y el rango de la correlación de Spearman, r_s , fórmula en la página 261 para calcular la respuesta.)

To be completed by the student. (Para ser completado por el estudiante.)

- c. You would expect the correlation between the 2010 ranking and the projected 2025 ranking to be larger than that between 2010 and 2050 because there is less time for population patterns to change. For example, the four largest countries maintained the same rank from 2010 to 2025 but are expected to change by 2050. (Usted esperaría que la correlación entre la clasificación de 2010 y la clasificación proyectada para 2025 ser más grande que entre 2010 y 2050 porque hay menos tiempo para los patrones de la población cambiar. Por ejemplo, los cuatro países más grandes mantienen el mismo rango de 2010 a 2025, pero se espera que cambien en el año 2050.)

The rank correlation between the 2010 ranking and the 2025 projected ranking is 0.939. (El rango de correlación entre la clasificación de 2010 y la clasificación proyectada de 2025 es 0,903.)

Lesson 1, Investigation 1, Connections Task 7 (p. 275)

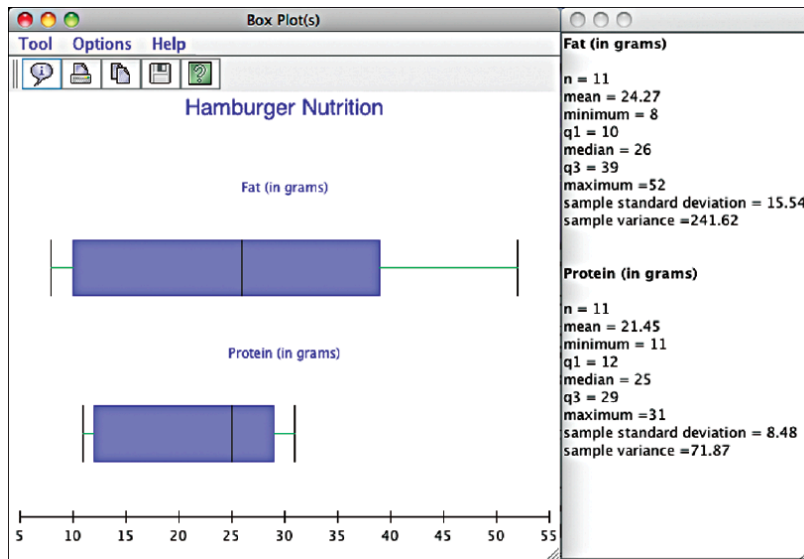
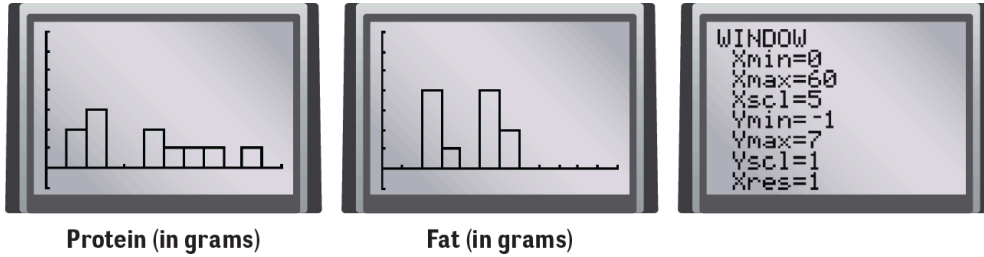
The distance formula was developed in Unit 3 of this course. Your student should have this formula in their toolkit. (La fórmula de distancia fue desarrollada en la Unidad 3 de este curso. Su estudiante debe tener esta fórmula en su caja de herramientas.)

- a. $\sqrt{(3 - (-5))^2 + (4 - 1)^2} = \sqrt{8^2 + 3^2} \approx 8.54$ units
- b. To be completed by the student. (Para ser completado por el estudiante.)
- c. Both involve a sum of squared differences. Both are meant to measure distance—one between two points in the plane and one between two rankings. (Ambos implican una suma de diferencias cuadradas. Ambos tienen por objeto medir una distancia—uno entre dos puntos en el plano y uno entre dos clasificaciones.)

Lesson 1, Investigation 1, Review Task 18 (p. 278)

These data are located in *CPMP-Tools* under Statistics, Data Analysis, Data>Unit 4 Regression and Correlation> Hamburger Nutrition II. (Se encuentra estos datos en *CPMP-Tools* debajo de Statistics, Data Analysis, Data>Unit 4 Regression and Correlation> Hamburger Nutrition II.)

- a. There are only 11 hamburgers, so students might see the most detail by making dot plots or histograms. (Solo hay 11 hamburguesas, entonces los estudiantes podrían ver los detalles por el uso de un gráfico de puntos o un histograma.)



- b-d. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 1, Investigation 1, Review Task 19 (p. 278)

- a. Two possible dimensions for the cloth are 2 yards by 1 yard, or 0.25 yards by 8 yards. (Dos dimensiones posibles para la tela son 2 yardas por 1 yarda o 0,25 yardas por 8 yardas.)
- b-c. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 1, Investigation 2, Applications Task 5 (p. 273)

Hint: As you do this task, remember that the city with the lowest crime rate, Minneapolis, has a rank of 1. (*Pista:* Cuando usted haga esta tarea, recuerde que la ciudad con la menor tasa de delincuencia, Minneapolis, tiene un rango de 1.)

a. Health Care is the variable graphed on the x -axis in the third column. It is the variable graphed on the y -axis in the third row. (El seguro médico es la variable trazada en el eje x en la tercera columna. Es la variable trazada en el eje y en la tercera fila.)

b, d, f. To be completed by the student. (Para ser completado por el estudiante.)

c. Education and health care have the strongest positive correlation. Some possible reasons: Cities that fund education as a high priority may also fund health care more generously. People who are better educated have the money to pay for better health care. People who are better educated insist on better health care. (La educación y el seguro médico tienen la correlación positiva más fuerte. Algunas posibles razones: las ciudades que financian la educación como alta prioridad también financian el seguro médico con más generosidad. Las personas con una nivel más alta de educación tienen el dinero para pagar un seguro médico mejor. Las personas que son más educadas insisten en un seguro médico mejor.)

e. *Hint:* To find the missing rank correlation, r_s , use Spearman's formula found on page 261 in the student book. (*Pista:* Para encontrar el rango de correclación que falta, utilice la fórmula de Spearman encontrada en la página 261 en el texto del estudiante.)

To be completed by the student. (Para ser completado por el estudiante.)

g. The entry in row i and column j of the correlation matrix gives the rank correlation between the two variables graphed in the scatterplot matrix in row i and column j . (La entrada en la fila i y la columna j de la matriz de correlación dan el rango de correlación entre las dos variables en la matriz de dispersión en la fila i y la columna j .)

Lesson 1, Investigation 2, Connections Task 9 (p. 275)

Students are looking for a pattern that will find the value for any number of variables k . (Los estudiantes están buscando un patrón que encontraría el valor para cualquier número de las variables de k .)

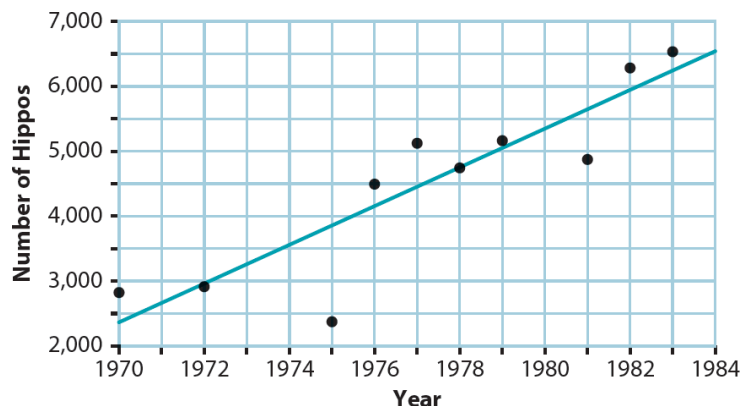
To find the number of scatterplot matrices for 3 variables, you need to calculate the area of a 3×3 square, which is 9, however, the main diagonal (the three boxes going from top-right to bottom-left) are not scatterplots because they are comparing a variable to itself. So to find the number of scatterplots, you would go $3^2 - 3$, or 6 scatterplots. (Para encontrar el número de las matrices de dispersión para 3 variables, es necesario calcular la área de un cuadro de 3×3 , que es 9, sin embargo, la diagonal principal (las tres cajas que van desde la parte superior derecha hasta la parte inferior izquierda) no son diagramas de dispersión, ya que están comparando una variable a sí misma. Por lo tanto, para encontrar el número de diagramas de dispersión, usted escribirá $3^2 - 3$, o 6 diagramas de dispersión.)

The variables of 4, 5, and k are left for the student to complete. (Las variables de 4, 5, y la k están dejadas para que el estudiante las completa.)

Lesson 2, Investigation 1, Applications Task 1 (p. 305)

The data for this task has been included in the *CPMP-Tools* software. Students can access the data and find the regression equation or they can put the data into their calculator. (Los datos para esta tarea están incluidos en el software *CPMP-Tools*. Los estudiantes pueden acceder a los datos y encontrar la ecuación de regresión o pueden poner los datos en su calculadora.)

- a. The regression equation is _____.





- b. To be completed by the student. (Para ser completado por el estudiante.)
- c. *Hint:* To find the residual, you need to take the actual value for 1975 and subtract the predicted value from that. (*Pista:* Para encontrar el residuo, tendrá que tomar el valor real para 1975 y restar el valor predicho de eso.)

To find the residual for 1975, take a good estimate from the scatterplot of the actual point (2,300) minus the predicted (3,800) to get $-1,500$. Use the regression equation to evaluate y when $x = 1975$ to find the predicted number on hippos. Now look at the table on page 305 and find that the actual number of hippos in 1975 is 2,342. The residual is $-1,495$. This means that there were 1,495 fewer hippos in 1975 than were predicted by the regression line. (Para encontrar el residuo para 1975, tome una buena estimación del diagrama de digresión del punto actual (2.300) menos el previsto (3.800) para obtener -1.500 . Utilice la ecuación de regresión para evaluar y cuando $x = 1975$ para encontrar el número predicho de los hipopótamos. Ahora mire la tabla en la página 305 y encuentre que el número actual de hipopótamos en 1975 es 2.342. El residuo es -1.495 . Esto significa que había 1.495 menos de hipopótamos en 1975 que la línea de regression predijo.)

- d. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 1, Connections Task 11 (p. 313)

You can get a nice picture of the squared error by using *CPMP-Tools*. Put the data into the spreadsheet and click on this icon . (Usted puede obtener una buena foto del error cuadrado mediante el uso de *CPMP-Tools*. Ponga los datos en la hoja de cálculo y haga clic en este icono )

- a. Use the equation to find the predicted y values for each of the x values in the list. Then find the residuals and the squared residuals. Calculate the sums. The following table shows the computations. (Utilice la ecuación para encontrar los valores de y predichos para cada uno de los valores de x en la lista. Luego encuentre los residuos y los residuos cuadrados. Calcule las cantidades. La tabla siguiente muestra los cálculos.)

x	y	Predicted $y = x + 0.75$	Residual	Squared Residual
1	3	1.75	1.25	1.5625
2	2	2.75	-0.75	0.5625
3	5	3.75	1.25	1.5625
6	5	6.75	-1.75	3.0625
Total			0	6.75

b–c. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 1, Reflections Task 18 (p. 316)

- a.** This question has to do with the true interpretation of slope. At first glance, students may believe that all are correct, but with closer analysis they can begin to eliminate two of the choices. The definition of slope is the rate of change between any two points; the change in the y values divided by the change in the x values. So, the correct interpretation of slope needs to compare the x and y values of the two points. (Esta cuestión tiene que ver con la verdadera interpretación de la pendiente. A primera vista, los estudiantes pueden creer que todos son correctos, pero con un análisis más cerca puedan comenzar a eliminar dos de las opciones. La definición de la pendiente es la tasa de cambio entre dos puntos cualesquiera, el cambio en los valores y dividido por el cambio de los valores x . Por lo tanto, la interpretación correcta de la pendiente debe comparar los valores de x y y de los dos puntos.)

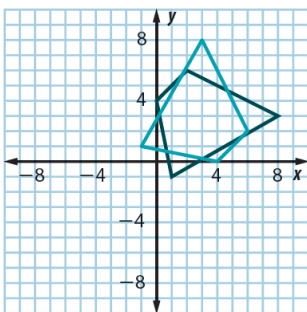
The first interpretation is not correct because it does not compare two students, but rather one student over time. The remainder of the solution is left to the student. (La primera interpretación no es correcta porque no compara los dos estudiantes, sino más bien un estudiante a través del tiempo. El resto de la solución es para ser completado por el estudiante.)

- b.** *Hint:* Look at the scatterplot and estimate the values of the residuals. Then estimate the error in prediction. (Vea el diagrama de dispersión y estime los valores de los residuos. Luego, estime la equivocación en la predicción.)

To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 1, Review Task 28 (p. 320)

- a.** $\begin{bmatrix} 4 & 6 & 3 & -1 \\ 0 & 2 & 8 & 1 \end{bmatrix}$



b–c. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 2, Applications Task 3 (p. 307)

These data are located in *CPMP-Tools* under Statistics, Data Analysis, Data>Unit 4 Regression and Correlation>Canines. (Se encuentran estos datos en *CPMP-Tools* debajo de “Statistics, Data Analysis, Data>Unit 4 Regression” y “Correlation>Canines”.)

- a. Yes, as there is no curvature, however, the pattern is not very elliptical. Another worry is the influence of the outlier. (Sí, no hay curvatura, pero, el patrón no es muy elíptico. Otra duda es la influencia de un valor extremo.)
- b. *Hint*: Remember when you are telling what the slope indicates, you need to represent it as a comparison of two pieces of data. (Recuerde cuando usted está diciendo lo que indica la pendiente, hay que representarla como una comparación de dos piezas de datos.)
To be completed by the student. (Para ser completado por el estudiante.)
- c. The predicted lifespan of a Cape fox is 14.712 years. There are two reasons not to have much faith in this prediction. (La esperanza de vida prevista para un Zorro Chama es 14.712 años. Hay dos razones no tener mucha confianza en esta predicción.)
To be completed by the student. (Para ser completado por el estudiante.)
- d. The error in prediction is quite large: *observed maximum longevity – predicted maximum longevity* = $14.712 - 7 = 7.712$ years. (El error en la predicción está bastante grande: *la longevidad máxima observada – la longevidad máxima prevista* = $14.712 - 7 = 7.712$ años.)
- e–f. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 2, Review Task 30 (p. 321)

- a, c, e. To be completed by the student. (Para ser completado por el estudiante.)
- b. $5x(x - 5)$
- d. $6x(3 - 2x)$

Lesson 2, Investigation 3, Applications Task 6 (p. 310)

These data are located in *CPMP-Tools* under Statistics, Data Analysis, Data>Unit 4 Regression and Correlation>Seal Sizes. (Se encuentran estos datos en *CPMP-Tools* debajo de “Statistics, Data Analysis, Data>Unit 4 Regression” y “Correlation>Seal Sizes”.)

- a. A good estimate would be between 0.7 and 0.9. (Una buena estimación sería entre 0,7 y 0,9.)
- b. If it is allowed to use a tool to find the correlation coefficient, you can find this on *CPMP-Tools*. If you need to find it by hand, then use the formula on page 291. To help you check your calculations, the mean of $x = 8.76$, mean of $y = 795$, $s_x = 1.324$, $s_y = 375.169$, and $n = 5$. (Si es permitido use una herramienta para encontrar el coeficiente de correlación, usted la puede encontrar en *CPMP-Tools*. Si usted necesita encontrarlo a mano y, luego, utilizar la fórmula en la página 291. Para ayudarlo a comprobar sus cálculos, la media de $x = 8,76$, con una media de $y = 795$, $s_x = 1,324$, $s_y = 375,169$, y $n = 5$.)

The remainder is to be completed by the student. (El resto es para ser completado por el estudiante.)

- c. From the plot, it is clear that a curve, not a line, might be a better model for the sizes of seals. This makes sense as, from the geometry of the situation, you would expect the relationship to be cubic because seals are somewhat similar in shape and linear measures are related by a cubic equation to measures depending on volume, such as weight. (Del gráfico, es evidente que una curva, no una línea, sería un mejor modelo para los tamaños de las focas. Esto tiene sentido ya que, a partir de la geometría de la situación, que se pueden esperar que la relación sea cúbica, porque las focas son algo semejantes en forma y medidas lineales están relacionadas por una ecuación cúbica a las medidas en función de volumen, como el peso.)
- d–e. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 3, Connections Task 15 (p. 315)

Students have transformed measurements in a previous course and found that if you multiply each data point by the same value, then the new mean and standard deviation are also multiplied by that value. However, if the same value is added to each piece of data, then the new mean is the previous mean plus the added value, but the standard deviation is unchanged.

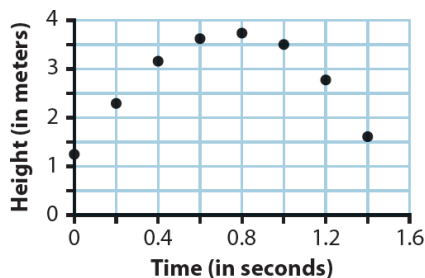
- a. Multiply each mean height by 2.54.
- b. The value of r does not change if values are all multiplied by the same constant.
- c. The y -intercept is multiplied by 2.54, but the slope does not change.
- d. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 3, Reflections Task 20 (p. 316)

This task serves as a warning to always plot the data you are working with before you conclude that a line is or is not the best fit for the data. The correlation coefficient can be weak or strong, but until the graph is plotted it is not obvious whether a line will fit the data well or not. In this particular task, the plot shows the data are going to be best fit by some type of curve. (Esta tarea sirve como una advertencia para siempre poner los datos en un gráfico con los cuales usted está trabajando antes de concluir que una línea es o no es el mejor ajuste para los datos. El coeficiente de correlación puede ser débil o fuerte, pero hasta que el gráfico este hecho no es obvio si una línea se ajusta bien a los datos o no. En esta tarea, el gráfico muestra que los datos serán mejores con una línea con curva.)

- a. To be completed by the student. (Para ser completado por el estudiante.)
- b. i. When the time is equal to zero, the height is 1.1 meters, so $h_0 = 1.1$. (cuando el tiempo es igual a cero, la altura es de 1,1 metros, entonces $h_0 = 1.1$.)
- ii. To be completed by the student. (Para ser completado por el estudiante.)
- iii. $h = 1.1 + 7.3t - 4.9t^2$
- iv. To be completed by the student. (Para ser completado por el estudiante.)
- c. As time increases, the height increases, reaches a maximum, then decreases. The equation is quadratic and the pattern on the graph below is parabolic. The correlation is very weak because it only indicates the degree to which points cluster about a line; it gives no indication of the degree to which they follow another relationship. There is no reason to expect a high correlation using a formula that measures how closely the points cluster about a line rather than about a parabola. (Mientras el tiempo aumenta, la altura aumenta, alcanza un máximo y luego disminuye. La ecuación es cuadrática y el patrón en el gráfico abajo es parabólica. La correlación es muy débil, ya que sólo indica el grado en

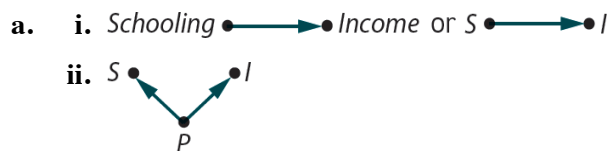
que los puntos agrupan en una línea, que no da ninguna indicación de la medida en que siga otra relación. No hay ninguna razón para esperar una correlación alta utilizando una fórmula que mide cuán cerca los puntos agrupan sobre una línea en lugar de una parábola.)



Lesson 2, Investigation 3, Review Task 32 (p. 321)

- a. Equivalent
- b. Equivalent
- c–f. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 4, Applications Task 7 (p. 310)



- b. i. The researchers chose twins for the study. The twins in the study probably grew up in the same household and so had identical economic status. By comparing the income of the twin with more education to that of the twin with less education, researchers can eliminate the effect of the parent’s economic status. Lurking variables controlled include such things as quality of parenting, quality of local schools, number of siblings, and other variables, as well as economic status. (Los investigadores eligieron gemelos para el estudio. Los gemelos en el estudio probablemente crecieron en el mismo hogar y así tuvieron una situación económica igual. Al comparar los ingresos del gemelo con más educación al gemelo con menos educación, los investigadores pueden eliminar el efecto del estado económico de los papás. Las variables merodeadas controladas incluyen tales cosas como la calidad de la crianza de los hijos, la calidad de las escuelas locales, el número de hermanos, y otras variables, y también la situación económica.)
- ii. To be completed by the student. (Para ser completado por el estudiante.)

Lesson 2, Investigation 4, Review Task 34 (p. 321)

- a. About 83.4 people per square mile (Unos 83.4 personas por milla cuadrada)
- b–c. To be completed by the student. (Para ser completado por el estudiante.)