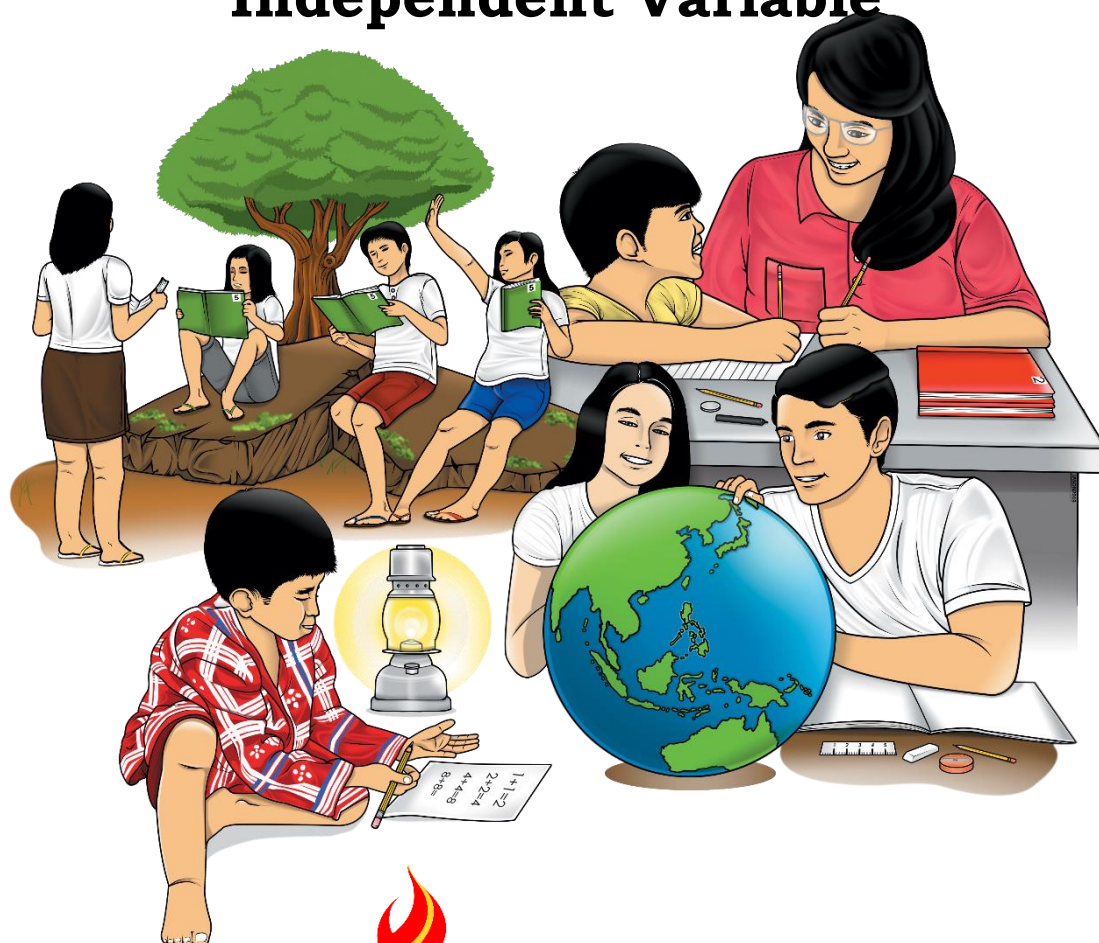**DepED**
DEPARTMENT OF EDUCATION

Statistics and Probability
Quarter 4 – Module 23
Predicting the Value of Dependent Variable Given the Value of Independent Variable

ALTERNATIVE DELIVERY MODE
ADM

GOVERNMENT PROPERTY
NOT FOR SALE

Statistics and Probability – Grade 11
Alternative Delivery Mode
Quarter 4 – Module 23: Predicting the Value of Dependent Variable Given the Value of Independent Variable
First Edition, 2021

Published by the Department of Education
Secretary: Leonor Magtolis Briones
Undersecretary: Diosdado M. San Antonio

Printed in the Philippines by _____

Department of Education – Region 4A CALABARZON

Statistics and Probability
Quarter 4 – Module 23
Predicting the Value of Dependent Variable Given the Value of Independent Variable

DepED
DEPARTMENT OF EDUCATION

Introductory Message

This Self-Learning Module (SLM) is prepared so that you, our dear learners, can continue your studies and learn while at home. Activities, questions, directions, exercises, and discussions are carefully stated for you to understand each lesson.

Each SLM is composed of different parts. Each part shall guide you step-by-step as you discover and understand the lesson prepared for you.

Pre-tests are provided to measure your prior knowledge on lessons in each SLM. This will tell you if you need to proceed on completing this module or if you need to ask your facilitator or your teacher's assistance for better understanding of the lesson. At the end of each module, you need to answer the post-test to self-check your learning. Answer keys are provided for each activity and test. We trust that you will be honest in using these.

In addition to the material in the main text, Notes to the Teacher are also provided to our facilitators and parents for strategies and reminders on how they can best help you on your home-based learning.

Please use this module with care. Do not put unnecessary marks on any part of this SLM. Use a separate sheet of paper in answering the exercises and tests. And read the instructions carefully before performing each task.

If you have any questions in using this SLM or any difficulty in answering the tasks in this module, do not hesitate to consult your teacher or facilitator.

Thank you.

What I Need to Know

In previous modules, you became familiar with the equation of the regression line. Your knowledge about the regression line, its equation, slope, and y-intercept will aid you in learning the content of this module.

This module will help you learn how to formulate predictions about dependent variables based on independent variables. Also, you will learn to apply previous learnings you gained about the nature of variables. Furthermore, this module gives you an opportunity to link and unify concepts learned about regression analysis and utilize these in the prediction and estimation of values based on the equation of a regression line. In time, you will be able to utilize your knowledge and skills in analysis and solving of regression analysis problems.

After going through this module, you are expected to:

1. explain the use of the regression equation in predicting the value of the dependent variable;
2. estimate the value of dependent variable from the value of the independent variable; and
3. reflect on the importance of making predictions in decision-making.

Before you proceed to the lesson, make sure to answer first the questions in the next page (*What I Know*).
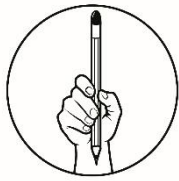
What I Know

Choose the letter of the best answer. Write the chosen letter on a separate sheet of paper.

For numbers 1 and 2, refer to the following:

Research shows that the academic performance of a person is related to his/her grit. Likewise, grit is related to a person's self-efficacy.

1. Based on the research, which of the variables corresponds to the dependent variable in relation to grit?
   a. grit
   b. self-efficacy
   c. emotional intelligence
   d. academic performance

2. Based on the research, which of the variables corresponds to the independent variable in relation to academic performance?
   a. grit
   b. self-efficacy
   c. emotional intelligence
   d. performance

For numbers 3 to 5, refer to the situation below:

Even before pandemic, Annie truly values health. As a fitness instructor and at the same time a researcher, she conducted a regression analysis between the weight in kilograms of her friends and the number of hours they exercise in a week. She found out that the regression equation that will predict the weight from number of hours spent exercising is $\hat{y} = 65 - 2x$.

3. Which among the following is the independent variable?
   a. math grades
   b. hours of exercise
   c. number of friends
   d. number of absences

4. What is the average weight of a friend who does not exercise??
   a. 45   b. 51   c. 59   d. 65

5. What is the average weight of a friend who exercises 7 hours a week?
   a. 45   b. 51   c. 59   d. 65

6. If x is temperature and y is coffee sales, use the following formula to predict the coffee sales on a 34° day: $y = -60x + 6{,}443$.
   a. 4,403   b. 8,483   c. 4,673   d. 6,443

7. The equation of the regression line on cycling is $\hat{y} = 1 + \frac{x}{2}$. How many laps can a bicycle make around the park in 23 minutes?
   a. 12.5 laps
   b. 23 laps
   c. 32 laps
   d. 44 laps

For numbers 8 to 11, refer to the situation below:

Rand is a member of a gym. He pays a monthly fee plus a per-visit fee. The equation representing the monthly amount Rand pays for his membership in the gym per month for x visits is $\hat{y} = 500 + 6x$.

8. What does the slope of the graph of this equation represent?
   a. the monthly membership fees
   b. the amount Rand pays per visit
   c. the total cost at the end of each month
   d. the number of times Rand goes to the gym

9. What does the y-value of the graph of this equation represent?
   a. the monthly membership fees
   b. the amount Rand pays per visit
   c. the total cost at the end of each month
   d. the number of times Rand goes to the gym

10. If Rand goes to the gym eight times in a month, how much would be his monthly fee?
    a. 458 pesos
    b. 542 pesos
    c. 548 pesos
    d. 845 pesos

11. If Rand is unable to visit the gym for a month, how much would he pay for that month?
    a. 450 pesos
    b. 500 pesos
    c. 506 pesos
    d. 456 pesos

For numbers 12 to 15, refer to the situation below:

Students in a group of chess teams submitted data about hours of weekly team practice and the number of matches the team won. The data gathered are represented by this linear regression equation: $\hat{y} = 3.52 + 3.14x$.

12. What is the independent variable?
    a. hours of weekly practice
    b. number of team members
    c. number of matches the team won
    d. increase in the team's number of wins per hour of practice

13. What does the y-value of the graph of this equation represent?
    a. hours of weekly practice
    b. number of team members
    c. number of matches the team won
    d. increase in the team's number of wins per hour of practice

14. What does the slope of the linear regression equation represent?
    a. hours of weekly practice
    b. number of team members
    c. number of matches the team won
    d. increase in the team's number of wins per hour of practice

15. If the team practices 4.5 hours per week next year, which is the best prediction of the number of chess matches the team could expect to win?
    a. about 8   c. about 18
    b. about 12   d. about 24

Lesson 1 — Regression Analysis: Prediction and Estimation

If you possess the ability to predict what happens in the future, you will probably use the ability all the time. Well, you do not need psychic powers to make predictions. Regression equations can be used to make predictions. As you already know, the field of statistics that deals with prediction is called regression analysis.

You have already learned to calculate and interpret the slope and y-intercept. In this lesson, you will learn how to estimate and predict the value of one variable in terms of the other variable of a regression equation.

First, let us try to connect your previous learnings with this new lesson for better understanding of the steps in linear regression. Answer the next activity to check your readiness for this module.
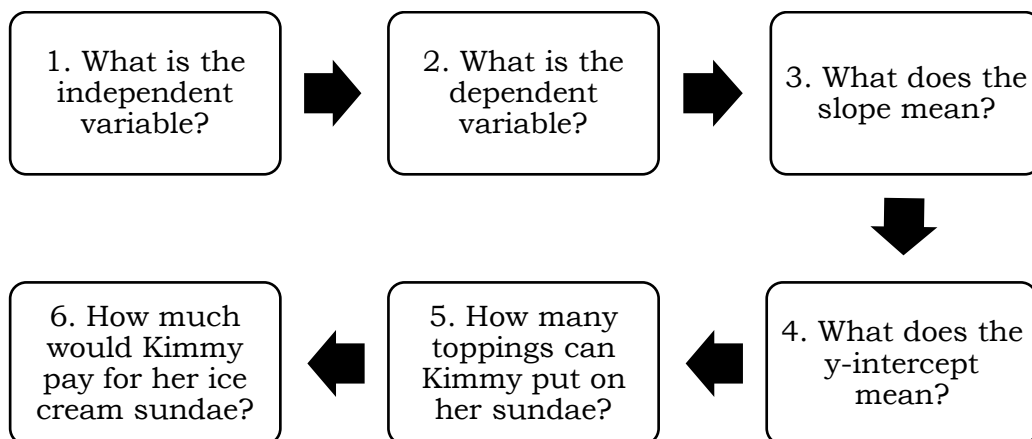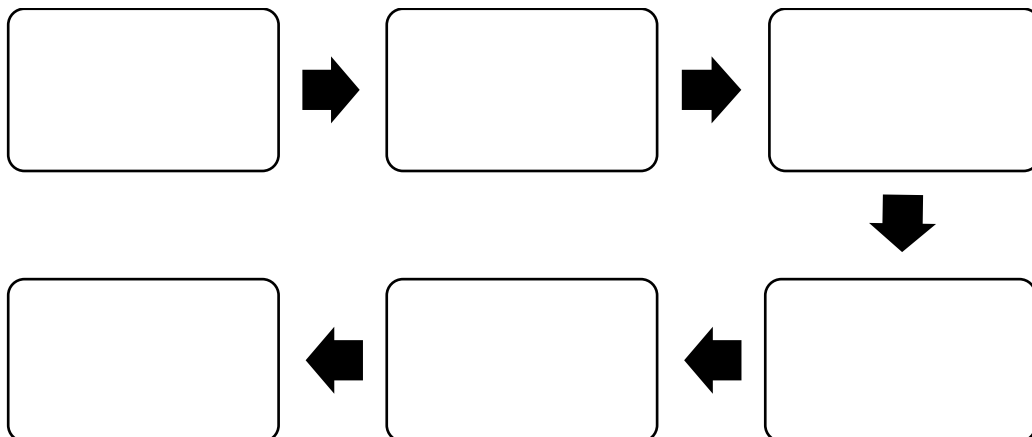
What's In

Read the situation and make a process chart like the given example on a separate sheet of paper. Use the questions as a guide.

Tayabas Ice, a local ice cream store, launches its new ice cream creation called Halo Halo Overload Sundae where customers get to personalize the flavors and toppings on their ice cream sundae. The following regression equation shows the cost of the ice cream sundae with different numbers of toppings: ŷ = 25 + 4x.

If Kimmy has Php50.00, what is the maximum number of toppings she can put on her Halo Halo Overload Sundae? How much would Kimmy pay for her ice cream sundae?

1. What is the independent variable?	→	2. What is the dependent variable?	→	3. What does the slope mean?

↓

6. How much would Kimmy pay for her ice cream sundae?	←	5. How many toppings can Kimmy put on her sundae?	←	4. What does the y-intercept mean?

Answer:

[ ] → [ ] → [ ]

↓

[ ] ← [ ] ← [ ]

In this activity, the concepts you recalled will play a role in learning how to make predictions based on the regression line. You can predict and estimate values that the dependent variable will take based on the value of the independent variable. Also, identifying the slope and y-intercept and using these to interpret the relationship based on the regression equation can further assist you in making predictions accurately.

Now that you have connected prior learnings to better grasp this lesson, please begin by doing the next activity.

Notes to the Teacher

This contains helpful tips or strategies that will help you in guiding the learners.
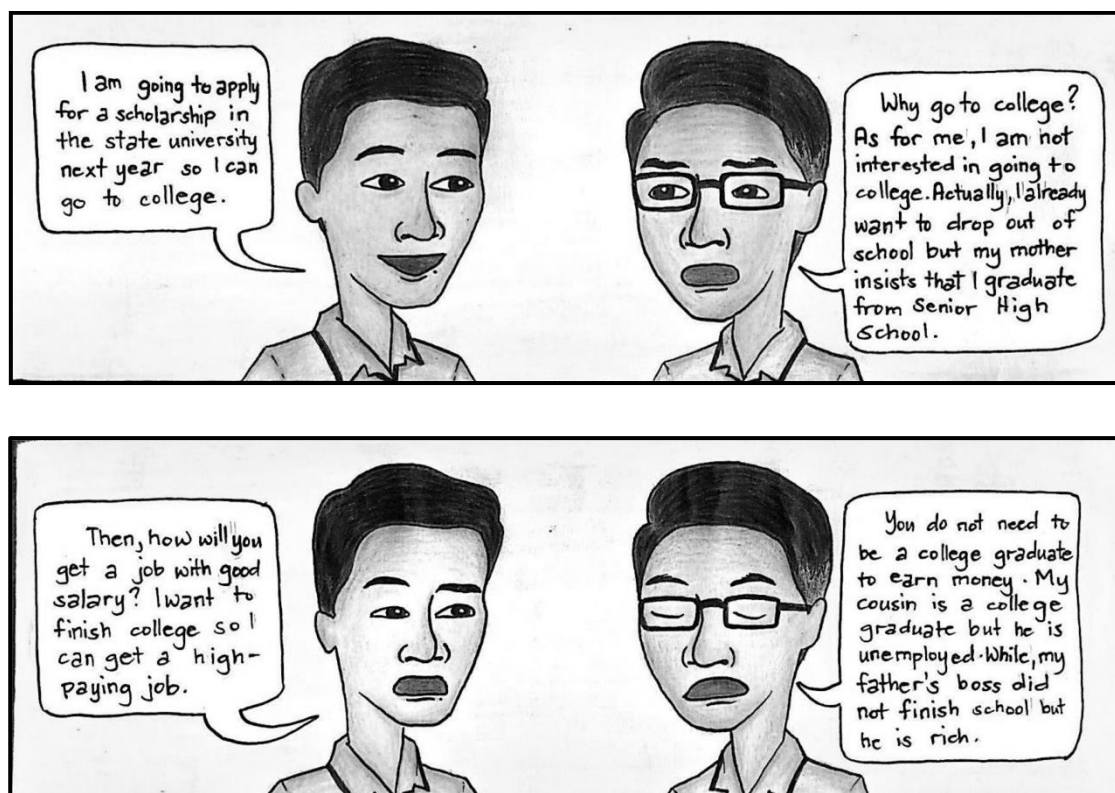
Will Carl Choose College?

Read the given situation and answer the guide questions that follow.

Carl and Michael, both Grade 11 students, are talking about their plans for the future.

Carl starts to rethink about his plans of going to college. Is it true that spending more years in school will not increase the chance of getting a higher salary?

Totally confused, he consults his class adviser. Miss Jean, a Statistics teacher, tells him about a study describing the relationship between years in school and salary. In the study, the regression equation was found to be $\hat{y} = 30x + 100$ where y refers to the person's daily salary and x refers to the number of years the person spent studying in school.

Guide Questions:

1. In the regression equation, what is the independent variable (*x*)?
2. In the regression equation, what is the dependent variable (*y*)?
3. Based on the slope, is the relationship between the variables positive or negative?
4. Can Carl predict his average salary in the future if he will finish college based on the regression equation? How?
5. Can the prediction of Carl's average salary in the future help in his decision making? How?
6. If you were Carl, would you continue your plans of graduating from college based on the given regression equation? Why?

What is It

Regression equations can be used to make predictions. The field of statistics that deals with prediction is called *regression analysis*. In this lesson, you will learn how to estimate and predict the value of one variable in terms of the other variable.

In predicting the dependent variable *y*, just substitute the given value of *x* into the equation $\hat{y} = bx + a$.

Let us look back at the situation in the previous activity:

What might happen if Carl chooses to go to college?

To help Carl decide if he is going to college, he can predict his average daily salary based on the regression equation $y = 30x + 100$.

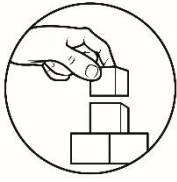How much is the possible salary if Carl takes a vocational course?

If Carl decides to take a 2-year course, his time spent on studying will be 15 years, so *x = 15*.

Solution: $\hat{y} = 30x + 100$ (Substitute 15 into *x*.)

         $\hat{y} = 30(15) + 100$ (Multiply 30 by 15.)

         $\hat{y} = 450 + 100$ (Add.)

         $\hat{y} = 550$ ₱ 550.00 is the average daily salary Carl may earn.

What's More

Activity 1.1 Continuing College with Carl

Answer completely by answering the questions in the *Steps* column.

Based on another study, Carl found that the regression equation is $\hat{y} = 40x + 95$ where y refers to the person's daily salary and x refers to the number of years the person spent studying in school.

How much is the possible salary if Carl finishes with an engineering degree?

	Given/Solution	Steps
		(1) What is the regression equation?
		(2) How many years does it take to finish engineering?
		(3) What is the total number of years Carl spent studying?
		(4) Substitute the answer in (3) into the regression equation (1).
		(5) Solve.
		(6) What is the average daily salary Carl may receive?

Activity 1.2 The Y's from Our X's

In this activity, you will practice your skill in substituting x to the regression equation. Compute the average value of y from each regression equation based on the given value of x.

| 1. $y = 3x + 17$, $x = 6$	2. $y = -2x + 9$, $x = 5$	3. $y = 7x + 2.5$, $x = 8$	4. $y = 3.5x + 6$, $x = 4.5$

Activity 1.3 Solving Substitutions

Substitute the value of the independent variable to predict the value of the dependent variable.

1. The cost of hiring a video camera for a number of days is calculated using the formula Hire Cost = 50 x Number of Days + 350.
   a. Calculate the cost of hiring a video camera for 4 days.
   b. Calculate the cost of hiring a video camera for 9 days.
   c. If the hire cost is Php1,000, how many days was the video camera hired?

2. The time it takes in minutes to cook a chicken is given in the formula:
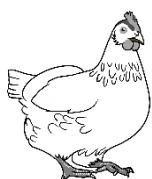
   Time = 30 minutes per kilogram plus 20 minutes

   a. Predict the time it takes to cook a 5kg chicken
   b. Predict the time it takes to cook a 2.5kg chicken

Activity 1.4 Looking into the Future

Predict the value of y (dependent variable) from the given value of x (independent variable). Solve on a separate piece of paper.

1. A student's grade in the first grading period (x) and grade in the second grading period (y) is described by the equation $\hat{y} = 0.5x + 45$. If Jake gets a grade of 90 in the first grading, what grade can he expect in the second grading?

2. A counselor has collected data on the number of times per year a SHS student goes out on a date (x) and the number of hours a student does his/her homework per week (y). She came up with the equation of the regression line: $\hat{y} = 30 - 0.2x$. What is the average number of hours a student gets his/her homework done per week if s/he goes out on 50 dates in a year?

3. The relationship between the number of a student's absences (x) to his/her general average (y) is $\hat{y} = 97 - 2x$. If a student has 5 absences, what is his predicted general average?

4. When smoking cigarettes, one by-product is carbon monoxide. Data are collected to determine if the carbon monoxide emission can be predicted by the nicotine level of the cigarette. Variables are measured in milligrams. The regression equation is $\hat{y} = 3 + 10.3x$. If the amount of nicotine in a cigarette is 4 milligrams, how much carbon monoxide in milligrams will be emitted?

What I Have Learned

Answer the following questions by completing the statements.

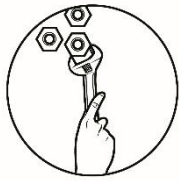1. What mathematical technique did you use to predict the value of y?

    I used _____

2. What difficulties did you encounter in predicting values?

    I had difficulty _____

3. Why is careful thinking important before making a decision?

    It is important because_____

What I Can Do

College Costs???

Investigate the relationship between school year and the cost of tuition fees of the university or college of your choice. Make a data table for the past seven (7) years and compute the slope and y-intercept. Analyze the relationship between year and tuition fee cost. Make interpretations about the slope and y-intercept based on the situation. Estimate the amount that you may need for tuition fees by the time you go to college.

Assessment

Multiple Choice. Choose the letter of the best answer. Write the chosen letter on a separate sheet of paper.

For numbers 1 and 2, refer to the following:

Research shows that the emotional intelligence of a person is related to his/her academic performance. Likewise, academic performance is related to job performance.

1. Based on the research, which of the variables corresponds to the dependent variable in relation to academic performance?
   a. academic intelligence
   b. emotional intelligence
   c. emotional state
   d. job performance

2. Based on the research, which of the variables corresponds to the independent variable in relation to academic performance?
   a. academic intelligence
   b. emotional intelligence
   c. emotional state
   d. job performance

For numbers 3 to 5, refer the situation below:

Kyle conducted a regression analysis between the Math grades of his classmates and the number of times they were absent in the subject. He found that the regression equation that will predict the grade based on the number of absences is $\hat{y} = 98 - 2.5x$.

3. Which variable is the dependent variable?
    a. favorite subject
    b. general average
    c. Math grades
    d. number of absences

4. What is the predicted grade of a student who has no absences?
    a. 93
    b. 95
    c. 97
    d. 98

5. What is the predicted grade of a student who has 10 absences?
    a. 73
    b. 75
    c. 77
    d. 78

6. If x is temperature and y is coffee sales, use the following formula to predict the coffee sales on a 28° day: $y = -60x + 6{,}443$.
    a. 4,403
    b. 8,483
    c. 4,763
    d. 6,443

7. The equation of the regression line on cycling is $\hat{y} = 1 + \frac{x}{2}$. How many laps can a bicycle make around the park in 44 minutes?
    a. 12.5 laps
    b. 23 laps
    c. 32 laps
    d. 44 laps

For numbers 8 to 11, refer to the situation below:

Rand is a member of a gym. He pays a monthly fee plus a per-visit fee. The equation representing the monthly amount Rand pays for his membership in the gym per month for x visits is $\hat{y} = 500 + 6x$.

8. What does the y-value of the graph of this equation represent?
    a. the monthly membership fees
    b. the amount Rand pays per visit
    c. the total cost at the end of each month
    d. the number of times Rand goes to the gym

9. What does the slope of the graph of this equation represent?
    a. the monthly membership fees
    b. the amount Rand pays per visit
    c. the total cost at the end of each month
    d. the number of times Rand goes to the gym

10. If Rand goes to the gym eight times in a month, how much would be his monthly fee?
    a. 458 pesos   c. 548 pesos
    b. 542 pesos   d. 845 pesos

11. If Rand is unable to visit the gym for a month, how much would he pay for that month?
    a. 450 pesos   c. 506 pesos
    b. 500 pesos   d. 456 pesos

For numbers 12 to 15, refer to the situation below:

Students in a group of chess teams submitted data about hours of weekly team practice and the number of matches the team won. The data gathered are represented by this linear regression equation: ŷ = 3.52 + 3.14x.

12. What does the y-value of the graph of this equation represent?
    a. hours of weekly practice
    b. number of team members
    c. number of matches the team won
    d. increase in the team's number of wins per hour of practice

13. What is the independent variable?
    a. hours of weekly practice
    b. number of team members
    c. number of matches the team won
    d. increase in the team's number of wins per hour of practice

14. What does the slope of the linear regression equation represent?
    a. hours of weekly practice
    b. number of team members
    c. number of matches the team won
    d. increase in the team's number of wins per hour of practice

15. If the team practices 4.5 hours per week next year, which is the best prediction of the number of chess matches the team could expect to win?
    a. about 8   c. about 18
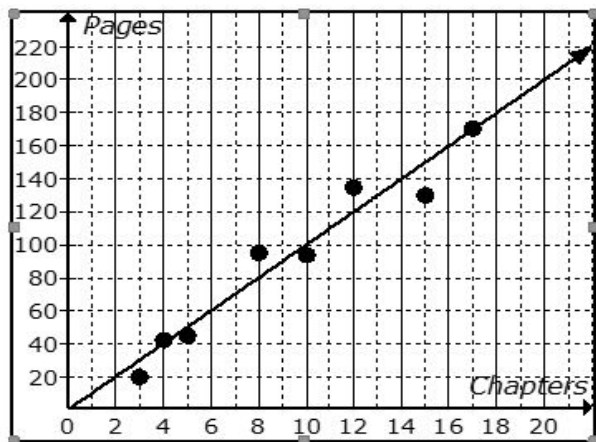   b. about 12   d. about 24
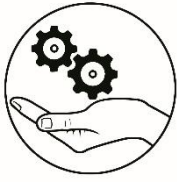
Additional Activities

Real-Life Solving!

1. The advertising expenditure (x) and sales in millions of pesos (y) of a retail business in its first eight years in operation are shown in the table.
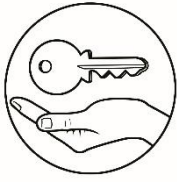
x	1.4	1.6	1.6	2.0	2.0	2.2	2.4	2.6
y	180	184	190	220	186	215	205	240

a. What is the equation of the regression line?
b. On average, for each million pesos spent on advertising, how does revenue change?
c. Predict the revenue if 2.5 million pesos is spent on advertising next year.

2. The scatter plot shows the relationship between the number of chapters and the total number of pages for several books. Find the regression equation and use the regression line to predict how many chapters a book with 180 pages will have.

x	3	4	5	8	10	12	15	17
y	20	40	42	98	98	138	130	170

References

What's In

I Scream for Ice Cream
1. number of toppings
2. cost of ice cream
3. price per topping
4. base price of ice cream
5. 6 toppings
6. Php 49.00

What I Know
1. D   6. A   11. B
2. A   7. A   12. A
3. B   8. B   13. C
4. D   9. C   14. D
5. B   10. C   15. C

Assessment
1. D   9. B
2. B   10. C
3. C   11. B
4. D   12. C
5. A   13. A
6. A   14. D
7. B   15. C
8. C

What's New

Will Carl Choose College?
1. number of years the person spent studying in school
2. person's daily salary
3. positive
4. Yes. The regression equation is a model to predict the daily salary. Just plug – in x, number of years in school and the result is the predicted amount of daily salary.
5. Yes. The regression model if proven significant can be used to predict values of y given a value of x. It means that generally, a person with a higher number of years in school tends to have a higher daily salary.
6. It depends upon the person or the situation. The regression analysis tells us that there is a correlation between number of years spent in school and person's daily salary. However, it does not tell causation. There are many other factors for having a higher daily salary.

What's More

1.1 Continuing College with Carl

$\hat{y} = 40x + 95$

5 years

18 years

$\hat{y} = 40(18) + 95$

$\hat{y} = 720 + 95$

PhP 815.00

1.2 The Y's from Our X's

1. $y = 35$
2. $y = -1$
3. $y = 58.5$
4. $y = 21.75$

What's More

1.3 Solving Substitutions

1. a. 550
   b. 800
   c. 13 days
2. a. 170 minutes
   b. 95 minutes

1.4 Looking into the Future

1. $y = 90$
2. $y = 20$
3. $y = 87$
4. $y = 44.2$

Additional Activities

1. a. $\hat{y} = 119.50 + 42.02x$
   b. an additional of 42.02 millions on revenues for every 1 million expenditure
   c. 224.55 million pesos
2. The regression equation is $\hat{y} = 10.04x - 0.85$. There would be 18 chapters.

References

Books

Albacea, Zita VJ., Mark John V. Ayaay, Isidoro P. David, and Imelda E. De Mesa. *Teaching Guide for Senior High School: Statistics and Probability*. Quezon City: Commision on Higher Education, 2016.

Caraan, Avelino Jr S. *Introduction to Statistics & Probability: Modular Approach*. Mandaluyong City: Jose Rizal University Press, 2011.

De Guzman, Danilo. *Statistics and Probability*. Quezon City: C & E Publishing Inc, 2017.

Punzalan, Joyce Raymond B. *Senior High School Statistics and Probability*. Malaysia: Oxford Publishing, 2018.

Sirug, Winston S. *Statistics and Probability for Senior High School CORE Subject A Comprehensive Approach K to 12 Curriculum Compliant*. Manila: Mindshapers Co., Inc., 2017.

Online Resources

Frost, Jim. "Choosing the Correct Type of Regression Analysis." Accessed May 25, 2020. https://statisticsbyjim.com/regression/choosing-regression-analysis/

Lane, David M. "Online Statistics Education: A Multimedia Course of Study." Accessed May 25, 2020. http://onlinestatbook.com/2/regression/intro.html

Parker, Mary. "Interpreting the Slope and Intercept in a Linear Regression Model." Accessed May 25, 2020. https://www.austincc.edu/mparker/1342/lessons/less5-8/interpret_slope.pdf

Quizziz. "Linear Regression Algebra I Quiz." Accessed May 25, 2020. https://quizizz.com/admin/quiz/5acae751c4daf70019c2369f/linear-regression

Rourke, Emily O. "Performance Based Learning and Assessment Task Tuition Cost Activity." Accessed May 25, 2020. https://www.radford.edu/rumath-smpdc/Performance/src/Emily O'Rourke - Tuition Cost Activity.pdf

themathtuber. "4.4 Interpret the Slope & Y-Intercept." *YouTube* video, 3:58. Oct 22, 2018. https://www.youtube.com/watch?v=fP2Dnq63kyg