NFL Quarterback Salaries

Jeanie Gibson
jgibson@hutchisonschool.org

Anna Bargagliotti
abargagl@lmu.edu

Project-SET
www.project-set.com

Published: July 2013

Overview of Lesson

In this activity students will set up a statistical question to explore how to interpret a linear regression equation and the correlation coefficient for a relationship between two quantitative variables. Students are divided into groups of three with each student being presented with a table of data showing statistics about the 30 top paid NFL quarterbacks in 2009-10. The data include total salary, pass completion percentage, total number of touchdowns, and average number of yards per game for each quarterback. Each person in the group then receives a separate table of data about the 30 highest paid NFL quarterback salaries in 2009-2010 and one of the three other variables listed above. Using these data, students set up a statistical question concerning the salary of a top paid NFL quarterback in 2009-2010 and the explanatory variable on their sheet. They use a software application (e.g., Fathom, graphing calculator, Excel) to create a scatterplot with the graph of the regression line superimposed and find a linear regression equation along with the correlation and the coefficient of determination. The three members of each group then interpret and compare their results and decide which of the three explanatory variables seemed to be the best predictor of a top paid 2009-2010 NFL quarterback's salary. The lesson includes questions that require students to demonstrate an understanding of the concept of correlation versus causation.

GAISE Components

This investigation follows the four components of statistical problem solving put forth in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*. The four components are: formulate a question, design and implement a plan to collect data, analyze the data by measures and graphs, and interpret the results in the context of the original question. This is a GAISE Level C activity.

Common Core State Standards for Mathematical Practice

4. Model with mathematics.
7. Look for and make use of structure.

Common Core State Standards Grade Level Content (High School)

S-ID. 6. Represent data on two quantitative variables on a scatterplot, and describe how the variables are related.

S-ID. 6a. Fit a function to the data; use functions fitted to the data to solve problems in the context of the data. *Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models.*

S-ID. 6c. Fit a linear function for a scatter plot that suggests a linear association.

S-ID. 7. Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

S-ID. 8. Compute (using technology) and interpret the correlation coefficient of a linear fit.

S-ID. 9. Distinguish between correlation and causation.

NCTM Principles and Standards for School Mathematics

Data Analysis and Probability Standards for Grades 9-12

Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them:

- understand the meaning of measurement data and categorical data, of univariate and bivariate data, and of the term variable;
- understand histograms, parallel box plots, and scatterplots and use them to display data.

Select and use appropriate statistical methods to analyze data:

- for bivariate measurement data, be able to display a scatterplot, describe its shape, and determine regression coefficients, regression equations, and correlation coefficients using technological tools;
- identify trends in bivariate data and find functions that model the data or transform the data so that they can be modeled.

Prerequisites

Prior to completing this activity students should be able to set up a statistical question that explores the relationship between two quantitative variables. They should have practiced defining the population and relationship of interest and should have experience with both determining and carrying out a data sampling method. They should be able to determine whether the data allows one to estimate causal effects. Students should have previous experience with the following bivariate data analysis techniques:

- plotting sample data in a scatterplot and visually determining and describing the trend of the scatterplot
- examining scatterplots and placing a model on the graph (e.g., drawing a line or curve)
- discussing where the "line of best fit" would be placed when a linear model is appropriate
- discussing mathematical ways with which the relationship could be represented/modeled
- finding the sample regression equation for a sample using technology

Learning Targets

In this lesson, students will gain continued experience with setting up a statistical question that explores a relationship between two quantitative variables. In addition to the prerequisite skills listed above, students will gain an understanding of what correlation represents on a scatterplot, including describing three characteristics of correlational data: direction, form, and strength or consistency. In their interpretation of the data, they will learn what the correlation and the squared correlation represent; and be able to explain, in the context of the posed question, that the squared correlation is the fraction of the sample variance explained by the explanatory variable.

Time Required

The time required for this activity is roughly 90 minutes. PowerPoint presentations will take additional class time.

Materials Required

For this activity, students will need a pencil, paper, the activity sheet (page 12), and software for producing a group presentation (e.g., PowerPoint). Students will also need some type of statistical software or graphing calculator capable of estimating regression equations.

Instructional Lesson Plan

The GAISE Statistical Problem-Solving Procedure

I. Formulate Question(s)

Before beginning the activity, the teacher may wish to review the prerequisite concepts with the entire class, reinforcing student understanding of the foundational concepts for the lesson. The class will then be divided into groups of three. Each student in the class will be given a copy of the activity sheet labeled "Activity Sheet – Introduction," (page 12) which includes a table of data showing the salary, pass completion percentage, total number of touchdowns, and average number of yards per game for the 30 top paid NFL quarterbacks in 2009-2010. Two introductory questions on the sheet will prompt students to observe the wide range of quarterback salaries and to consider what factors might determine how much an NFL quarterback is paid. These two questions will be discussed as a class before the students begin working within groups.

After the class discussion, the teacher will distribute Activity Sheets 1, 2 and 3 to each group for Group Members #1, 2 and 3 respectively. Each of these activity sheets will contain data and questions about the relationship between quarterback salaries and one of the three possible explanatory variables in the original table of data. The first question on the student activity sheet instructs students to review the data and write a statistical question concerning the salary of a top paid NFL quarterback in 2009-2010 and the explanatory variable on their sheet. Example

student questions include: is there an association between the 30 top NFL quarterback salaries in 2009-2010 and the quarterback's pass completion percentage; can we use the total number of touchdowns of the 30 top paid NFL quarterbacks in 2009-2010 to predict their salaries; or is there an association between the 30 top NFL quarterback salaries in 2009-2010 and the quarterback's average number of yards per game?

II. Design and Implement a Plan to Collect the Data

After formulating the question to be answered, students will enter the bivariate data on their activity sheet into their graphing calculator or other software such as Excel, Fathom, or SPSS. The data for the explanatory and response variables of interest are thus previously collected for the students. As the students complete this process, they will be prompted to distinguish between the explanatory and response variables.

Student Activity Sheets 1, 2 and 3 include some instructions for how to enter the data on a TI-83 or TI-84 graphing calculator.

III. Analyze the Data

After entering the data, students will be required to define and graph a scatterplot and sketch it on the activity sheet. This process will require them to consider how to best represent the data on a scatterplot, including decisions about how to scale and label axes for the explanatory and response variables.

Students will then use the data to determine the least-squares regression equation, correlation r, and coefficient of determination r^2. This information will be used to perform further analysis of the linear relationship between the explanatory and response variables. Students will then be instructed to add the least-squares line to the scatterplot and sketch this on the scatterplot on their activity sheet.

Student Activity Sheets 1, 2, and 3 include instructions for how to perform the steps in the analysis on the TI-83 or TI-84 graphing calculator.

CORRECT RESPONSES FOR DATA ANALYSIS

- **Use two decimal places for all results***

Group Member #1 (Activity Sheet 1)

Sample Regression Equation: $\hat{y} = -4{,}443{,}755.85 + 244{,}351.32x$

$\hat{y}$ = predicted salary of 2009-2010 top paid NFL quarterback in $

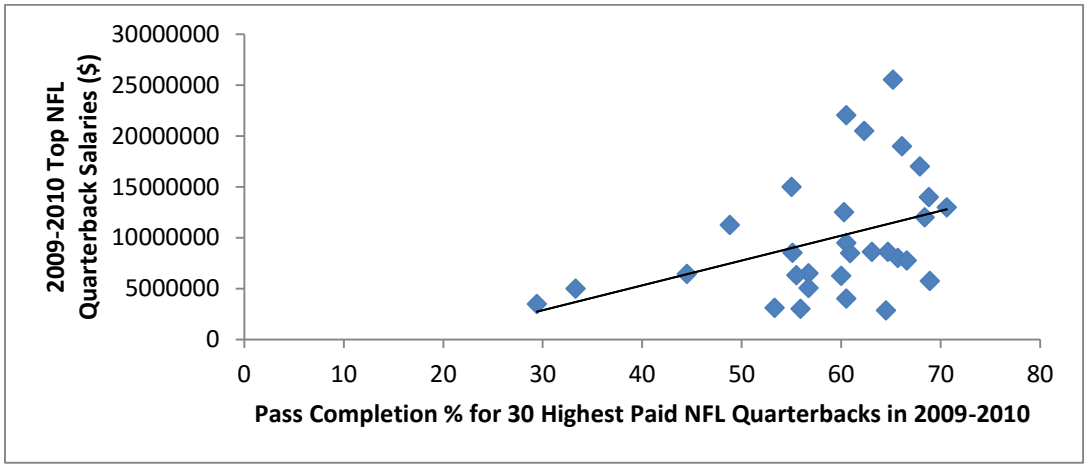
x = pass completion % for top paid NFL quarterback in 2009-10

It is important that students use $\hat{y}$ instead of y for the response variable and define it to be the **predicted** salary of a 2009-2010 top paid NFL quarterback in $.

Correlation *r*: $r = 0.3940$

Coefficient of Determination r^2: $r^2 = 0.1553$

Scatterplot with Least-Squares Regression Line:



It is important that students label the two axes, clearly indicating the explanatory variable on the horizontal axis and the response variable on the vertical axis.

Group Member #2 (Activity Sheet 2)

Sample Regression Equation: $\hat{y} = 4{,}393{,}649.84 + 320{,}510.26x$

$\hat{y}$ = predicted salary of 2009-2010 top paid NFL quarterback in $

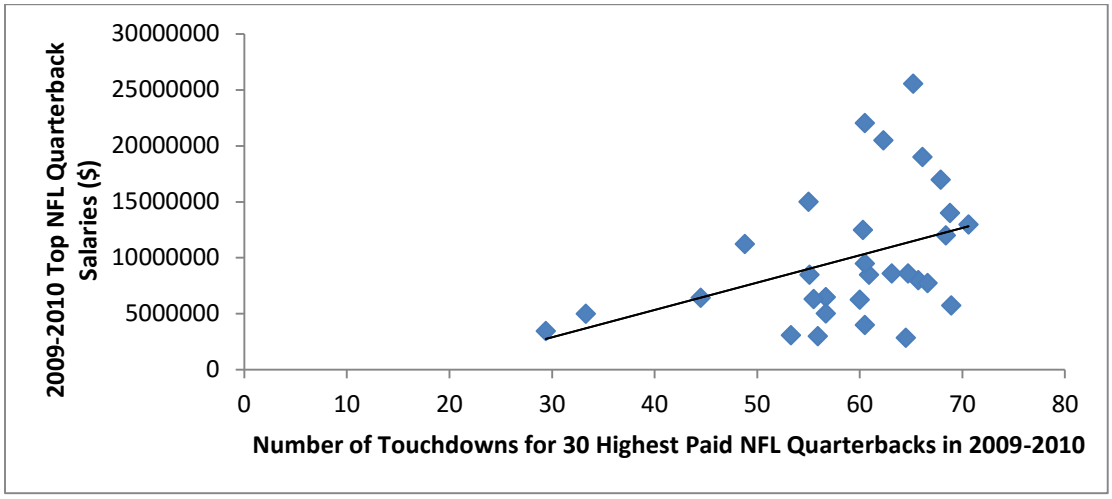
x = total number of touchdowns in 2009-10 of top paid NFL quarterback

It is important that students use $\hat{y}$ instead of y for the response variable and define it to be the **predicted** salary of a 2009-2010 top paid NFL quarterback in $.

Correlation *r*: $r = 0.5954$

Coefficient of Determination r^2: $r^2 = 0.3545$

Scatterplot with Least-Squares Regression Line:



It is important that students label the two axes, clearly indicating the explanatory variable on the horizontal axis and the response variable on the vertical axis.

Group Member #3 (Activity Sheet 3)

Sample Regression Equation: $\hat{y} = 2{,}242{,}948.15 + 38{,}047.53x$

$\hat{y}$ = predicted salary of 2009-2010 top paid NFL quarterback in $

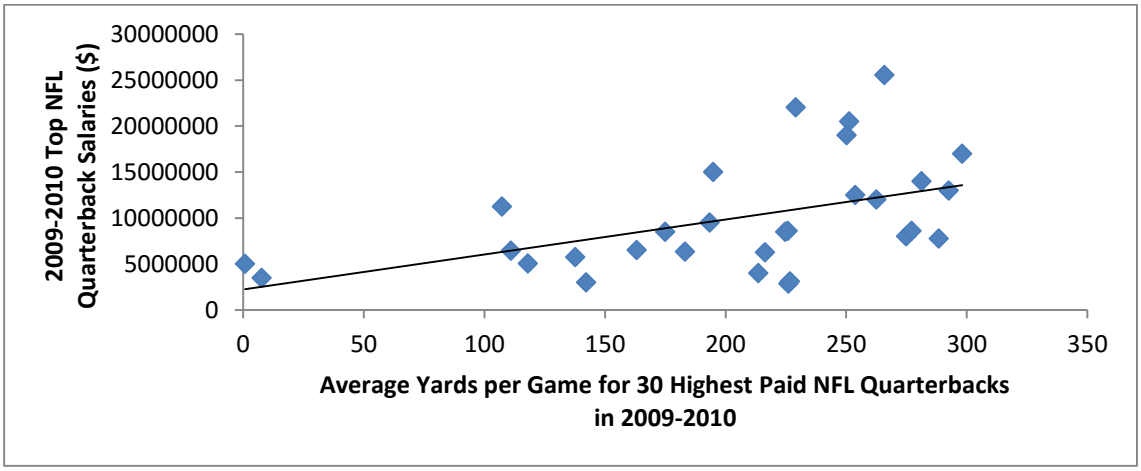
x = average yards per game in 2009-2010 for top paid NFL quarterback

It is important that students use $\hat{y}$ instead of y for the response variable and define it to be the **predicted** salary of a 2009-2010 top paid NFL quarterback in $.

Correlation *r*: $r = 0.4886$

Coefficient of Determination r^2: $r^2 = 0.2387$

Scatterplot with Least-Squares Regression Line:



It is important that students label the two axes, clearly indicating the explanatory variable on the horizontal axis and the response variable on the vertical axis.

IV. Interpret the Results

During the activity students were asked to write a statistical question concerning the salary of a top paid NFL quarterback in 2009-2010 and a given explanatory variable (pass completion percentage, total number of touchdowns, or average yards per game). They perform an analysis of data they are given, which includes finding and interpreting the correlation r and coefficient of determination r^2. Foundational concepts that will be used for the interpretation of r and r^2 are:

- The correlation r measures the direction and strength of the linear relationship between two quantitative variables.
- The correlation r is always a number between -1 and 1, with $r > 0$ indicating a positive association between the variables, and $r < 0$ indicating a negative association.
- Values of r near 0 indicate a very weak linear relationship, with the strength of the linear relationship increasing as r becomes closer to -1 or 1.
- The values $r = -1$ and $r = 1$ indicate a perfect linear relationship between two quantitative variables.
- r^2 is the coefficient of determination which gives the proportion of variation in the dependent variable y that can be attributed to the least-squares regression line of the dependent variable y on the independent variable x.

The correct student responses are given below.

Group Member #1 (Activity Sheet 1)

Correlation *r*: $r = 0.3940$

Interpretation: There is a weak to moderate positive linear relationship between the pass completion percentage of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries. Students need to include the strength (weak to moderate), direction (positive), and form (linear) in order for the answer to be complete.

Coefficient of Determination r^2: $r^2 = 0.1553$

Interpretation: 15.53% of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their pass completion percentage and salaries. This means that 84.47% of the variation in salaries is explained by factors other than the quarterbacks' pass completion percentages.

Group Member #2 (Activity Sheet 2)

Correlation *r*: $r = 0.5954$

Interpretation: There is a moderate positive linear relationship between the total number of touchdowns of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries. Students need to include the strength (moderate), direction (positive), and form (linear) in order for the answer to be complete.

Coefficient of Determination r^2: $r^2 = 0.3545$

Interpretation: 35.45% of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their total number of touchdowns and their salaries. This means that 64.55% of the variation in salaries is explained by factors other than their total number of touchdowns.

Group Member #3 (Activity Sheet 3)

Correlation *r*: $r = 0.4886$

Interpretation: There is a moderate positive linear relationship between the average yards per game of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries. Students need to include the strength (moderate), direction (positive), and form (linear) in order for the answer to be complete.

Coefficient of Determination r^2: $r^2 = 0.2387$

Interpretation: 23.87% of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their average yards per game and their salaries. This means that 76.13% of the variation in salaries is explained by factors other than their yards per game.

After interpreting the results of their analysis, students are asked to consider a hypothetical situation in which the correlation had instead been $r = 0.98$. They are then asked to answer the following two questions:

(1) What would this value of *r* tell you about the nature of the association between the salaries of top paid NFL quarterback in 2009-2010 and the specific explanatory variable considered by the student?

Answer: A value of $r = 0.98$ would tell us that there is a strong, positive, linear association between the salaries of the 30 top paid NFL quarterbacks in 2009-2010 and the specific explanatory variable considered by the student.

(2) Would this have been evidence that a quarterback's high pass completion percentage caused his salary to increase? Why or why not?

Sample Answer: The strong association would not have been evidence that a quarterback's large number of touchdowns caused his salary to increase. Although a correlation $r = 0.98$ would indicate a strong linear relationship between a quarterback's number of touchdowns and his salary, we cannot conclude that an increase in a quarterback's salary is **caused** by his large number of touchdowns. There could be other variables that contribute to the relationship between the two variables. A strong association between two variables is not enough to draw conclusions about cause and effect. Student answers will vary, but need to clearly indicate that association does not imply causation.

To answer Questions 9 through 11 on the activity sheets, each group will discuss and summarize their individual results. As a group, they will decide which of the three explanatory variables seems to be the best predictor of a 2009-2010 top paid NFL quarterback's salary. Students should conclude that the quarterback's total number of touchdowns seems to be the best predictor of a top paid 2009-2010 NFL quarterback's salary and should give the following justifications.

(1) The correlation $r = 0.5954$ for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is higher than the correlation r values for the linear relationships between the salary of a top paid NFL quarterback in 2009-2010 and the other two explanatory variables. This tells us that the linear relationship between the salary and total number of touchdowns is stronger than the linear relationships between the salary and the other two explanatory variables.

(2) The coefficient of determination for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is $r^2 = 0.3545$. This is higher than the r^2 values for quarterback salaries versus the other two explanatory variables. This tells us that a larger percentage of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their total number of touchdowns and their salaries than by the straight-line relationships between the other two explanatory variables and their salaries.

Question 10 requires that students do critical thinking to apply the concepts of the lesson to a given scenario involving the relationship between shoe size and height. Students continue to discuss and work as a group to formulate a response to this question. Students conclude the activity by preparing a group PowerPoint presentation which summarizes their analysis and presents their conclusion along with appropriate justification.

Assessment

1. The correlation *r* measures what two characteristics of the linear association between two quantitative variables?

2. How does one use the correlation *r* to determine the direction of the linear association between two quantitative variables?

3. How does one use the correlation *r* to determine the strength of the relationship between two quantitative variables?

4. Explain what the coefficient of determination r^2 tells you about how well a regression line fits a set of data.

Answers

1. The correlation *r* measures the direction and strength of the linear association between two quantitative variables.
2. $r > 0$ indicates a positive association and $r < 0$ indicates a negative association.
3. The correlation *r* always takes on values $-1 \le r \le 1$ and indicates the strength of a relationship by how close it is to -1 or 1.
4. The coefficient of determination r^2 tells us the fraction of the variation in the values of the response variable y that is accounted for by the least-squares regression line of y on the explanatory variable x.

Possible Extensions

1. Students use "real-life" data to investigate the linear relationship between two variables of interest to them. Students could be given suggested topics such as Voting Age Population vs. Electoral Votes for States in the U.S., Year vs. Tuition at All Public Colleges, or Decade vs. Total Population of the World. A list of suggested websites could be given to students to assist them in their search for data.

2. After students have mastered the skills necessary to interpret a linear regression equation and the correlation for a relationship between two quantitative variables measured on an entire population, they explore the relationship between two quantitative variables for which you **cannot** get data for the entire population. They use the concept of random sampling from the population of interest to obtain a sample regression equation and correlation as estimates of population parameters. This would then be extended to estimating the true regression equation from a population from which multiple random samples are taken using the regression equation computed from the repeated random samples.

References

1. *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) *Report*, ASA, Franklin et al., ASA, 2007 http://www.amstat.org/education/gaise/.

2. http://content.usatoday.com/sportsdata/football/nfl/salaries/position/qb/2009

3. www.pro-football-reference.com/years/2009/passing.htm.

What Makes NFL Quarterbacks Worth Their Salaries? Activity Sheet

Introduction

How does an NFL team decide how much to pay the quarterback? What makes one quarterback worth a total salary of $25,556,630 and another only around $3,000,000? The data below show statistics about the 30 top paid NFL quarterbacks in 2009-10. Their total salary is shown, along with their pass completion percentage, total number of touchdowns, and average number of yards per game.

2009-10 Statistics for 30 Highest Paid NFL Quarterbacks

Player	Salary	Pass Completion %	Touchdowns	Yards Per Game
1. Derek Anderson	6,450,000	44.5	3	111.0
2. Tom Brady	8,007,280	65.7	28	274.9
3. Drew Brees	12,989,500	70.6	34	292.5
4. Marc Bulger	6,507,280	56.7	5	163.2
5. Jason Campbell	2,864,780	64.5	20	226.1
6. Matt Cassel	15,005,200	55.0	16	194.9
7. Kerry Collins	8,507,280	55.1	6	175.0
8. Daunte Culpepper	5,050,000	56.7	3	118.1
9. Jay Cutler	22,044,090	60.5	27	229.1
10. Jake Delhomme	6,325,000	55.5	8	183.2
11. Brett Favre	12,000,000	68.4	33	262.6
12. Ryan Fitzpatrick	2,995,590	55.9	9	142.2
13. Joe Flacco	8,601,760	63.1	21	225.8
14. David Garrard	8,500,000	60.9	15	224.8
15. Matt Hasselbeck	6,256,240	60.0	17	216.4
16. Eli Manning	20,500,000	62.3	27	251.3
17. Peyton Manning	14,005,720	68.8	33	281.3
18. Luke McCown	5,006,760	33.3	0	0.7
19. Donovan McNabb	12,507,280	60.3	22	253.8
20. Carson Palmer	9,500,000	60.5	21	193.4
21. Chad Pennington	5,750,000	68.9	1	137.7
22. Philip Rivers	25,556,630	65.2	28	265.9
23. Aaron Rodgers	8,600,000	64.7	30	277.1
24. Ben Roethlisberger	7,751,560	66.6	26	288.5
25. JaMarcus Russell	11,255,440	48.8	3	107.3
26. Matt Schaub	17,000,000	67.9	29	298.1
27. Chris Simms	3,466,500	29.4	0	7.7
28. Alex D. Smith	4,007,280	60.5	18	213.6
29. Matthew Stafford	3,100,000	53.3	13	226.7
30. Kurt Warner	19,004,680	66.1	26	250.2

Data was taken from the following sources:

www.pro-football-reference.com/years/2009/passing.htm,
http://content.usatoday.com/sportsdata/football/nfl/salaries/position/qb/2009

Problem:

Use the data collected to investigate the relationship between the 30 top NFL quarterback salaries in 2009-2010 and the following three variables:

1. Pass Completion Percentage in 2009-2010
2. Total Number of Passing Touchdowns in 2009-2010
3. Average Number of Yards per Game in 2090-2010

Instructions

Your class will be divided into groups of three. Each person in the group will receive data about the 30 highest paid NFL quarterback salaries in 2009-2010 and one of three other variables listed above. Using the data, your group will explore the relationship between the salary of a 2009-2010 top paid NFL quarterback and each of these variables.

Activity Sheet 1 (Group Member #1) – page 1

2009-10 Pass Completion % and Salary for 30 Highest Paid NFL Quarterbacks

Player	Pass Completion %	Salary ($)
1. Derek Anderson	44.5	6,450,000
2. Tom Brady	65.7	8,007,280
3. Drew Brees	70.6	12,989,500
4. Marc Bulger	56.7	6,507,280
5. Jason Campbell	64.5	2,864,780
6. Matt Cassel	55.0	15,005,200
7. Kerry Collins	55.1	8,507,280
8. Daunte Culpepper	56.7	5,050,000
9. Jay Cutler	60.5	22,044,090
10. Jake Delhomme	55.5	6,325,000
11. Brett Favre	68.4	12,000,000
12. Ryan Fitzpatrick	55.9	2,995,590
13. Joe Flacco	63.1	8,601,760
14. David Garrard	60.9	8,500,000
15. Matt Hasselbeck	60.0	6,256,240
16. Eli Manning	62.3	20,500,000
17. Peyton Manning	68.8	14,005,720
18. Luke McCown	33.3	5,006,760
19. Donovan McNabb	60.3	12,507,280
20. Carson Palmer	60.5	9,500,000
21. Chad Pennington	68.9	5,750,000
22. Philip Rivers	65.2	25,556,630
23. Aaron Rodgers	64.7	8,600,000
24. Ben Roethlisberger	66.6	7,751,560
25. JaMarcus Russell	48.8	11,255,440
26. Matt Schaub	67.9	17,000,000
27. Chris Simms	29.4	3,466,500
28. Alex D. Smith	60.5	4,007,280
29. Matthew Stafford	53.3	3,100,000
30. Kurt Warner	66.1	19,004,680

Data was taken from the following sources:
www.pro-football-reference.com/years/2009/passing.htm
http://content.usatoday.com/sportsdata/football/nfl/salaries/position/qb/2009

Activity Sheet 1 (Group Member #1) – page 2

1. **Determine a Statistical Question That Involves the Data Given**
   After reviewing the data, write a statistical question concerning the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's pass completion %.

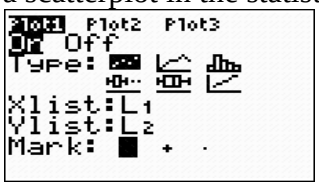
   What is the population of interest in your question?

   What is the relationship of interest?

NOTE: On #2 through #7, instructions are given for the TI-83 or TI-84 graphing calculator. You may use other software as instructed by your teacher.

2. **Enter Data Values Into Lists on Graphing Calculator (or other software)**
   Clear lists L1 (List 1) and L2 (List 2) on your calculator. Enter the pass completion % (explanatory variable) in L1 and the quarterback salaries (response variable) in L2.

3. **Using Data Entered in #2, Make a Scatterplot**
   - Define a scatterplot in the statistics plot menu. Specify the settings shown.



   - Use ZoomStat to obtain a graph. The calculator will set the window dimensions automatically by looking at the values in L1 and L2.

   Sketch the graph below. Make sure that you scale and label the axes. (Hint: You can use TRACE on your calculator to help you label the axes.)

4. **Find Sample Regression Equation**
   To determine the least-squares regression equation for the data in L1 and L2, carry out the following commands on your graphing calculator:
   - Press the STAT key; choose CALC and then 8:LinReg(a + bx).
   - Finish the command to read LinReg(a + bx) L1, L2, Y1 and press ENTER.

   Write the least-squares regression equation you found, defining the variables in the context of the problem. Round off the values of the slope and y-intercept to two places after the decimal since they represent salaries.

5. **Graph Regression Line on Scatterplot**
   Turn off all other equations in the Y= screen and press GRAPH to add the least-squares line to the scatterplot. Add this line to the sketch of your scatterplot in #3.

6. **Find and Interpret Correlation *r***
   To determine the correlation *r* for the linear relationship, you will again use the data entered in L1 and L2. You will carry out the following commands as you did in #4.
   - Press the STAT key; choose CALC and then 8:LinReg(a + bx).
   - Finish the command to read LinReg(a + bx) L1, L2, Y1 and press ENTER.

   Write the value of the correlation *r* that appears at the bottom of the list on the screen.

   NOTE: If the value of *r* does not appear on the screen, enter 2^{ND} 0 to access the catalog on your calculator. Scroll down to DiagnosticOn and press ENTER two times. You should see the word "Done". The value of the correlation *r* should now appear on the screen every time you calculate the linear regression equation.

   What does the value of *r* tell you about the association between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's pass completion %?

7. **Find and Interpret the Coefficient of Determination r^2**
   In #6, the value of r^2 was listed on your calculator with the linear regression equation. (You could also find this value by squaring the value for *r* you found in #6.) Write the value for r^2 below and explain what this tells you about the association between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's pass completion %.

8. Correlation vs. Causation

(a) In #6, you found a value for the correlation *r* that indicated a weak to moderate positive, linear relationship between the pass completion % of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries. Suppose the correlation had instead been $r = 0.98$. What would this value of *r* tell you about the nature of the association between the salaries of top paid NFL quarterback in 2009-2010 and their pass completion percentages?

(b) As you have just seen, if the correlation had been $r = 0.98$, there would have been a very strong association between the salaries of top paid NFL quarterbacks in 2009-2010 and their pass completion percentages. Would this have been evidence that a quarterback's high pass completion percentage caused his salary to increase? Why or why not?

9. After the three members of your group have completed #1-8, compare your results and answer the following question.
Based on the results of your investigations, which of the three explanatory variables seemed to be the best predictor of a top paid 2009-2010 NFL quarterback's salary?
Justify your answer.

10. As a group, discuss and answer the following question: Suppose that you wish to predict a person's height from their shoe size. What steps would you take to determine if this is possible? Concepts from this lesson should be included in your steps.

11. As a group, prepare a 10-15 slide PowerPoint presentation which includes the following:

- a summary of the results of your investigation about NFL quarterback salaries, including any tables and graphs used in the data analysis
- your group's answer to Question 9 above, including justification for your answer

Be prepared, as a group, to present this to the class and discuss and answer questions about your investigation. Every group member must participate in the presentation.

Activity Sheet 2 (Group Member #2) – page 1

2009-2010 Number of Touchdowns and Salary for 30 Highest Paid NFL Quarterbacks

Player	# of Touchdowns	Salary
1. Derek Anderson	3	6,450,000
2. Tom Brady	28	8,007,280
3. Drew Brees	34	12,989,500
4. Marc Bulger	5	6,507,280
5. Jason Campbell	20	2,864,780
6. Matt Cassel	16	15,005,200
7. Kerry Collins	6	8,507,280
8. Daunte Culpepper	3	5,050,000
9. Jay Cutler	27	22,044,090
10. Jake Delhomme	8	6,325,000
11. Brett Favre	33	12,000,000
12. Ryan Fitzpatrick	9	2,995,590
13. Joe Flacco	21	8,601,760
14. David Garrard	15	8,500,000
15. Matt Hasselbeck	17	6,256,240
16. Eli Manning	27	20,500,000
17. Peyton Manning	33	14,005,720
18. Luke McCown	0	5,006,760
19. Donovan McNabb	22	12,507,280
20. Carson Palmer	21	9,500,000
21. Chad Pennington	1	5,750,000
22. Philip Rivers	28	25,556,630
23. Aaron Rodgers	30	8,600,000
24. Ben Roethlisberger	26	7,751,560
25. JaMarcus Russell	3	11,255,440
26. Matt Schaub	29	17,000,000
27. Chris Simms	0	3,466,500
28. Alex D. Smith	18	4,007,280
29. Matthew Stafford	13	3,100,000
30. Kurt Warner	26	19,004,680

Data was taken from the following sources:
www.pro-football-reference.com/years/2009/passing.htm
http://content.usatoday.com/sportsdata/football/nfl/salaries/position/qb/2009

Activity Sheet 2 (Group Member #2) – page 2

1. **Determine a Statistical Question That Involves the Data Given**
   After reviewing the data, write a statistical question concerning the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns.

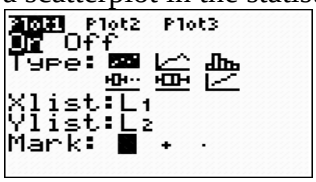
   What is the population of interest in your question?

   What is the relationship of interest?

NOTE: On #2 through #7, instructions are given for the TI-83 or TI-84 graphing calculator. You may use other software as instructed by your teacher.

2. **Enter Data Values Into Lists on Graphing Calculator**
   Clear lists L1 (List 1) and L2 (List 2) on calculator. Enter the number of touchdowns (explanatory variable) in L1 and the quarterback salaries (response variable) in L2.

3. **Using Data Entered in #2, Make a Scatterplot on Graphing Calculator**
   - Define a scatterplot in the statistics plot menu. Specify the settings shown.



   - Use ZoomStat to obtain a graph. The calculator will set the window dimensions automatically by looking at the values in L1 and L2.

   Sketch the graph below. Make sure that you scale and label the axes. (Hint: You can use TRACE on your calculator to help you label the axes.)

4. **Find Sample Regression Equation**
   To determine the least-squares regression equation for the data in L1 and L2, carry out the following commands on your graphing calculator:
   - Press the STAT key; choose CALC and then 8:LinReg(a + bx).
   - Finish the command to read LinReg(a + bx) L1, L2, Y1 and press ENTER.

   Write the least-squares regression equation you found, defining the variables in the context of the problem. Round off the values of the slope and y-intercept to two places after the decimal since they represent salaries.

5. **Graph Regression Line on Scatterplot**
   Turn off all other equations in the Y= screen and press GRAPH to add the least-squares line to the scatterplot. Add this line to the sketch of your scatterplot in #3.

6. **Find and Interpret Correlation *r***
   To determine the correlation *r* for the linear relationship, you will again use the data entered in L1 and L2. You will carry out the following commands as you did in #4.
   - Press the STAT key; choose CALC and then 8:LinReg(a + bx).
   - Finish the command to read LinReg(a + bx) L1, L2, Y1 and press ENTER.

   NOTE: If the value of *r* does not appear on the screen, enter 2^{ND} 0 to access the catalog on your calculator. Scroll down to DiagnosticOn and press ENTER two times. You should see the word "Done". The value of the correlation *r* should now appear on the screen every time you calculate the linear regression equation.

   Write the value of the correlation *r* that you see at the bottom of the list on the screen. What does the value of *r* tell you about the association between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns?

7. **Find and Interpret the Coefficient of Determination r^2**
   In #6, the value of r^2 was listed on your calculator with the linear regression equation. (You could also find this value by squaring the *r* value you found in #6.) Write the value for r^2 below and explain what this tells you about the association between the salary of a top paid quarterback in 2009-2010 and the quarterback's total number of touchdowns.

8. Correlation vs. Causation

(a) In #6, you found a value for the correlation *r* that indicated a moderate positive, linear relationship between the number of touchdowns of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries. Suppose the correlation had instead been $r = 0.98$. What would this value of *r* tell you about the nature of the association between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's number of touchdowns?

(b) As you have just seen, if the correlation had been $r = 0.98$, there would have been a very strong association between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's number of touchdowns. Would this have been evidence that a quarterback's large number of touchdowns caused his salary to increase? Why or why not?

9. After the three members of your group have completed #1-8, compare your results and answer the following question.
Based on the results of your investigations, which of the three explanatory variables seemed to be the best predictor of a top paid 2009-2010 NFL quarterback's salary?
Justify your answer.

10. As a group, discuss and answer the following question.
Suppose that you wish to predict a person's height from their shoe size. What steps would you take to determine if this is possible? Concepts from this lesson should be included in your steps.

11. As a group, prepare a 10-15 slide PowerPoint presentation which includes the following:

- a summary of the results of your investigation about NFL quarterback salaries, including any tables and graphs used in the data analysis
- your group's answer to Question 9 above, including justification for your answer

Be prepared, as a group, to present this to the class and discuss and answer questions about your investigation. Every group member must participate in the presentation.

Activity Sheet 3 (Group Member #3) – page 1

2009-10 Average Yards Per Game and Salary for 30 Highest Paid NFL Quarterbacks

Player	Average Yards Per Game	Salary
1. Derek Anderson	111.0	6,450,000
2. Tom Brady	274.9	8,007,280
3. Drew Brees	292.5	12,989,500
4. Marc Bulger	163.2	6,507,280
5. Jason Campbell	226.1	2,864,780
6. Matt Cassel	194.9	15,005,200
7. Kerry Collins	175.0	8,507,280
8. Daunte Culpepper	118.1	5,050,000
9. Jay Cutler	229.1	22,044,090
10. Jake Delhomme	183.2	6,325,000
11. Brett Favre	262.6	12,000,000
12. Ryan Fitzpatrick	142.2	2,995,590
13. Joe Flacco	225.8	8,601,760
14. David Garrard	224.8	8,500,000
15. Matt Hasselbeck	216.4	6,256,240
16. Eli Manning	251.3	20,500,000
17. Peyton Manning	281.3	14,005,720
18. Luke McCown	0.7	5,006,760
19. Donovan McNabb	253.8	12,507,280
20. Carson Palmer	193.4	9,500,000
21. Chad Pennington	137.7	5,750,000
22. Philip Rivers	265.9	25,556,630
23. Aaron Rodgers	277.1	8,600,000
24. Ben Roethlisberger	288.5	7,751,560
25. JaMarcus Russell	107.3	11,255,440
26. Matt Schaub	298.1	17,000,000
27. Chris Simms	7.7	3,466,500
28. Alex D. Smith	213.6	4,007,280
29. Matthew Stafford	226.7	3,100,000
30. Kurt Warner	250.2	19,004,680

Data was taken from the following sources:
www.pro-football-reference.com/years/2009/passing.htm
http://content.usatoday.com/sportsdata/football/nfl/salaries/position/qb/2009 (page unavailable as of October 2015. Similar data found at theredzone.org/salaryshowposition.aspx?Position=QB)

Activity Sheet 3 (Group Member #3) – page 2

1. **Determine a Statistical Question That Involves the Data Given**
   After reviewing the data, write a statistical question concerning the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's average number of yards per game.

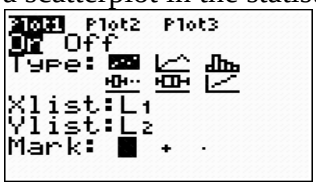
   What is the population of interest in your question?

   What is the relationship of interest?

NOTE: On #2 through #7, instructions are given for the TI-83 or TI-84 graphing calculator. You may use other software as instructed by your teacher.

2. **Enter Data Values Into Lists on Graphing Calculator**
   Clear lists L1 (List 1) and L2 (List 2) on calculator. Enter the yards per game (explanatory variable) in L1 and the quarterback salaries (response variable) in L2.

3. **Using Data Entered in #2, Make a Scatterplot**
   - Define a scatterplot in the statistics plot menu. Specify the settings shown.



   - Use ZoomStat to obtain a graph. The calculator will set the window dimensions automatically by looking at the values in L1 and L2.

   Sketch the graph below. Make sure that you scale and label the axes. (Hint: You can use TRACE on your calculator to help you label the axes.)

4. **Find Sample Regression Equation**
   To determine the least-squares regression equation for the data in L1 and L2, carry out the following commands on your graphing calculator:
   - Press the STAT key; choose CALC and then 8:LinReg(a + bx).
   - Finish the command to read LinReg(a + bx) L1, L2, Y1 and press ENTER.

   Write the least-squares regression equation you found, defining the variables in the context of the problem. Round off the values of the slope and y-intercept to two places after the decimal since they represent salaries.

5. **Graph the Regression Line on Scatterplot**
   Turn off all other equations in the Y= screen and press GRAPH to add the least-squares line to the scatterplot. Add this line to the sketch of your scatterplot in #3.

6. **Find and Interpret Correlation *r***
   To determine the correlation *r* for the linear relationship, you will again use the data entered in L1 and L2. You will carry out the following commands as you did on #4.
   - Press the STAT key; choose CALC and then 8:LinReg(a + bx).
   - Finish the command to read LinReg(a + bx) L1, L2, Y1 and press ENTER.

   NOTE: If the value of *r* does not appear on the screen, enter 2^{ND} 0 to access the catalog on your calculator. Scroll down to DiagnosticOn and press ENTER two times. You should see the word "Done". The value of the correlation *r* should now appear on the screen every time you calculate the linear regression equation.

   Write the value of the correlation *r* that you see at the bottom of the list on the screen. What does the value of *r* tell you about the association between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's average number of yards per game?

7. **Find and Interpret the Coefficient of Determination r^2**
   In #6, the value of r^2 was listed on your calculator with the linear regression equation. (You could also find r^2 by squaring the value of *r* you found in #6.) Write the value of r^2 below and explain what this tells you about the association between the salary of a top paid NFL quarterback in 2009-10 and the quarterback's average number of yards per game.

8. Correlation vs. Causation

(a) In #6, you found a value for the correlation *r* that indicated a weak to moderate positive, linear relationship between the average number of yards per game of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries. Suppose the correlation had instead been $r = 0.98$. What would this value of *r* tell you about the nature of the association between the salaries of the 30 top paid NFL quarterback in 2009-2010 and their average number of yards per game?

(b) As you have just seen, if the correlation had been $r = 0.98$, there would have been a very strong association between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's average number of yards per game. Would this have been evidence that a quarterback's high average number of yards per game caused his salary to increase? Why or why not?

9. After the three members of your group have completed #1-8, compare your results and answer the following question.
Based on the results of your investigations, which of the three explanatory variables seemed to be the best predictor of a top paid 2009-2010 NFL quarterback's salary?
Justify your answer.

10. As a group, discuss and answer the following question.
Suppose that you wish to predict a person's height from their shoe size. What steps would you take to determine if this is possible? Concepts from this lesson should be included in your steps.

11. As a group, prepare a 10-15 slide PowerPoint presentation which includes the following:

- a summary of the results of your investigation about NFL quarterback salaries, including any tables and graphs used in the data analysis
- your group's answer to Question 9 above, including justification for your answer

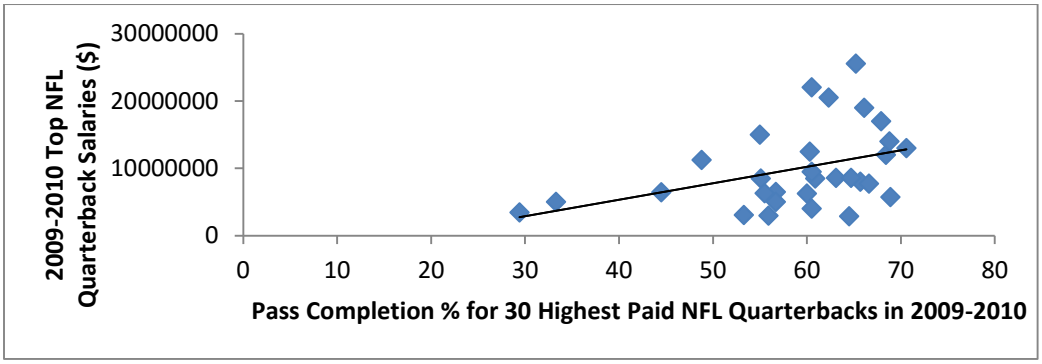
Be prepared, as a group, to present this to the class and discuss and answer questions about your investigation. Every group member must participate in the presentation.

Answers to Activity Sheet 1 (Group Member #1)

1. **Statistical Question** – Answers will vary, but should be similar to the following:
   Is there an association between the 30 top NFL quarterback salaries in 2009-2010 and the quarterbacks' pass completion percentages?
   OR
   Can we use the pass completion percentages of the 30 top paid NFL quarterbacks in 2009-2010 to predict their salaries?
   Population of Interest – 30 top paid NFL quarterbacks in 2009-2010
   Relationship of Interest – Relationship between the salaries of the top 30 NFL quarterbacks in 2009-2010 and the quarterbacks' pass completion percentages.

2. Students will enter the pass completion % data in L1 and the salary data in L2 in their graphing calculators or use other software such as Excel, Fathom or SPSS to enter data.

3. Students should sketch the following graph (without the regression line). The regression line will be added on question #5.



4. **Sample Regression Equation**
   $\hat{y} = -4{,}443{,}755.85 + 244{,}351.32x$
   $\hat{y}$ = predicted salary of 2009-2010 top paid NFL quarterback in $
   x = pass completion % for top paid NFL quarterback in 2009-2010

5. **Graph of Regression Line –** See graph on #3

6. **Correlation *r***
   $r = 0.3940$ There is a weak to moderate positive linear relationship between the pass completion % of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries.

Answers to Activity Sheet 1 (continued)

7. **Coefficient of Determination r^2**
   $r^2 = 0.1553$
   15.53% of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their pass completion % and salaries. This means that 84.47% of the variation in salaries is explained by factors other than the quarterbacks' pass completion percentages.

8. (a) A value of $r = 0.98$ would tell us that there is a strong, positive, linear association between the salaries of the 30 top paid NFL quarterbacks in 2009-2010 and the quarterbacks' pass completion percentages.
   (b) The strong association would not have been evidence that a quarterback's high pass completion percentage caused his salary to increase. Although a correlation $r = 0.98$ indicates a strong linear relationship between a quarterback's pass completion percentage and his salary, we cannot conclude that an increase in a quarterback's salary is **caused** by his high pass completion percentage. There could be other variables that contribute to the relationship between the two variables. A strong association between two variables is not enough to draw conclusions about cause and effect.

9. The explanatory variable that seems to be the best predictor of a top paid 2009-2010 NFL quarterback's salary is the quarterback's total number of touchdowns. This decision is justified by the following:
   - The correlation $r = 0.5954$ for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is higher than the correlation r values for the linear relationships between the salary of a top paid NFL quarterback in 2009-2010 and the other two explanatory variables. Thus, the linear relationship between salary and total number of touchdowns is stronger than the linear relationships between salary and the other two explanatory variables.
   - The coefficient of determination for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is $r^2 = 0.3545$. This is higher than the r^2 values for quarterback salaries versus the other two explanatory variables. This tells us that a larger percentage of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their total number of touchdowns and their salaries than by the straight-line relationships between the other two explanatory variables and their salaries.

Answers to Activity Sheet 1 (continued)

10. Students should give steps which include:
    - A method of collecting data on shoe size and height that indicates an attempt to randomly select a large number of subjects
    - The following elements from this lesson:
      – enter bivariate data into calculator or other software with shoe size as the explanatory variable and height as the response variable
      – plot a scatterplot and check for approximate linearity
      – graph regression line on scatterplot
      – use calculator or other software to find the linear regression equation, correlation coefficient *r*, and coefficient of determination r^2
      – an explanation of how they would interpret *r* and r^2 to determine the strength of the linear association between height and shoe size

11. Students conclude the activity by preparing a PowerPoint presentation with supporting charts and graphs which summarize their analyses and presents their conclusions, along with appropriate justifications. Every member of the group should participate in the presentation.

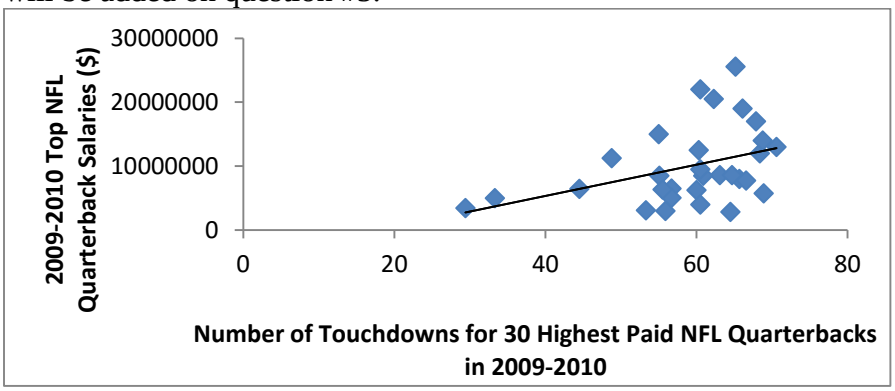
Answers to Activity Sheet 2 (Group Member #2)

1. **Statistical Question** – Answers will vary, but should be similar to the following:
   Is there an association between the 30 top NFL quarterback salaries in 2009-2010 and the quarterbacks' total number of touchdowns?
   OR
   Can we use the total number of touchdowns of the 30 top paid NFL quarterbacks in 2009-2010 to predict their salaries?
   Population of Interest – 30 top paid NFL quarterbacks in 2009-2010
   Relationship of Interest – Relationship between the salaries of the top 30 NFL quarterback in 2009-2010 and their total number of touchdowns.

2. Students will enter the total number of touchdowns in L1 and the salary data in L2 in their graphing calculators or use other software such as Excel, Fathom or SPSS to enter data.

Answers to Activity Sheet 2 (continued)

3. Students should sketch the following graph (without the regression line). The regression line will be added on question #5.



4. **Sample Regression Equation**
   $\hat{y} = 4{,}393{,}649.84 + 320{,}510.26x$
   $\hat{y}$ = predicted salary of 2009-2010 top paid NFL quarterback in $
   x = total number of touchdowns in 2009-2010 of top paid NFL quarterback

5. **Graph of Regression Line –** See graph on #3

6. **Correlation *r***
   $r = 0.5954$ There is a moderate positive linear relationship between the total number of touchdowns of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries.

7. **Coefficient of Determination r^2**
   $r^2 = 0.3545$
   35.45% of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their total number of touchdowns and their salaries. This means that 64.55% of the variation in salaries is explained by factors other than their total number of touchdowns.

8. (a) A value of $r = 0.98$ would tell us that there is a strong, positive, linear association between the salaries of the 30 top paid NFL quarterbacks in 2009-2010 and the quarterbacks' total number of touchdowns.
   (b) The strong association would not have been evidence that a quarterback's large number of touchdowns caused his salary to increase. Although a correlation $r = 0.98$ would indicate a strong linear relationship between a quarterback's number of touchdowns and his salary, we cannot conclude that an increase in a quarterback's salary is **caused** by his large number of touchdowns. There could be other variables that contribute to the relationship between the two variables. A strong association between two variables is not enough to draw conclusions about cause and effect.

Answers to Activity Sheet 2 (continued)

9. The explanatory variable that seems to be the best predictor of a top paid 2009-2010 NFL quarterback's salary is the quarterback's total number of touchdowns. This decision is justified by the following:
   - The correlation $r = 0.5954$ for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is higher than the correlation r values for the linear relationships between the salary of a top paid NFL quarterback in 2009-2010 and the other two explanatory variables. Thus, the linear relationship between salary and total number of touchdowns is stronger than the linear relationships between salary and the other two explanatory variables.
   - The coefficient of determination for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is $r^2 = 0.3545$. This is higher than the r^2 values for quarterback salaries versus the other two explanatory variables. This tells us that a larger percentage of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their total number of touchdowns and their salaries than by the straight-line relationships between the other two explanatory variables and their salaries.

10. Students should give steps which include:
    - A method of collecting data on shoe size and height that indicates an attempt to randomly select a large number of subjects
    - The following elements from this lesson:
      – enter bivariate data into calculator or other software with shoe size as the explanatory variable and height as the response variable
      – plot a scatterplot and check for approximate linearity
      – graph regression line on scatterplot
      – use calculator or other software to find the linear regression equation, correlation coefficient r, and coefficient of determination r^2
      – an explanation of how they would interpret r and r^2 to determine the strength of the linear association between height and shoe size

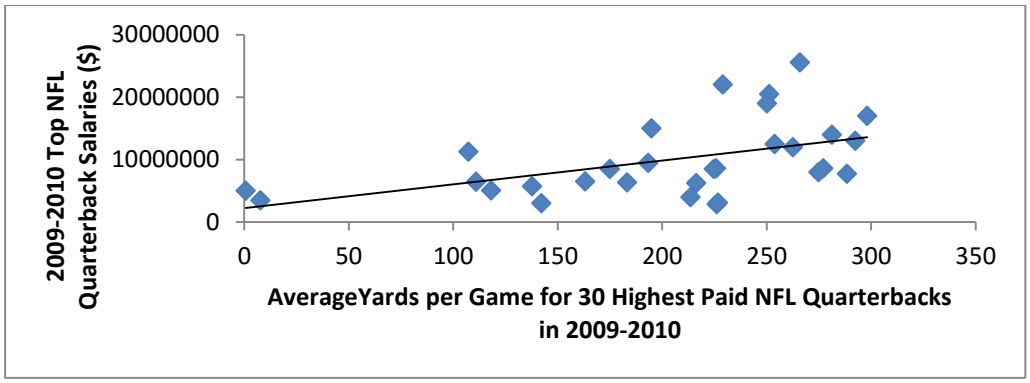
11. Students conclude the activity by preparing a PowerPoint presentation with supporting charts and graphs which summarizes their analysis and presents their conclusion, along with appropriate justification. Every member of the group should participate in the presentation.

Answers to Activity Sheet 3 (Group Member #3)

1. **Statistical Question** – Answers will vary, but should be similar to the following:
   Is there an association between the 30 top NFL quarterback salaries in 2009-2010 and the quarterbacks' average number of yards per game?
   OR
   Can we use the average number of yards per game of the 30 top paid NFL quarterbacks in 2009-2010 to predict their salaries?
   Population of Interest – 30 top paid NFL quarterbacks in 2009-2010
   Relationship of Interest – Relationship between the salaries of the top 30 NFL quarterbacks in 2009-2010 and the quarterbacks' average number of yards per game.

2. Students will enter the average yards per game in L1 and the salary data in L2 in their graphing calculators or use other software such as Excel, Fathom or SPSS to enter data.

3. Students should sketch the following graph (without the regression line). The regression line will be added on question #5.



4. **Sample Regression Equation**
   $\hat{y} = 2{,}242{,}948.15 + 38{,}047.53x$
   $\hat{y}$ = predicted salary of 2009-2010 top paid NFL quarterback in $
   x = average yards per game in 2009-2010 for top paid NFL quarterback

5. **Graph of Regression Line –** See graph on #3

6. **Correlation *r***
   $r = 0.4886$ There is a moderate positive linear relationship between the average yards per game of the 30 top paid NFL quarterbacks in 2009-2010 and their salaries.

Answers to Activity Sheet 3 (continued)

7. **Coefficient of Determination r^2**
   $r^2 = 0.2387$
   23.87% of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their average yards per game and their salaries. This means that 76.13% of the variation in salaries is explained by factors other than their average yards per game.

8. (a) A value of $r = 0.98$ would tell us that there is a strong, positive, linear association between the salaries of the 30 top paid NFL quarterbacks in 2009-2010 and the quarterbacks' average number of yards per game.
   (b) The strong association would not have been evidence that a quarterback's high average number of yards per game caused his salary to increase. Although a correlation $r = 0.98$ indicates a strong linear relationship between a quarterback's average number of yards per game and his salary, we cannot conclude that an increase in a quarterback's salary is **caused** by his high average number of yards per game. There could be other variables that contribute to the relationship between the two variables. A strong association between two variables is not enough to draw conclusions about cause and effect.

9. The explanatory variable that seems to be the best predictor of a top paid 2009-2010 NFL quarterback's salary is the quarterback's total number of touchdowns. This decision is justified by the following:
   - The correlation $r = 0.5954$ for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is higher than the correlation r values for the linear relationships between the salary of a top paid NFL quarterback in 2009-2010 and the other two explanatory variables. Thus, the linear relationship between salary and total number of touchdowns is stronger than the linear relationships between salary and the other two explanatory variables.
   - The coefficient of determination for the linear relationship between the salary of a top paid NFL quarterback in 2009-2010 and the quarterback's total number of touchdowns is $r^2 = 0.3545$. This is higher than the r^2 values for quarterback salaries versus the other two explanatory variables. This tells us that a larger percentage of the variation in the salaries of the 30 top paid NFL quarterbacks in 2009-2010 is explained by the straight-line relationship between their total number of touchdowns and their salaries than by the straight-line relationships between the other two explanatory variables and their salaries.

Answers to Activity Sheet 3 (continued)

10. Students should give steps which include:
    - A method of collecting data on shoe size and height that indicates an attempt to randomly select a large number of subjects
    - The following elements from this lesson:
        - enter bivariate data into calculator or other software with shoe size as the explanatory variable and height as the response variable
        - plot a scatterplot and check for approximate linearity
        - graph regression line on scatterplot
        - use calculator or other software to find the linear regression equation, correlation coefficient r, and coefficient of determination r^2
        - an explanation of how they would interpret r and r^2 to determine the strength of the linear association between height and shoe size

11. Students conclude the activity by preparing a PowerPoint presentation with supporting charts and graphs which summarizes their analysis and presents their conclusion along with appropriate justification. Every member of the group should participate in the presentation.