Name: _____ Class: _____

Applications of Linear Regressions

1. We are going to revisit some data from yesterday. Below is a survey to determine the effect of time spent on studying and grade point average. The table below shows the results for 10 students randomly selected.

Study time (Hours per week)	2	4	5	7	10	12	14	17	19	20
GPA (out of 100)	64	71	69	74	81	86	84	94	91	96

(a) Enter the data in your calculator and use it to generate the equation for the line of best fit. Round your slope to the nearest tenth and round your y-intercept to the nearest integer.

(b) According to the linear regression model from part (a), what GPA, to the nearest integer, would result from studying for 15 hours in a given week? Justify your answer.

(c) A passing average is defined as a 65% or above. Does the model predict a passing average if the student spends no time studying in a given week? Justify your answer.

(d) For each additional hour that a student studies per week, how many points does the model predict a GPA will rise? Explain how you arrived at your answer.

(e) Create a scatter plot of this data on your calculator, along with the linear regression. Change the window to get a good view and sketch what you see.

2. The mean annual temperature of a location generally depends on its elevation above sea level. A collection of nine locations in Nevada were chosen and had their elevation and mean annual temperature recorded. The data is shown below.

Elevation (feet)	1200	4125	6230	2378	5625	6328	4375	1864	3160
Mean Temperature (°F)	62	45	36	51	48	32	40	58	49

(a) Use your calculator to determine the equation for the line of best fit. Round your slope to the nearest *thousandth*. Note that it will be a small number. Round your y-intercept to the nearest integer.

(b) What does the y-intercept tell you about the temperature in Nevada?

(c) Using correct units, give an interpretation of the slope of this line.

(d) Using your model from part (a), what would be the predicted mean temperature at an elevation of 3000 feet above sea level?

(e) Would you characterize this correlation as being positive or negative? How can you tell this from the equation itself?

(f) Create a scatter plot of the data and graph the line of best fit on it as well. Are there any data points from the table above that are significantly "missed" by the model? If so, which data point?

REVIEW: FREQUENCY TABLES AND CORRELATION COEFFICIENTS

Frequency Tables

Model

Number of coffee cups drank by workers in a week

Number of people	10	30	30	50	80
Cups of coffee	1	2	3	4	5

Find the **mean** number of coffee cups drank:

Find the **median** number of coffee cups:

Find the **mode** of coffee cups:

If you graphed the frequency table in a dot plot, would it be symmetric, skew left, or skew right?

Practice

Number of goals scored by soccer players

Number of players	50	70	30	10	5
Goals scored	1	2	3	4	5

Find the **mean** number of goals scored:

Find the **median** number of goals scored:

Find the **mode** of goals scored:

If you graphed the table in a dot plot, would it be symmetric, skew left, or skew right?

Correlation Coefficients: From r-values to words

Model

For each, indicate the type of correlation (strong/weak, positive/negative).

(a) Running on a treadmill and burning calories have a correlation coefficient of $r = 0.97$

(b) Microwaves and radiation poisoning have a correlation coefficient of $r = 0.23$

(c) Temperature and magazine subscriptions have a correlation coefficient of $r = -0.34$

(d) Temperature and air conditioning costs have a correlation coefficient of $r = -0.92$

Practice

For each, indicate the type of correlation (strong/weak, positive/negative).

(a) Miles of running and likelihood of heart attack have a correlation coefficient of $r = -0.92$

(b) Hours studying and rate of diabetes have a correlation coefficient of $r = 0.235$

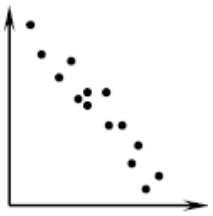
(c) Distance from home and deer population have a correlation coefficient of $r = -0.05$

(d) Gasoline in your car and distance you can travel have a correlation coefficient of $r = 0.85$

Correlation Coefficients: From scatter plots to words

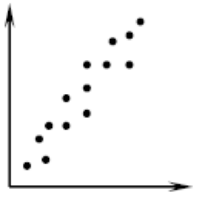
Model

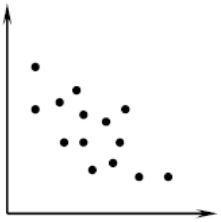
For each scatter plot, indicate the type of correlation and a possible correlation coefficient

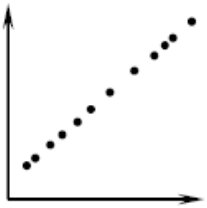


Practice

For each scatter plot, indicate the type of correlation and a possible correlation coefficient







Correlation Coefficients: From tables to lines of best fit

Model

The table below shows the attendance at a museum in select years from 2007 to 2013.

Attendance at Museum

Year	2007	2008	2009	2011	2013
Attendance (millions)	8.3	8.5	8.5	8.8	9.3

State the linear regression equation represented by the data table when $x = 0$ is used to represent the year 2007 and y is used to represent the attendance. Round all values to the *nearest hundredth*. State the correlation coefficient to the *nearest hundredth* and determine whether the data suggest a strong or weak association.

Practice (almost certainly a quiz question)

Eric collected data on the high temperature of coffee and on the sales of coffee.

High Temperature, t	54	50	62	67	70
Coffee Sales, f(t)	$2900	$3080	$2500	$2380	$2200

State the linear regression function, $f(t)$, represented by the data table when $t = 0$ is used to represent a high temperature of 50 degrees. Round to the *nearest hundredth*. State the correlation coefficient and determine whether the data suggests a strong or weak correlation.