# Scatterplots & Correlation Coefficient Worksheet (Answer Key)

#### **KEY TERMS**

- <u>Scatterplot:</u> the primary tool for examining two-variable relationships, when both variables are numerical
  - In a scatterplot, each point represents one observation and the location of the point depends on the values of the two variables.
- <u>Trend:</u> the general tendency of the scatterplot scanning from left to right • Usually, trends are either increasing (uphill) or decreasing (downhill).
- <u>Strength:</u> the amount of spread the scatterplot has in the vertical direction
  - Weak associations result in a large amount of scattering in the scatterplot, while strong associations have little vertical variation.
- Shape: the shape that the scatterplot plots form
  - The simplest shape for a trend is linear.
- **<u>Correlation Coefficient:</u>** a numerical value that provides a measure of strength and direction of the linear association between the independent variable x and the dependent variable y
  - $\circ \quad r = (\Sigma Z x Z y) / (n-1)$
  - The value of r is always between -1 and +1 and indicates the strength of the linear relationship between x and y. Values of *r* close to -1 or to +1 indicate a stronger linear relationship between x and y.

Illowsky, B., & Dean, S. (2018). Introductory statistics Gould, R., & Ryan, C. N. (2015). Introductory statistics: Exploring the world through data. Pearson.

## **EXERCISE 1:** Salinity vs. Electrical Conductivity

The below scatterplot contains a sample of 500 randomly selected salinity vs. electrical conductivity data points from the Mokauea loko i'a data collected in June 2020.



June Salinity vs. Electrical Conductivity

The above data is real data collected from research at Mokauea loko i 'a. Additional Mokauea loko i `a data can be explored here: <u>http://grogdata.soest.hawaii.edu/</u> What is a research question that could ask based on variables used in the scatter plot?

Does electrical conductivity in the Mokauea loko i'a depend on the pond's salinity?

What is the trend, strength, and shape of the scatterplot?

Trend: Positive

Strength: Strong association

Shape: Linear

What could be determined about the two-variable based on the trend, strength, and shape of the scatterplot?

Electrical conductivity in the Mokauea loko i'a depends on the pond's salinity.

#### EXERCISE 2: Salinity vs. Dissolved Oxygen

The below scatterplot contains a sample of 500 randomly selected salinity vs. electrical conductivity data points from the Mokauea loko i'a data collected in June 2020.





The above data is real data collected from research at Mokauea loko i 'a. Additional Mokauea loko i`a data can be explored here: <u>http://grogdata.soest.hawaii.edu/</u> What is a research question that could ask based on variables used in the scatter plot?

Does dissolved oxygen in the Mokauea loko i'a depend on the pond's salinity?

What is the trend, strength, and shape of the scatterplot?

Trend: No trend

Strength: Weak association

Shape: N/A

What could be determined about the two-variable based on the trend, strength, and shape of the scatterplot?

Dissolved oxygen in the Mokauea loko i'a does not depend on the pond's salinity.

## **EXERCISE 3:** Correlation Coefficient

Describe what are likely the trend and strength based on the following correlation coefficients found from comparing two Mokauea loko i'a parameters from March 2020. Assume that the shape is linear.

a) Surface Temperature (°C) vs. Bottom Temperature (°C): r = 0.99799799599

Trend: Positive trend

Strength: Strong association

b) Salinity (PSU) vs. Electrical Conductivity (mS/cm): r = 0.97621718894

Trend: Positive trend

Strength: Strong association