**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Per.\_\_\_\_\_\_\_**

**U7 CWK #1** *Statistics - Scatter Plots*

In 8th grade math, we learn the difference between various types of data.

|  |  |
| --- | --- |
| **Day** | **Money Made in Tips (dollars)** |
| 1 | 120 |
| 2 | 75 |
| 3 | 80 |
| 4 | 100 |
| 5 | 115 |
| 6 | 100 |
| 7 | 55 |
| 8 | 90 |
| 9 | 100 |
| 10 | 120 |
| 11 | 90 |
| 12 | 105 |
| 13 | 105 |
| 14 | 75 |
| 15 | 100 |

Jenny is a hair stylist. She decides to record the amount of money she makes in tips over a 15-day period. She records the following data:

To better visualize the data, Jenny makes a dot plot of the data.



 Make some observations about the data shown in the dot plot.

Jenny then asks herself the following question: “I wonder if the amount I make in tips is associated to the number of clients I have each day?” She looks back through her appointment book and records the number of clients she had on each of the 15 days. She records the following data.

|  |  |  |
| --- | --- | --- |
| **Day** | **Number of Clients** | **Money Made in Tips (dollars)** |
| 1 | 12 | 120 |
| 2 | 8 | 75 |
| 3 | 10 | 80 |
| 4 | 12 | 100 |
| 5 | 11 | 115 |
| 6 | 9 | 100 |
| 7 | 6 | 55 |
| 8 | 8 | 90 |
| 9 | 10 | 100 |
| 10 | 14 | 120 |
| 11 | 10 | 90 |
| 12 | 10 | 105 |
| 13 | 3 | 105 |
| 14 | 9 | 75 |
| 15 | 11 | 100 |

To better visualize the data, Jenny makes a **scatter plot** of the data. A **scatter plot** is a graph in the coordinate plane of the set of all $\left(x,y\right)$ ordered pairs of bivariate data.



* 1. Make some observations about the scatter plot.

The examples of data to the left show univariate data and bivariate data.

1. Which example shows univariate data? Why did you pick that example? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

b. Which example shows univariate data? Why did you pick that example? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Based on the examples and your thoughts above, what could be the definition for univariate and bivariate data?

Univariate data: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Bivariate data: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Directions**: Determine if the following scenarios represent univariate or bivariate data.

1. Lucas conducts an experiment where he records the number of speeding tickets issued in Iron County in a given year along with the average price of gasoline for that same given year. He collects this data from the year 1972 through 2012.
2. Lea conducts an experiment where she records the heights of all the NBA basketball players on the Miami Heat’s roster for the 2014 season.
3. Adel conducts an experiment where she records the selling price of several homes in a neighborhood.
4. Adel conducts an experiment where she records the selling price and square footage of homes in a neighborhood.
5. Lisa conducts an experiment on the number of times a person works out a week and the person’s weight.

As 8th graders, we will focus our study on bivariate data sets and we will explore the relationship between two variables of interest. (Which we have already started doing in our study of linear equations and the relationship between y and x.)

Izumi is the score keeper for her school’s basketball team. Izumi’s responsibilities as score keeper are to keep a record for several plays during the 2012-2013 season.

|  |  |  |
| --- | --- | --- |
| **Player** | **Baskets Attempted** | **Baskets Made** |
| Amber Carlson | 34 | 15 |
| Casey Corbin | 368 | 134 |
| Joan O’Connell | 94 | 23 |
| Monique Ortiz | 102 | 36 |
| Maria Ferney | 91 | 32 |
| Amelia Krebs | 310 | 137 |
| Tonya Smith | 56 | 25 |
| Juanita Martinez | 58 | 17 |
| Sara Garcia | 151 | 61 |
| Alicia Mortenson | 67 | 26 |
| Parker Christiansen | 94 | 29 |
| Rachel Reagan | 183 | 66 |
| Paula Lyons | 276 | 108 |
| Thao Ho | 221 | 94 |
| Jessica Geffen | 127 | 54 |

The table given shows the record that Izumi made.

As Izumi examines the data she wonders, “Is there is an association between the number of baskets made and the number of baskets attempted?” To further investigate the relationship between these two random variables, “Baskets Made” and “Baskets Attempted” Izumi makes a **scatter plot** of the data as shown below.

1. Izumi ran out of time while creating her scatter plot and did not plot the data for the last two players in the table, Thao Ho and Jessica Geffen. Help Izumi finish the scatter plot by plotting the data for these players and labeling the points with these players’ initials.
2. Which player does the circled data point represent?
3. Casey Corbin sees Izumi’s graph and asks which point on the scatter plot represents her data. Put Casey’s initials by the point that represents his data.

 

d. Using the scatter plot, determine if there is a relationship between baskets attempted and baskets made.

 Describe any trends or patterns you observe in the data.

|  |  |  |
| --- | --- | --- |
| **Player** | **Assists** | **Rebounds** |
| Amber Carlson | 82 | 64 |
| Casey Corbin | 6 | 170 |
| Joan O’Connell | 43 | 37 |
| Monique Ortiz | 50 | 54 |
| Maria Ferney | 89 | 42 |
| Amelia Krebs | 25 | 193 |
| Tonya Smith | 70 | 39 |
| Juanita Martinez | 3 | 26 |
| Sara Garcia | 100 | 73 |
| Alicia Mortenson | 33 | 152 |
| Parker Christiansen | 64 | 93 |
| Rachel Reagan | 45 | 67 |
| Paula Lyons | 59 | 117 |
| Thao Ho | 15 | 179 |
| Jessica Geffen | 30 | 113 |

In addition to data about baskets, Izumi is curious about the relationship between the number of assists and the number of rebounds a player makes in a season. In order to study this relationship, Izumi gathers data on the number of assists and rebounds each player makes during the season.

Izumi’s Assist and Rebound data are given in the following table.

Izumi made the scatter plot of assists and rebounds shown below to help her better visualize the data.

1. Again, Izumi ran out of time while creating her scatter plot and did not plot the data for the last two players in the table, Thao Ho and Jessica Geffen. Help Izumi finish the scatter plot by plotting the data for these players and labeling the points with these players’ initials.
2. Which player does the circled data point represent?
3. Locate the data points for 3 different players and put the initials of the players next to their data point.

1. Izumi notices the circled data point stands out noticeably from the general behavior of the data set. We call this point an **outlier**. Provide an explanation as to why this player’s data does not fit with the rest of the data.
2. Using the scatter plot, determine if there is a relationship between number of assists and number of rebounds. Describe any trends or patterns you observe in the data.

Which data sets Izumi made appears to have a stronger association: the relationship between number of baskets made and attempted or the relationship between number of rebounds and assists? Explain your reasoning.

There are many ways that we can describe the association (if there is one) between two variables. Common ways to talk about the association of two variables are shown in the table below. (Think about the slope of a line or the relationship between the x and y variables in linear relationships.)

 **Sketch scatter plots that correspond to each of the four associations described.**

|  |  |
| --- | --- |
| **Positive Linear Association**  | **Negative Linear Association** |
| **No Apparent Association** | **Nonlinear Association** |

If the variables show a linear association, we can determine whether that relationship is strong, weak, or perfect. Imagine drawing a line through the center of the points—EYEBALLING the line.

* If the data points are closely packed around your line, the linear relationship is a strong one.
* If the data points are more spread out from the line, the linear relationship is a weak one.
* If your data points fall on a straight line, the linear association is perfect.

We may also observe the following patterns in our data:

* Clusters - A cluster is a set of points that are in close proximity to each other.
* Outliers - An outlier is a data point that noticeably stands out from the general behavior of the data set.

Now that we have defined some of the typical associations, patterns, and trends seen in scatterplots, go back and check your answer in d. for the previous example. Can you add any math terms to your description? Can you add any details to your interpretation?

Directions: Describe the association between x and y using the terms from the previous page.

Circle any clusters in the data and put a star by any points that appear to be outliers.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

**Directions:** Examine the following scatter plots. *Describe the association between the two variables. Circle any clusters in the data. Put a star by any points that appear to be outliers.*

Use the context to give possible explanations as to why these trends, patterns, and associations exist.

The scatter plot given below shows the temperature of a cup of tea sitting on the counter for 30 minutes. The cup of tea is sitting in a room that is 70 degrees.



The Paradise Pool records the average daily temperature and the number of visitors to their pool for 18 days throughout the month of July. On July 24th, to celebrate Pioneer Day, admission is half off. The average daily temperature on that day is 90 degrees.

Visitors vs. Temperature at a Swimming Pool



The scatter plot below shows the population (in millions) and number of area codes for some states in the United States.

Area Codes and State Population



Holly’s math teacher asks her to conduct her own survey to study different types of association. She chooses to investigate the number of pets a person has and their shoe size.



Shoe Size vs. Number of Pets