Worksheet 10/23/2023

Two variables X and Y are **correlated** if higher values of X tend to associate with higher values of Y (**positive correlation**), or tend to associate with lower values of Y (**negative correlation**).

A **scatterplot** is a graph in which each point represents the values of two variables.

In Problem 1 and 2, make a scatterplot for the data. State whether the two variables appear to be correlated, and if so, state whether the correlation is positive, negative, strong, or weak. Suggest a reason for the correlation or lack of correlation.

1) The following table gives per capita personal income and percentage of the population below the poverty level for ten states in 2019.

State	Average income (dollars)	Percentage of population below poverty level
Connecticut	77,289	10.0%
Colorado	61,157	9.3%
Illinois	58,764	11.5%
Iowa	51,865	11.2%
Minnesota	58,834	9.0%
Montana	49,747	12.6%
Nevada	51,161	12.5%
New Hampshire	63,502	7.3%
Utah	48,939	8.9%
Mississippi	38,914	19.6%

2) The table below shows the percentage increase in average men's and women's heights between 1914 and 2014 in nine selected countries.

Country	% Increase in average men's heights	% Increase in average women's heights
Bangladesh	3.3	6.6
Canada	4.3	4.0
China	6.7	6.4
Iran	10.5	7.5
Philippines	5.5	1.2
South Korea	9.5	14.2
Switzerland	6.5	3.8
United Kingdom	6.4	7.2
United States	3.5	3.2

Problem 3,4,5 are statements about a correlation. In each problem, state the correlation (for example, there is a positive correlation between variable A and variable B). Then state whether the correlation is most likely due to coincidence, a common underlying cause, or a direct cause. Explain your answer.

3) In recent years, the population of bald eagles in the United States has increased along with major stock market indicators.

4) Automobile gas mileage decreases with tire pressure.

5) In the past three decades, levels of atmospheric carbon dioxide have increased while the volume of Arctic ice has decreased.

6) The following table gives the ages of the Presidents of the United States (in order) at the time they were inaugurated. Make a frequency table of the data for the ages using 5-year age bins (e.g., 40 to 44, 45 to 49). Then draw a histogram to display the binned data.

Order	1	2	3	4	5	6	7	8	9	10
Age	57	61	57	57	58	57	61	54	68	51
Order	11	12	13	14	15	16	17	18	19	20
Age	49	64	50	48	65	52	56	46	54	49
Order	21	22	23	24	25	26	27	28	29	30
Age	51	47	55	55	54	42	51	56	55	51
Order	31	32	33	34	35	36	37	38	39	40
Age	54	51	60	62	43	55	56	61	52	69
Order	41	42	43	44	45	46				
Age	64	46	54	47	70	78				