Name: Date: Class:

Pre-Activity Worksheet Answer Key

Section 1: Statistics Review: Summarizing Data

Data Distribution

Circle the correct answer:

Sample #	Data Set A	Data Set B
1	5	2
2	4	3
3	7	2
4	5	14
5	4	1

Which data set has a higher mean? (A) B

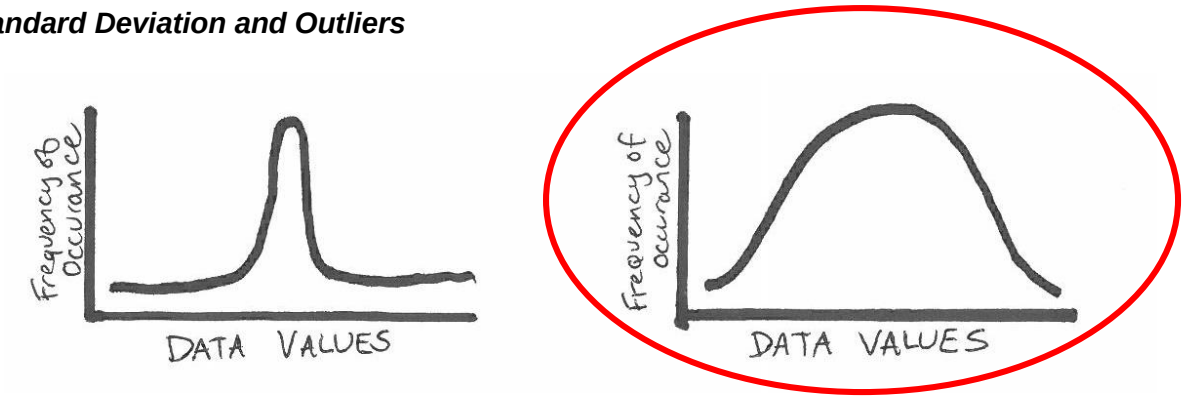
Which data set has a higher median? (A) B

Which data set has a larger range? A (B)

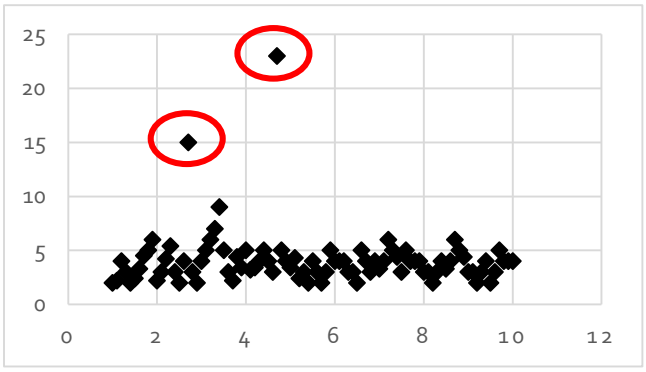
If you were to collect more samples and the mean and median for the above data remained the same, which data set would you expect to be normally distributed? (A) B

(Hint: In a normal distribution, mean = median. In a skewed distribution, mean ≠ median.)

Standard Deviation and Outliers



Circle the figure that has the higher standard deviation.



In the chart above, circle any points you suspect to be outliers.

Name: Date: Class:

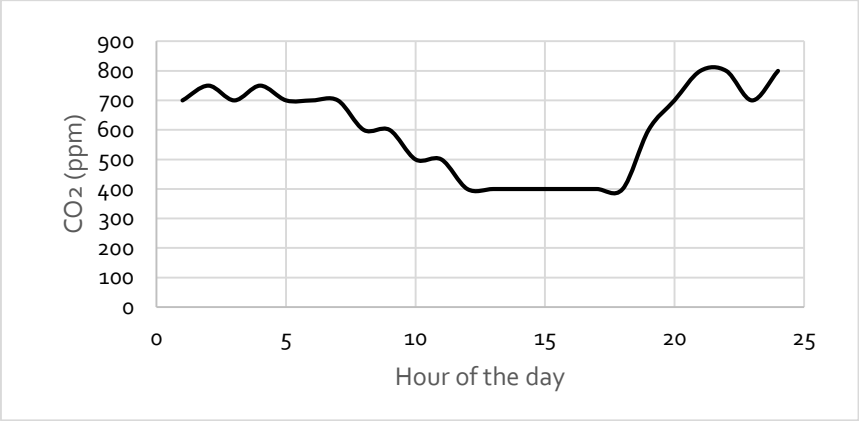
Section 2: Visualizing Data, Graphing

Imagine you have collected air quality data inside your home, and now you want to analyze the data from one 24-hour period. Focus on the pollutant—carbon dioxide (CO_2). *What type of plot would you choose?*

Next, make a sketch of what you might expect this plot to look like. Feel free to annotate your plot with activities such as sleeping, left home, returned home, etc. *(Hint: Consider where CO_2 comes from, and how these "sources" might change throughout the day.)*

Plot type: **Time series**

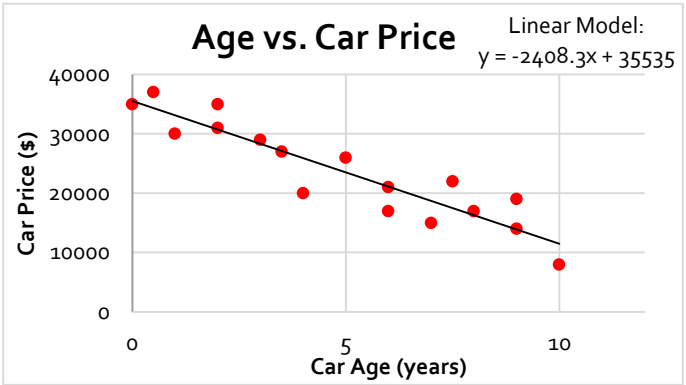
Plot sketch:



Interpretation: Higher steady CO_2 overnight, drops when I leave home, sharper rise when I come back home (cook dinner, etc.) and then becomes more steady again over night.

Section 3: Comparing Data Sets

Take a look at the plot below of hypothetical data on car ages and their prices. Do you see a relationship in the data? Does this make sense? Why or why not? Estimate the R-squared for this data set. (Remember R^2 is explained in the *Pre-Activity Reading* as a value between 0 and 1.)



The graph shows a negative relationship between car age and value, which makes physical sense as cars depreciate in value over time. I estimate the R-squared to be approximately .85, because the data follow the linear trend fairly well, but there is still some variance.

Bonus Activity

Google "air quality infographic" and click on the image results. Skim through these and find one that interests you. Be prepared to share a one-sentence summary of the infographic and why you liked it.