

1.1 Statistics: The Science and Art of Data

Individuals The person or thing described in the data set

Variables Any attribute for the individuals that can vary there are two types of variables

Categorical Variable Has a label (favorite color)

Quantitative Variable Numerical Value and makes sense to find an average (age)

Frequency table Shows the count of each data point (Blue: 10)

Relative Frequency table Shows the percentage of each data point.

Percentage Formula $(\text{Part/Whole}) \times 100$

Dot Plots Each dot represents one data point (don't skip values that don't have data points)

1.3-4: Displaying Quantitative Data (cont)

Described in context The dot plot represents the number of books read last summer. The data is skewed to the right. The center is around around 1-2 books read last summer. The number of books read varies from 0 to 9 books. There is an outlier of 9 books.

Stemplots Stem and leaf plots organize quantitative data using the digits of the values

Leaf The last number/digit of the data point

Stem the other numbers before the last number/digit

Split Stemplot when each stem has too many data point the stems can be split in two: leaves 0-4 and leaves 5-9

Info for stemplots
- Don't skip stems even if they don't have any leaves!
- Always include a KEY!!

Back-to-Back Stemplot To compare two groups of data you can create a stemplot with leaves on either side. The left represents one group of data and the other side represents the other group

1.3-4: Displaying Quantitative Data

Skewed Left The tail is on the left (more data is on the right)

Skewed Right The tail is on the right (more data is on the left)

Symmetrical The data is split evenly (if you folded it, it would be similar)

Shape Skewed left/right or symmetric

Center The typical value in the data set

Variability How spread out the data is (Variability from __ to __)

Outliers Values significantly far from the others

In context Always connect the description to the actual scenario or context of the data

1.7: Measuring Variability

Standard Deviation (S_x) The average distance from the mean

Formula: S_x For the sake of your sanity PLEASE use your calculator

Interquartile Range (IQR) The range of the middle 50% of the data

Formula for IQR $IQR = Q_3 - Q_1$



By **Lisa Dance**
(SarahLisaDance)

cheatography.com/sarahlisadance/

Published 17th October, 2025.

Last updated 21st October, 2025.

Page 1 of 2.

Sponsored by **ApolloPad.com**

Everyone has a novel in them. Finish

Yours!

<https://apollopad.com>

1.2: Displaying Categorical Data

Bar Charts	each bar represents one category and the frequency or relative frequency (Helps compare data side by side)
Pie Chart	the circle represents the whole data set and each wedge represents the relative frequency of a category. Represents data as a part of a whole
Deceptive graphs	Some representation of data are created to manipulate the perception of the data (Always check the scales on a chart and beware of pictographs)

1.5: Displaying Quantitative Data

Step 1	Divide your data into equal intervals
Step 2	Create a frequency table for each interval
Step 3	Label the axis. Label the horizontal axis with the intervals
Step 4	Draw the bars for each interval. (no gaps)
Notes on Histograms	The interval contains the first value but not the last
Relative Frequency histogram	- Use the same steps but with a relative frequency table - When made correctly, all the bars in a histogram should add up to 100% or 1
Shape	Skewed left/right or symmetric
Center	The typical interval in the data set
Variability	How spread out the data is
Outliers	intervals significantly far from the others

1.6: Measuring Center

Mean	The average value of a data set
Median	The value in the middle of a data set
Symmetric	Mean $\approx$ Median - use the mean for center
Skewed Right	Mean $>$ Median - use the median for center
Skewed Left	Mean $<$ Median - use the median for center
Notes on Mean vs. Median	If there are outliers, use median for center since median is resistant to outliers. Mean is not resistant to outliers
Mode	The number that appears the most
Range	The difference between the highest and lowest data values
Quartiles	Dividing a data set into four intervals
First (Lower) Quartile	The middle of the lower half of the data set. The first quarter
Second Quartile (Median)	Middle Value. The second Quarter of half
Third (Upper) Quartile	Middle of the upper half of the data set. The third quarter

1.8: Boxplots and Outliers

Boxplots	displays data using Min, Max, Q ₁ , Q ₃ , median, and outlier values
Step 1	Create a numberline from the variability of the data
Step 2	Make the Q ₁ , Median, and Q ₃ . draw a box between Q ₁ and Q ₃ with a line through the median
Step 3	Mark the Min and Max (excluding the outlier values) and connect them to the box with a line
Step 4	Add * to mark the high and low outliers



By **Lisa Dance**
(SarahLisaDance)

cheatography.com/sarahlisadance/

Published 17th October, 2025.
Last updated 21st October, 2025.
Page 2 of 2.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
<https://apollopad.com>