Statistics Honors – Histograms (Day 1)   Name ___**KEY**___

HISTOGRAMS

Rows and rows of data provide little information. For example, below are thickness measurements, in millimeters, from a sample of 25 polished wafers used in the manufacture of microchips. Notice that it is difficult to extract much information from staring at these numbers. The numbers need to be organized, summarized, and displayed graphically in order to unlock the information they contain.

0.402	0.496	0.533	0.387	0.384
0.528	0.411	0.367	0.462	0.499
0.539	0.546	0.425	0.457	0.586
0.558	0.588	0.425	0.437	0.479
0.427	0.485	0.443	0.441	0.658

A *frequency distribution* is one method of *organizing and summarizing data in a table*. The basic idea behind a frequency distribution is to set up categories (class intervals), classify data values into the categories, and then determine the frequency with which data values are placed into each category.

Creating a frequency distribution table

Step 1: Identify an interval that is wide enough to contain all the data. (Find your lowest and highest data value.)

Step 2: Subdivide the interval identified in Step 1 into class intervals of equal width. The class intervals will serve as the categories.

Step 3: Set up a table with three columns for the following: class interval, tally, and frequency.

Step 4: To complete the table, determine the frequency with which data values fall into each class interval.

Convention: Any data value that falls on a class interval boundary is placed in the class interval to the right.

For example: The value of 0.35 would fall in the class interval of [0.35, 0.40). ✱ interval notation

---

Step 1: Interval: ___.367___ to ___.658___   ✱ Put data into List & sort
                                                            STAT, edit, enter data

Step 2: Class width: ___.05___   What would be the first class? [.35, .40)

Steps 3 & 4: Construct and complete Frequency Table.

Frequency Distribution Table

Thickness (mm)	Tally	Frequency
0.30 – 0.35		0
0.35 – 0.40	III	3
0.40 – 0.45	HH III	8
0.45 – 0.50	HH I	6
0.50 – 0.55	IIII	4
0.55 – 0.60	III	3
0.60 – 0.65		0
0.65 – 0.70	I	1

➢ The frequency distribution table reveals more information about the data.

➢ From the frequency distribution table, where do the most measurement fall ? .40 - .45

➢ Do you see any gaps in the data? If so, in what class interval? .60 - .65

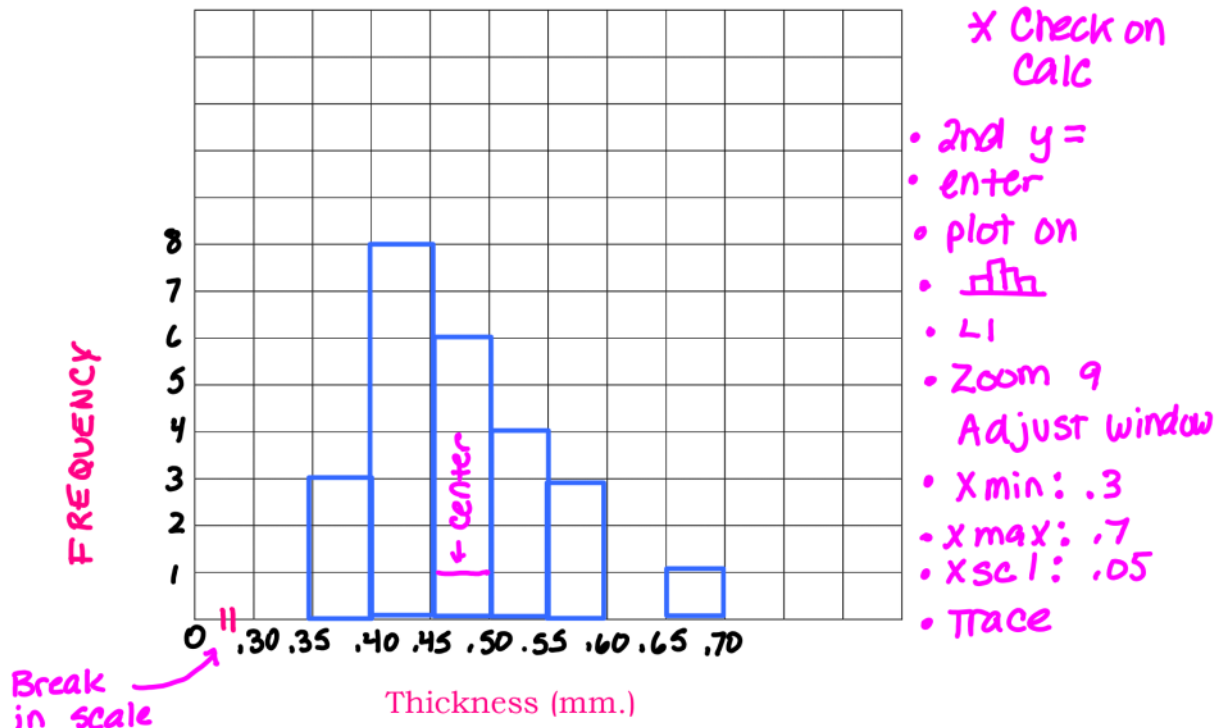### Creating a histogram from a frequency distribution

Step 1: Draw a set of axes and **include a title!** On the horizontal axis, mark the boundaries of the class intervals. On the vertical axis, set up a scale appropriate for the frequencies. You may need to use a "break in scale" indicator.

Step 2: Label the horizontal axis with the name of the variable being measured and the units. Label the vertical axis with the word "frequency".

Step 3: Over each interval, draw a rectangle with the interval as its base. The height of the rectangle should match the frequency of data contained in that interval.

Construct the Histogram:

Polished Wafer Thickness Measurements

FREQUENCY — Thickness (mm.)

Break in scale

← center

X Check on Calc
• 2nd y =
• enter
• plot on
• (histogram icon)
• L1
• Zoom 9
  Adjust window
• Xmin: .3
• Xmax: .7
• Xscl: .05
• Trace

In describing a histogram, we first look for the overall pattern of the distribution. In sizing up the overall pattern, look for the following: (SOCS)

Outliers: An individual observation that lies clearly outside the overall pattern.

There is a possible outlier in the [.65 -.70) class.

Center: The class in which the midpoint falls. *Count to the 12.5th box*
*25 data entries. 25÷2 = 12.5 * what class does it fall in?* *[.45-.50)*

Spread: The range of data. (lowest to highest value)

.35 to .70

Shape: A regular shape, such as symmetric or skewed; one peak or several. (unimodal or multimodal) **Be sure to include OUTLIERS when determining shape.**

unimodal and roughly skewed to the right.

TAIL   *TAIL*

Skewed left   Roughly Symmetric   Skewed right
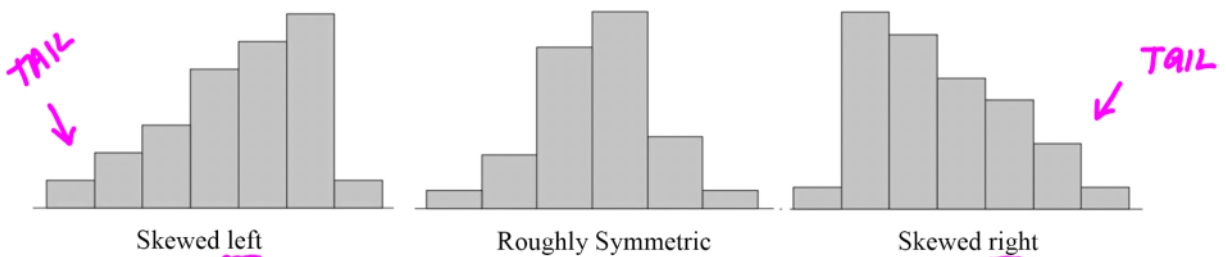
Homework #5A: Histogram Vocabulary Check.
MUST BE COMPLETED IN HW NOTEBOOK

Fill in the blanks using the word bank

Histogram	Skewed to the left (negatively skewed)
Symmetric	Frequency Distribution
Skewed to the right (positively skewed)	

1. A *FREQUENCY DISTRIBUTION* provides a means of organizing and summarizing data by classifying data values into class intervals and recording the number of data that fall into each class interval.

2. A *Histogram* is a graphical representation of a frequency distribution. Bars are drawn over each class interval on a number line. The areas of the bars are proportional to the frequencies with which data fall into the class intervals.

3. The shape of a unimodal distribution of a quantitative variable may be *Symmetric* (right side close to a mirror image of left side) or skewed to the right or left.

4. A distribution is *skewed right* if the right tail of the distribution is longer than the left and is *skewed left* if the left tail of the distribution is longer than the right.