

Name: _____ Stats Review

1. Find:
 Mean 11345.9 Mode NONE
 Q1 8459 Median 12,664
 Q3 13,038 Outliers NONE
 Standard Deviation 2234.21
 Type of distribution Skewed left

Day	Number of Births
Sunday	7,374
Monday	11,704
Tuesday	13,169
Wednesday	13,038
Thursday	13,013
Friday	12,664
Saturday	8,459

2. Given the number of siblings that students in my class have, find the following:
 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 5, 6

Mean 2
 Mode 1
 Q1 1
 Median 2
 Q3 3

Outliers NONE
 Standard Deviation 1.55
 Type of distribution skewed right
 Which measure of center is better and why?
Median b/c data is skewed

3. Barry Bonds set the major League record by hitting 73 home runs in a single year. In his 21 seasons he hit the following number of home runs:

16 25 24 19 33 25 34 46 37 33 42
 40 37 34 49 73 46 45 45 26 28

Min=16 Q₁=25.5 Med=34 Q₃=45 Max=73

- Find the 5 number summary.
- Are there any outliers? If so name them NONE
- Find his mean number of home runs 30
- Find the standard Deviation 12.4
- What does the standard deviation mean? Most runs fall b/w $\bar{x}-\sigma$ & $\bar{x}+\sigma$
- What type of distribution is it? Skewed right
- Which is the best choice for the measure of center and why?
Median b/c skewed.

4. Below are tables with results from the 2014 ACT.

States in which all graduates were tested	CO	IL	KY	LA	MI	MS	MT	NC	ND	TN	UT	WY		
Average composite score	20.6	20.7	19.9	19.2	20.1	19.0	20.5	18.9	20.6	19.8	20.8	20.1		
States in which less than 70% of graduates were tested	ME	RI	DE	PA	NH	MD	WA	MA	NJ	NY	VA	CA	CT	VT
Average composite score	23.6	22.9	23.2	22.7	24.2	22.6	23	24.3	23.1	23.4	22.8	22.3	24.2	23.2

Source: 2014 ACT National and State Scores <http://www.act.org/newsroom/data/2014/states.html>

- Describe the distribution of each set of data.
- Compare the appropriate measures of central tendency and variation.
- What conclusions, if any, can you draw from your findings?

The number of pages a print cartridge can print before needing to be replaced is normally distributed. The mean of a certain printer cartridge is 480 pages before needing to be replaced with a standard deviation of 20 pages. A ge office building places a bulk order for 300 of those print cartridges.

- How many of the 300 print cartridges should be expected to print between 460 and 500 pages before needing to be replaced? $2^{nd} \rightarrow Var \rightarrow normal\ cdf \rightarrow (upper, lower, \bar{x}, sd)$
- How many of the 300 print cartridges should be expected to print between 440 and 520 pages? $ncdf(440, 500, 480, 20) = .68 \rightarrow 300(.68) = 204$

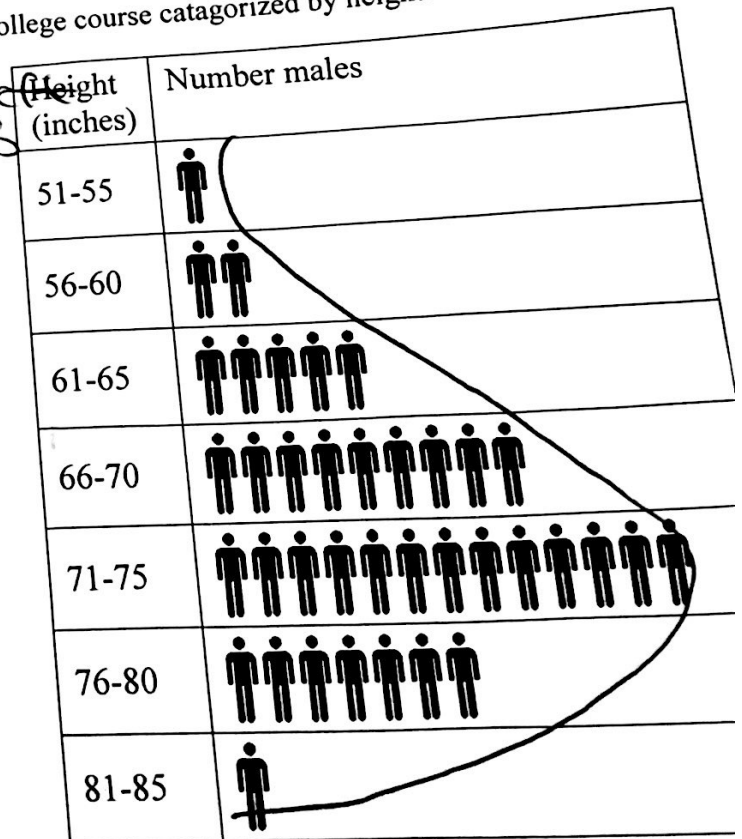
$ncdf(440, 520, 480, 20) = .95 \rightarrow 300(.95) = 285$

The frequency chart below shows the number of males in a college course categorized by height.

- What is the shape of the distribution? *skewed left*
- Estimate the mean and the median. *blw 66-75*
- How might this chart and distribution be effected if the data for the females were included?

~~The graph is skewed~~

Free response



Match each vocab word with the correct definition.

Name: _____

Sample	Median	Data
Population	Element	Mean

1. Collection of information in context data
2. Subset of the population where data has been collected sample
3. Sum of a data set divided by the number of data collected mean
4. Set of all elements that someone wishes to describe or make predictions about population
5. One specific item from a set of data is referred to as element
6. Middle value, or mean of the middle two values, when data is in numerical order median

Fill in the blank:

- Quantitative Data is data that measures a numerical value/characteristic about an object.
- Categorical Data is data that records or names qualities/characteristics about an object.

Determine which of the following are Categorical (C) or Quantitative (Q) types of data

1. Eye color C
2. Type of Shoe C
3. Number of cousins Q
4. Phone Number C
5. Number of skittles in a package Q
6. Year & Model of a car C
7. Area Code C
8. Length of your foot Q
9. Number of cookies in the cookie jar Q
10. Weight of a shark Q
11. Type of Hair C
12. Length of Hair Q
13. Number of Hairs Q
14. Amount of Change in Pocket Q
15. Zip Code C
16. Candidate for upcoming election C

Determine which measure of spread should be used for each of the following scenarios (Range, Interquartile Standard Deviation):

There is a boxplot displaying the average cost of homes in Franklin County. IQR

Exact data has been collected regarding the number of siblings for everyone in the school. Standard Dev.

A histogram shows the frequency of people who are within certain age groups. Range

Describing the center of a set of data when should you use mean or median?

When there are extreme values (outliers), use Median.

When there are no extreme values, use Mean.

A math student is interested in figuring out the average price of phones at Heritage High School. She takes a sample of 100 students in the school's cafeteria and finds the average value to be \$650.

a. What is the population?

All students at Heritage

b. What are the elements?

Each student's individual phone price

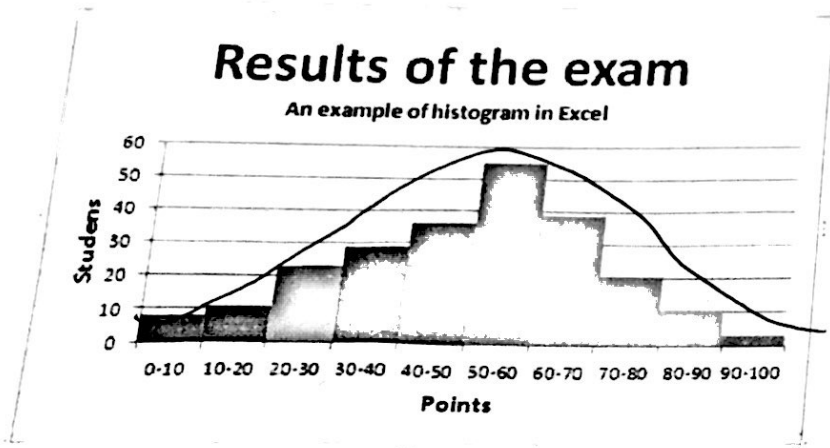
c. What data is being collected? (Include units if applicable)

Average price of phones

d. What type of data is it (categorical or quantitative)? How do you know?

Quantitative because it is measuring \$.

Describe that data in terms of center, spread, shape, and outliers. State all in terms of the context.

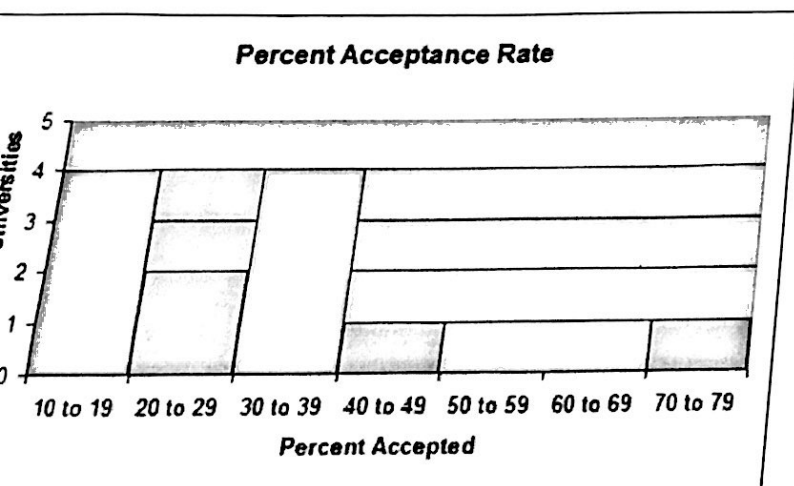


Shape- symmetrical

Center- b/w 40-60

Spread- Range 100

Outliers- NONE



Shape- skewed right

Center- b/w 20-39

Spread- Range 69

Outliers-

Potentially b/w 70-79