1. In the following study, identify the (implied) population, the sample, and any statistics and parameters:

   A survey (reported in *USA Today*) asked 1000 adults where they hide their messes when unexpected company comes. The survey showed that 68% tossed their mess in the closet, 23% shoved things under the bed, 6% put things in the bathtub, and 3% put their mess in the freezer.

   Population: All adults (or all US adults)

   Sample: The 1000 adults who were surveyed

   Statistic(s): 68%, 23%, 6%, 3%

   Parameter(s): none

2. Identify the data set's level of measurement (nominal, ordinal, interval, ratio), and whether the data set is qualitative or quantitative:

   (a) The highest level of education you have received ("some high school," "high school diploma," "some college," "bachelor's degree," "some graduate school," "graduate degree")

   Ordinal, qualitative

   (b) The highway fuel consumption (in miles per gallon) of a passenger vehicle.

   Ratio, quantitative

   (c) The high temperature at Alta for a day in the ski season (in degrees Fahrenheit).

   Interval, quantitative

   (d) The location where you hide your mess when company comes ("closet," "under the bed," "bathtub," "freezer," "other")

   Nominal, qualitative

   (e) The year in which you first enrolled at the University of Utah.

   Interval, quantitative

3. For each of the following studies, identify which method of data collection would be best (observational study, experiment, simulation, survey):

   (a) Determining how often people eat at fast food restaurants.

   Survey

   (b) Determining the average gas mileage that a new car gets.

   Observational study

   (c) Determining the effects of stopping the cooling process of a nuclear reactor.

   Simulation

   (d) Determining the nitrogen concentration in a body of water.

   Observational study

   (e) Determining the effect on bone mass of a calcium supplement given to young children.

   Experiment

4. Modern Managed Hospitals (MMH) is a national chain of hospitals. Management wants to survey patients discharged this past year to obtain information on patient satisfaction. For each of the following methods, identify the sampling technique used (random, cluster, stratified, systematic, convenience):

   (a) Obtain a list of patients discharged from all MMH facilities. Divide the patients according to length of stay (2 days of less, 3–7 days, 8–14 days, more than 14 days). Draw a simple random sample from each group.

   Stratified

   (b) Obtain lists of patients discharged from all MMH facilities. Number these patients, and then use a random-number table to obtain the sample.

   Random

   (c) Randomly select some MMH facilities from each of the five geographic regions, and then include all the patients on the discharge lists of the selected hospitals.

   Cluster (stratified sampling is used to choose which MMH facilities are surveyed)

   (d) At the beginning of the year, instruct each MMH facility to survey every 500th patient discharged.

   Systematic

(e) Instruct each MMH facility to survey 10 discharged patients this week and send in the results.
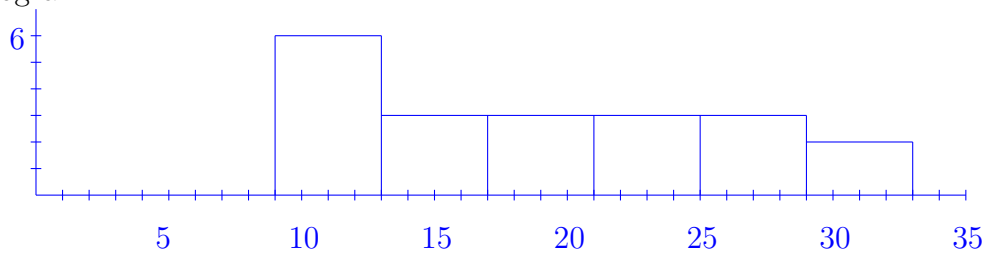
Convenience

5. Construct a frequency distribution, a histogram, and a cumulative frequency graph for the following data using 6 classes:
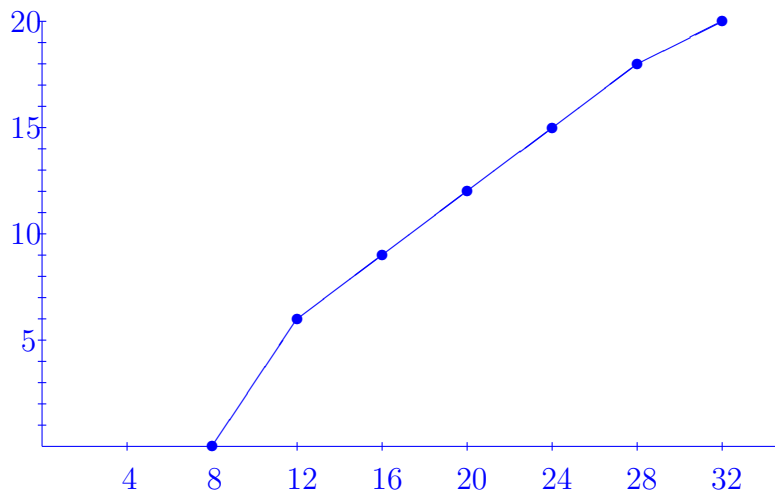
9 9 9 10 11 12 14 14 16 17 19 20 21 21 21 25 26 28 29 29

| Class	Frequency	Relative Frequency	Cumulative Frequency
9–12	6	0.3	6
13–16	3	0.15	9
17–20	3	0.15	12
21–24	3	0.15	15
25–28	3	0.15	18
29–32	2	0.1	20

Histogram:

Cumulative frequency graph:

6. Sketch a Pareto Chart for the following data: According to the M&M/Mars Department of Consumer Affairs, the distribution of colors for plain M&M candies is:

Color	Red	Orange	Yellow	Green	Blue	Brown
Percentage	13	20	14	16	24	13
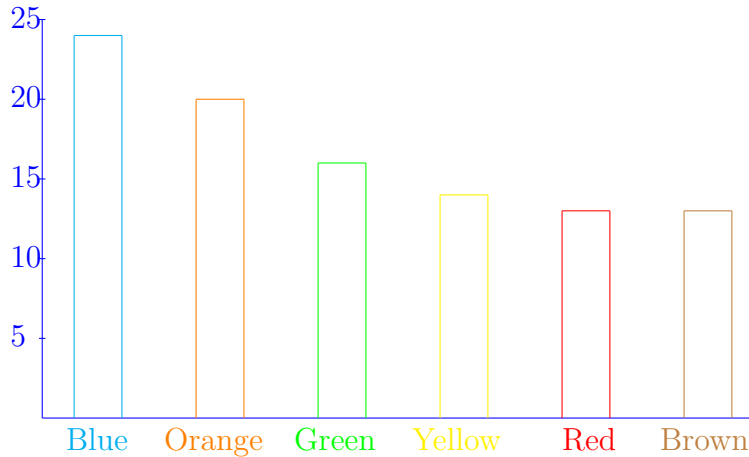
Note: Since Red and Brown had the same percentages, either one could be put first.

7. Use a scatter plot to display the data shown: A sociologist is interested in the relation between the number of job changes and the annual salary (in thousands of dollars) of people living in the Nashville area. A random sample of 10 people employed in the Nashville area provided:

Number of job changes	4	7	5	6	1	5	9	10	10	3
Salary	33	37	34	32	32	38	43	37	40	33
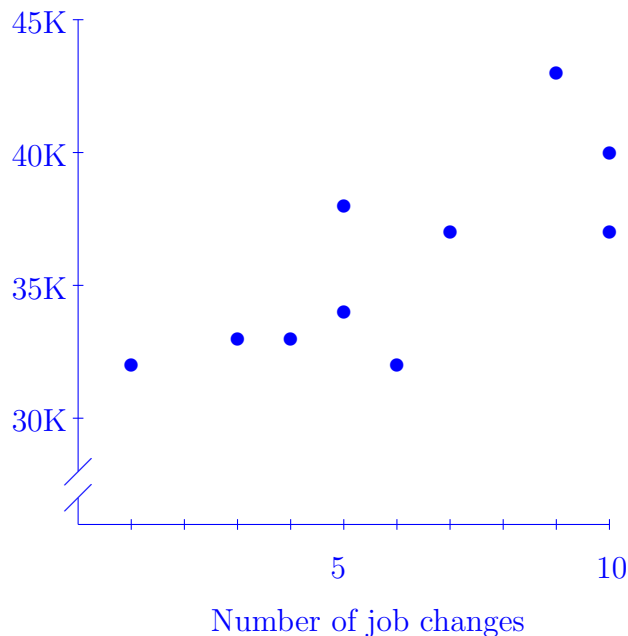
8. Use a Time Series Graph to display how many miles Jill walked/jogged in 30 minutes for her first 8 weeks after starting an exercise program:

Week	1	2	3	4	5	6	7	8
Distance	1.5	1.4	1.7	1.6	1.9	2.0	1.8	2.0
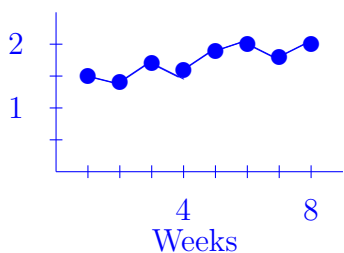
(Vertical axis is distance walked/jogged)

9. Find the mean, median, and mode of the following data. A random sample of professional football players' ages gave:

23 23 24 25 26 28 29 30

$$\bar{x} = \frac{23 + 23 + 24 + 25 + 26 + 28 + 29 + 30}{8} = \frac{208}{8} = 26$$
$$\text{Median} = \frac{25 + 26}{2} = 25.5$$
$$\text{Mode} = 23$$

10. Calculate the range, mean, variance, and standard deviation of the following data set. Five ground temperatures in degrees Fahrenheit in Furnace Creek (part of Death Valley) taken between May and November are:

146 152 168 174 180

$$\text{Range} = 180 - 146 = 34$$
$$\bar{x} = \frac{146 + 152 + 168 + 174 + 180}{5} = \frac{820}{5} = 164$$
$$s^2 = \frac{(146 - 164)^2 + (152 - 164)^2 + (168 - 164)^2 + (174 - 164)^2 + (180 - 164)^2}{5 - 1}$$
$$= \frac{840}{4} = 210$$
$$s = \sqrt{s^2} = \sqrt{210} = 14.491$$

11. Sunshine Stereo cassette decks have a deck life that can be represented by a bell-shaped curve with a mean of 2.3 years and a standard deviation of 0.4 years. Use the Empirical Rule to determine the percentage of cassette decks that break down between 1.9 and 2.7 years after purchase. What is the percentage of cassette decks that break down between 1.9 and 3.1 years?

By the Empirical Rule, the percentage of cassette decks that will break down between 1.9 and 2.7 years (within one standard deviation of the mean) is 68%.

The percentage that will break down within one standard deviation less than and two standard deviations above the mean is $\frac{68}{2} + \frac{95}{2} = 81.5\%$

12. For the following data set, find the five-number summary and draw a box-and-whisker plot: The weights of 10 people's carry-on luggage in pounds are:

0 17 19 22 27 30 32 33 38 43

$\text{Min} = 0$
$Q_1 = 19$
$Q_2 = 28.5$
$Q_3 = 33$
$\text{Max} = 43$
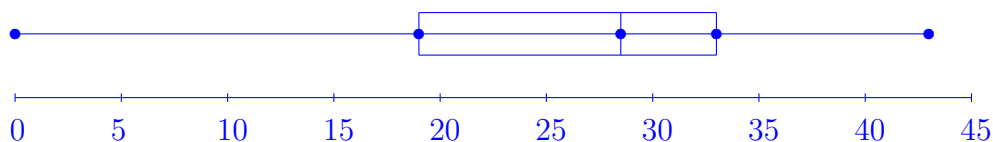
Formulas to Know

- Range = Maximum − Minimum

- For frequency distributions and histograms: Class width = $\dfrac{\text{Range}}{\text{number of classes}}$, then round up

- Mean: $\bar{x} = \dfrac{\sum x}{n}$

- Standard Deviation: $s = \sqrt{s^2} = \sqrt{\dfrac{\sum (x - \bar{x})^2}{n - 1}}$, where s^2 is the variance (the formula for variance will be given)

- Also know how to find the median, mode, and quartiles.