Ex 1. The individuals are Bob, Sue, Bill and John. The variables are Gender, Age and Number of siblings. Gender is categorical, Age & number of siblings are quantitative.

Ex 2. The distribution of height for adult men and women in the US is as follows:

- Mean height for men: 69.1 inches, Standard deviation: 29 inches
- Mean height for women: 63.7 inches, Standard deviation: 27 inches
The distribution of Columbia students probably has the same general shape and center. One might expect that the histogram over all Columbia students would be bimodal but this is probably not the case.

It isn't the case for all adults in the US which takes the following form:

![Histogram Graph]

Ex 3 (a) \[6, 3, 2, 4, 21\]

\[\bar{x} = \frac{6 + 3 + 2 + 4 + 21}{5} = 7.2\]

2 3 [4] 6 21

\(M = 4\)

The mean is larger because the value 21 has a large effect on its value. It does not affect the median.
(b) \( \{1, 1, 2, 2, 2, 4, 6\} \)
\[
\bar{x} = \frac{1+1+2+2+2+4+6}{7} = \frac{18}{7}
\]
1 1 2 2 2 4 6 \( M = 2 \)

(c) \( \{1, 2, 3, 6, 8, 8, 10, 14, 15, 17\} \)
\[
\begin{array}{cccccccc}
1 & 2 & 3 & 6 & 8 & 8 & 10 & 14 & 15 & 17 \\
\hline
M = \frac{8+8}{2} = 8
\end{array}
\]

Find \( Q_1 \): 1 2 3 6 8 \( Q_1 = 3 \)

Find \( Q_3 \): 8 10 14 15 17 \( Q_3 = 14 \)

\[ IQR = Q_3 - Q_1 = 14 - 3 = 11 \]

(d) \( \{2, 4, 3, 7\} \)
\[
\bar{x} = \frac{15}{4} = 4 \quad x_1 - \bar{x} = -2 \quad (x_1 - \bar{x})^2 = 4 \\
x_2 - \bar{x} = 0 \quad (x_2 - \bar{x})^2 = 0 \\
x_3 - \bar{x} = -1 \quad (x_3 - \bar{x})^2 = 1 \\
x_4 - \bar{x} = 3 \quad (x_4 - \bar{x})^2 = 9 \\
\]
\[ s = \sqrt{\frac{1}{4-1} \left(4 + 0 + 1 + 9\right)} = \sqrt{\frac{14}{3}} \]
Exercise: The distance between the median and $Q_1$ is small compared to the distance between the median and $Q_3$. Similarly, the distance between the minimum and $Q_1$ and $M$ is smaller than the distance between the maximum and $Q_3$ and $M$.

The distribution is skewed to the right.

It could possibly look something like this, but we don’t know for sure.

(b) The mean will be higher than the median since the distribution is skewed to the right.

(c) $IQR = 3.6 - 2.8 = 0.8$

$1.5 \times IQR = 1.2$

Anything below $2.8 - 1.2 = 1.6$ is an outlier.
Above $3.6 + 1.2 = 4.8$ is an outlier.

4.9 is an outlier.
Smallest value not considered an outlier - 2.3 lbs
Largest value - 4.7 lbs