11.12. Here is how we would interpret the histogram in terms of our four characteristics:

- Center: The center of race horse speed distribution is around 36 mph.
- Variability: Almost all of the speeds are between 33.75 and 37.75 mph.
- Shape: This distribution has a single peak (slightly below 37 mph) and is **skewed to the left**.
- Deviations: There is an outlier slightly below 33.5 mph.

(b) Justify's speed of 36.23 mph in 2018 falls right in the center of the distribution. It is a typical speed when compared to the speeds of other winning horses.

11.13. Below is a stemplot of the gas mileage data. The "stem" is the tens digit (hundreds if needed) and the "leaf" is the units digit.

```
> stem(mileage,scale=2)

   1 | 34588
   2 | 122222344579
   3 | 001117
   4 | 27
   5 | 0
   6 |
   7 |
   8 | 488
   9 | 5
  10 |
  11 | 4
```

Here is how we would interpret the stemplot in terms of our four characteristics:

- Center: The center of gas mileage distribution is around 30 mpg.
- Variability: All of the gas-driven cars have mileage between 13 and 50 mpg.
- Shape: This distribution has a single peak (around 29 mpg) and is **skewed to the right**.
- Deviations: The values on the high (right) side correspond to the electric vehicles (84, 88, 88, 95, and 114 mpg). They are outliers when compared to the rest of the distribution.

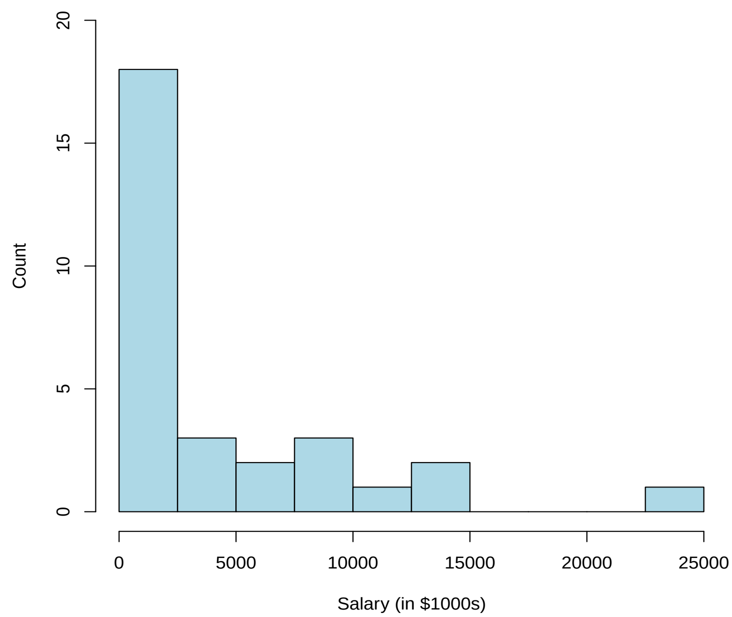
Here is the R code I used to prepare the stemplot above:

```
mileage = c(24,22,27,13,21,22,14,22,29,31,18,88,88,42,31,47,37,22,18,30,24,25,
    30,84,23,31,114,15,50,95,22)
stem(mileage,scale=2)
```

11.15. The histogram of the Minnesota Twins salaries is shown below. The first thing I did was change the salaries to be in \$1000s. For example,

$$23,000,000 \text{ was recorded as } 23,000 \text{ (in \$1000s).}$$

I did this so I wouldn't have to deal with really large numbers. Because we are doing this for each observation, it won't change the shape of the distribution. Also, I didn't like R's default selections for interval widths (it was 5000), so I changed them to 2500. This gave a better picture of the distribution (see code below).



Here is how we would interpret the histogram in terms of our four characteristics:

- Center: The center of the salary distribution is around 2500 (in \$1000s, so this is really \$2,500,000).
- Variability: Almost all of salaries are between 545 and 13500 (in \$1000s).
- Shape: This distribution has a single peak (around 1000, in \$1000s) and is **skewed to the right**.
- Deviations: There is an outlier at 23000 (in \$1000s). This is Joe Mauer's salary.

Here is the R code I used to prepare the histogram above:

```
salary = c(23000,13500,13200,12000,9000,8250,8000,6500,6300,4850,4500,4200,2150,
    2000,2000,1600,1000,650,602,602,587,585,580,575,570,565,555,547,545,545)
bins = seq(0,25000,2500)
hist(salary,breaks=bins,xlab="Salary (in $1000s)",ylab="Count",main="",
    col="lightblue",xlim=c(0,25000),ylim=c(0,20))
```

11.18. Below is a stemplot of the Asian population data (in state population percentages). The "stem" is the units digit and the "leaf" is the tenths digit.

```
> stem(percent)
  The decimal point is at the |

  0 | 79
  1 | 01133467
  2 | 2234479
  3 | 228
  4 | 6
  5 | 355
  6 |
  7 | 3
  8 | 3
```

Here is how we would interpret the stemplot in terms of our four characteristics:

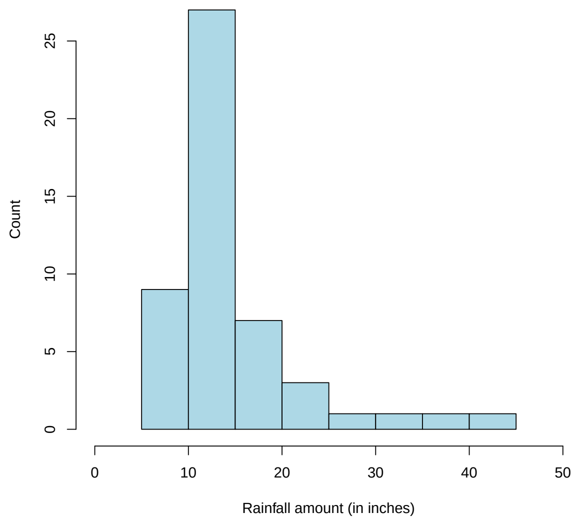
- Center: The center of Asian population percentage distribution is around 2.5 percent.
- Variability: Most of the percentages are between 0.7 and 5.5.
- Shape: This distribution has a single peak (around 1.7 percent) and is **skewed to the right**.
- Deviations: There are two outliers (7.3 and 8.3 percent). These are the percentages for New York and New Jersey.

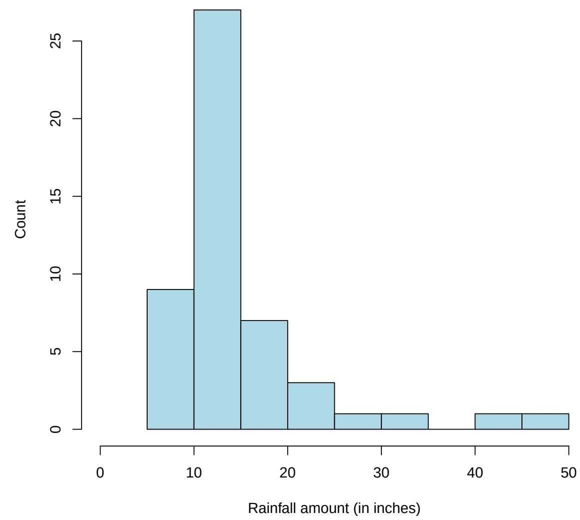
Here is the R code I used to prepare the stemplot above:

```
percent = c(1.1,2.4,1.6,5.5,0.9,7.3,2.7,1.4,0.7,3.8,3.2,1.1,5.3,2.2,2.2,2.9,1.3,
    2.3,3.2,4.6,1.0,2.4,8.3,1.7,1.3,5.5)
stem(percent)
```

11.23. I prepared the histogram in R as requested for the maximum rainfall amount given in the table. This histogram is shown on the top of the next page (left). Here is how we would interpret this histogram in terms of our four characteristics:

- Center: The center of the rainfall amount distribution is around 15 inches.
- Variability: All of the rainfall amounts are between 5 and 45 inches.
- Shape: This distribution has a single peak (around 12.5 inches) and is **skewed to the right**.
- Deviations: There are no obvious outliers.





When we change the "38.00" inches for Hawaii to "49.69," the histogram (on the right) changes slightly in appearance. Now, it appears as though there are two outliers: Texas (42.00 inches) and Hawaii (49.69 inches).

Here is the R code I used to make the histogram on the left (obvious changes to prepare the one on the right):

```
precipitation = c(32.52,15.2,11.4,14.06,25.83,11.08,12.77,8.5,12.28,21.1,38,
    7.17,16.91,10.5,13.18,13.53,10.4,22,13.32,14.75,18.15,9.78,15.1,15.68,
    18.18,11.5,13.15,7.78,11.07,14.81,11.28,11.15,22.22,8.1,10.75,15.68,11.77,
    13.5,12.13,14.8,8.74,13.6,42,5.08,9.92,14.28,14.26,12.02,11.72,6.06)
hist(precipitation,xlab="Rainfall amount (in inches)",ylab="Count",main="",
    col="lightblue",xlim=c(0,50))
```

11.24. (a) The average price distribution is not skewed to the left; it is skewed to the right side (the high side).
(b) It looks like 18 fruits and vegetables are between \$1.00 and \$1.50 per pound (18 is the frequency; i.e., the height of the histogram bar in this interval). This is not a majority. Note that

$$\frac{18}{58} \approx 0.31 \text{ (or 31\%)}.$$

A majority would be something over 50%.
(c) No, the horizontal axis does not measure time (like in line graphs). The horizontal axis gives the price for all 58 fruits and vegetables at one point in time (in 2016).
(d) This could be true, but it probably isn't. We do know there are 5 fruits and vegetables with average prices between \$2.50 and \$3.00. There are also 5 fruits and vegetables with average prices between \$3.00 and \$3.50.