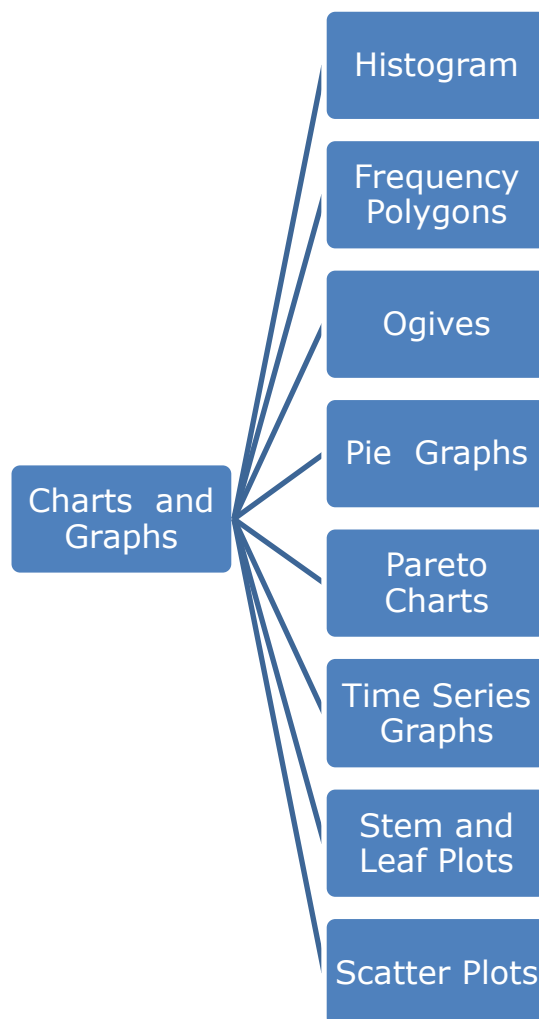


Chapter 2: Frequency Distributions and Graphs

Presenting frequency distributions as graphs

In a statistical study, researchers gather data that describe the particular variable under study. To present the data in a meaningful way a frequency distribution may be constructed. A *frequency distribution* is the raw data organized in a table. The raw data may also be manipulated into classes of grouped data; however, this process results in a loss of accuracy in subsequent calculations on the data classes. The most useful method to present the raw data is by constructing statistical charts and graphs.



A *Histogram* is a graph that displays the data by using *contiguous* vertical bars of various heights to represent the frequencies of the classes. Note that the height of a vertical bar may be zero.

A *Frequency Polygon* is a graph that displays the data by using lines that connect points plotted for the frequencies at the *midpoints* of the classes. The frequencies are represented by the heights of the points.

An *Ogive* is a graph that represents the *cumulative* frequencies at the *high-end class-boundary* of the classes by points. The points are then connected by a line.

A *Pie Graph* is a circle that is divided into sections or wedges according to the *proportion* of the frequency in the category to the sum of all frequencies.

A *Pareto Chart* displays *categorical* frequency data. The frequencies are represented by the height of vertical bars and are ordered in *descending* order.

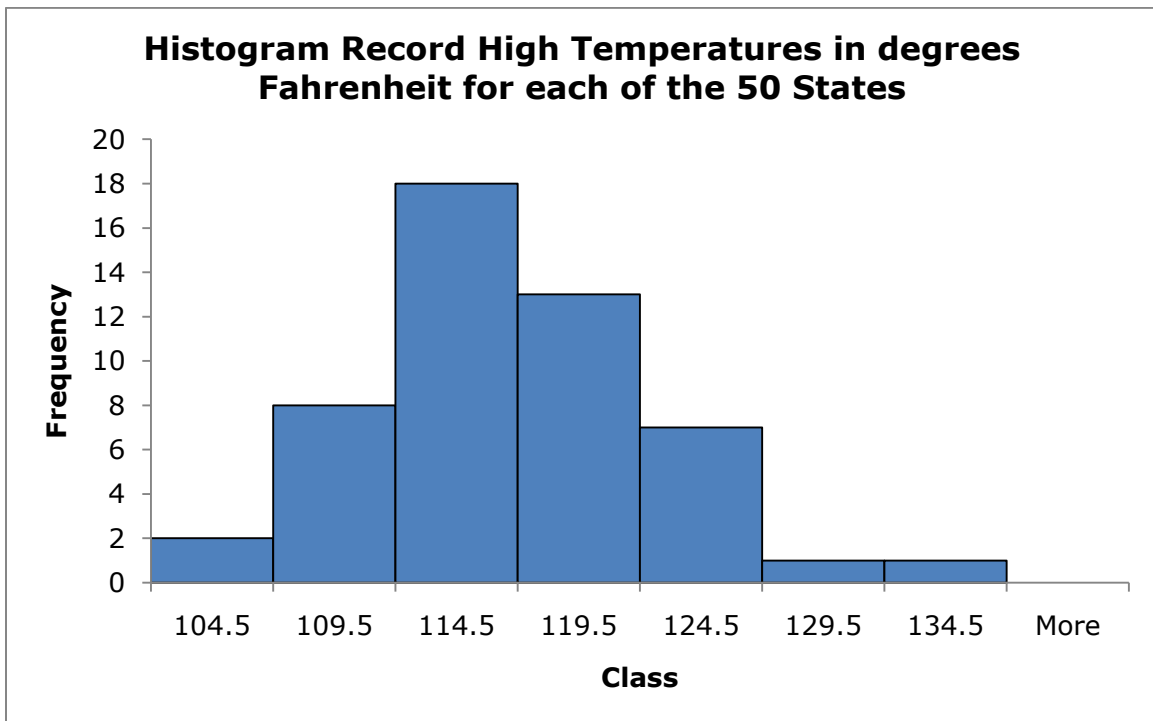
A *Time Series Graph* displays data that is referenced to *time*. The data values are represented by points. The points are then connected by a line.

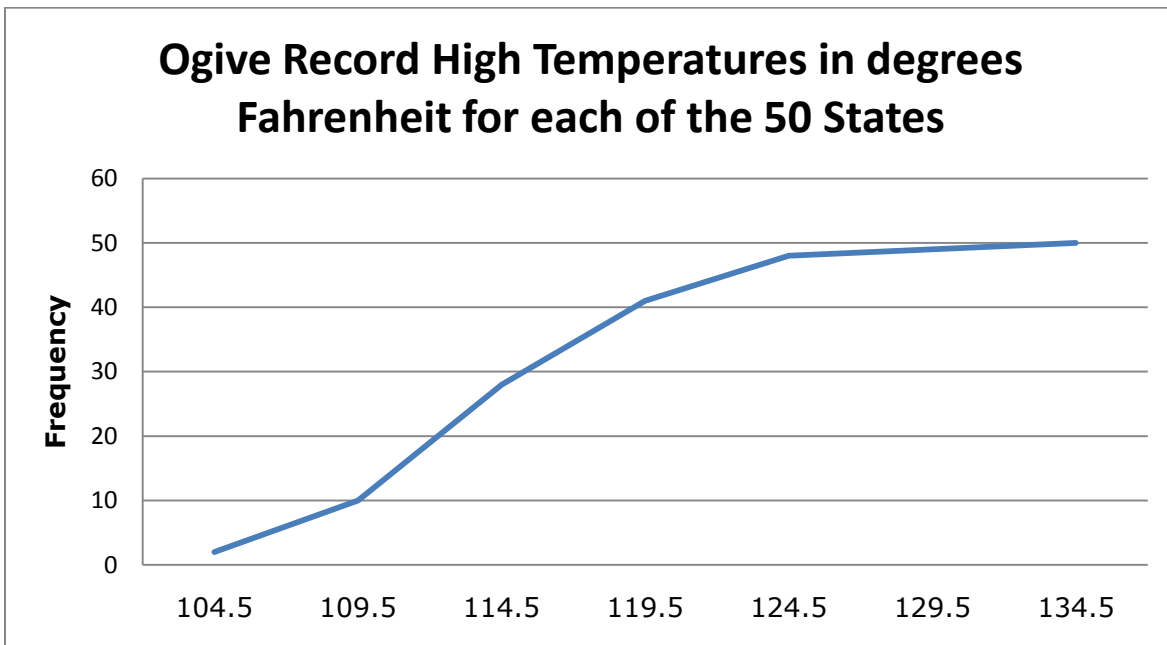
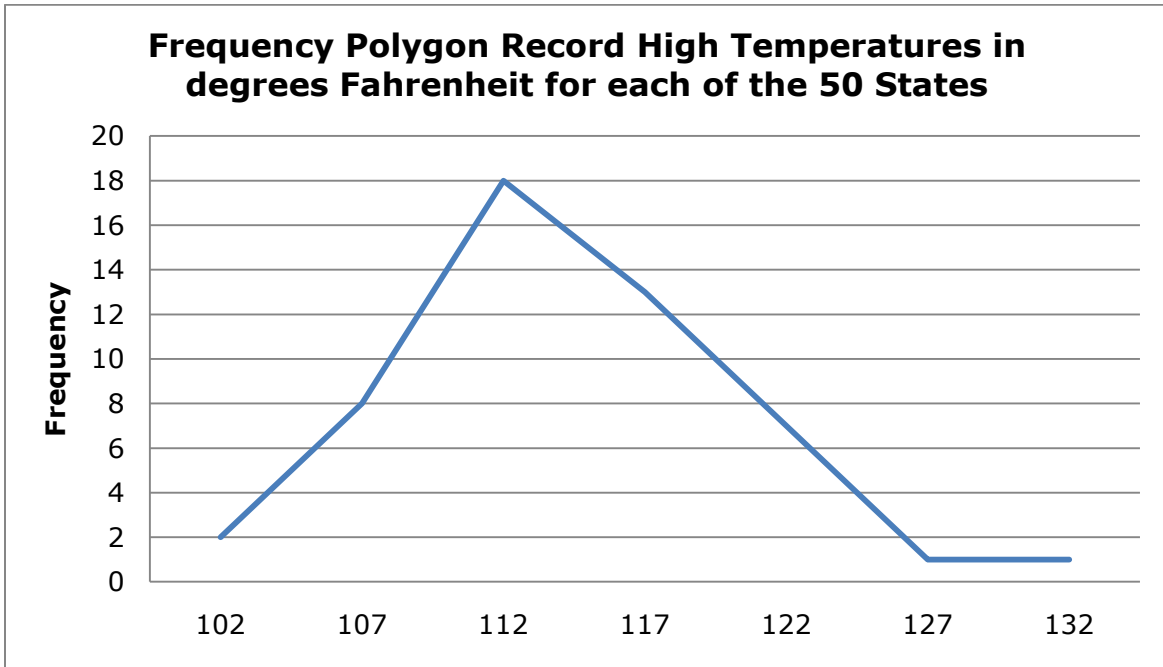
A *Stem and Leaf Plot* displays frequency data in classes while *retaining* the actual data value.

A *Scatter Plot* is a graph of ordered pairs of data values that is used to determine if a *relationship* exists between two variables.

These data represent the record high temperatures in degrees Fahrenheit for each of the 50 States. Construct a Histogram using seven classes, construct a Frequency Polygon, and construct an Ogive for the data.

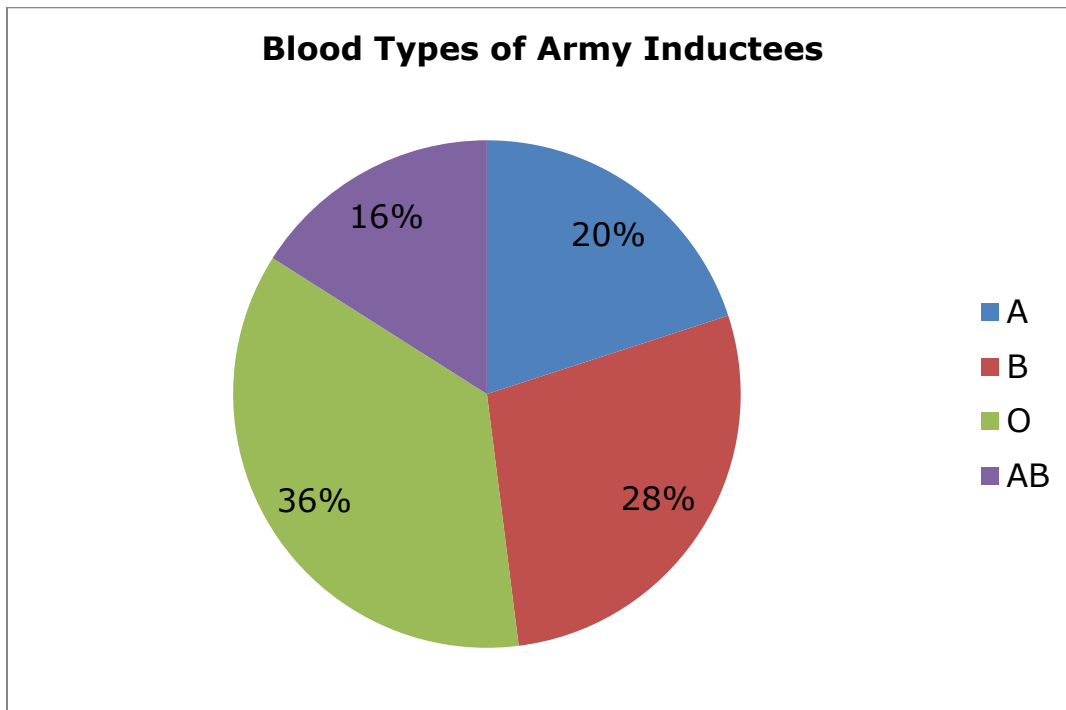
112	100	127	120	134	118	105	110	109	112
110	118	117	116	118	122	114	114	105	109
107	112	114	115	118	117	118	122	106	110
116	108	110	121	113	120	119	111	104	111
120	113	120	117	105	110	118	112	114	114





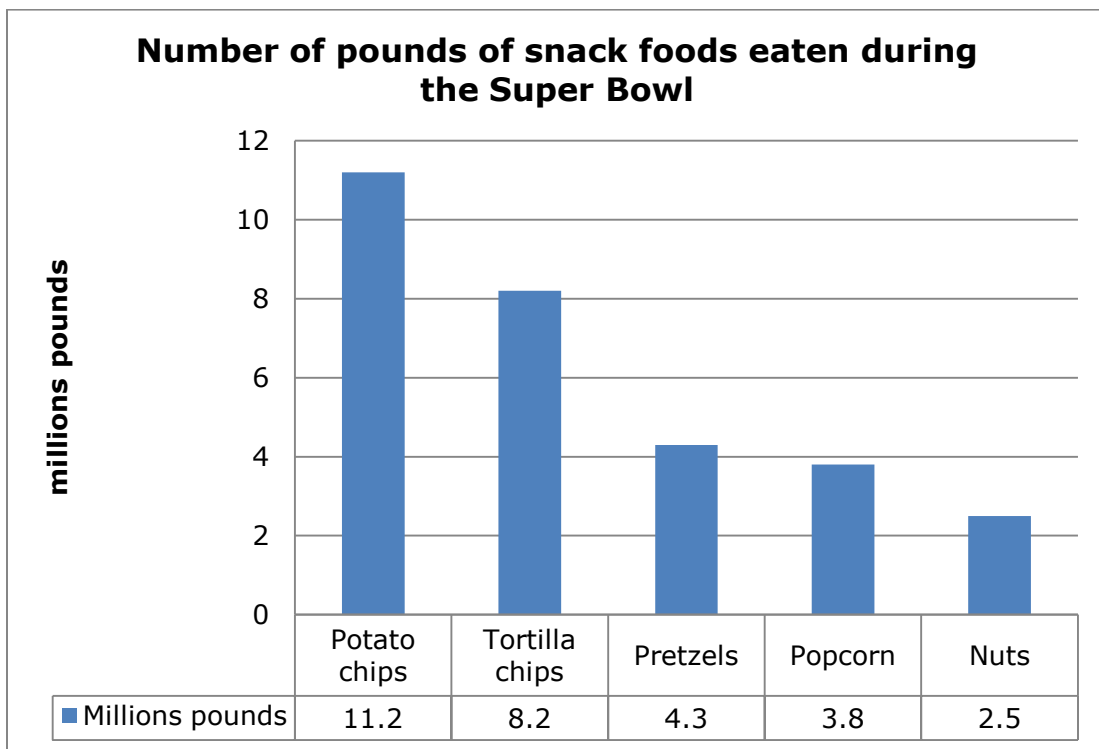
These data represent the blood types of army inductees. Construct a Pie Chart for the data.

Class	Frequency	Percent
A	5	$5/25 = 20$
B	7	$7/25 = 28$
O	9	$9/25 = 36$
AB	4	$4/25 = 16$
Total	25	100



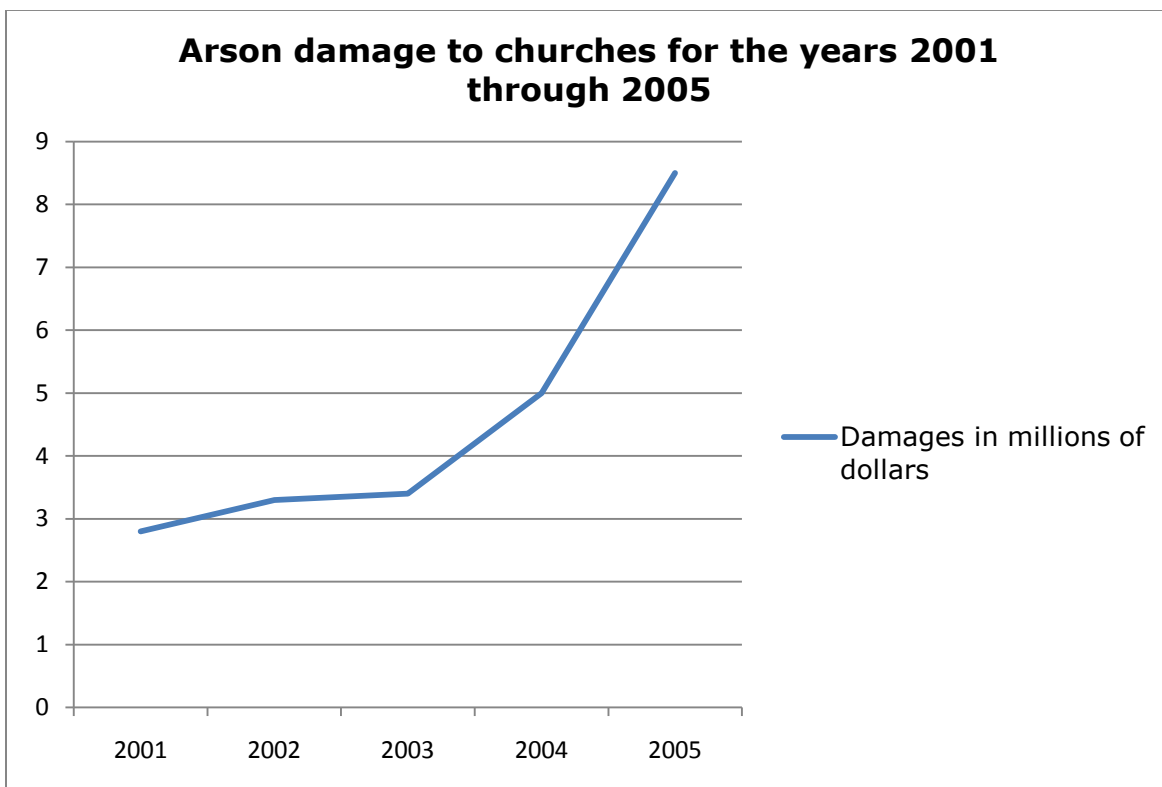
These data represent the number of pounds of each snack food eaten during the Super Bowl. Construct a Pareto Chart for the data.

Snack	Pounds (millions)
Potato chips	11.2
Tortilla chips	8.2
Pretzels	4.3
Popcorn	3.8
Nuts	2.5
Total	30.0



These data represent the arson damage to churches for the years 2001 through 2005. Construct a Time Series Graph for the data.

Year	Damage in millions of dollars
2001	2.8
2002	3.3
2003	3.4
2004	5.0
2005	8.5



These data represent the number of cardiograms performed each day in an outpatient testing center in a timeframe of 20 days. Construct a Stem and Leaf Plot for the data.

25	31	20	32	13
14	43	02	57	23
36	32	33	32	44
32	52	44	51	45

Stem Plot Cardiogram data

```
0 | 2
1 | 3 4
2 | 0 3 5
3 | 1 2 2 2 2 3 6
4 | 3 4 4 5
5 | 1 2 7
```

These data represent the number of wet bike accidents and the number of wet bike fatalities for a ten-year timeframe. Construct a Scatter Plot for the data.

Number of accidents (x)	Number of fatalities (y)
376	5
650	20
884	20
1162	28
1513	26
1650	34
2236	35
3002	56
4028	68
4010	55

