

How Can We Visualize Data?

Representing Data Using Graphs

J. Garvin



Slide 1/25

Bar Charts and Histograms

A *bar chart* displays the frequency of occurrences of categorical data.

Bar charts are typically used for *discrete* data.

A *histogram* is a special type of bar chart, with the area of each bar proportional to the frequency of a particular value.

Histograms are used for variables whose values can be arranged in some numerical order.

Continuous data lends itself easily to histograms.

J. Garvin — How Can We Visualize Data?
Slide 2/25

Bar Charts and Histograms

Example

A standard die is rolled 50 times, producing the following outcomes.

2	3	3	2	3	1	4	3	2	3
1	2	1	6	2	4	5	5	6	1
3	5	2	1	5	5	3	2	6	3
2	2	5	4	4	1	4	2	4	5
4	5	6	5	4	2	2	3	1	2

Use a *frequency table* to organize the data, then represent the data using a bar chart.

J. Garvin — How Can We Visualize Data?
Slide 3/25

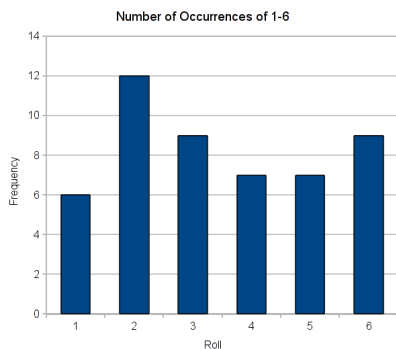
Bar Charts and Histograms

A frequency table records the number of occurrences of each outcome (or groups of outcomes).

Score	Frequency
1	6
2	12
3	9
4	7
5	7
6	9

J. Garvin — How Can We Visualize Data?
Slide 4/25

Bar Charts and Histograms

J. Garvin — How Can We Visualize Data?
Slide 5/25

Frequency Polygons

Another common way to represent data is by using a *frequency polygon*.

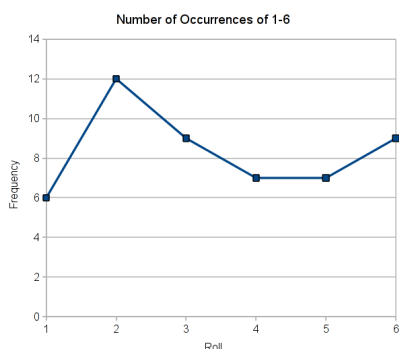
Frequency Polygon

Displays frequencies of outcomes using connected points, rather than bars.

Functionally, there is no difference between frequency polygons and histograms. It is typically a matter of taste.

J. Garvin — How Can We Visualize Data?
Slide 6/25

Frequency Polygons



J. Garvin — How Can We Visualize Data?
Slide 7/25

Cumulative-Frequency Graphs

Another way to represent data is to use a *cumulative-frequency polygon*.

Cumulative-Frequency Polygon

Displays a running total of frequencies, beginning from the minimum value. Also called an *ogive*.

Cumulative-frequency polygons are useful when one wants to see an overall increase in the data.

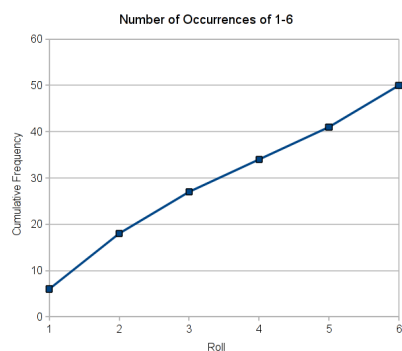
J. Garvin — How Can We Visualize Data?
Slide 8/25

Cumulative-Frequency Graphs

Score	Frequency	Cumulative Frequency
1	6	6
2	12	18
3	9	27
4	7	34
5	7	41
6	9	50

J. Garvin — How Can We Visualize Data?
Slide 9/25

Cumulative-Frequency Graphs



J. Garvin — How Can We Visualize Data?
Slide 10/25

Grouped Data

Sometimes it is easier, or desirable, to group data in intervals instead of using individual outcomes.

Large amounts of data, for example, may be better organized by analyzing the frequency of values within certain *ranges*.

Typically, 5-20 intervals are used, depending on the spread of the data and the desired ranges.

J. Garvin — How Can We Visualize Data?
Slide 11/25

Grouped Data

30 scores from a data management test are recorded below.

92	67	89	30	63	95	64	80	73	80
57	70	84	91	31	85	62	86	92	60
69	52	61	89	50	77	85	88	58	63

Use a frequency table to organize the data, then represent the data using both a histogram and a cumulative frequency polygon.

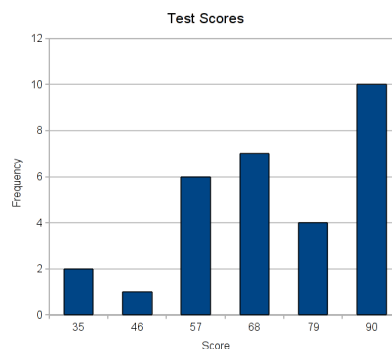
J. Garvin — How Can We Visualize Data?
Slide 12/25

Grouped Data

Score	Midpoint	Frequency	Cumulative
29.5-40.5	35	2	2
40.5-51.5	46	1	3
51.5-62.5	57	6	9
62.5-73.5	68	7	16
73.5-84.5	79	4	20
84.5-95.5	90	10	30

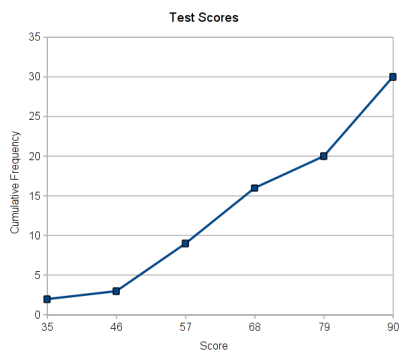
J. Garvin — How Can We Visualize Data?
Slide 13/25

Grouped Data



J. Garvin — How Can We Visualize Data?
Slide 14/25

Grouped Data



J. Garvin — How Can We Visualize Data?
Slide 15/25

Relative-Frequency Distributions

A relative-frequency distribution represents the frequency of a data group as a percentage of the entire data set.

Both histograms and frequency polygons can be used.

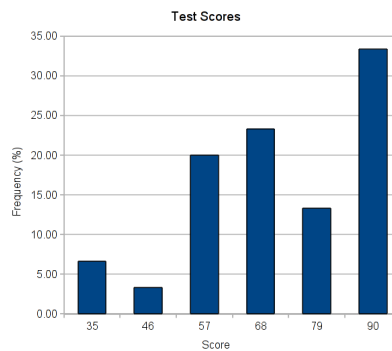
J. Garvin — How Can We Visualize Data?
Slide 16/25

Relative-Frequency Distributions

Score	Midpoint	Frequency	Rel. Freq.
29.5-40.5	35	2	6.67%
40.5-51.5	46	1	3.33%
51.5-62.5	57	6	20.00%
62.5-73.5	68	7	23.33%
73.5-84.5	79	4	13.33%
84.5-95.5	90	10	33.33%

J. Garvin — How Can We Visualize Data?
Slide 17/25

Relative-Frequency Distributions



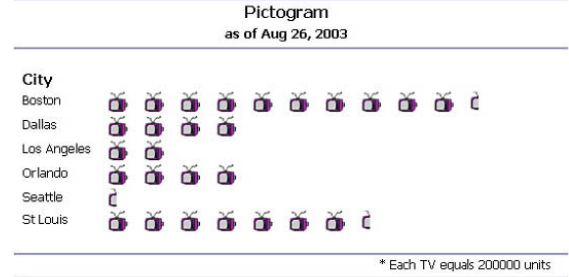
J. Garvin — How Can We Visualize Data?
Slide 18/25

Other Graphs and Charts

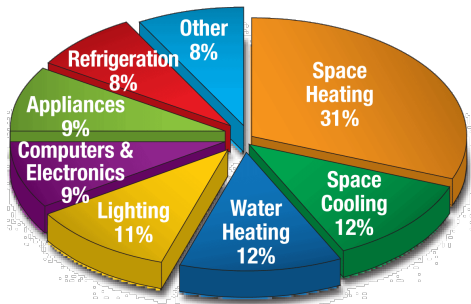
Pie charts represent data as a *percentage* of the data set. The percentages of all sectors must total 100%.

Pictograms represent data graphically, using pictures relevant to the category. They are not as widely used as other types of graphs.

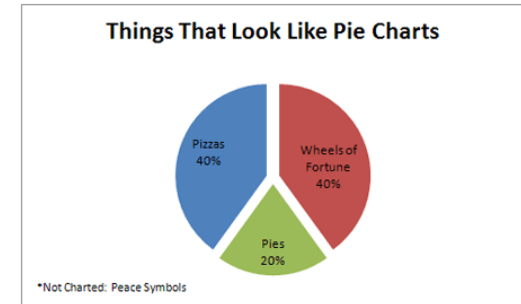
Other Graphs and Charts



Other Graphs and Charts

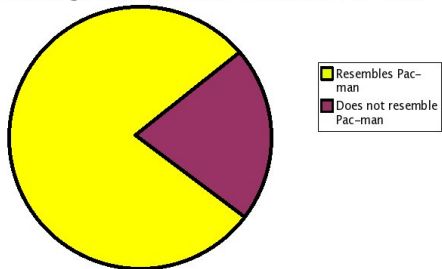


Other Graphs and Charts



Other Graphs and Charts

Percentage of Chart Which Resembles Pac-man



Other Graphs and Charts

