Graphing

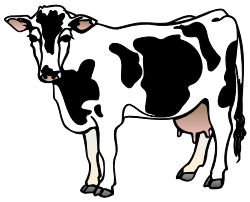
Types of graphs:

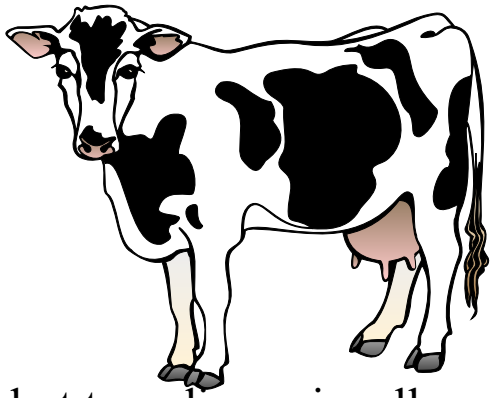
Pictogram: numeric data are represented by pictures, usually only nominal data are depicted in this way

Example: milk production increases by 200%

Before **After**

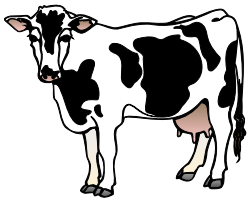
Biassed way:

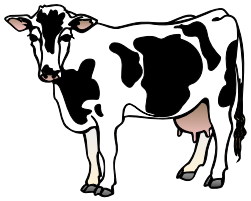


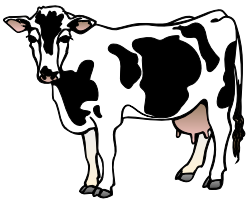


- height of cow is doubled but two-dimensionally cow is four times bigger, three-dimensionally it is eight times bigger

Unbiassed way:







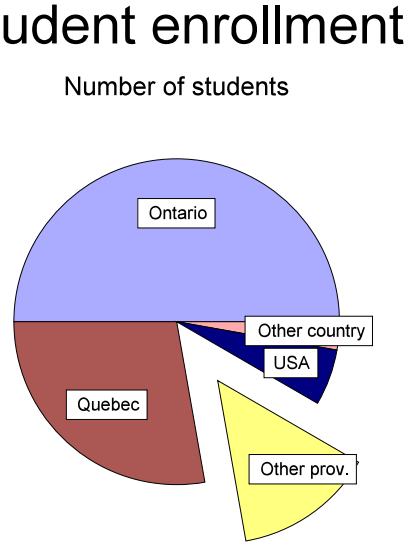
- increase is correctly depicted as two times greater

Graphing: Pie Chart

Pie chart: used with nominal or frequency data

Example: number of students by province and country

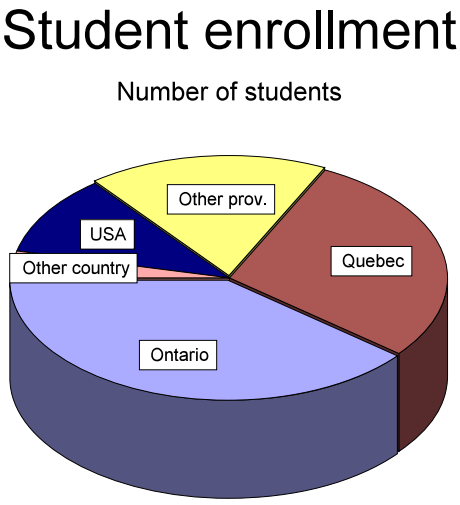
- a segment can be emphasized by separating it
- two-dimensional pie cannot create a biassed view

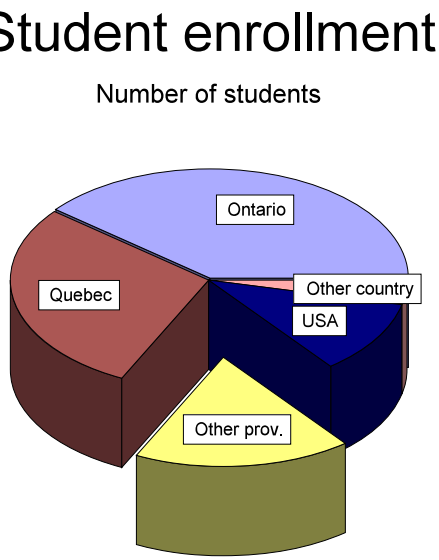


- three-dimensional pies can bias a slice depending on its position

- put a slice in the back to reduce its size

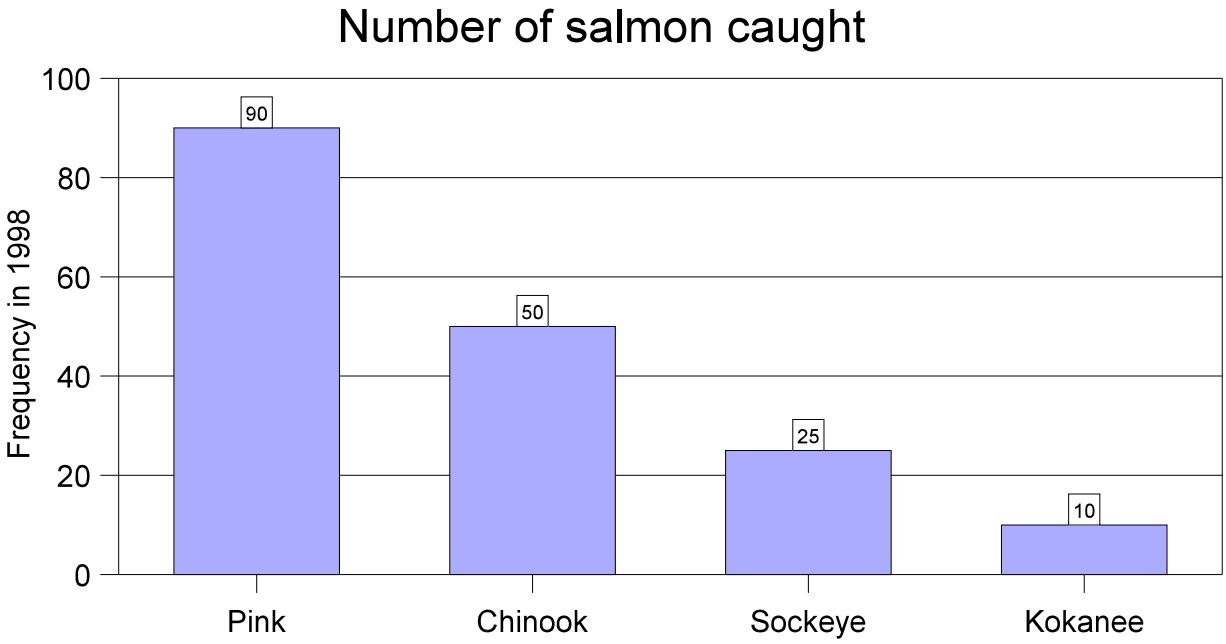
- a slice in front appears larger
- separating it creates emphasis





Graphing: Bar Chart and Histogram

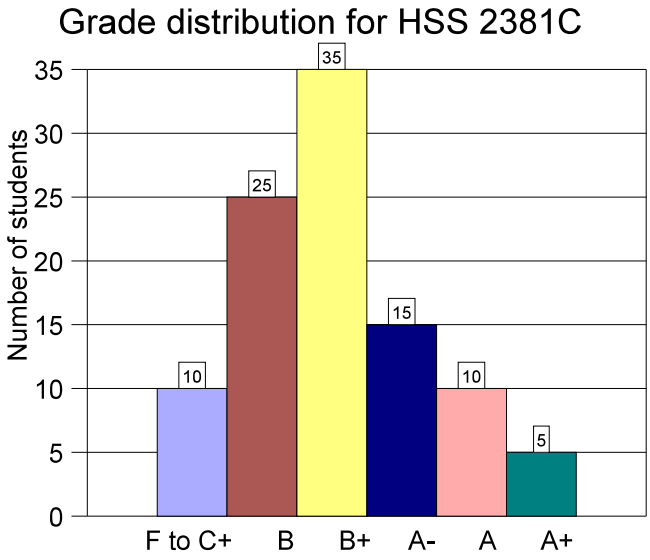
Bar graph: used for nominal data, usually frequency counts, are depicted by bars proportional to their magnitudes



- bars are separated
- extreme length bars can be split

Histogram: used for ordinal data

- bars are adjacent, no gaps
- one axis is ordered, first or last bars may include extremes

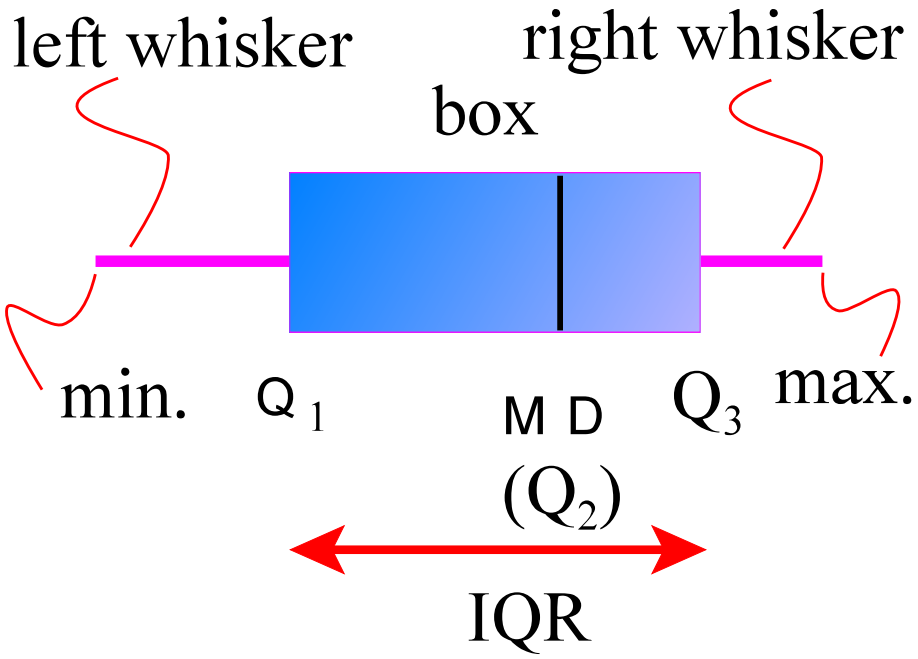


Graphing: Box-Plot

Box-and-whisker plot or Box-plot: used with interval/ratio data

- useful for investigating new types of data called Exploratory Data Analysis
- needs maximum, minimum and quartiles (Q_1, Q_2,Q_3)
- data needs to be sorted, which is difficult for large data sets

Example: body weights of a sample



Graphing: Stem-plot

Stem-and-leaf plot or Stem-plot: used with discrete interval/ratio data

- like a frequency graph but the actual numbers are preserved
- first make a list (stem) of all the first digits
- next for each stem digit, list last digit of all matching data (leaf)
- graph the results, usually horizontally
- can combine stem numbers e.g., 0-1, 2-3, etc.

Example: data of no. of cardiographs from an outpatient clinic:

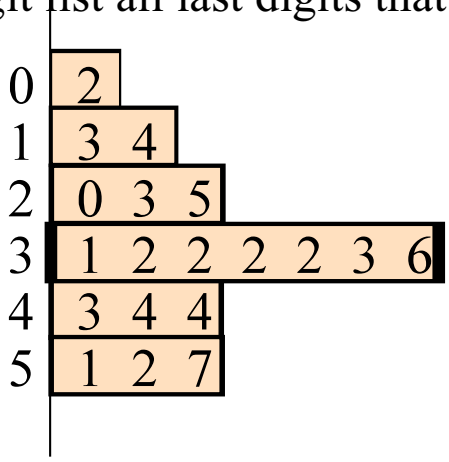
25	31	20	32	13	14	43	2
57	23	36	32	33	32	44	32
52	44	51	45				

First arrange in order:
2, 13, 14, 20, 23, 25, 31, 32, 32, 32, 32, 33, 36,
43, 44, 44, 45, 51, 52, 57

Next separate by first digits. Note, first digit for 2 is 0. Then beside each first digit list all last digits that start with the same first digit.

0 - 2
1 - 3, 4
2 - 0, 3, 5
3 - 1, 2, 2, 2, 2, 3, 6
4 - 3, 4, 4
5 - 1, 2, 7

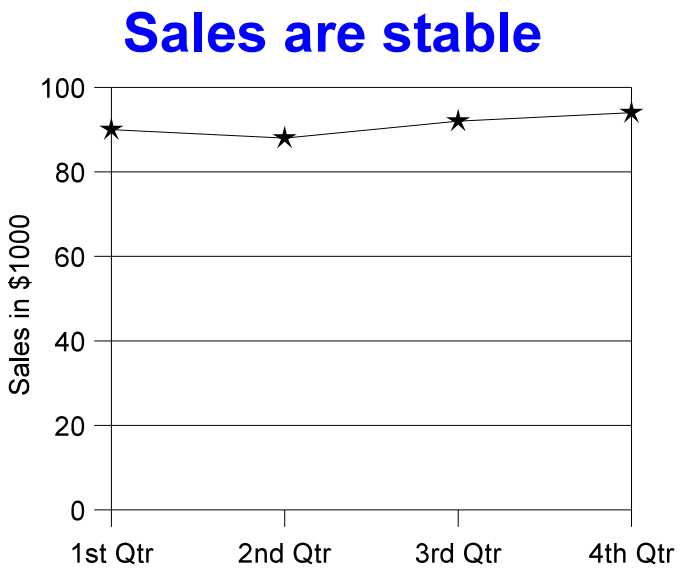
Then graph:



Graphing: Line Graph

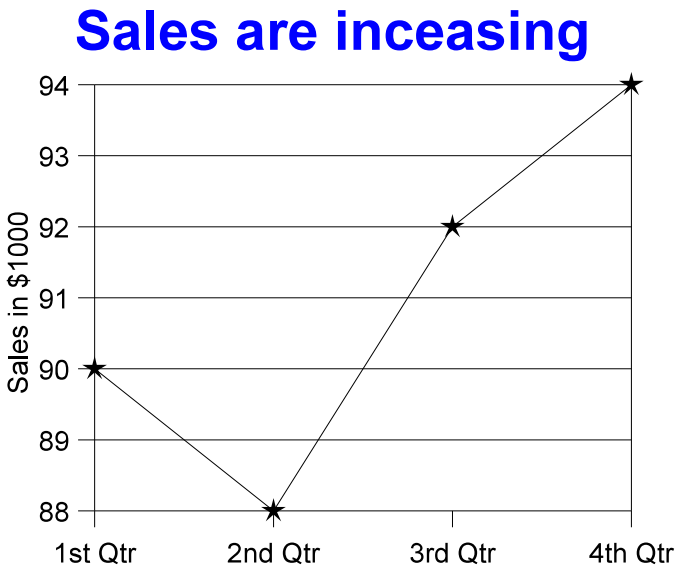
Line graph: used with interval and ratio data

- scaling can create a bias
- use large scales to hide changes



- truncated axis reduces "white space"

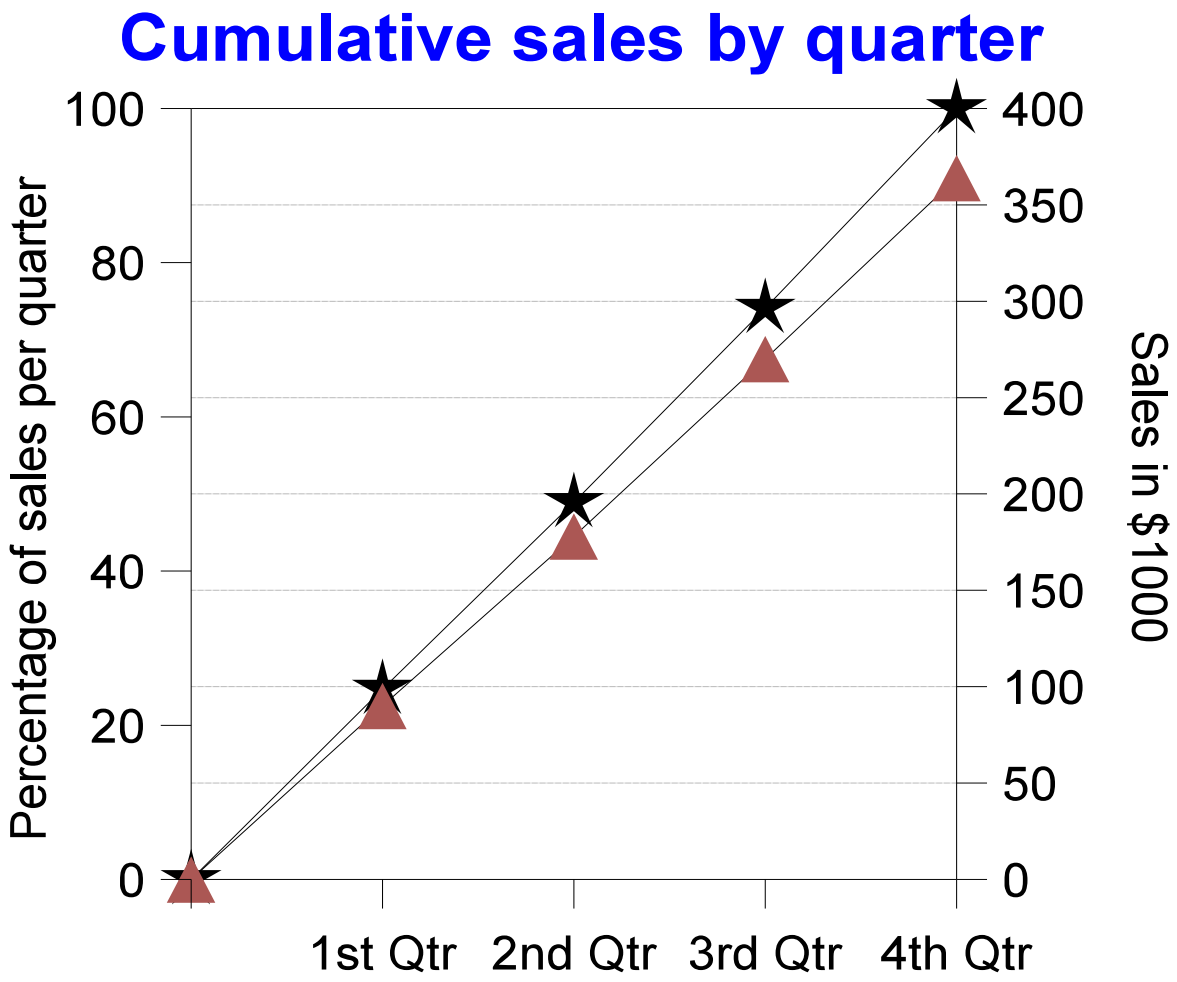
- scaling to minimum and maximum emphasizes changes



Graphing: Ogive

Ogive or cumulative frequency: (pronounced o-jive) line starts at zero and accumulates to 100%

- useful for determining percentages (by interpolation)



Rules for Constructing a Frequency Histogram

1. There should be between **5 and 20 classes**.
    - this is strictly for aesthetic purposes

2. The class width should be an **odd number**.
    - this ensures that the midpoint has the same number of decimal places as the original data

$$X_{midpoint} = \frac{lower\ limit + upper\ limit}{2}$$

3. The classes must be **mutually exclusive**.
    - each datum must fall into one class and one class only

4. The classes must be **continuous**.
    - there should be no "gaps" in the number line even if a class has no members

5. The classes must be **exhaustive**.
    - all possible data must fit into one of the classes

6. The classes must have **equal width**.
    - if not there will be a bias among the classes
    - you can have open-ended classes at the ends (i.e., for ages you may use 10 and under or 65 and over, etc.)

Types of Frequency Distributions

Categorical - for nominal types of data

Ungrouped - for numerical data with few scores

Grouped - for numerical data with many scores

Example: Distribution of the number of hours that boat batteries lasted.

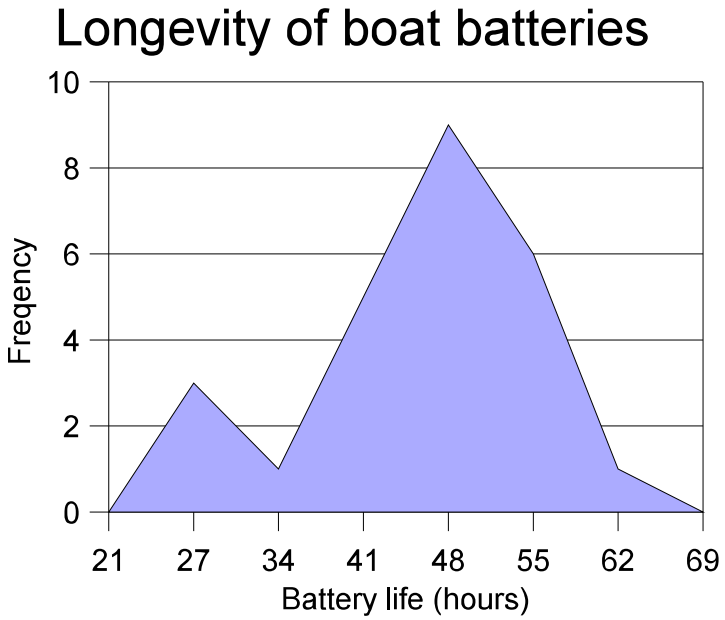
Class Limits	Class Bounds	Tally	Frequency	Cumulative frequency	Cumulative percentages
24-30	23.5-30.5	///	3	3	3/25*100= 12%
31-37	30.5-37.5	/	1	1+3= 4	4/25*100= 16%
38-44	37.5-44.5	~~////~~	5	5+4= 9	9/25*100= 36%
45-51	44.5-51.5	~~////~~ ////	9	9+9= 18	" 72%
52-57	51.5-57.5	~~////~~ /	6	6+18= 24	" 96%
58-64	57.5-64.5	/	1	1+24= 25	" 100%
Total			25	25	100%

Use these numbers for frequency polygon.

Use these numbers for constructing cumulative frequency polygon, also called an **ogive**.

Frequency Polygon and Ogive

Frequency polygon:



Cumulative frequency or ogive: