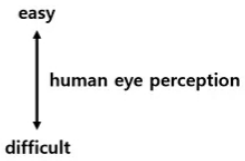
5. Optionally **add a text or an arrow** at interesting data points.
6. Use some **sizes and colors of the data** to make the plot more informative.

A visualization consisting of differently sized and colored **bubbles is more difficult for the human eye** to discern than a bar chart (position along a common scale). For effective visualization, reference to the following theory may be helpful.

Visual characteristics for human eye perception (by *Cleveland, William S., and Robert McGill. 1985)*

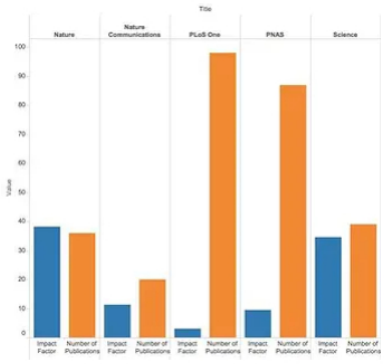
1. position along a common scale
2. position along a non-aligned scale
3. length
4. angle and slope
5. area
6. volume, density, and color saturation
7. color hue



"Graphical perception and graphical methods for analyzing scientific data."

Step 3 : Choose type of visualization

1. Bar chart



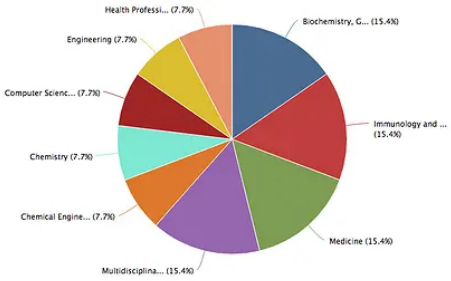
Detail

- **Intuitive** to read.
- The **simplest** type : one string and one numeric variable.
- **Easiest for human eye** perception : it uses alignment and length. (“Position along a common scale”)
- Good for **showing exact values**.

Things to consider

- Became **difficult** to read when **over-labeled** or **incorrectly** labeled.
- Horizontal or vertical bars.
- Pay attention to the numerical axis of the chart : **best to start at zero**.
- Order of bars : alphabetical, numerical, etc.

2. Pie chart



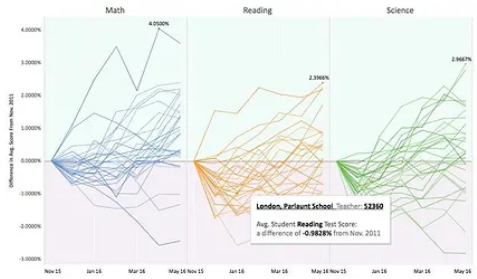
Detail

- When the **total** amount is one of your variables and you’d like to show the **subdivision** of variables.
- Best used with just **one string** and **one numeric variable like the pie chart above.**
- Show a **part-to-whole relationship**.

Things to consider

- The **more variables** you have, the more **difficult** the pie chart becomes to **read**.
- **Area is difficult for the eye to read**.
- If wedges are **similarly sized, try picking a different** visualization.
- **Avoid 3D** versions : notorious for causing distortion.
- **Use 2D** : easier to read while visually less stimulating.

3. Line chart



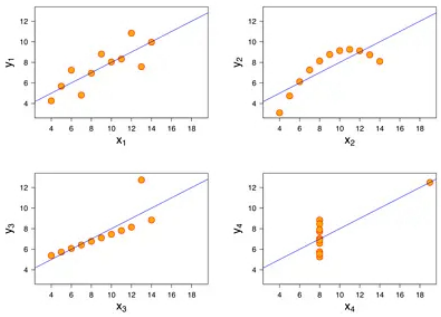
Detail

- An excellent way to show **change over time**.
- Use when there are **one date** variable and **one numeric** variable.
- Can show the **continuity** : better than bar charts.

Things to consider

- **Difficult** to read when there are **too many lines**
- **Avoid** giving **each line its own color** : the viewer has to scan back and forth from the key to the graph
- It might be **difficult** to see where the your **data points** are.
- It's best to **start with "zero" on the y-axis** to avoid distorting the data

4. Scatter chart



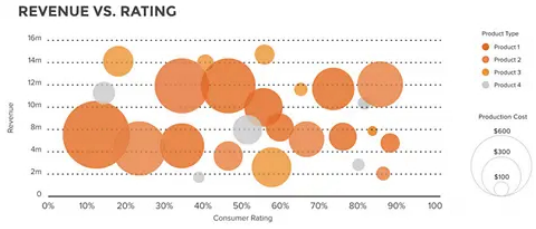
Detail

- Useful for showing **precise** and data **dense** visualizations , **correlations**, and **clusters** between **two numeric** variables.

Things to consider

- Not commonly used : more **difficult** for most people to read.
- **Large datasets** do not work well because **dots cover each other up**.

5. Bubble chart



Detail

- A variation to the scatter plot.
- Each dot is a different size, representing an additional variable.

Things to consider

Things to consider

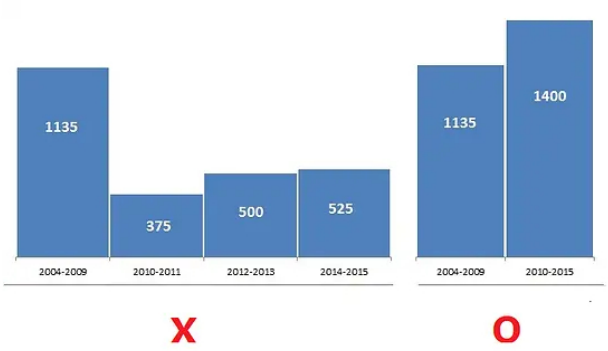
- Area of a circle is **difficult for the eye** to interpret.

Step 4 : Consider design

1. Color

- Use color meaningfully : **only use color when needed** to communicate something about the data.
- Choose the right color scheme for your data : categorical, diverging, sequential
- For categorical data, avoid using too many different colors : **no more than 6 colors is best; 12 colors max.**
- For sequential data, **don't use rainbows, use white to highly saturated.**
- Consider the format of your visualization : displayed on a projector, in print, copied in grey scale, etc.
- Be mindful of the potential color-deficiencies of your audience
- There are tools to help choose or test color schemes that are accessible for color deficient vision.
- You may also want to **consider the cultural connotations of particular colors**. (What Colors Mean in Other Cultures)

2. Scale



- Use **consistent scale divisions** when graphing data that involve continuous series.