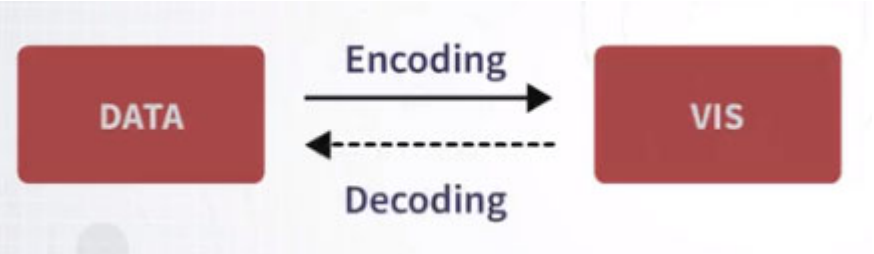
Graphical decoding

Thursday, November 1, 2018 12:41



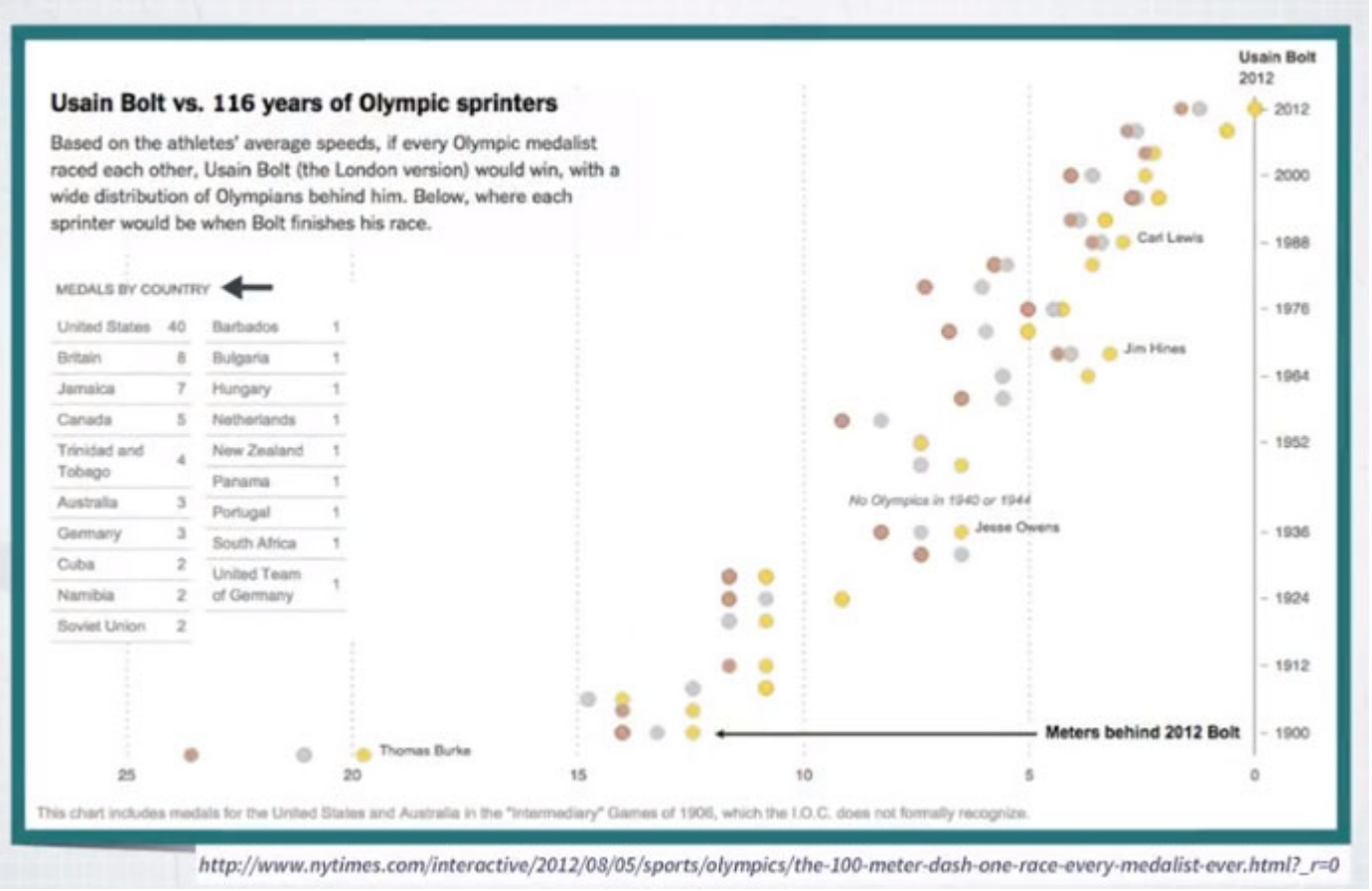
Decoding is king of a *reverse engineering* to figure out what the mapping rules and the graphical components are.

"*Algorithm*"
Identify explicitily:

1. Graphical components
    a. Marks -> what are the data items? (Nodes & links in network graphics)

2. Mapping rules
    a. Channels -> what attributes represent?

Examples:



1. Marks
    - Points

Representing medals and not necessarily runners, because there can be a multi medalist runner.

2. Channels
    - Color
      Medal type

    - Y-Position
      Time: Olympics year

    - X - Position
      Distance from Usain Bolt

I just wanted to point that in the video about "Graphical Decoding", the exercise examples are missing at the end after analyzing Bolt's visualization.

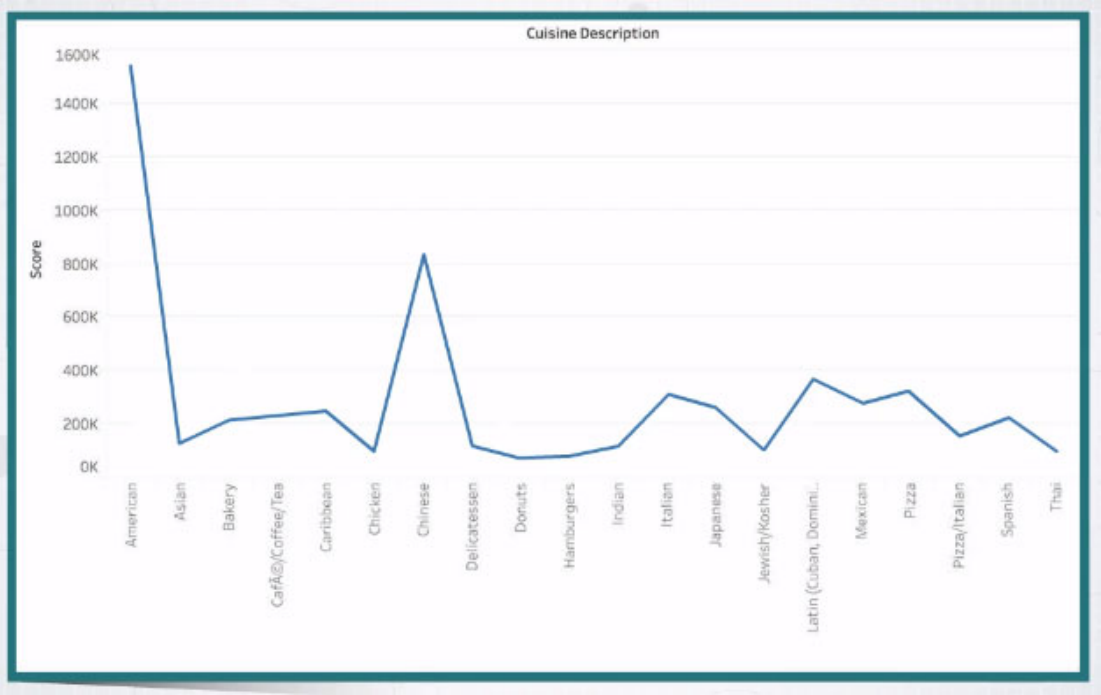
From <https://www.coursera.org/learn/information-visualization-fundamentals/discussions/weeks/4>

How do you know how good a visual encoding is?

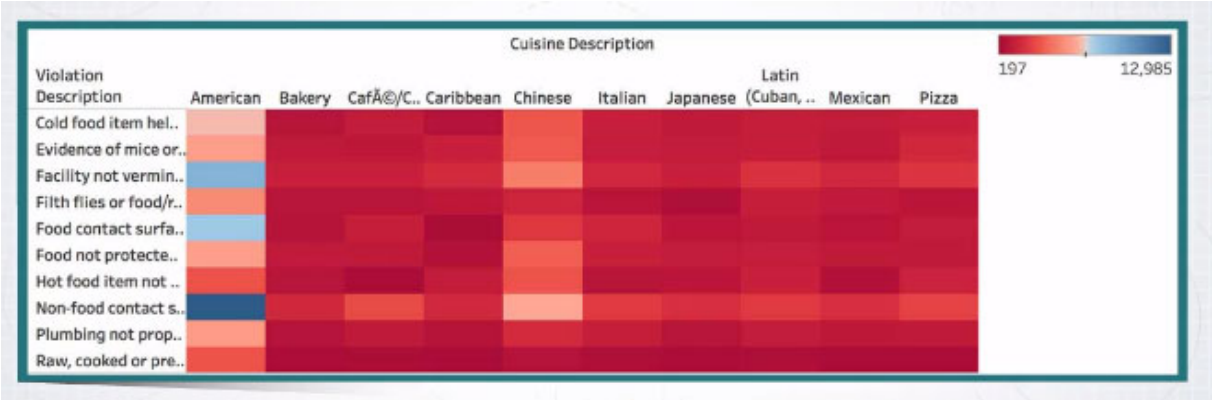
Guideline principles

1. **Expressiveness Principle**
   The visual representation should represent all and only the relationships that exist in the data. -> It shouldn't convey (transmitir/expresar/verbalizar) information that is actually not contained in the data.

   Examples:
- Ordered data should not appear unordered (and the other way round).



Line charts shouldn't be used in categorical data, because it will be perceived like something that is ordered, which is not the real case. Moreover, the patterns generated in the line chart are completely random and misleading.



Marks:

Areas: representing the violations of cuisines.

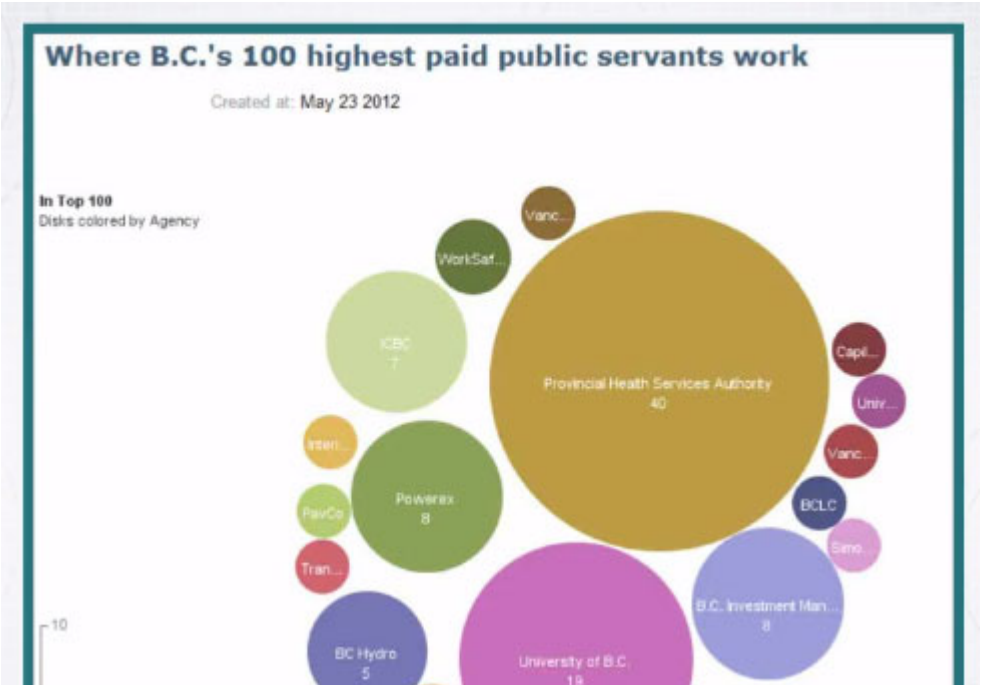
Channels:

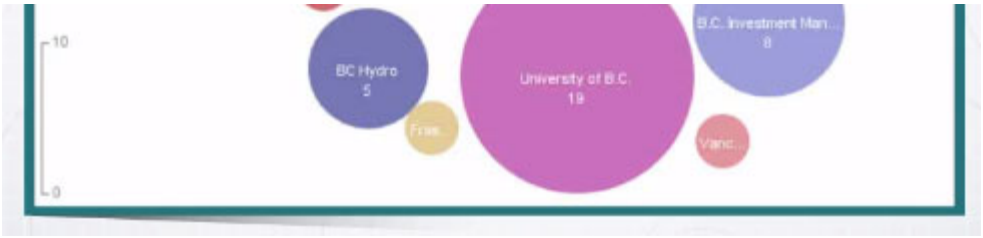
X-position: cusisine description / categorical data

Y-Position: violation / categorical data

Color hue + instensity: -> Matrix: Number of violations / statistics

A diverging color palette makes sense only when there is a zero value or a middle value. **The visual representation communicates something that IS NOT PRESENT on the data!**





Bubble chart

Mark:

Bubbles: represent some quantity

Channels:

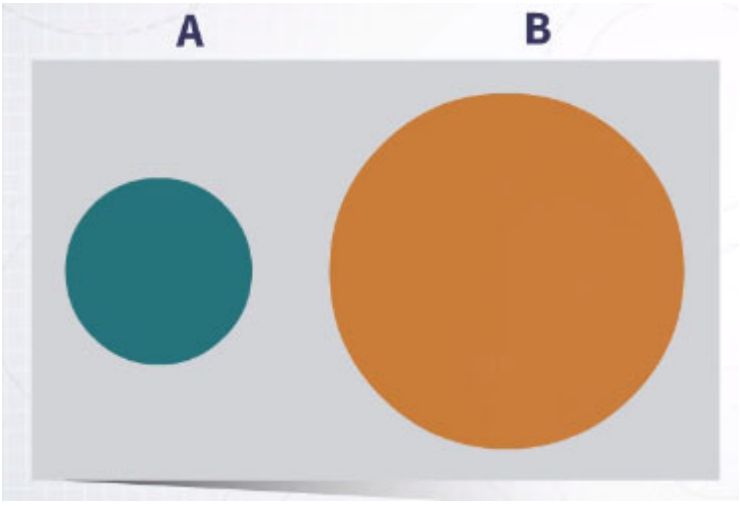
Color & position: don't encode any information -> **They don't encode any information therefore they're breaking the expressiveness principle.**

2. **Effectiveness Principle**

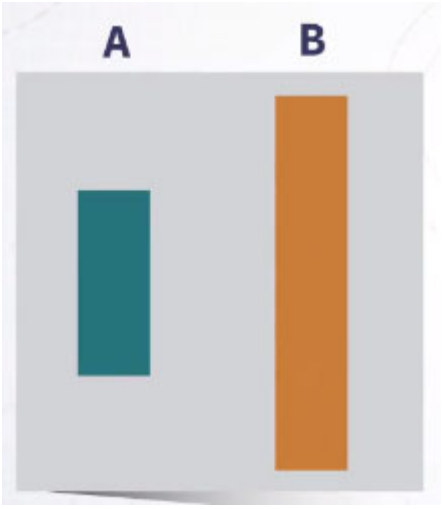
The relevance of information should match the effectiveness of the channel
-> *Represent important information with more effective channels!*

Example: try to estimate how much bigger is B compared to A

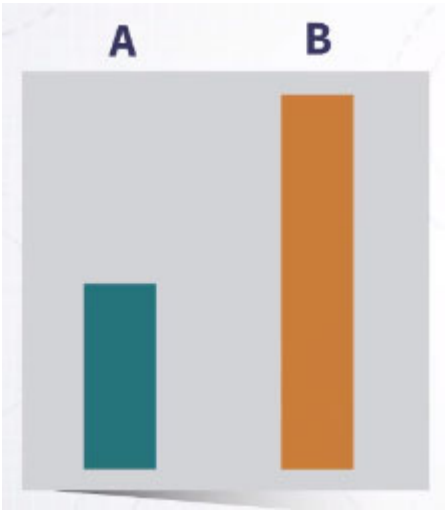
Doesn't make sense



It's hard to make a precise estimation



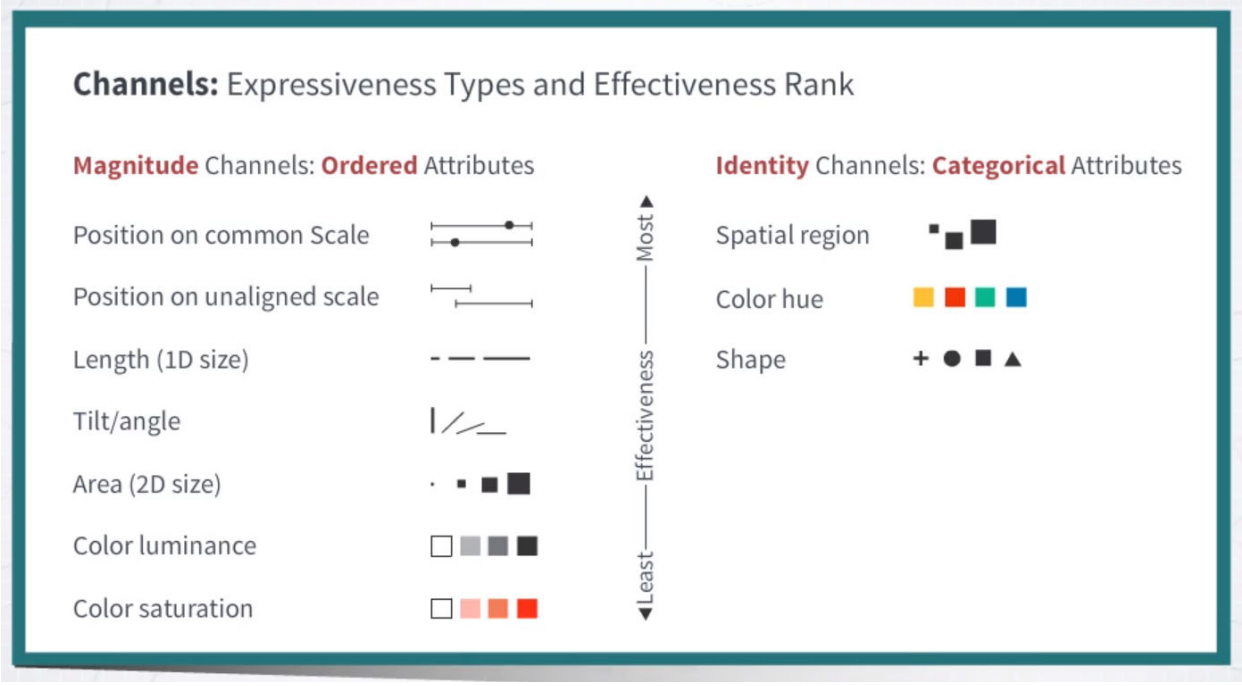
It's somewhat easier



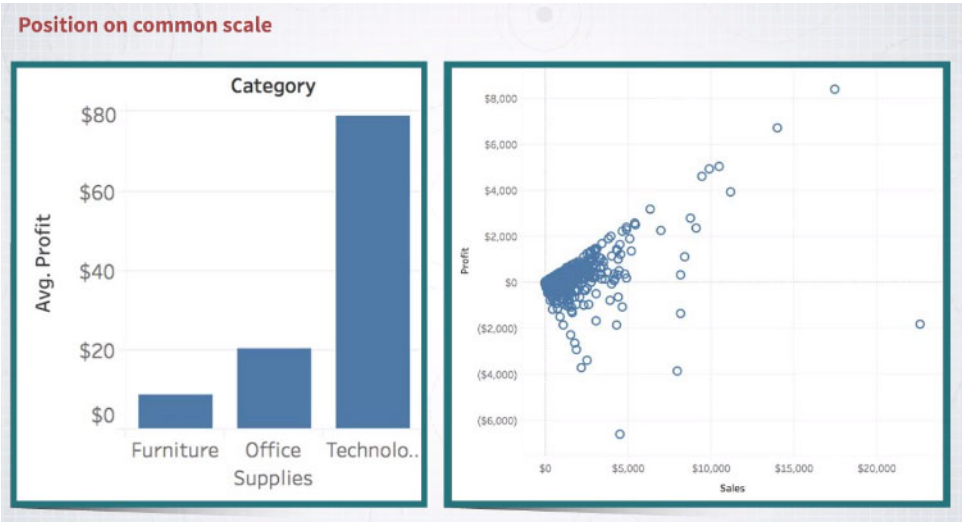
Easiest -> because your eyes are using also position

Different channels convey quantitative information more or less effectively!

Relevant information should be encoded and mapped with channels that are more effective and accurate at presenting information!

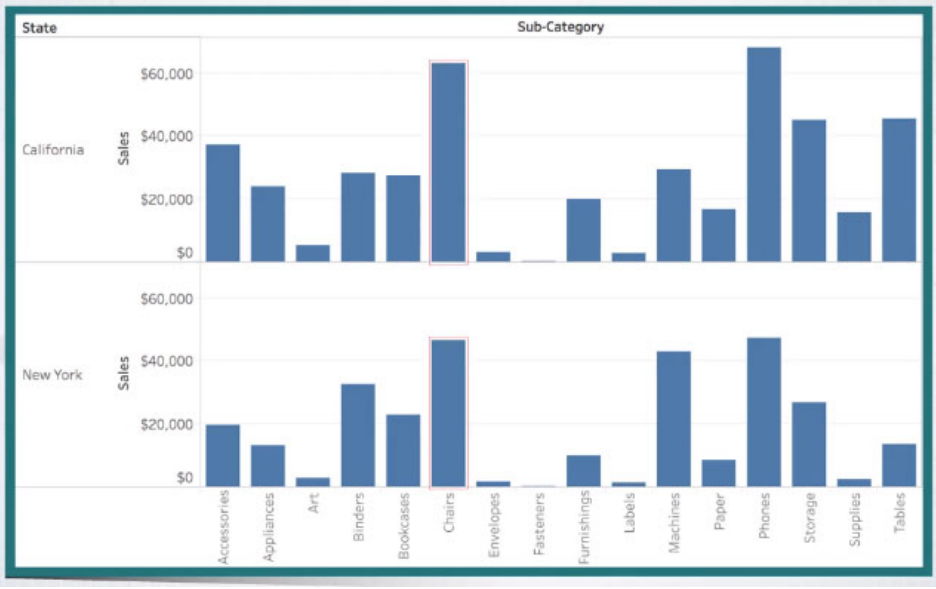


Color luminance: amount of brightness emitted by the object
Color saturation: vividness of the color (amount of black/white)

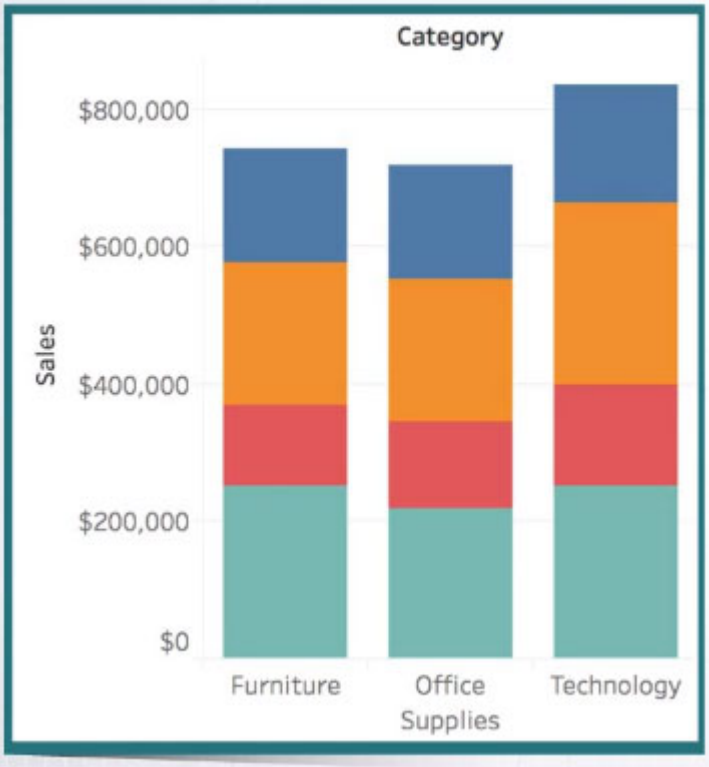


We only analyze the position on the end of the bar

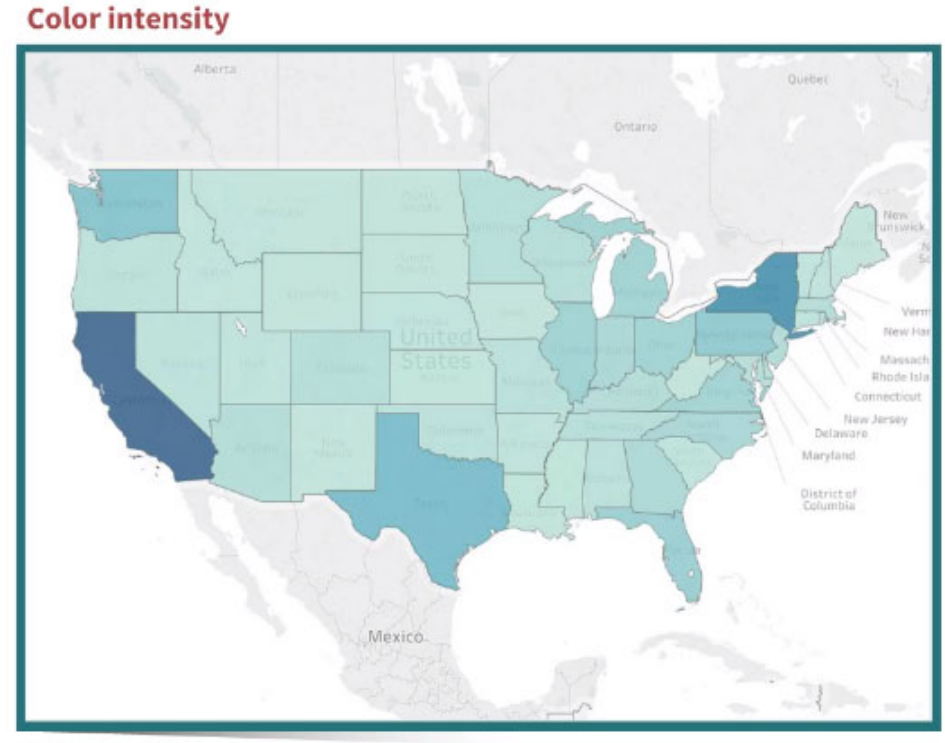
Position on unaligned scale



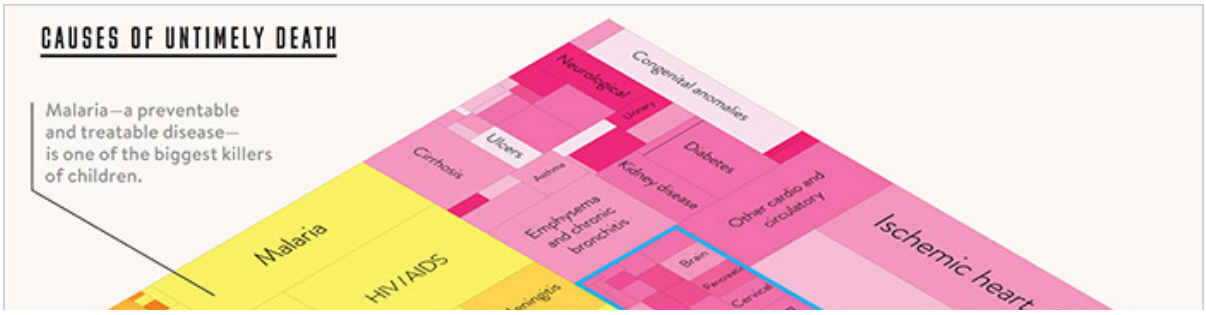
Length

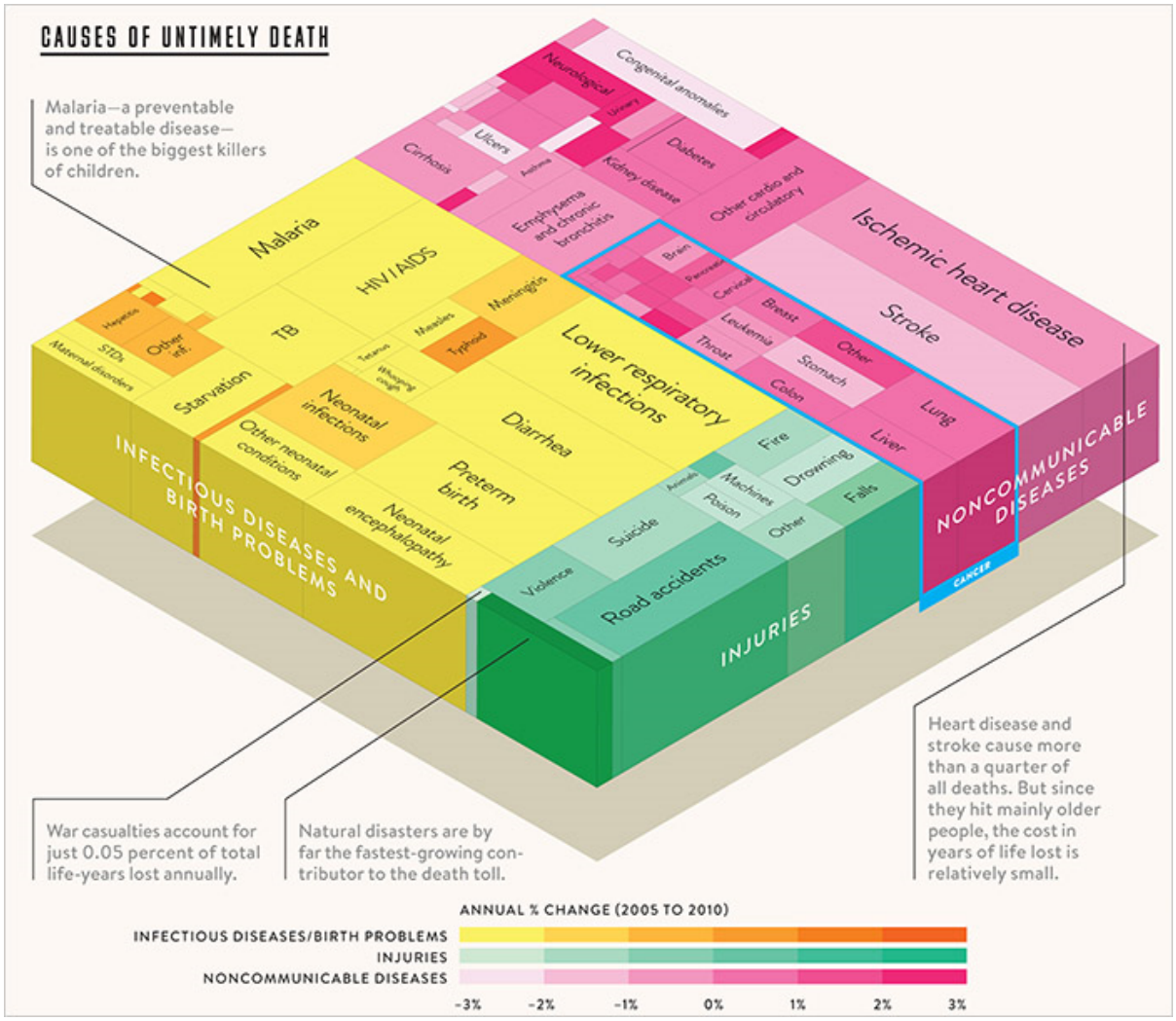


Tilt/Angle



Applying the principles for evaluating visualizations





Rectangle size: years of life loss for death specific cause
Color hue: group of causes of death
Color intensity: change from one year to another

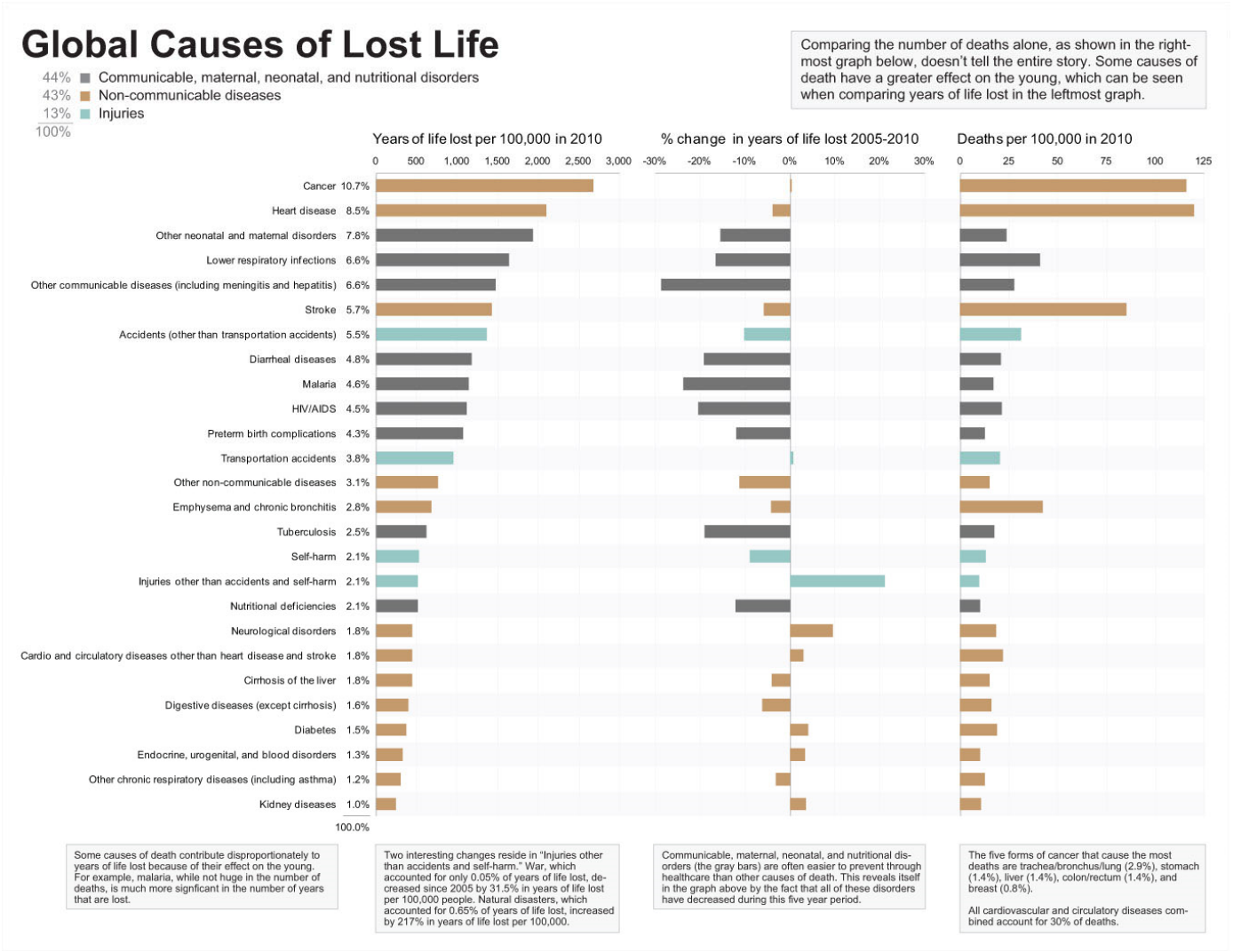
Improvements

- 3D projection skews the perspective and makes it difficult to read
- Area size is not a particular good channel to represent quantity
- Rectangles of different proportions between width and height make it difficult to read
- Color encoding used to represent yearly change is not particularly effective: is difficult to read which changes are positive and which are negative -> a divergent color scale will be better
- It is hard to label the small rectangles

A Solution

The information that appears in the treemap can be easily shown in two side-by-side bar graphs in a way that tells the story clearly and accurately and is just as visually engaging without resorting to gimmickry. In fact, by using a third variable to display information about the death rate for each cause, instead of solely showing the information in terms of years of life lost, the story can be enriched to give a clearer picture of the world. Here is our redesign:

From <http://www.perceptualedge.com/example20.php>

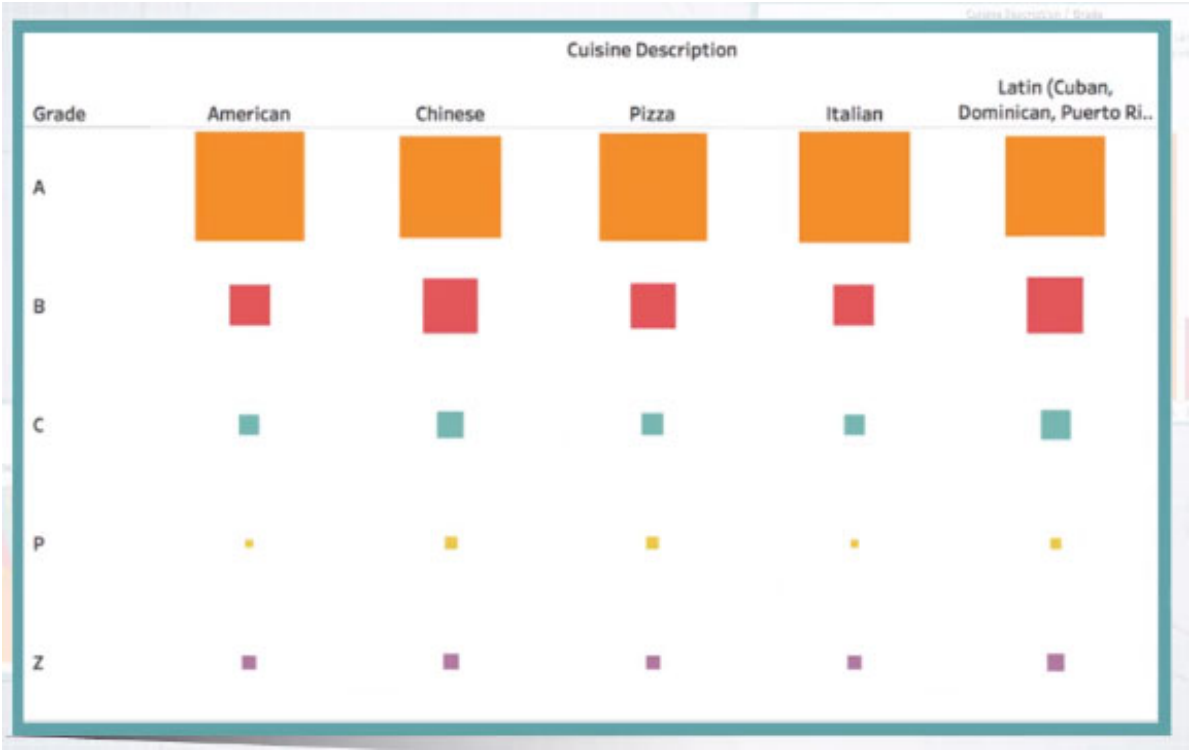


Using the Principles to Design Visualization

Visualizing restaurant inspection data…

Which cuisine type has the best distribution of grades?

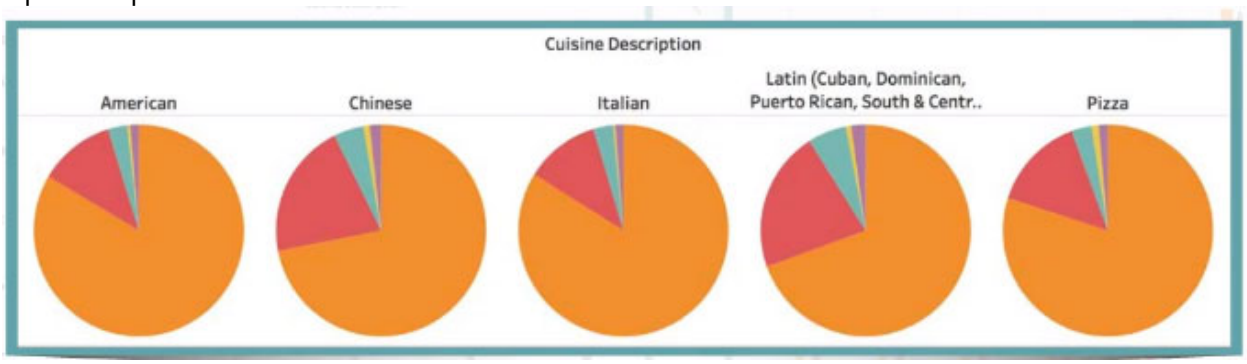
Option 1: matrix



Columns: cuisine types
Rows: grades

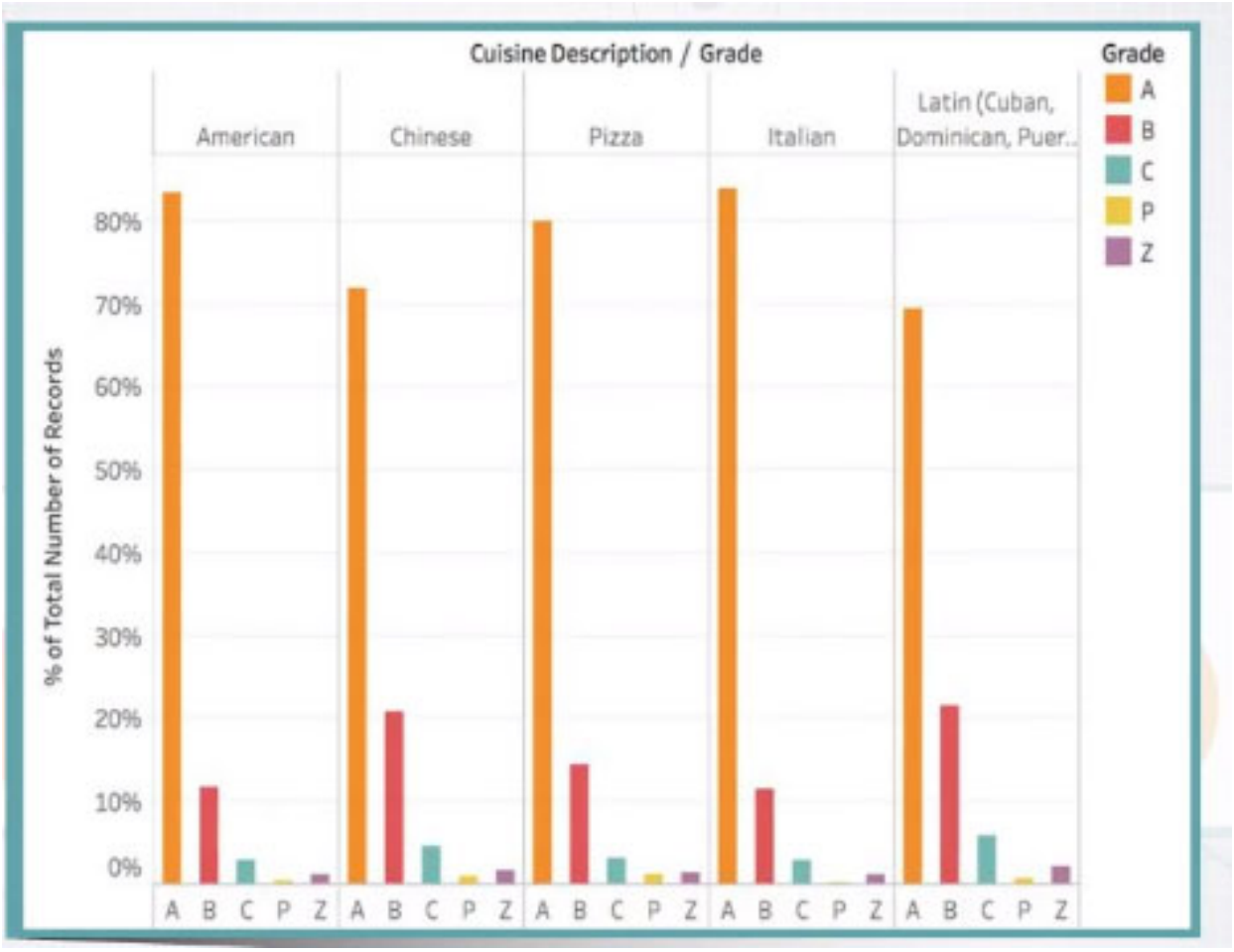
Channel:
Area size: % of grades

Option 2: pie charts



Segments: proportions of the grade

Option 3: grouped bar chart



Channels:

- Group: cuisine type
- Color: Grade
- Length: % of grade

Comparing bar charts is much easier.