Heuristic is a process or a rule that is meant to guide us in decision making, not optimal, perfect decisions but practical decisions.

**DATA INK RATIO**

One of the first heuristics of Graphics described by Edward Tufty is **Data Ink Ratio**.

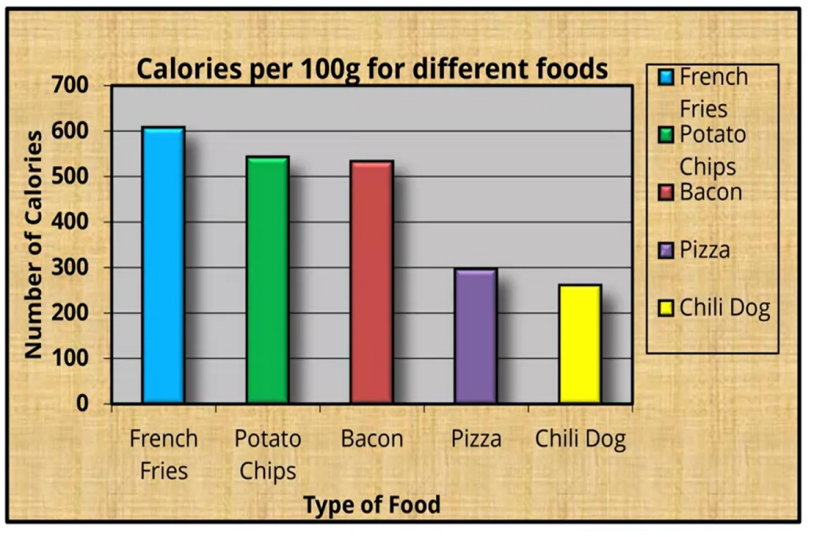
Mathematically DATA INK RATIO is given by the expression,

Amount of Data Ink

Total ink required to print the Graphic

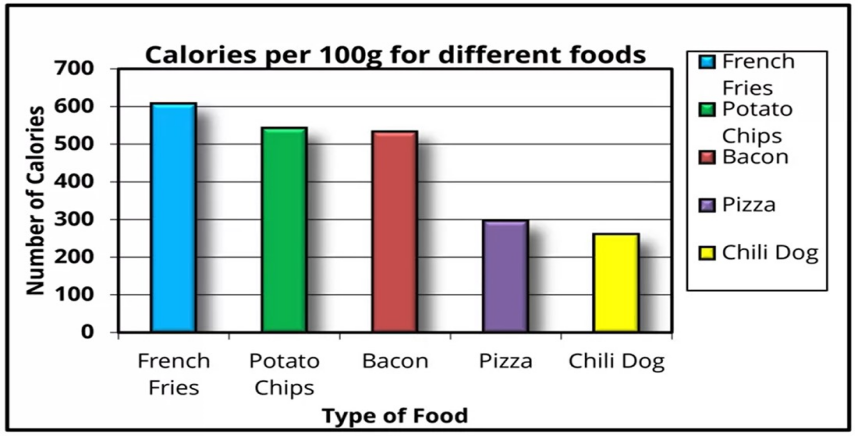
Mathematically its what it is, but in practicality, the general idea is to remove those elements from the data graphics that do not add value to the graphics. We will understand this with the help of an example.

Lets see this Data Graphics:



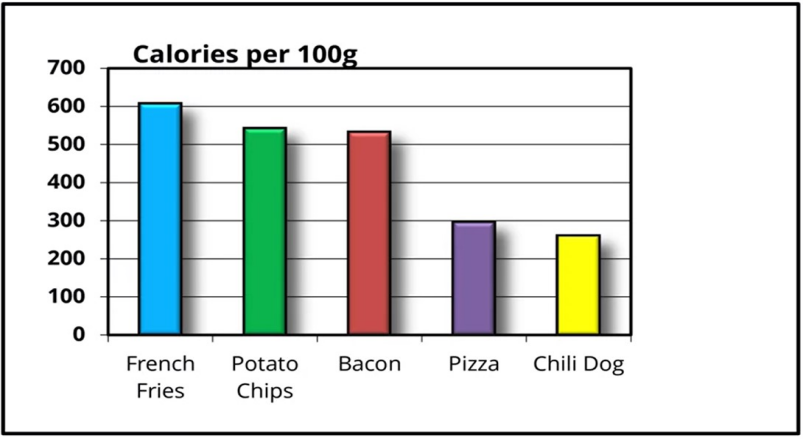
As its observable there are to many extra information available available to us, which when removed would not make any difference to the Data conveyed by the graph.

1. Lets start with removing the first unwanted thing. The background.

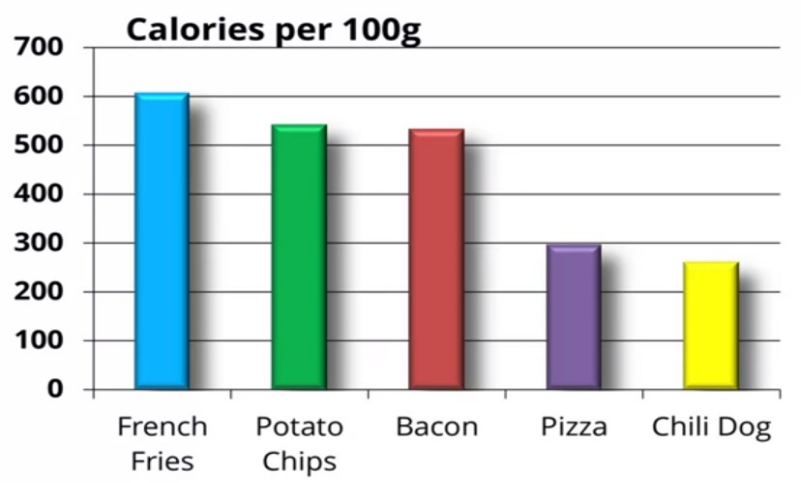


As observable, the grey and yellow background is removed.

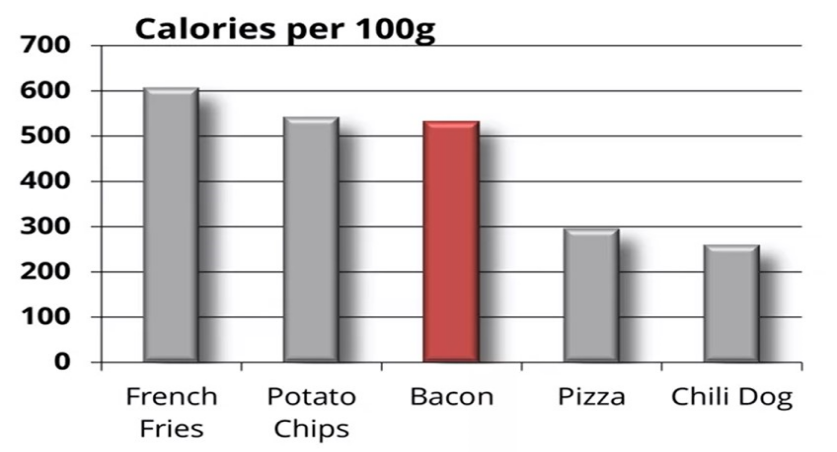
1. Let remove the next not so important thing. The key on the right is not required as the X-axis already gives a good idea of what represents what. Also we will remove the Y-axis label as the title gives enough idea about what the bar graphs actually represent.



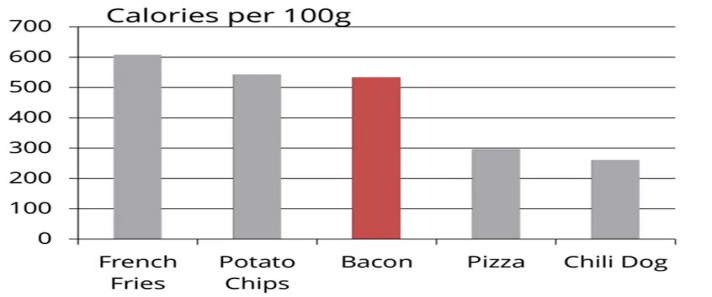
1. Lets look onto the inner and outer borders. They are no significance though are still there. Just remove them.



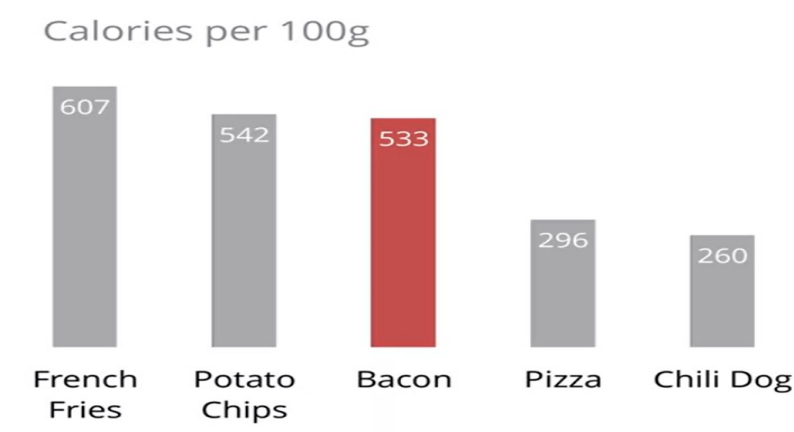
1. Next issue is, the colour. They do not relay any important information here and also, are a challenge when the data deals with colour blind people. Now there are two ways of dealing with this problem. One is replace colour with patterns, though this has issues of its own and we will talk about it later. Another way is making the complete graph colourless and only emphasizing on one of the bars and make it coloured.



1. Now the bold handwriting and the 3D bars don’t add much value to the graphic so they will be removed.



1. The grid format of data representation in the graphic is not much usable as it lacks accuracy. It can be clearly replaced by this.



See how change has come from where we started to where we are ending.

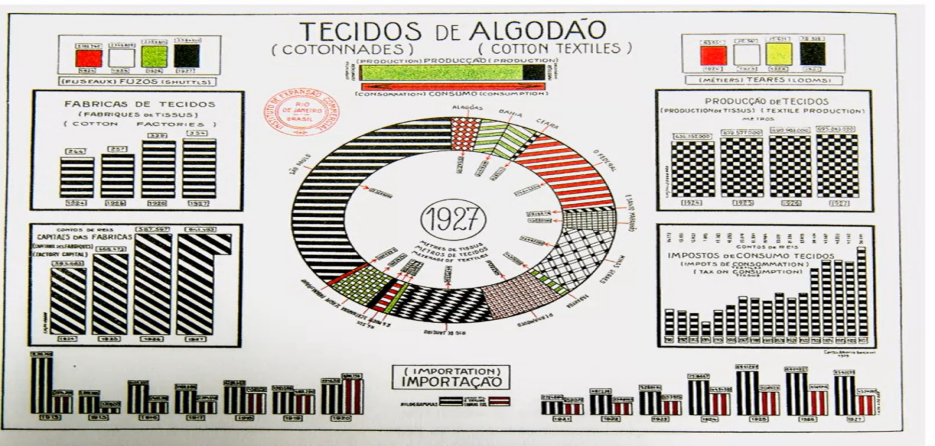
This graph has a higher data ink ratio than the one we started with.

**CHART JUNK**

Edward Tufte described artistic decorations on Data Graphics as weed. There are generally 3 kinds of data junk.

1. **Untended Optical Art**

Examples include shading, patterns and etc of chart features. An example can be observed below.



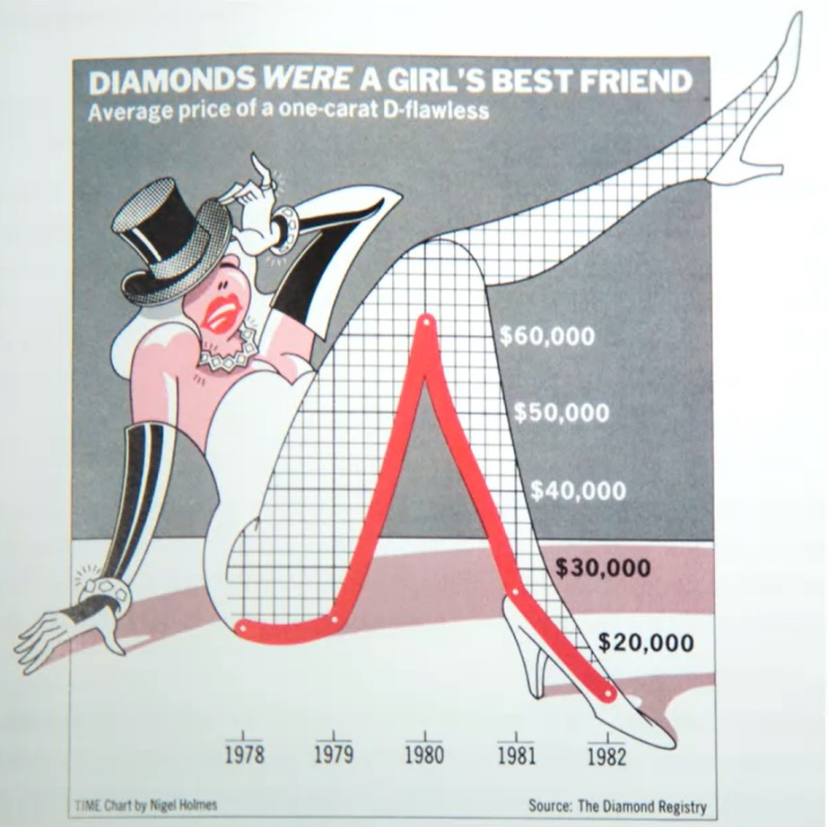
Patterns like these make people’s eye jump and create visual fatigue, and are referred as Moore’s Patterns. These can be perfected by labeling of data like we did in the previous case.

1. **The Grid**

Tufte describes grids to be unnecessary and says they actually cause problems in interpretation of data that is being actually shared. Hence thinning, removing or desaturation of grid makes the issue less pointing. Though, direct labeling as we say in the previous case is the best practice here to counter the problems being raised.

1. **The Duck**

Duck refers to non-data creative graphics. This form of graphic representation is commonly used in newspapers and magazines. Lets see an example of duck.

As observable we see an example of duck beside us. Well it is giving us the historic price trends of diamond, but rather than being you, a normal plot indicating the trends, it’s a graphic of the women’s leg that displays the trend.

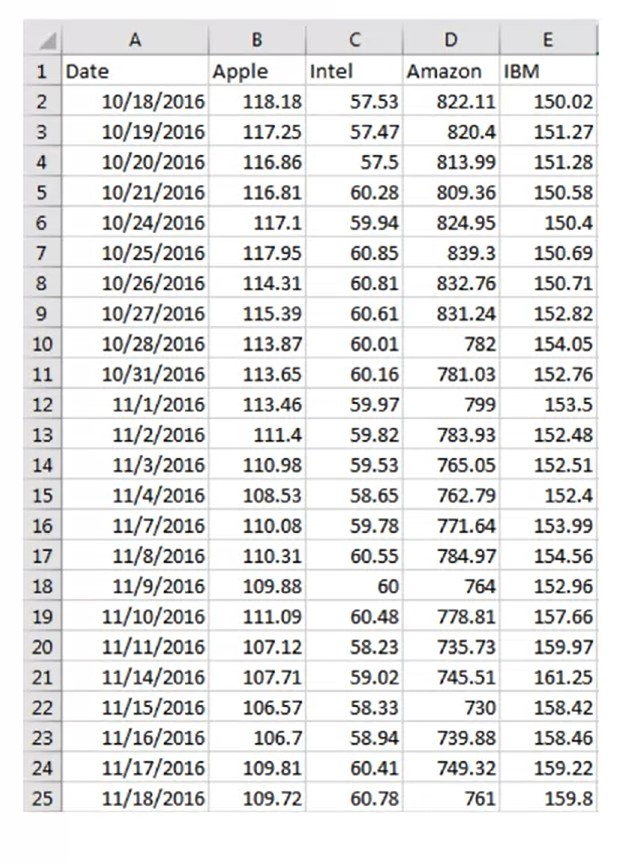
Though how advantageous or disadvantageous usage of ducks could be? Well it turns out, the data retention capacity of people Is higher when they intake data via ducks rather than you know, high ink ratio Data Graphics.

**SPARK LINES**

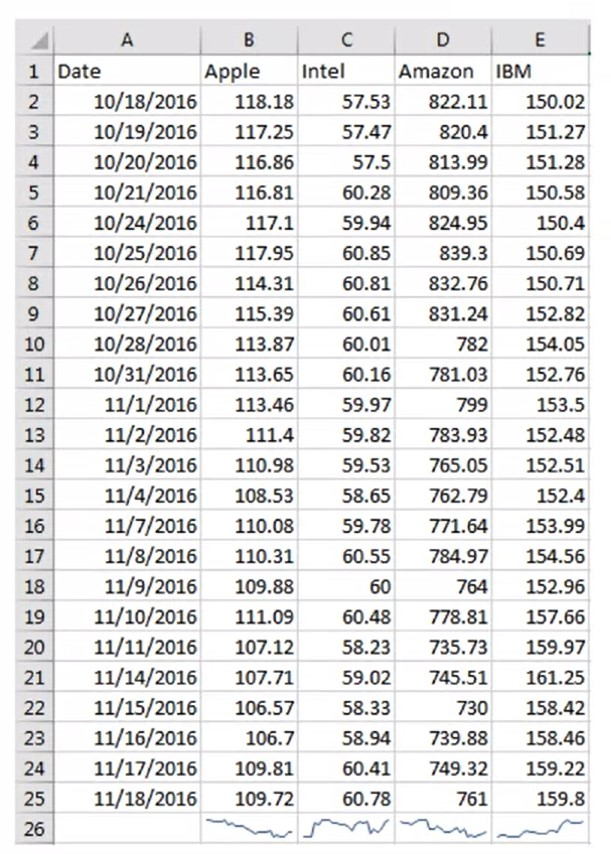
Instead of making Data Graphics such as charts a standalone entity to referred and studied on its own, Tufte suggested that they should be embedded within the context of the Data Being discussed. For example, a time series graph, can convey so much information to the users when studied under the context of the Data, rather then them being referred separately. These items were what he referred as Spark Line. He referred to them as Data Words to reduce gaps between texts and figures.

The spark lines can not only be embedded in texts, but can be direct used in tables along with the data they describe.

Lets see the difference that Spark line make.



The simple question that will be asked here is that given there are 4 major tech stocks, one need to find the price trends as soon as possible looking from the data besides. Yeah that’s hard.

Now here we have a Spark Line item at the bottom that help us predict the trends rather easily in comparison to the previous case.

Hence Spark Lines generally used in areas where trends and distributions are necessary to studied.

Spark lines are used in various spaces. From google finance to video games to spark tweets.

**LIE FACTOR**

Mathematically, Lie Factor is given by

Size of Effect On the Graphic

Size Of the Effect On the Actual Data

Generally it’s a method to find how the graphic shown to the audience is misleading in respect to its data set.

The closer it has its value to 1, the more accurate is the Data Visual.