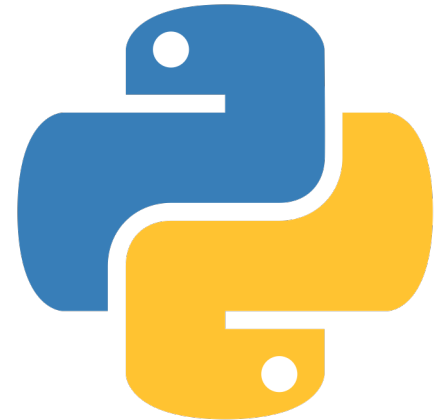


# Python programming and data visualization for beginners

Dr Joel Martin



# Week 4

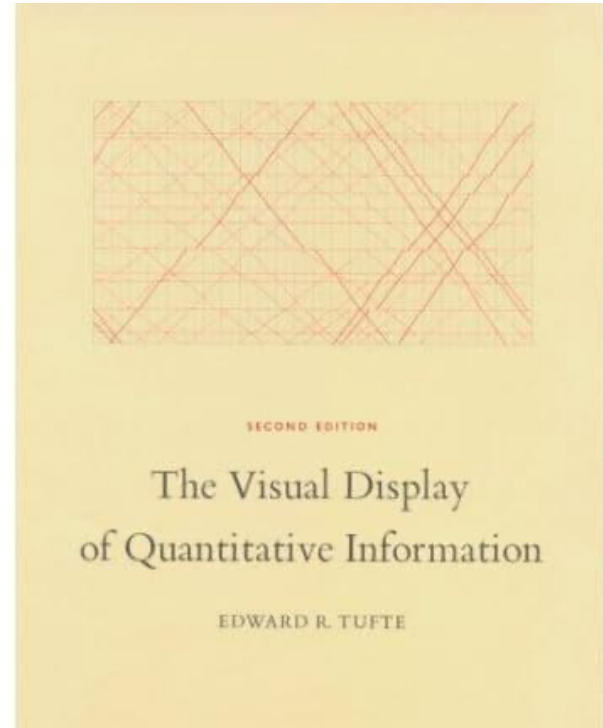
## Data visualization with NumPy and matplotlib

- Data visualisation
  - Edward Tufte
  - Misleading charts
- NumPy
- Matplotlib

# Data visualisation

## Guiding principles

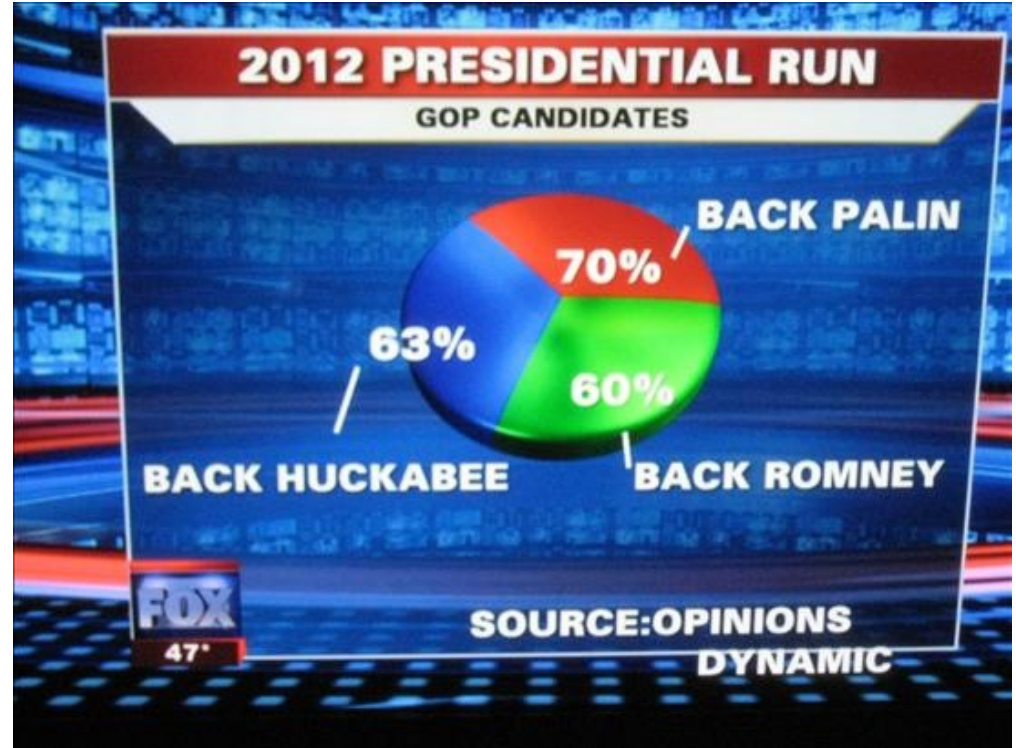
- Edward Tufte
  - Leading expert on data visualisation
  - Book: The Visual Display of Quantitative Information
- Key concepts:
  - *Chartjunk*: Useless, non-informative, or information-obscuring elements of quantitative information displays
  - *Lie factor*: The relation between the size of effect shown in a graphic and the size of effect shown in the data
  - *Data-ink ratio*: Use 'ink' for data, avoid unnecessary decorations
  - *Data density*: Number of data points divided by size of graphic (maximise, within reason)



# Data visualisation

## Misleading charts

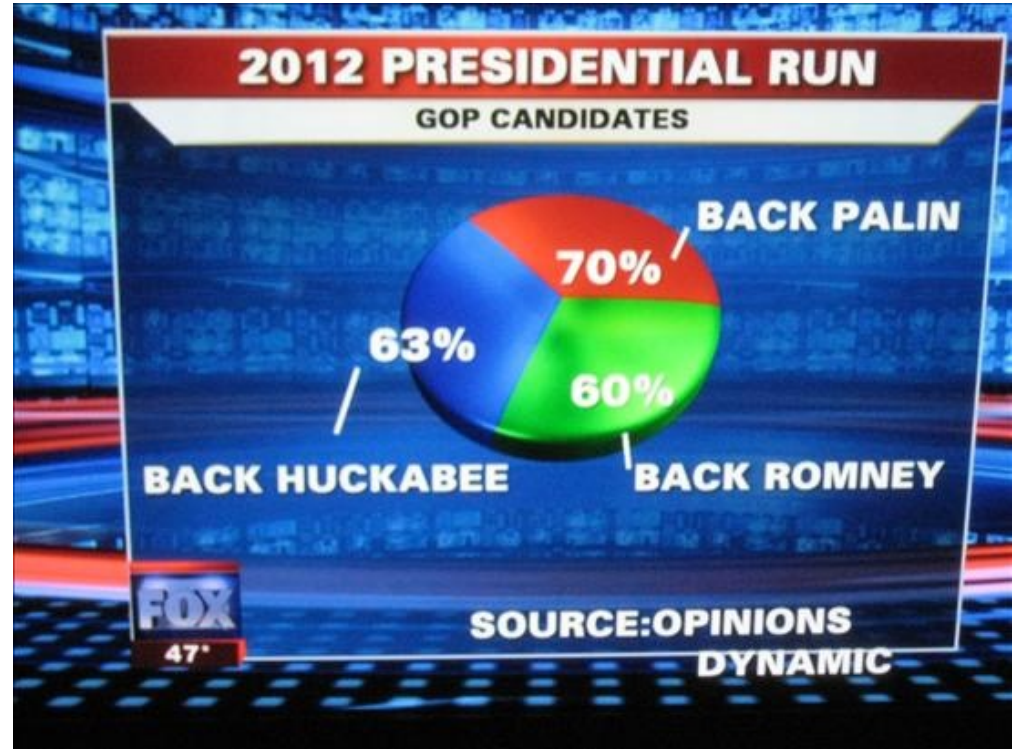
- Fox News
- What is wrong with this chart?



# Data visualisation

## Misleading charts

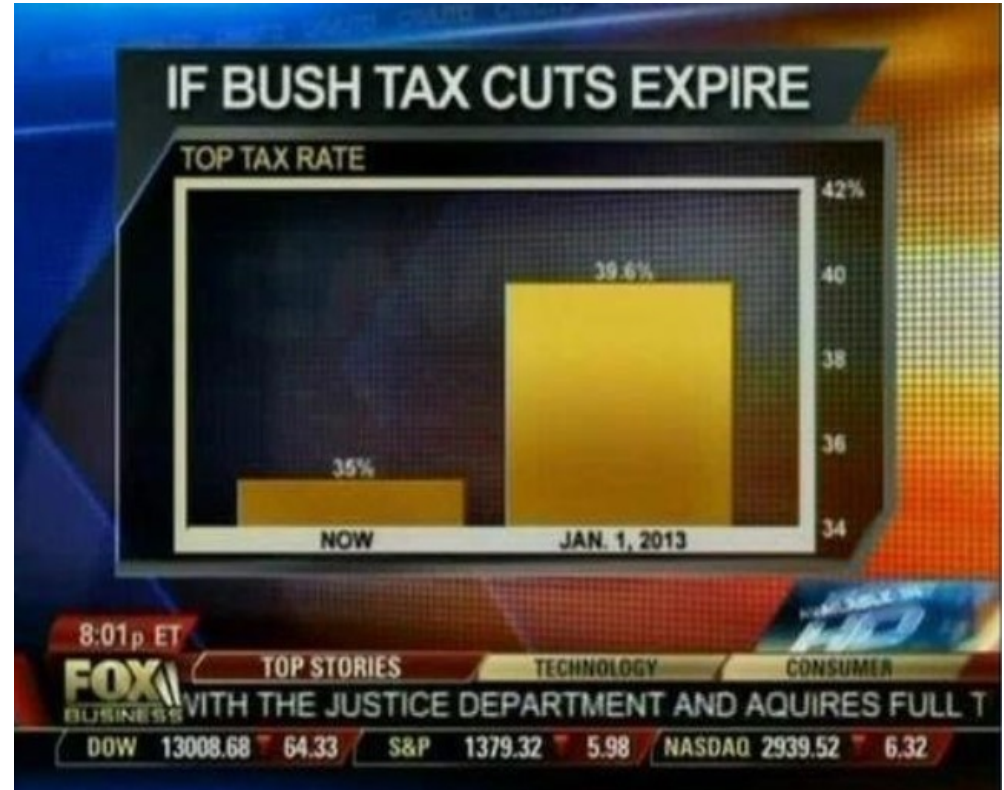
- Fox News
- What is wrong with this chart?
- Percentages do not add up to 100%
- Pointless 3D serves only to make some slices look bigger than they are



# Data visualisation

## Misleading charts

- Fox News
- What is wrong with this chart?

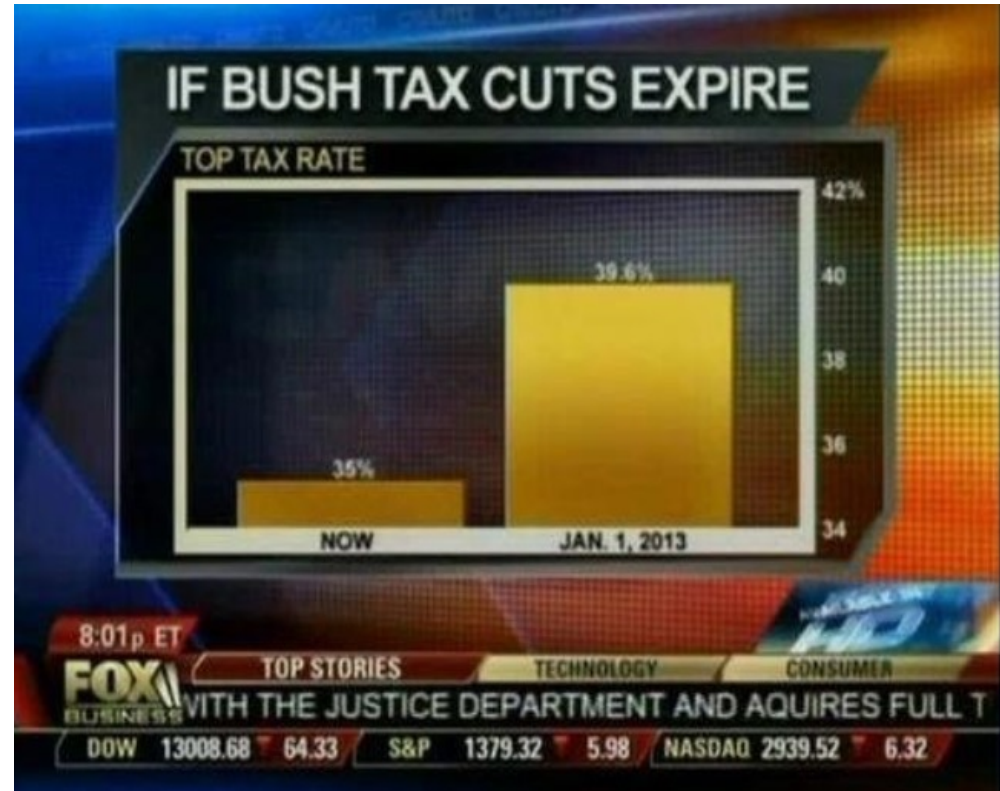
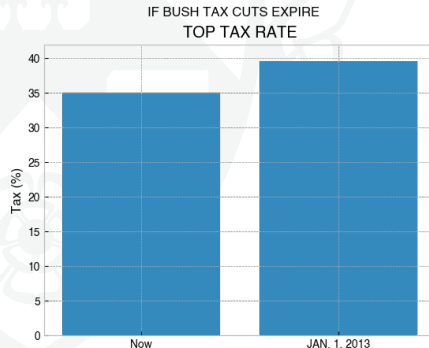




# Data visualisation

## Misleading charts

- Fox News
- What is wrong with this chart?
- Scale of y-axis is misleading



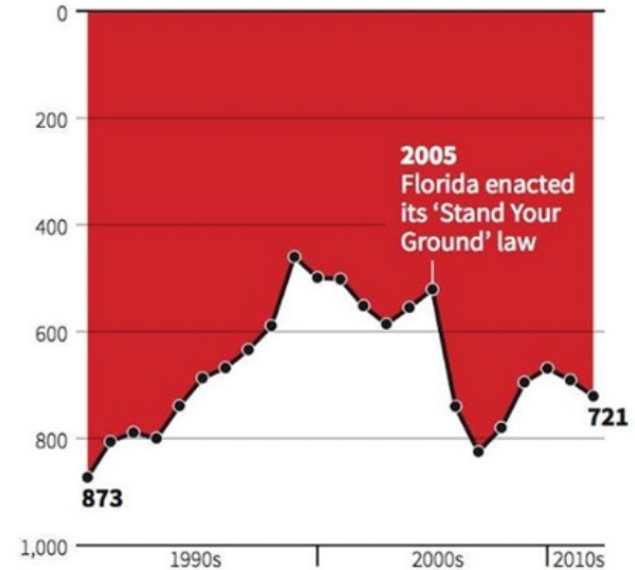
# Data visualisation

## Misleading charts

- Florida Department of Law Enforcement
- What is wrong with this chart?

### Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

REUTERS

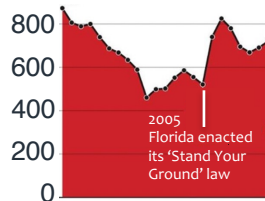




# Data visualisation

## Misleading charts

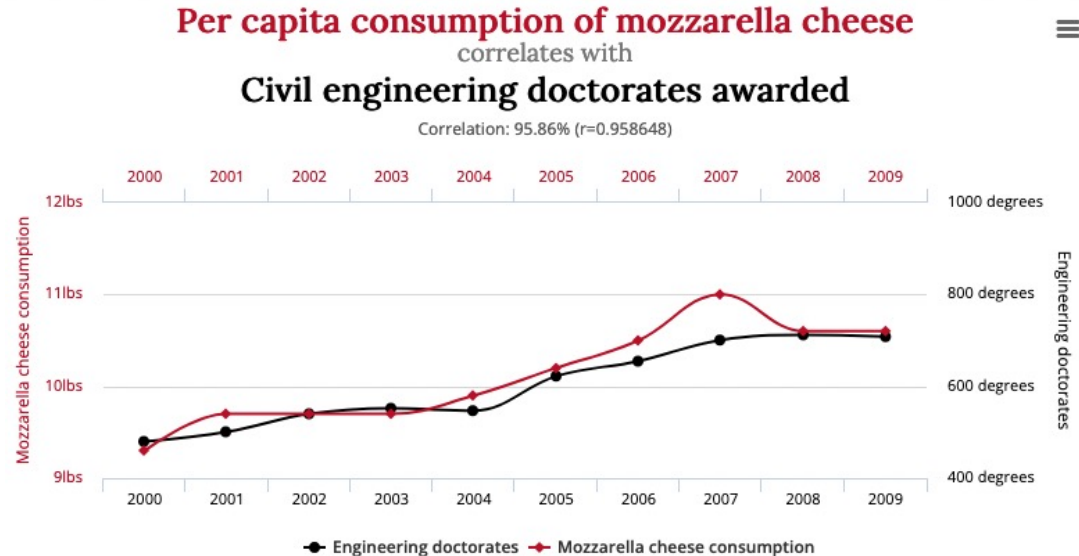
- Florida Department of Law Enforcement
- What is wrong with this chart?
- Inverted Y-axis!
- Cynical attempt at misleading people into thinking that the 'Stand Your Ground' law led to a reduction of gun deaths in Florida
- What that data really show →



# Data visualisation

## Misleading charts

- From the [www.tylervigen.com](http://www.tylervigen.com)
- What's wrong with this chart?



# Data visualisation

## Misleading charts

- From the [www.tylervigen.com](http://www.tylervigen.com)
- What's wrong with this chart?
- Twin axis scaling gives the impression of a causal relationship for two unrelated variables
- Correlation does not imply causation!
- More examples at <https://tylervigen.com/spurious-correlations>

