

Quantitative Data Analysis

Text durch Klicken hinzufügen

Quantitative Data Analysis

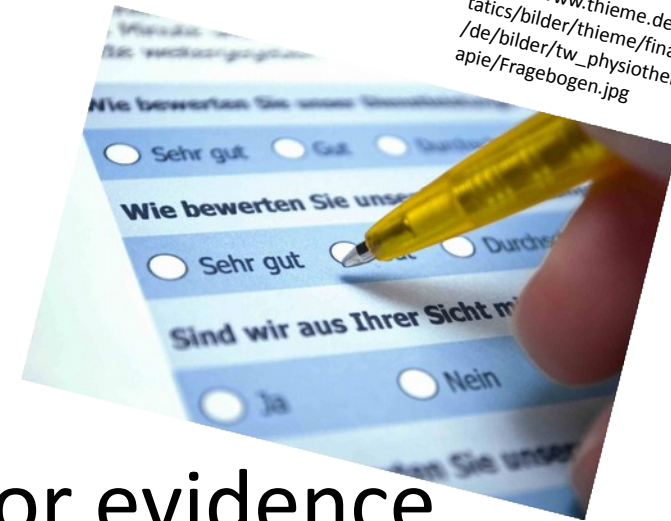
Text durch Klicken hinzufügen

Definition

[http://www.deutsche-apotheker-zeitung.de/uploads/tx_cronaz/agesnews-image/Umfrage-Bogen_-_cirquedesprit_-_Fotolia.jpg](http://www.deutsche-apotheker-zeitung.de/uploads/tx_cronaz/tagesnews-image/Umfrage-Bogen_-_cirquedesprit_-_Fotolia.jpg)



https://www.thieme.de/s-tactics/bilder/thieme/final/de/bilder/tw_physiotherapie/Fragebogen.jpg



„Quantitative data means data, or evidence, based on numbers.“

Selbsterkennung Depression: WHO-5-Fragebogen zum Wohlbefinden

Zutreffendes bitte ankreuzen

In den letzten beiden Wochen

	die ganze Zeit	meistens	über die Hälfte der Zeit	weniger als die Hälfte der Zeit	ab und zu	zu keinem Zeitpunkt
1. war ich froh und guter Laune	5	4	3	2	1	0
2. habe ich mich ruhig und entspannt gefühlt	5	4	3	2	1	0
3. habe ich mich aktiv und voller Energie gefühlt	5	4	3	2	1	0
4. habe mich beim Aufwachen frisch und ausgeruht gefühlt	5	4	3	2	1	0
5. war mein Alltag voller Dinge, die mich interessieren	5	4	3	2	1	0

Gesamtpunkte: _____

Wenn Sie in der Summe weniger als 13 Punkte erreichen, sprechen Sie bitte mit Ihrem Arzt.

Quelle: adaptiert nach WHO, Version 1998

OGS-Gesundheit

http://www.nein-zur-depression.at/Portals/5/WHO_fragebogen.jpg

Die neue Doodle Flip Flop Kollektion ist da!

Entdecken Sie die neuen Farben und lassen Sie Ihre Füße tanzen!

Doodle

Welche FlipFlop Farbe ist Ihr Favorit?

Eine Umfrage von Miss Summer | 27 | 0 | vor einem Tag

	Limettengrün	Canarygelb	Zitronengelb	Himmelblau
7 Teilnehmer				
Tom	✓			
Luisa				✓
Cloé				
Philipp			✓	
Raphaela				
Joy				
Stefan				
Ihr Name	2	1	1	3

<http://doodlede.files.wordpress.com/2012/07/doodle-deluxe-small.jpeg>

Kinds of quantitative data

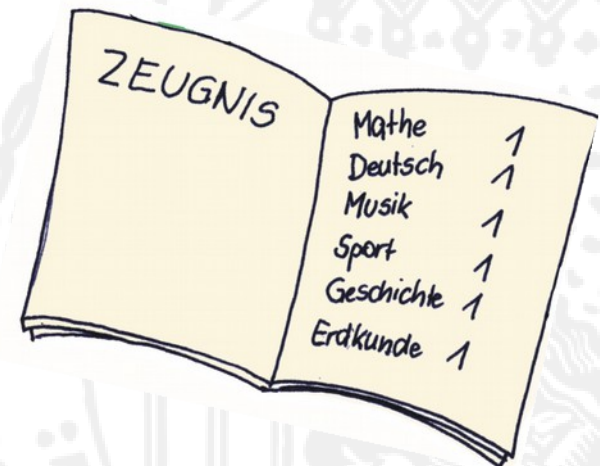
- Nominal data
 - Description of categories – „*categorical data*“
 - Unordered set
 - No numeric value
 - No arithmetical operations
 - Example: Gender (female, male)

0 + 1 = 2 ?



Kinds of quantitative data

- Ordinal data
 - Possess a natural order – „ranked data“
 - Arithmetical operations (<,>)
 - Example: ratings, school grades



<http://www.mindstation.de/wp-content/uploads/2014/04/1-zeugnis-Kopie1.jpg>

Kinds of quantitative data

- Interval data
 - Measurements
 - Distance between units
 - Ranking of categories is proportionate
 - Arithmetical operations (+,-)
 - Example: interval of years (1914,1918) same as (2000, 2004)
- Ratio data
 - Interval data with true zero
 - Arithmetical operations (+,-,*,/)
 - Example: age, height, weight can be 0



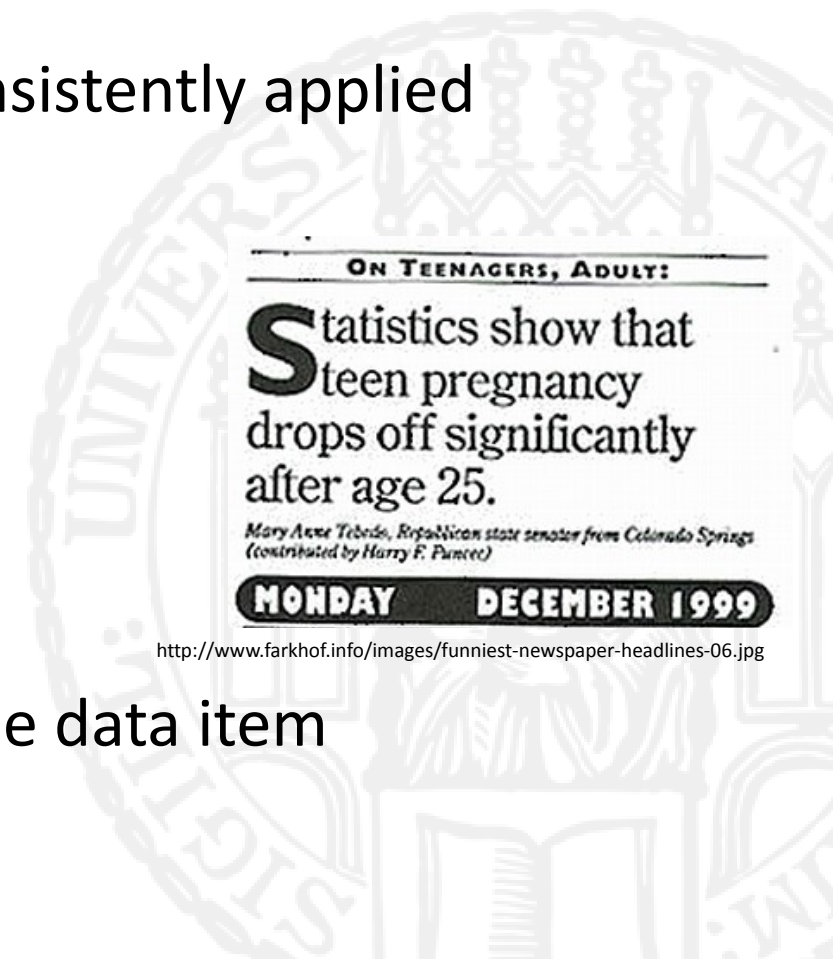
Kinds of quantitative data

- Discrete data
 - Measurement leads to whole number
 - Example: „number of children“
- Continuous data
 - Measurements in greater accuracy
 - Example: age in years, days, minutes, seconds...



Preparation for analysis

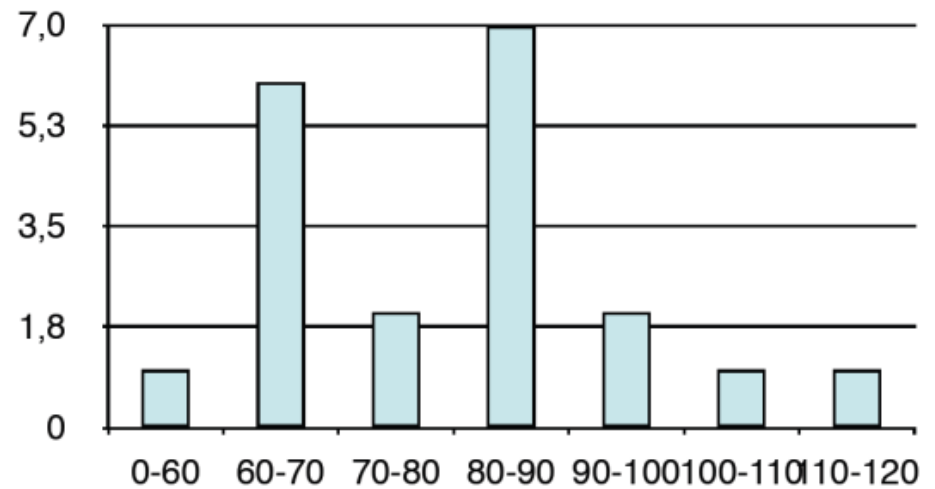
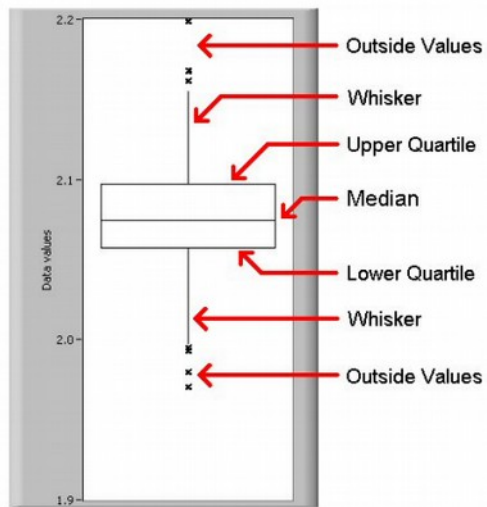
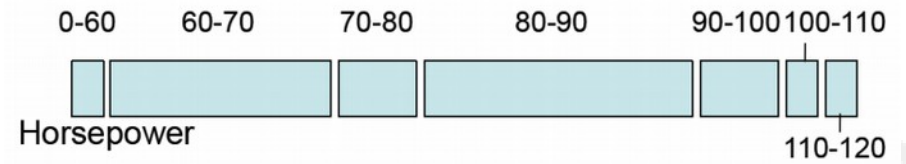
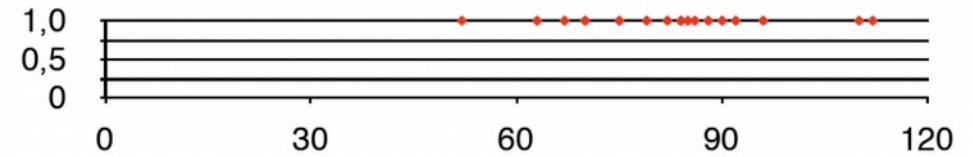
- Encoding data
 - predefined answers
 - Themes in answers to open questions
 - Mutually exclusive, exhaustive, consistently applied
- Error and outlier detection
 - logical inconsistencies
 - Bad values
 - Note: record what is wrong with the data item



Visual aids

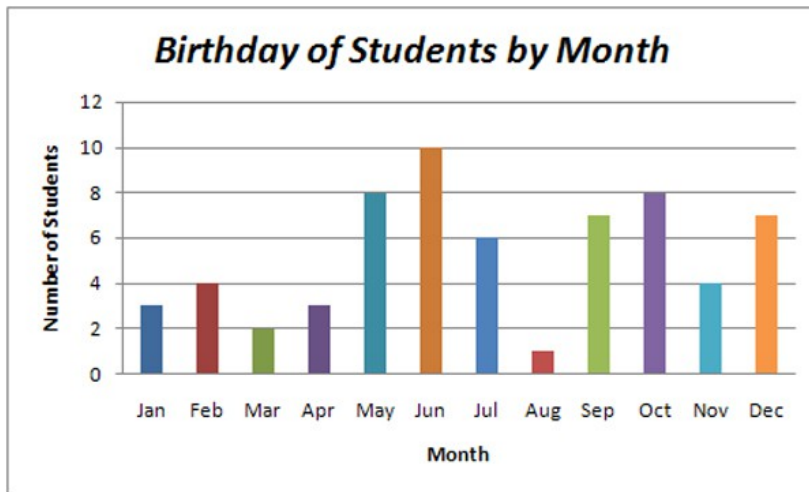
- Univariate data

- Plot
- Boxplot
- Histogram
- Bargram



Visual aids

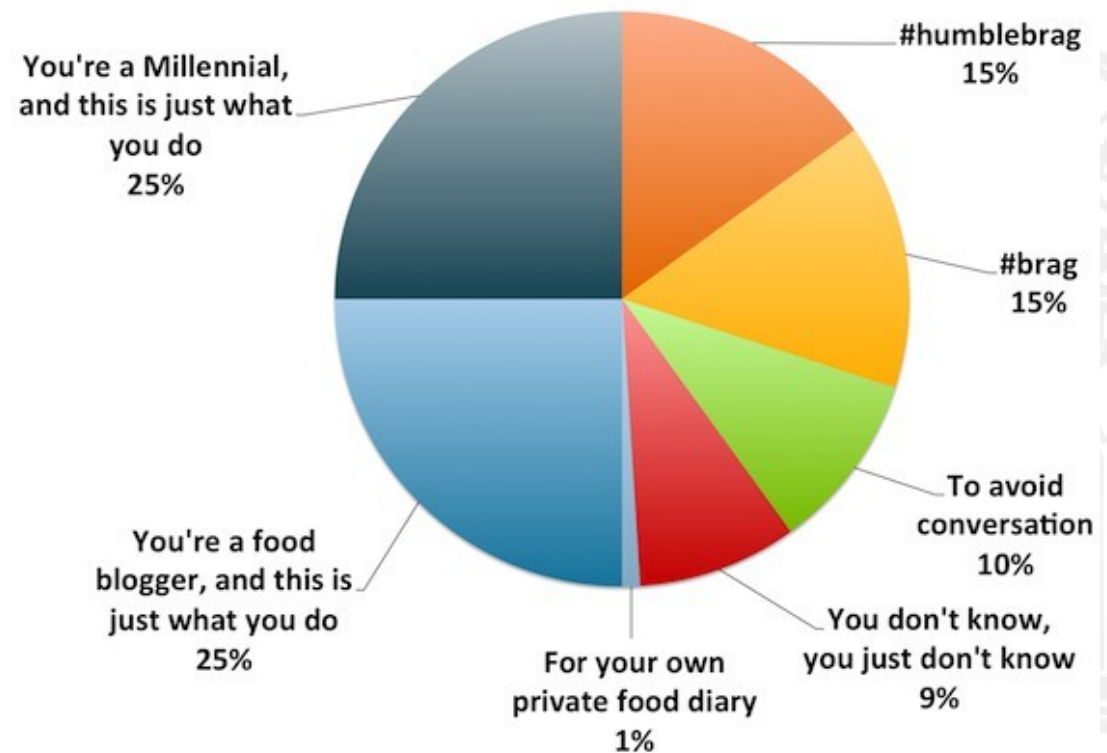
- Bar charts
 - Frequencies
 - Nominal data



<http://wikieducator.org/images/c/c7/Bargraph.jpg>

- Pie Charts

REASONS WHY YOU'RE TAKING PHOTOS OF YOUR FOOD



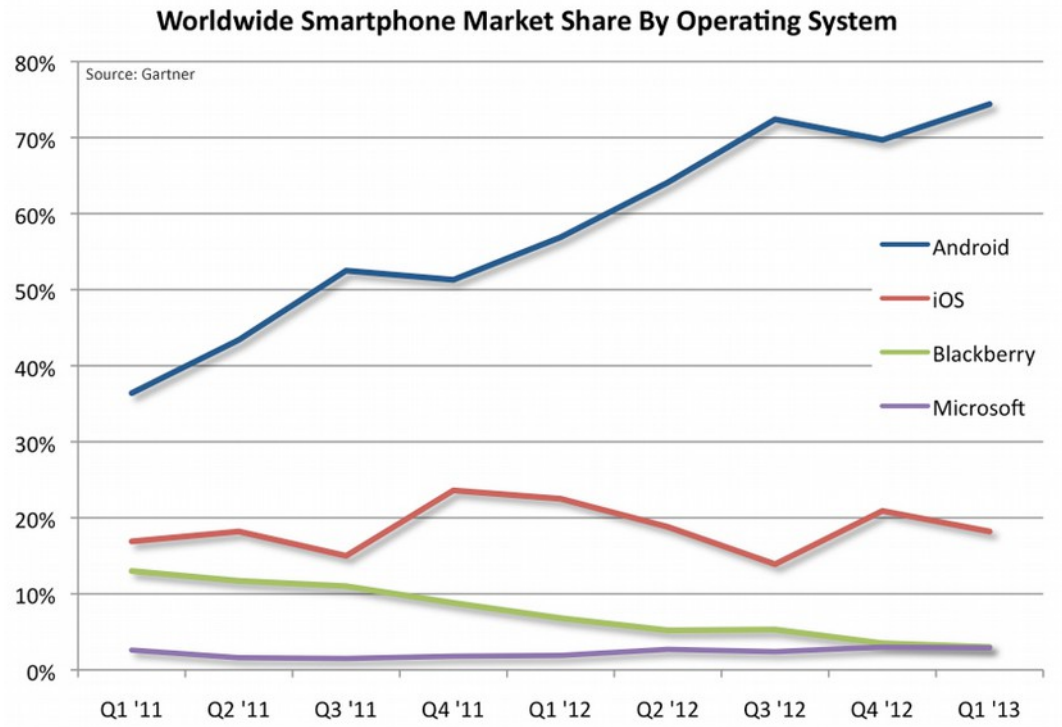
<http://www.laweekly.com/squidink/2013/07/15/pie-chart-reasons-why-youre-taking-photos-of-your-food>

Visual aids

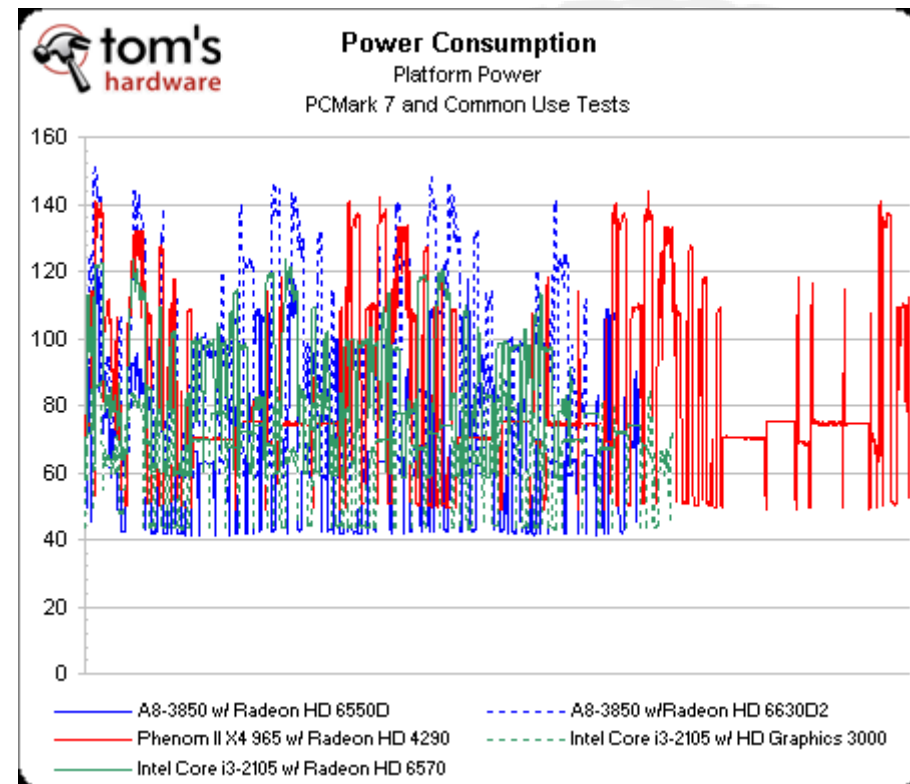
– Line graphs

- Trends in data
- Continuous data

„Be careful not to overload the reader.“



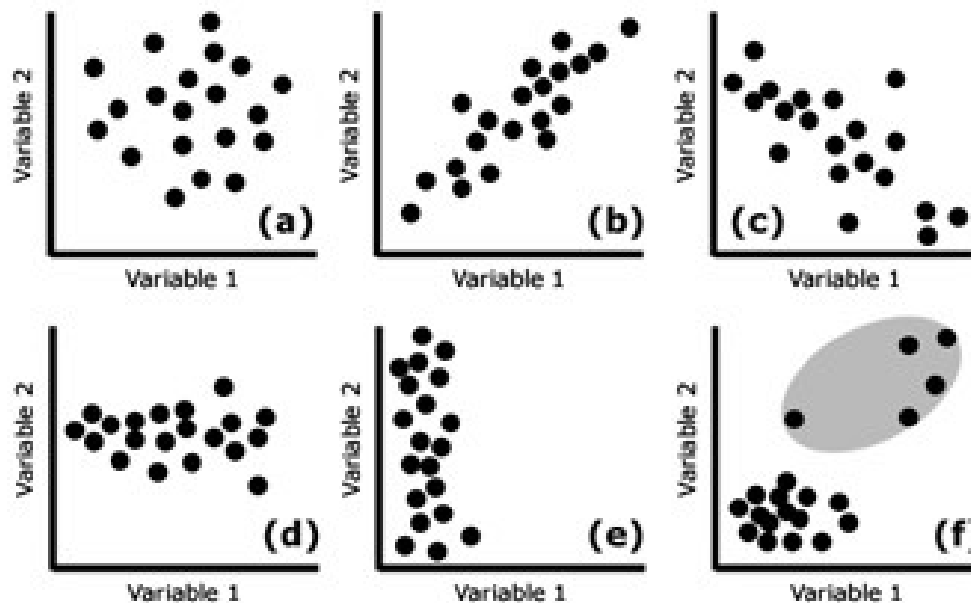
<http://static3.businessinsider.com/image/5192a95969bedd702200000a/chart-of-the-day-the-iphones-market-share-is-dead-in-the-water.jpg>



<http://media.bestofmicro.com/7/6/298338/original/power-consumption-line-graph.png>

Visual aids

- Bivariate data
 - Scatterplot
 - Relationship between variables
 - Many values required



http://web.anglia.ac.uk/numbers/common_folder/graphics/fig9_scatter_plots.jpg

Visual aids

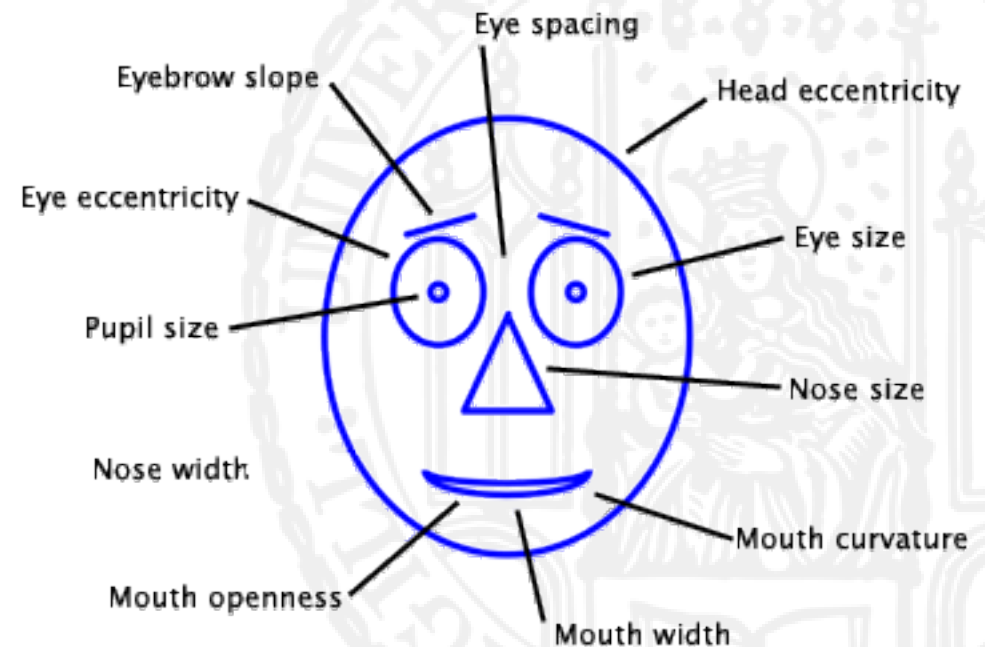
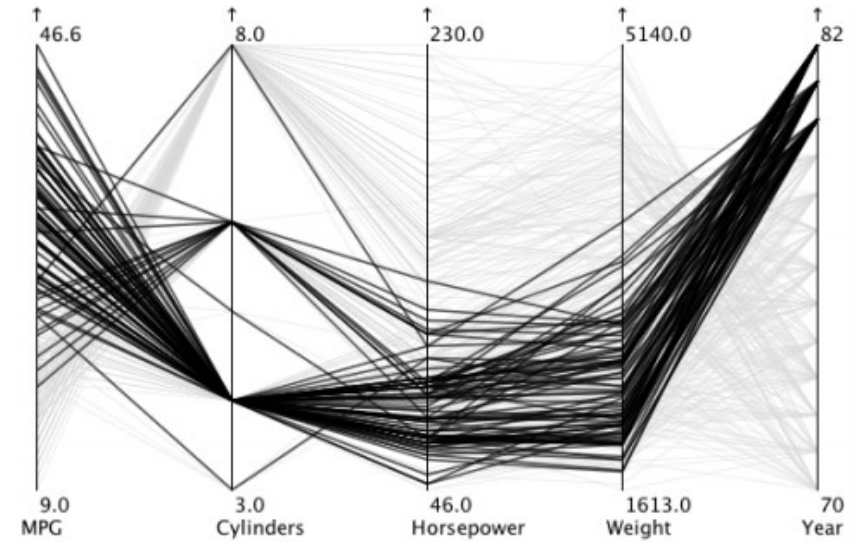
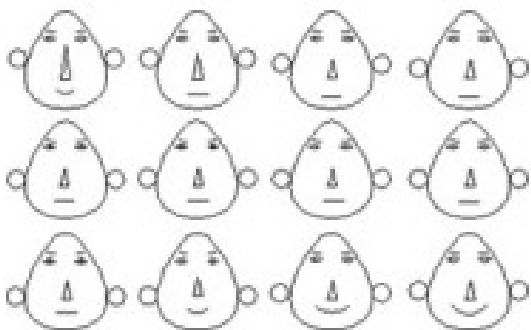
- Multivariate data

- Geometric Transformation

- Scatterplot Matrix
- Parallel Coordinate Plot

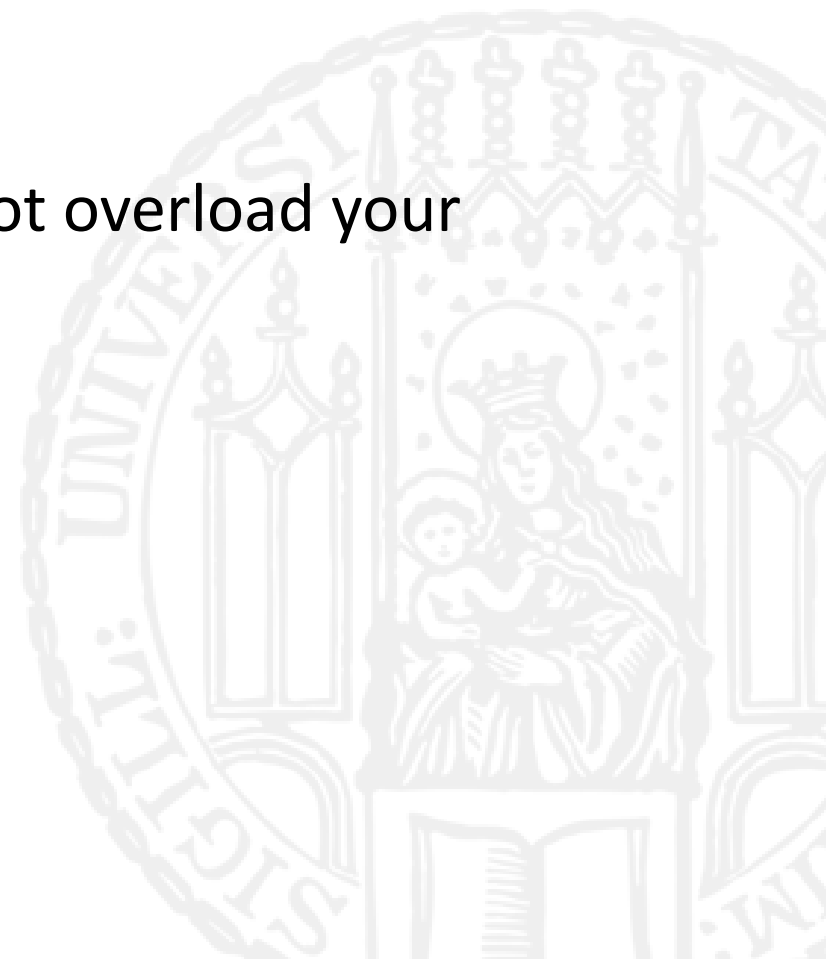
- Glyph-based Visualizations

- Chernoff Faces
- Shape Coding



Visual aids

- Guidelines
 - Make it easy to read
 - Entitle it
 - Show information about the units
 - Give enough information, but do not overload your visualization



Statistics

- Describing the central tendency

- Mean

$$\bar{x}_{\text{arithm}} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

- “average”
- add up a set of scores and divide by the number of scores

- Median

- Mid point in a range of scores
- List values in order – read off the middle value

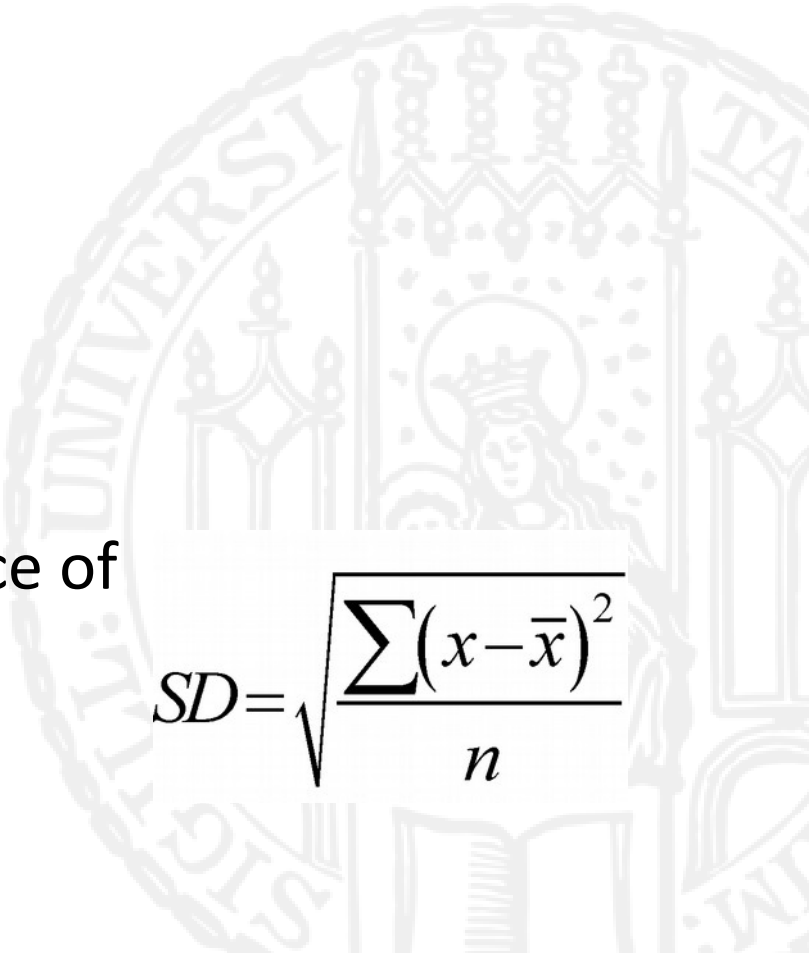
$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & n \text{ ungerade} \\ \frac{1}{2} (x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & n \text{ gerade.} \end{cases}$$

- Mode

- Value which occurs most frequently

Statistics

- Describing the distribution
 - Range: Highest and lowest data value
 - Fractile: Dividing up the data
 - Quartiles
 - Deciles
 - Percentiles
 - Standard deviation: Average distance of each data value from the mean


$$SD = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

Statistics

- Describing correlations

- Null hypothesis and significance

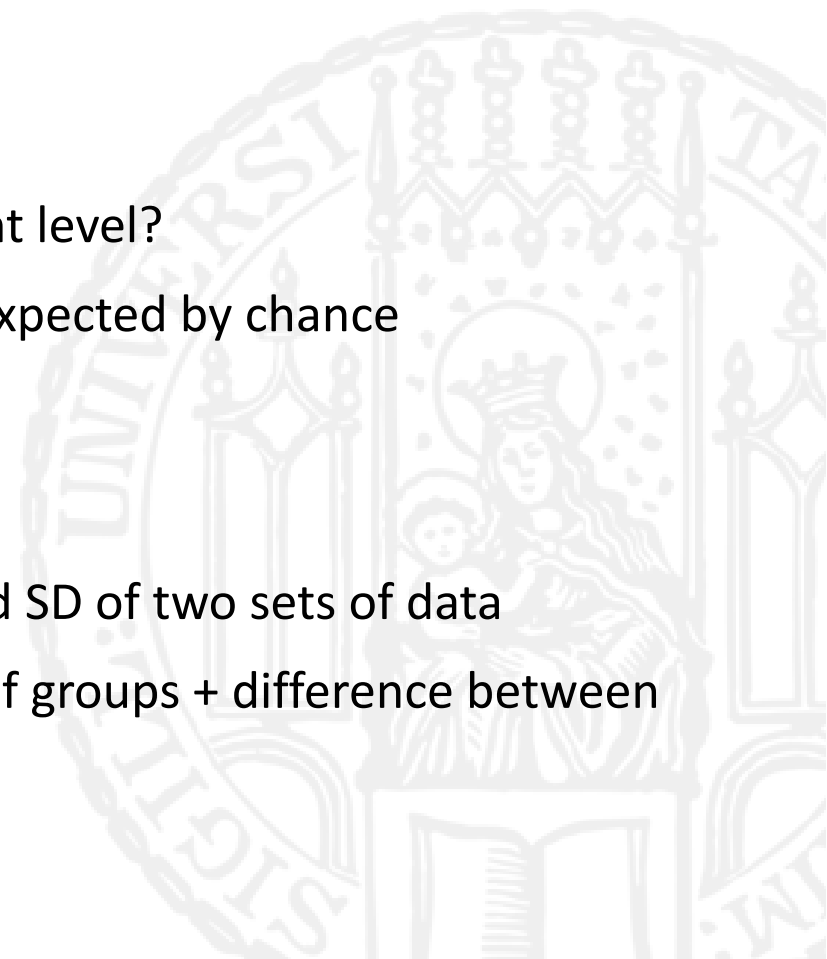
- Premise: No true relationship between variables
 - Probability

- Chi-square test

- Two variables are associated to a significant level?
 - Comparison between observed in data – expected by chance

- T-tests

- Significant difference?
 - Independent groups T-test: Uses mean and SD of two sets of data
 - Dependent / matched pairs T-test: Mean of groups + difference between scores



Analysis and Evaluation

- Analysis:
 - What do your results show?
 - What do they imply? What is important?
 - How do they relate to other reported research?
 - Do your findings confirm those of other researchers?
 - What relevance do they have?



Analysis and Evaluation

- Evaluation:
 - + Scientifically respectable
 - + Based on well-established techniques
 - + Not based on subjective impressions
 - + quick analysis using software
 - danger of going overboard
 - dependent on quality of initial data
 - forced to plan ahead
 - beware of hidden subjectivity



Take away message

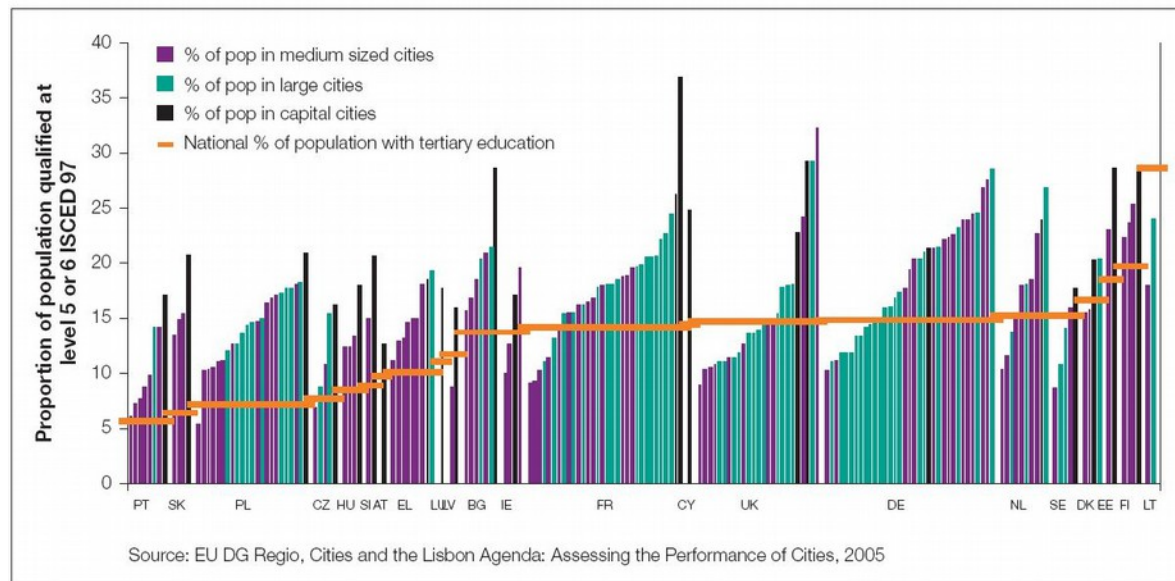
- Algorithmical evaluation of large datasets
- Numbers don't lie
- Don't draw conclusions you can't back up



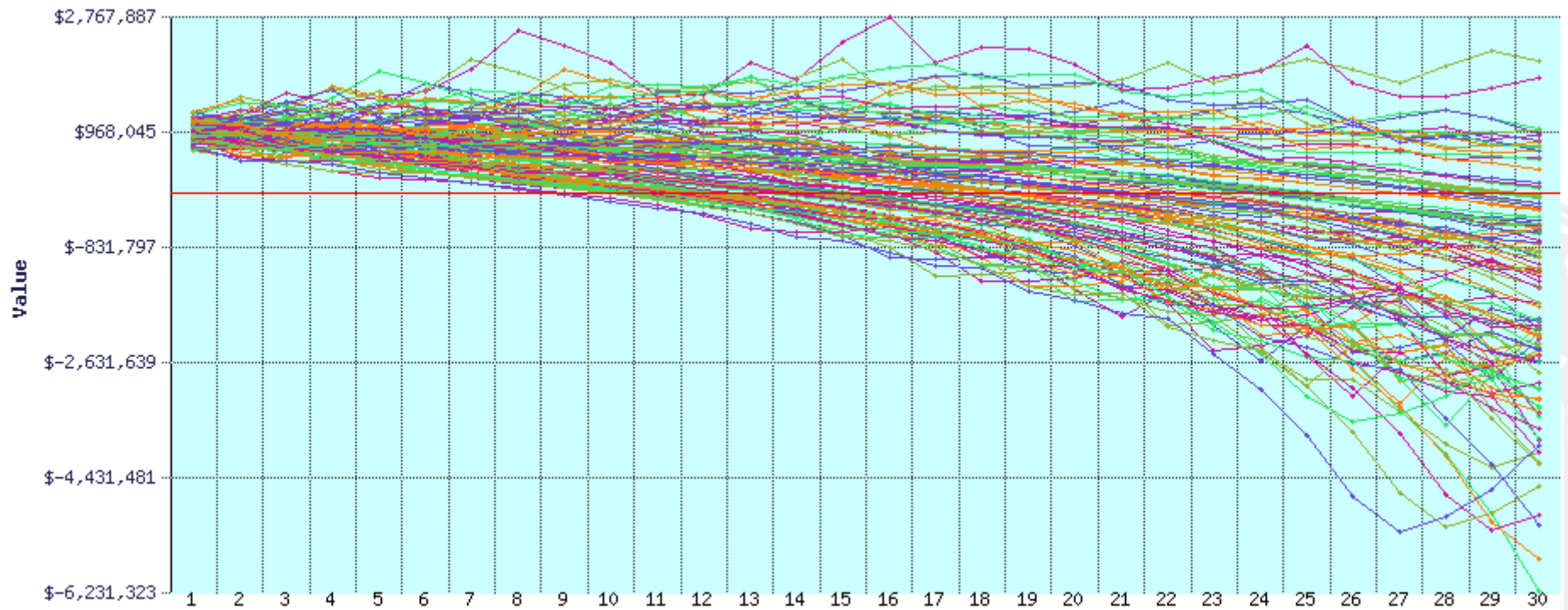
Discussion



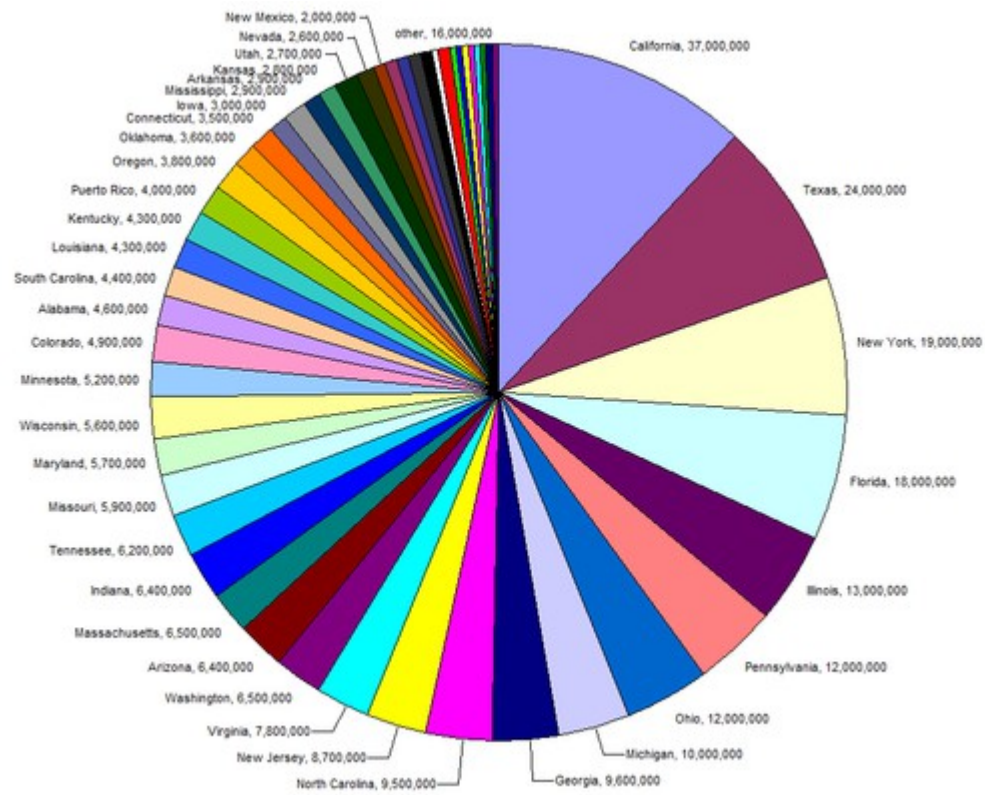
Figure 8.6: Proportion of the population with tertiary education, 2001.



http://www.improving-visualisation.org/img_uploads/2009-05-07_Thu/200956164538.jpg



http://genywealth.com/wp-content/uploads/2010/03/line-graph.php_png



U.S. States by population

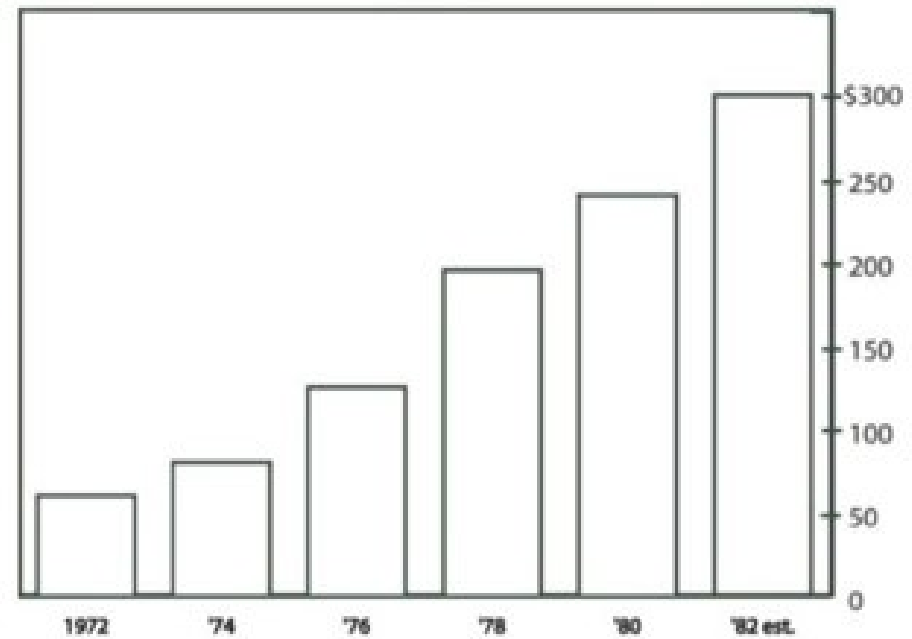
http://eagereyes.org/media/2010/501px-Pie_chart_of_US_population_by_state.png

MONSTROUS COSTS

Total House and Senate campaign expenditures, in millions



MONSTROUS COSTS
Total House and Senate campaign expenditures, in millions



<http://www.24hrco.com/images/articles/html/image3.jpg>