Lecture 9 Pre-attentive attributes, gestalt, illusions

Data visualization · 1-DAV-105 Lecture by Broňa Brejová

Human visual perception

What happens when we look at the figure?

- The light from the screen / projector hits the retinas of our eyes
- Photoreceptor cells transduce (convert) this signal into nerve impulses
- In the brain:
 - detection of basic features
 - recognition of patterns
 - interpretation, assignment of meaning

Human visual perception

What happens when we look at the figure?

- The **light** from the screen / projector hits the retinas of our **eyes**
- Photoreceptor cells transduce (convert) this signal into nerve impulses
- In the brain:
 - detection of basic features
 - recognition of patterns
 - interpretation, assignment of meaning

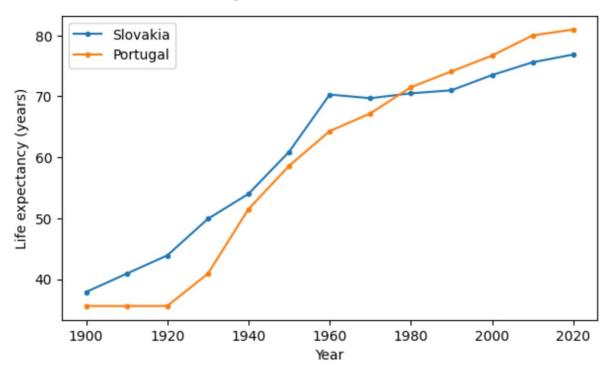
Today: Detection of features and patterns, use for visualization

Note: Human visual perception is very good for detecting **motion** (danger/prey). This is relevant for animated visualization, but not covered today.

In which period of time was life expectancy higher in Slovakia than in Portugal?

		1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020
	Country													
;	Slovak Republic	37.9	40.9	43.9	49.9	54.0	60.9	70.3	69.7	70.5	71.0	73.5	75.6	76.9
	Portugal	35.6	35.6	35.6	40.9	51.5	58.6	64.3	67.2	71.5	74.1	76.7	80.0	81.0

In which period of time was life expectancy higher in Slovakia than in Portugal?



How many copies of digit six do you see?

What about now?

- **6**53
- 821550

What about Slovakia vs Portugal in this table?

	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020
Slovakia	37.9	40.9	43.9	49.9	54.0	60.9	70.3	69.7	70.5	71.0	73.5	75.6	76.9
Portugal	35.6	35.6	35.6	40.9	51.5	58.6	64.3	67.2	71.5	74.1	76.7	80.0	81.0

Pre-attentive attributes

- Features of the seen objects detected by our brain very fast
- Prior to and without the need of conscious awareness
- Brain uses them to guide attention / gaze to interesting parts of the scene
- Their correct use allows fast and effortless understanding of our visualizations

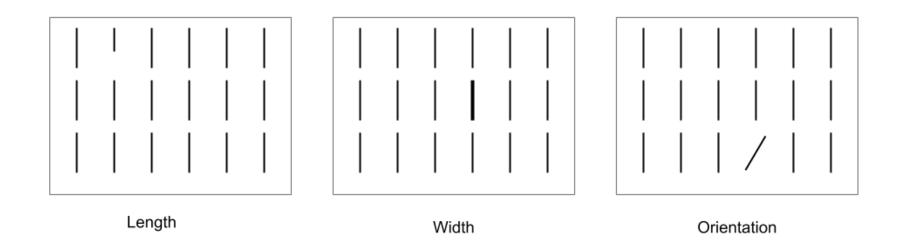
Next:

Examples of important pre-attentive attributes (form, color, position) following Few 2009

See also https://www.csc2.ncsu.edu/faculty/healey/PP/

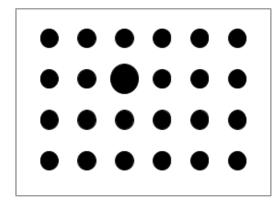
Pre-attentive attributes: form

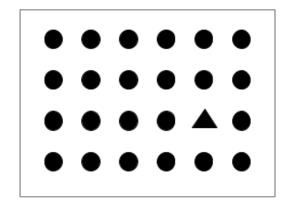
We can quickly distinguish one object that differs from the rest



Pre-attentive attributes: form

We can quickly distinguish one object that differs from the rest

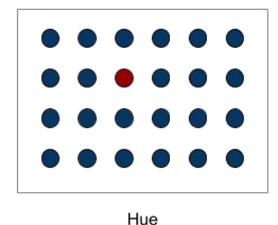


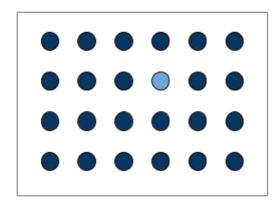


Size Shape

Pre-attentive attributes: color

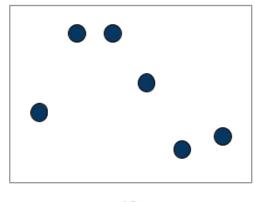
We can quickly distinguish one object that differs from the rest

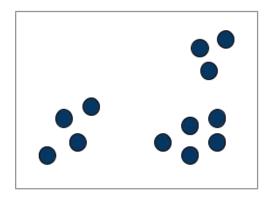




Lightness

Pre-attentive attributes: position





2D

Groups

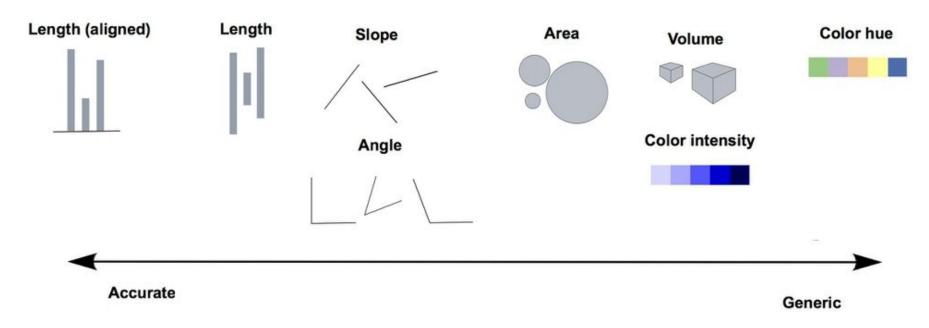
Hierarchy of graph elements

Cleveland, McGill 1985

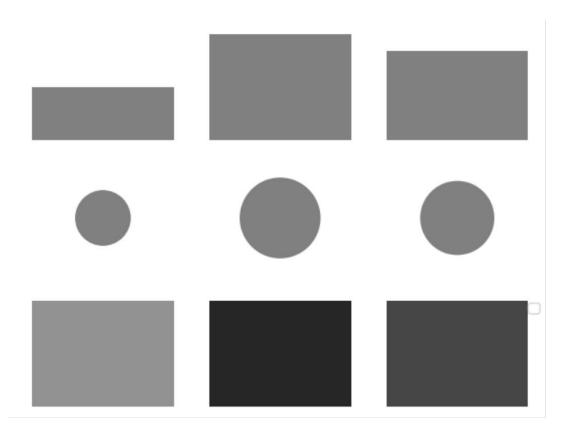
Experiments with volunteers of how well they **judge ratios** between values graphically encoded in different ways.

Not all pre-attentive attributes are equally good for quantitative reasoning.

Prefer elements on the left side for accuracy



The same data with length / area / color / angle



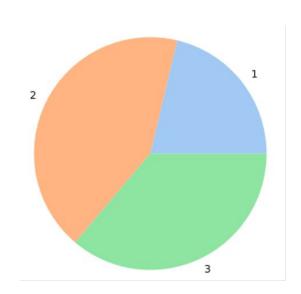


Chart selection tools

In lecture 3 and later, we have seen many types of graphs

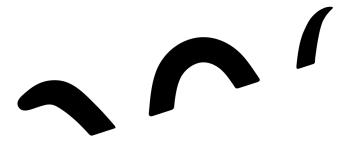
Some websites list them based on variable type and purpose for easier selection:

- https://www.data-to-viz.com/
- https://extremepresentation.typepad.com/blog/2006/09/choosing_a_good.html

Let us look at some the suggestions from the first website in terms of the hierarchy of graph elements

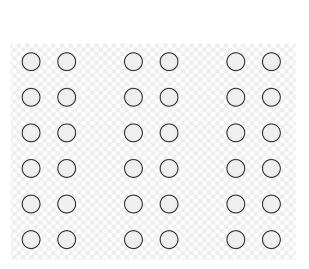
From parts to the whole: gestalt

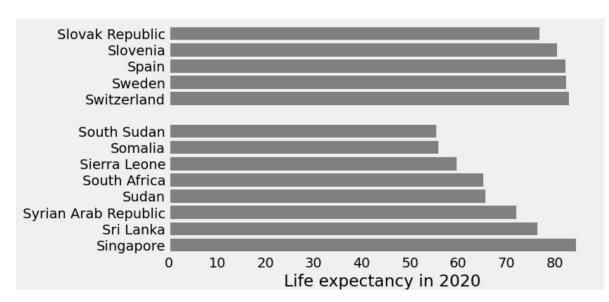
- Gestalt psychology (early 20th century, Austria and Germany)
- Gestalt means pattern
- Our brains group individual shapes into larger patterns
- The brain favors speed to precision (illusions, errors)
- It also favors symmetry and simplicity
- Several gestalt principles are relevant in data visualization



Principle of proximity

- Objects located close to each other are perceived as a group
- Good use of space in plots / tables / presentations can improve readability

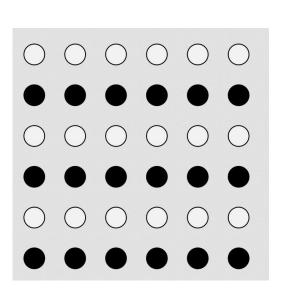


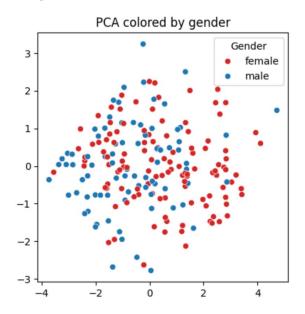


https://commons.wikimedia.org/wiki/File:Gestalt_proximity.svg

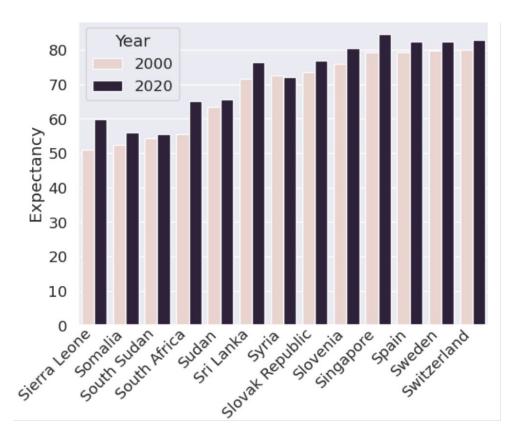
Principle of similarity

- Similar objects are perceived as a group even of not close by
- Various plots use color / shape to distinguish groups





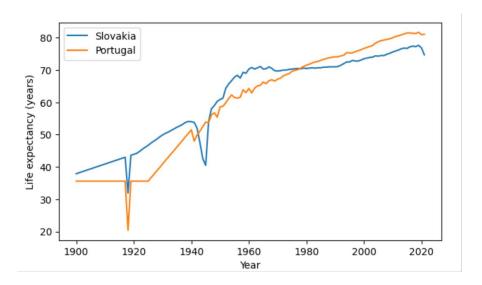
How are both principles used here?

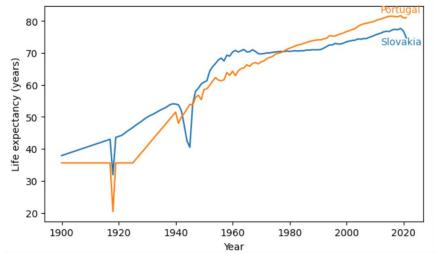


Example

separate legend vs marking lines with text in the same color

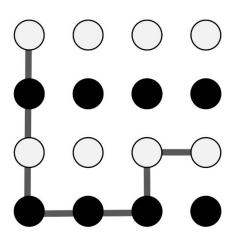
using principles of proximity and similarity

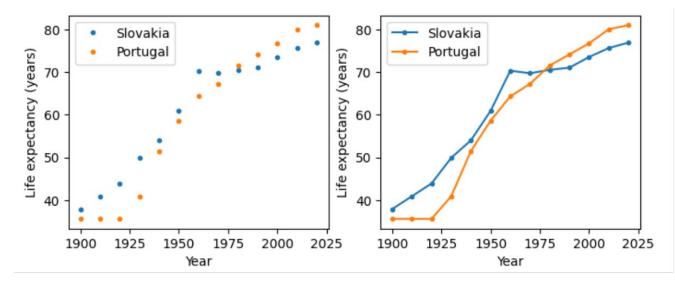




Principle of connection

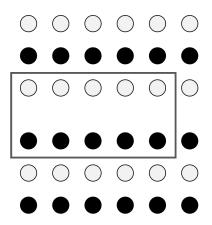
- Connected objects are perceived to form a group
- Stronger than proximity and similarity
- Consider carefully when to use line graph vs. scatter plot

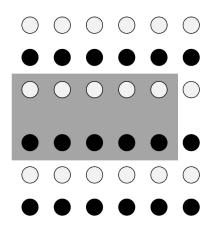


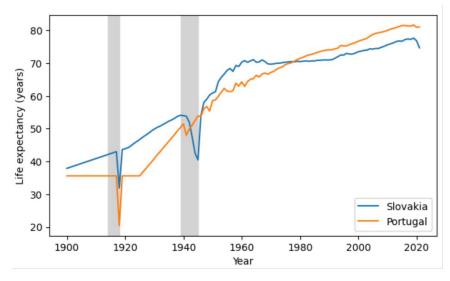


Principle of enclosure

- Enclosed objects are perceived as a member of the group
- Stronger than proximity and similarity
- Useful for highlighting in plots; little is enough (e.g. light background)

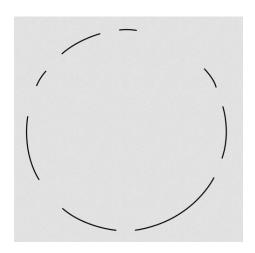


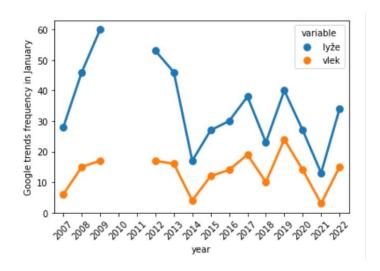




Principle of closure

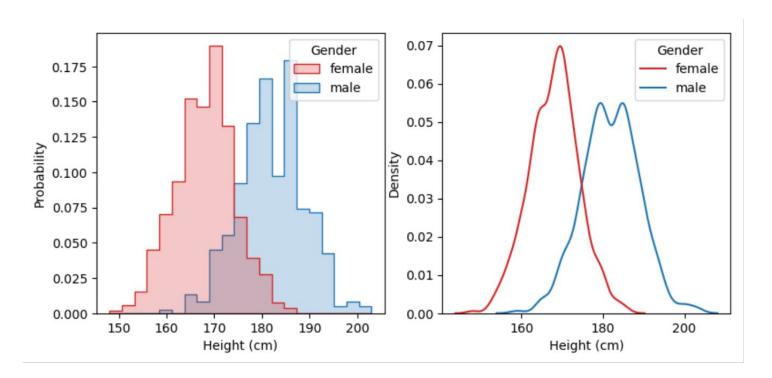
- Our brain fills gaps in figures, connects interrupted lines
- Useful / dangerous when interruptions by design





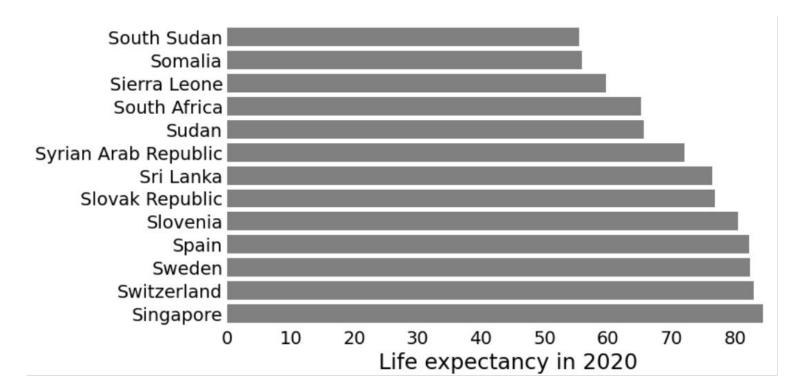
Principle of continuity

Smooth lines are easier to follow than angular and sharp



Frames not necessary, gestalt principles fills them in

(principles of closure and continuity)

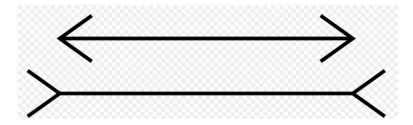


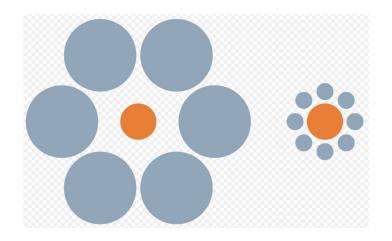
Illusions

- Fast visual processing leads to errors
- These are demonstrated by many optical illusions
- Beware not to create illusions in your plots

Illusions: length and size

Müller-Lyer and Ebbinghaus illusions

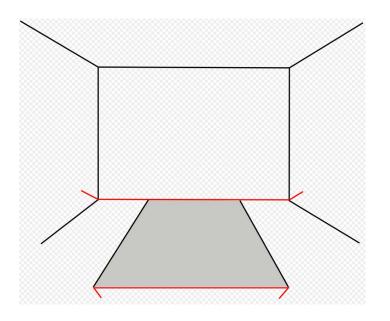


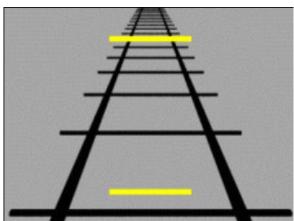




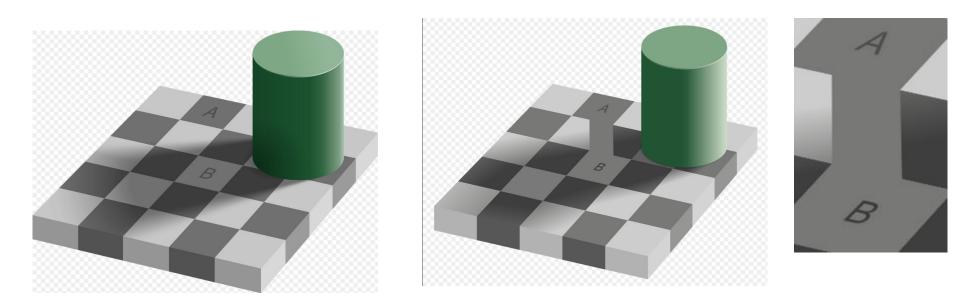
https://commons.wikimedia.org/wiki/File:M%C3%BCller-Lyer_illusion.svg https://commons.wikimedia.org/wiki/File:Mond-vergleich.svg

Illusions: length, perspective, spatial compensation





Illusions: color



https://en.wikipedia.org/wiki/File:Checker_shadow_illusion.svg https://commons.wikimedia.org/wiki/File:Grey_square_optical_illusion_proof2.svg

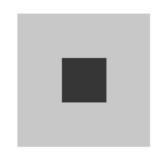
Illusions: color

Mach bands: when bands touch, the edge effect exaggerates their difference



Illusions: color

In heatmap the perception of colors influenced by their surroundings





```
a = np.array([
   [10, 10, 10, 0, 0, 0, 0],
   [10, 37, 10, 0, 0, 0, 0],
   [10, 10, 10, 0, 0, 0, 0],
   [0, 0, 0, 0, 0, 0, 0],
   [ 0, 0, 0, 0, 40, 40, 40],
   [ 0, 0, 0, 0, 40, 37, 40],
   [ 0, 0, 0, 0, 40, 40, 40],
pal = sns.color palette("dark:white r", as cmap=True)
axes = sns.heatmap(data=a, square=True,
                  cmap=pal, cbar=False)
axes.axis('off')
```

Working memory

- Iconic memory: extremely short-term (<1s), simple pre-attentive processing
- Visual short-term memory: many seconds, but very small capacity (only 3-5 items)
- Long-term memory: store and recall selected information

Since working memory is small, looking at a plot with many colors requires back-and-forth between legend and plot

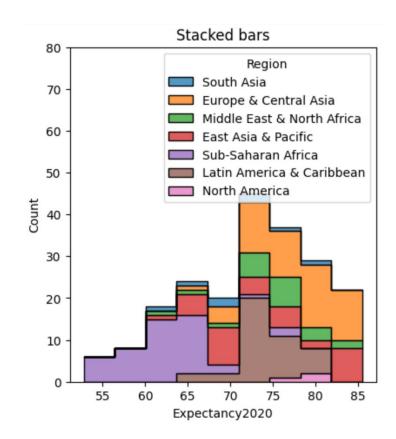


Chart and table junk

- Chart junk: elements of plots not necessary to convey information
- They unhelpfully catch our attention through pre-attentive attributes
- Most visualization can be improved by simplification
- Some redundancy can be useful

Nice visualizations of the simplification process:

- https://www.darkhorseanalytics.com/blog/data-looks-better-naked
- Also <u>tables</u>, <u>maps</u> and the unpopular <u>pie charts</u>

Summary

- Pre-attentive attributes are processed by our brains very fast
- Choosing the right attributes from the hierarchy allows accurate quantification
- Principles of gestalt describe how the brain connects part to the whole
- The brain can also make errors in visual processing as seen in illusions
- Removing chart junk concentrates our attention to the important elements

Additional sources

- <u>Utilizing Gestalt Principles to Improve Your Data Visualization Design</u>
- http://daydreamingnumbers.com/blog/gestalt-laws-data-visualization/
- Albert Cairo: The Functional Art
- C.N. Knaflic: Storytelling with Data
- Stephen Few: Now You See it

Visualizing text data

Visualizing text data

Working with natural text is difficult

- Complex grammar, ambiguous meaning, synonyms, etc.
- Lot of machine learning research
- Nonetheless sometimes simple statistics on frequencies of words or groups of words can be useful
- Usually we remove stop words (frequent words such as "and", "is"...) and apply lemmatization (convert inflected words to canonical form, such as "seen" -> "see")

State of the Union Address, 2002 vs. 2011

Word clouds

afghanistan american attack best budget children citizens Congress continue corps Country create danger depend destruction develop economy encourage enemies evil extend fight free freedom government health help history home homeland hope increase islamic JODS join lives mass military moment months nation opportunity peace people police power protect rebuild regimes resolve retirement Security states tax terror terrorists thousands together tonight training true united war was weapons women work workers world

President Bush, January 29, 2002

afghan ago already american behind believe best better building business Care century challenge chance change child children clean college company compete congress Country create cuts deficit democrats different don done dream economy education energy family future generation **Government** health help home idea nation republicans research responsibility schools spending states step students success support sure tax teachers technology things together tonight troops willing Win WORK workers world years

https://commons.wikimedia.org/wiki/File:State_of_the_union_word_clouds.png

Word clouds

- Display the most common words from a text
- Size of words grows with frequency
- Arranged to be visually pleasing
- Not the best option for understanding/comparing word frequencies
- You can also display word frequencies using bar graphs and other plot types

Tag cloud

- Endings of German city names typical for individual regions
- Combination of a word cloud and map
- Figure from <u>Reckziegel et al 2018</u>



Word tree

Shows with words most often follow or precede a given word using a hierarchy

Text: <u>Introduction</u> to The Origin of Species by Charles Darwin, 1859, 1872

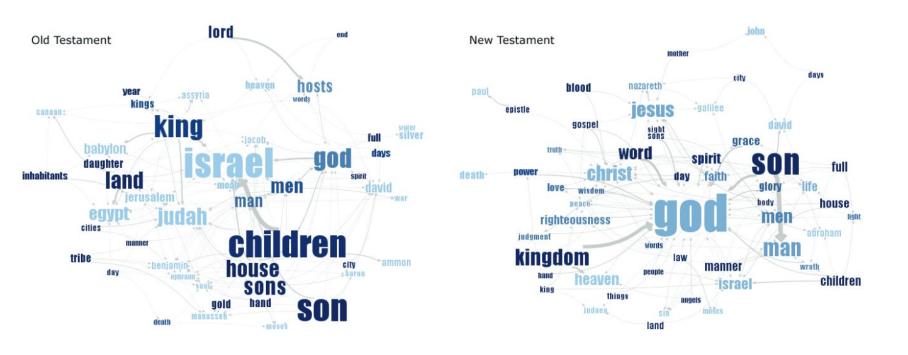
Figure source

emed to throw some light on cal conclusions that I have the origin his excellent judgment. In considering of aining as yet unexplained in regard to ating by his selection successive slight variations. I will then pass on to an can be changed and perfected into a highly developed being or into cal succession, and other such facts, might come to the conclusion that ntly entertained, and which I formerly entertained - namely, that acknowledged varieties of any one species are the descendants of s, applied to the whole animal and vegetable kingdoms. As many each tly entertained, and which I formerly entertained - namely, that ions of the many beings which live around us. Who can explain why one t species, in the same manner as the acknowledged varieties of any uccession, and other such facts, might come to the conclusion that species had not uccession, and other such facts, might come to the conclusion that species had not he mutual relations of the many beings which live around us. Who can explain why called the same genera are lineal descendants of some other and generally extinct

species

Phrase nets

Phrases of type "X of Y", X connected to Y in a graph; source van Ham et al 2009



Text visualization: additional sources

- Courses Data management (2L), Principles of Data Science (3Z)
- Text visualization browser https://textvis.lnu.se/
- Lecture from Univ. of Washington
- <u>Drawing Elena Ferrante's Profile</u>: Finding out who is Elena Ferrante, bestselling Italian author (My Brilliant Friend) by comparing word frequencies etc. (see e.g. page 100)