

Zuzana Nevěřilová
Hana Žižková
2024/25

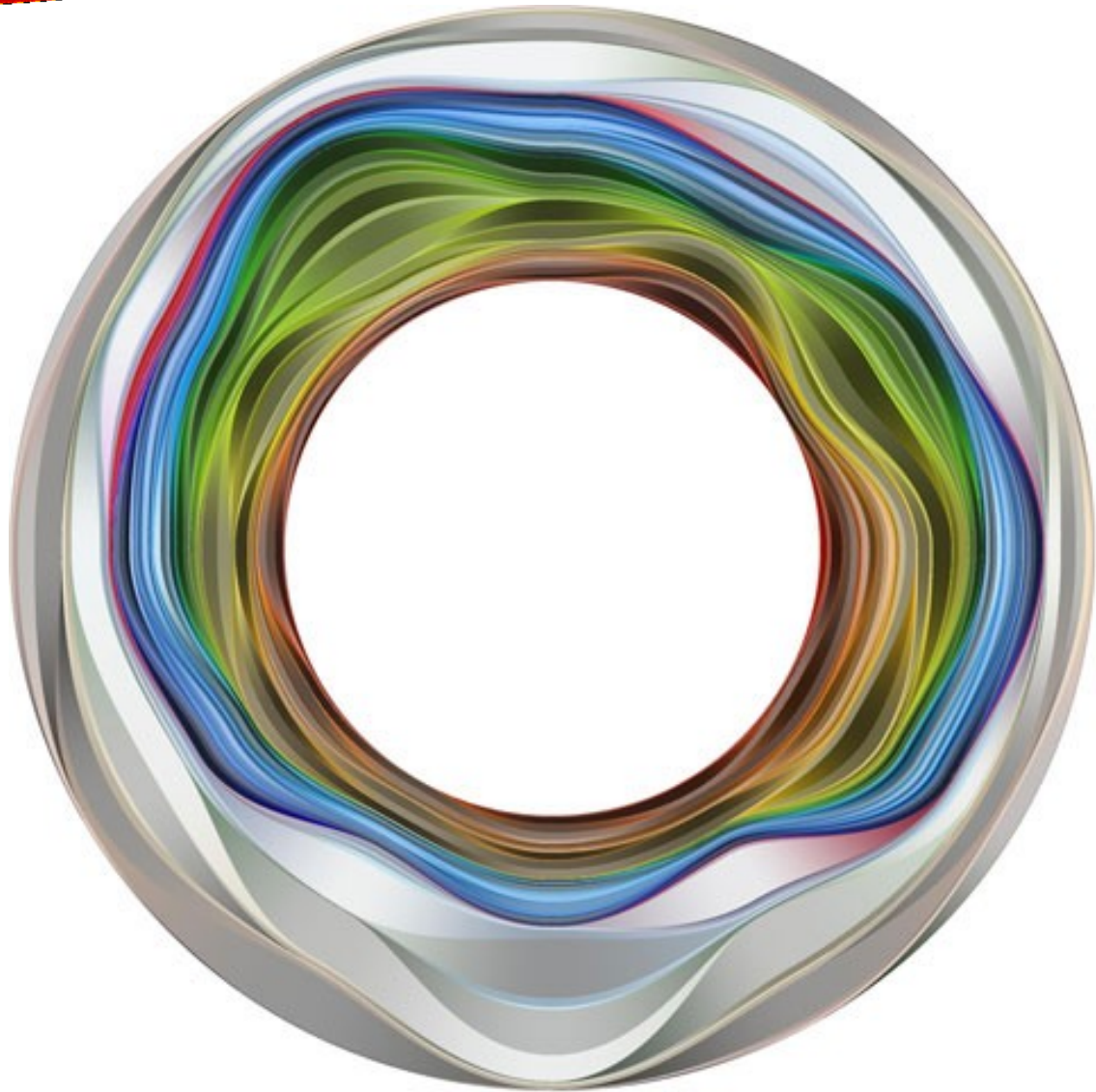
CORE147

DIGITÁLNÍ DATA
V HUMANITNÍCH
A SOCIÁLNÍCH



FLICKR FLOW

- Barvy roku ve fotografiích
- <http://hint.fm/projects/flickr/>



VIZUALIZACE A PŘÍBĚHY

VISUALIZATION AND STORYTELLING

Vizualizace

- Informace nedávají smysl, když je nemůžeme porovnat
- Vizualizace je informační mapa

Storytelling

- Nahlížení na komplexní data z více perspektiv
- Scénář

Komprimace informace
Změna úhlu pohledu
Edukace

MYŠLENKY DAVIDA MCCANDLESE

- Všichni jsme už viděli dobré a špatné vizualizace, infografiky atd.
- Najít dobrou vizualizaci je jako najít klidný palouček v informační džungli
- Dobrá vizualizace může pomoci při přesycení informacemi
- https://www.ted.com/talks/david_mccandless_the_beauty_of_data_visualization
- <https://informationisbeautiful.net/>

VIZUALIZACE DAT

- **Rysy** (features): barva, orientace, textura
- **Vzory** (patterns): oblasti se stejným rysem
- **Objekty**: to, co si uchová dlouhodobá paměť

Bottom-up information drives pattern building



Top-down attentional processes reinforce relevant information

VIZUALIZACE PRAKTICKY

- Obecné principy: Edward Tufte
- Praktické aspekty: Isabel Meirelles
- Příprava dat pro vizualizaci
- Volba vizualizačního nástroje
- [Data Basic](#)
- [Timeline JS](#)
- [Palladio](#) – prostorová data
- [Gephi](#) – sítě
- [Voyant Tools](#) - texty
- matplotlib
- <https://guides.nyu.edu/digital-humanities/tools-and-software/visualization>
- <https://guides.temple.edu/c.php?g=425229&p=9365463>

OBECNÉ PRINCIPY

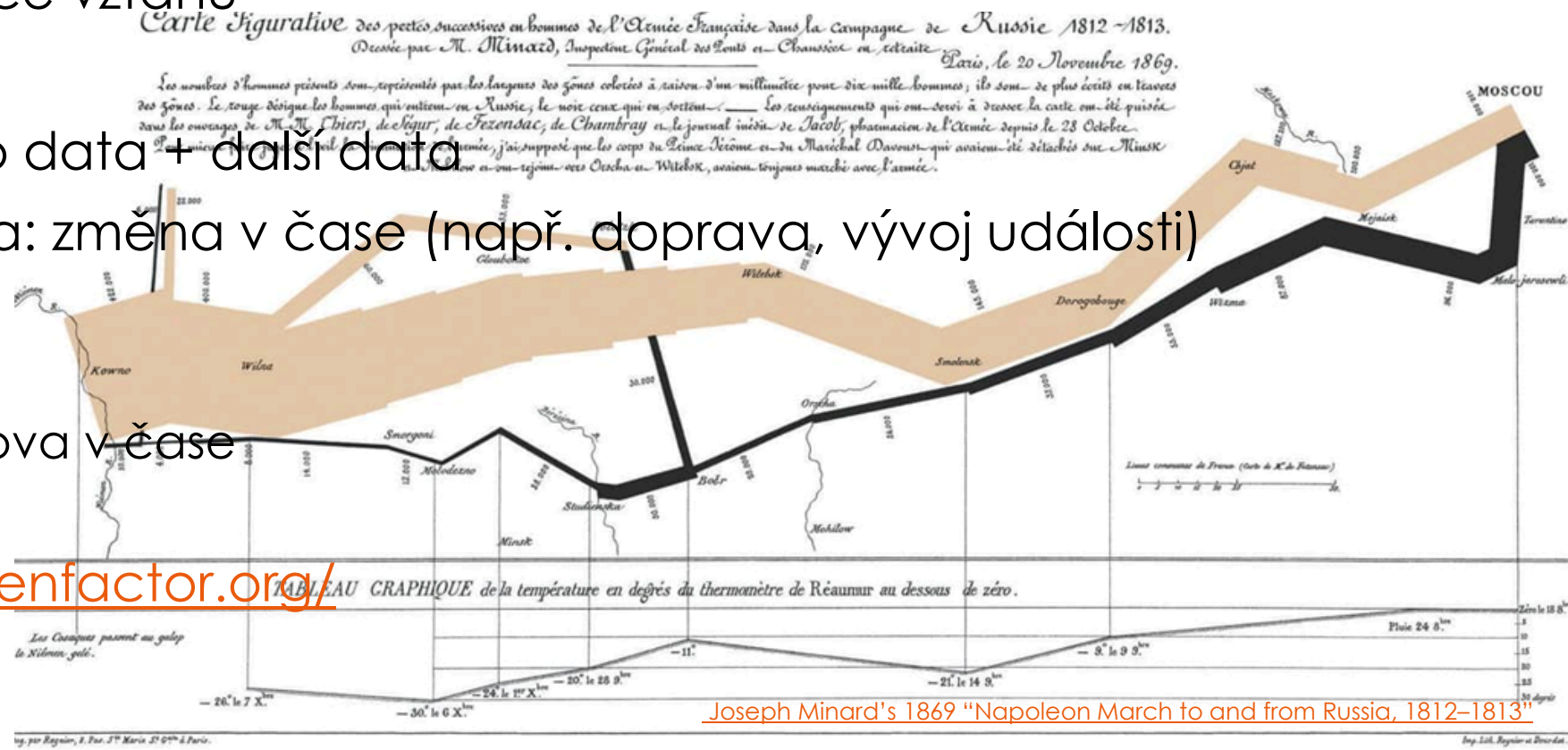
- Čtyři aspekty vizualizace (Edward Tufte)
 - **Graphical excellence**
“the greatest number of ideas, in the shortest time, using the least amount of ink, in the smallest space.”
 - **Visual integrity**
neither distort the underlying data nor create a false impression or interpretation of that data
 - **Maximize the data-ink ratio**
all superfluous elements that distract the user should be removed
 - **Aesthetic Elegance**
simplicity

MYŠLENKY EDWARDA TUFTEHO



TYPY DAT PODLE ISABEL MEIRELLES

- Stromy (hierarchie), taxonomie: (např. genealogie)
- Sítě (grafy): vizualizace vztahů
- Časové řady
- Prostory (mapy): geo data + další data
- Časoprostorová data: změna v čase (např. doprava, vývoj události)
- Textová data
 - komunikace (sítě)
 - Slova, kolokace, slova v čase



<http://well-formed.eigenfactor.org/>

PŘÍPRAVA DAT DATA V TABULKÁCH

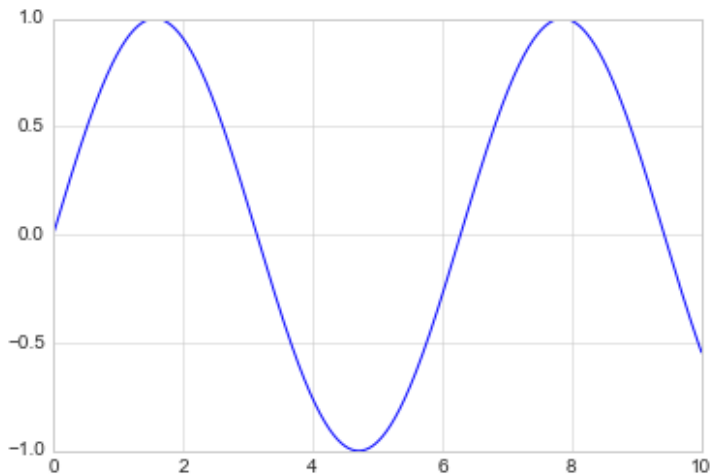
- Konverze z jiných dat
 - Např. Extrakce jmen osob z textů
 - Strukturovaná a „nestrukturovaná“ data
 - Ztráta informace (román vs. databázová tabulka)
- Zdroje tabulek
 - Formáty (existují převodníky)
 - Excel, TSV, CSV, pandas, JSON
 - Které sloupce, které osy zobrazit
 - `pandas.melt()`, `pandas.pivot()`, ...

MATPLOTLIB

ZÁKLADNÍ VIZUALIZACE: 2D DATA

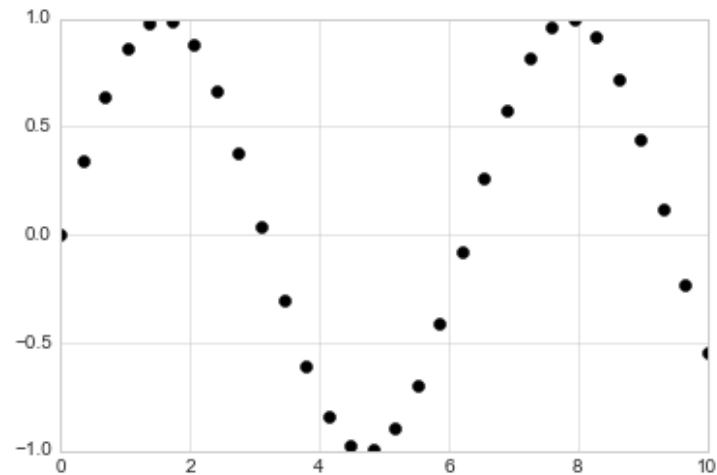
Continuous data

Function



Discrete data

Table



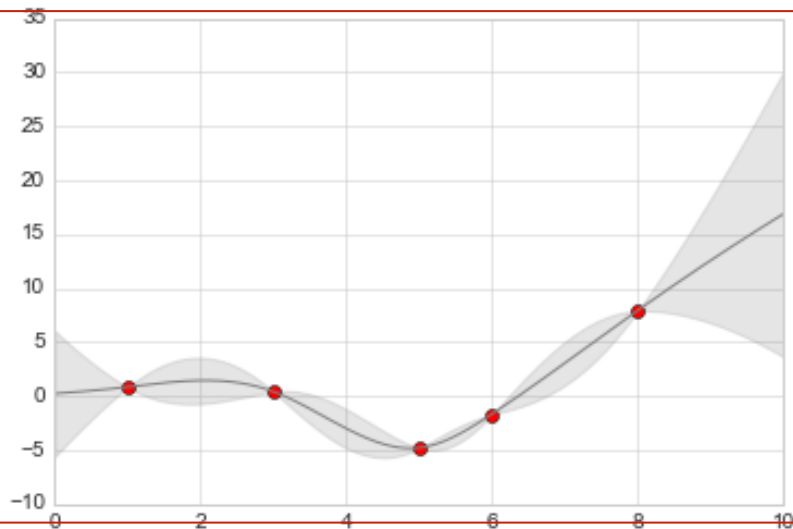
<https://jakevdp.github.io/PythonDataScienceHandbook/04.00-introduction-to-matplotlib.html>

MATPLOTLIB

ZÁKLADNÍ VIZUALIZACE: ROZPTYL

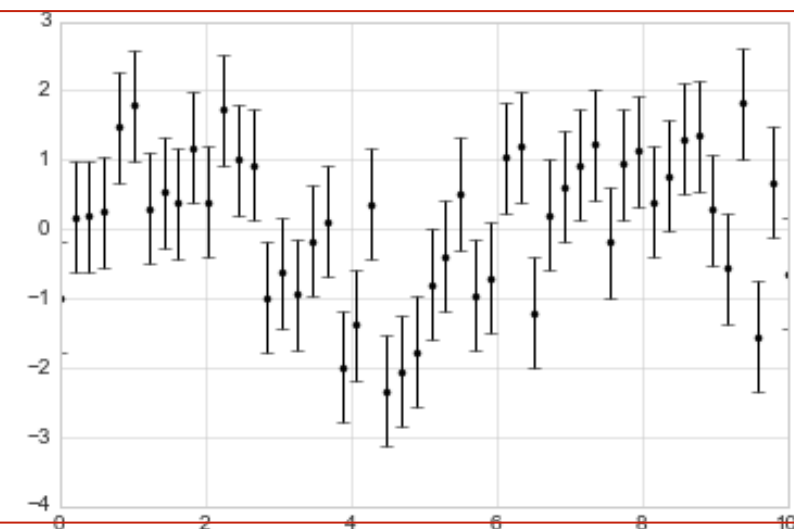
Continuous data

Function



Discrete data

Table

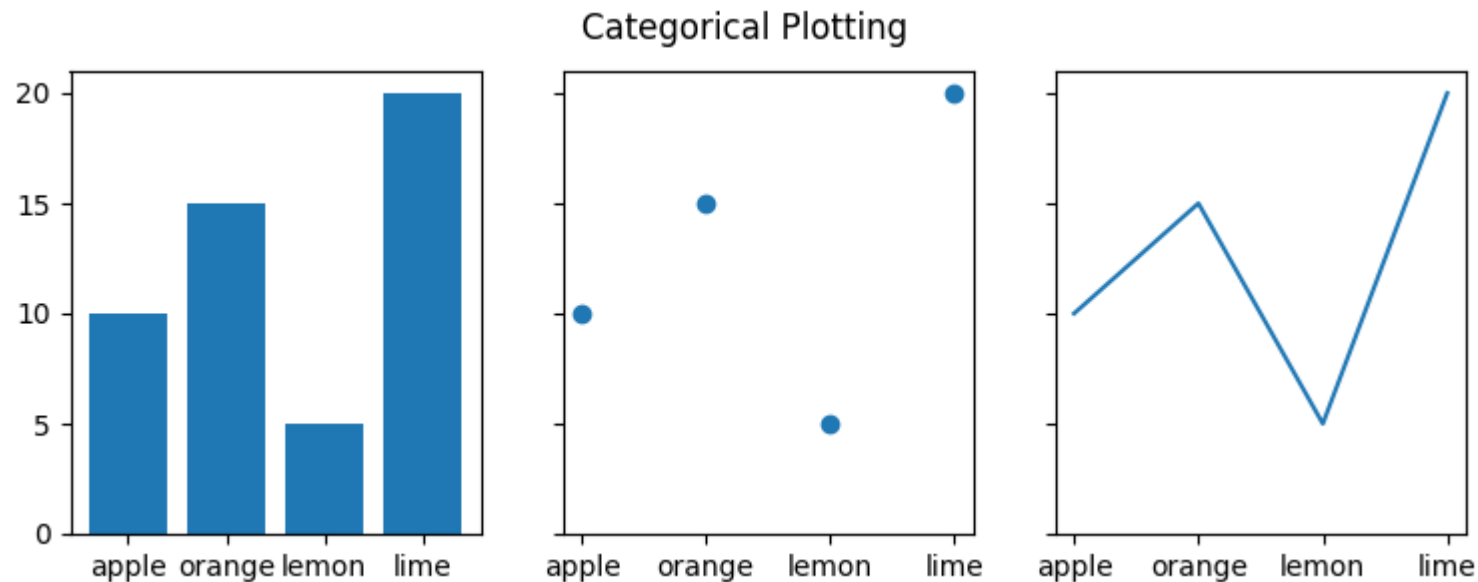


<https://jakevdp.github.io/PythonDataScienceHandbook/04.00-introduction-to-matplotlib.html>

MATPLOTLIB

ZÁKLADNÍ VIZUALIZACE: KATEGORIE

Vzdálenosti mezi kategoriemi nehrají roli

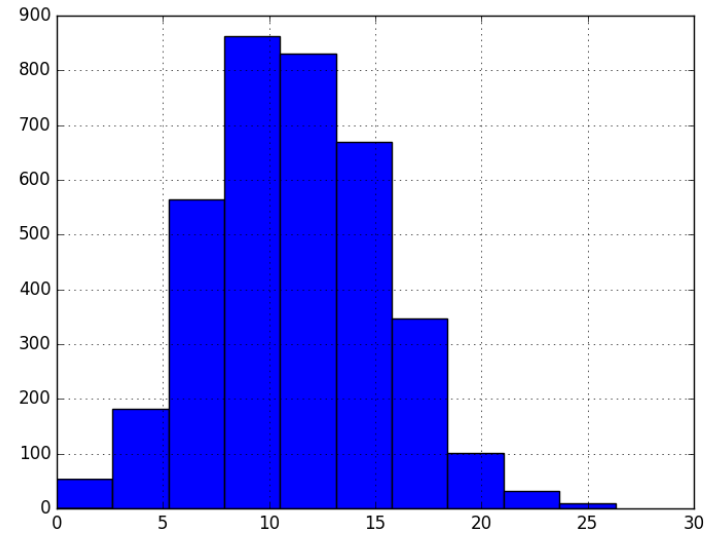
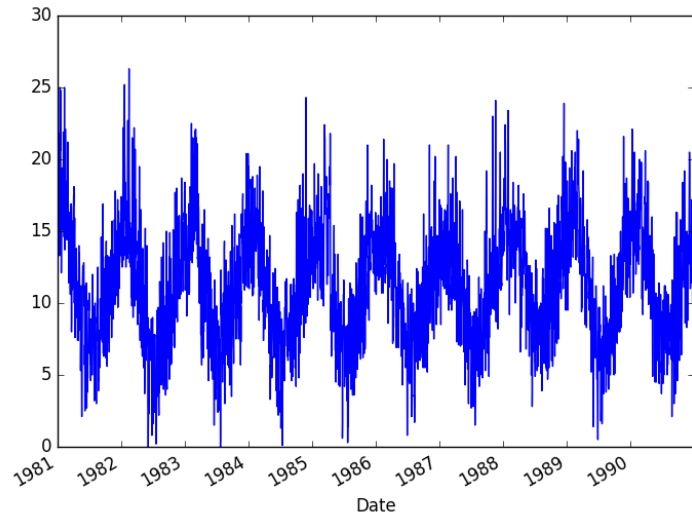


https://matplotlib.org/gallery/lines_bars_and_markers/categorical_variables.html

MATPLOTLIB

ZÁKLADNÍ VIZUALIZACE: ČASOVÁ DATA

Osa x = čas

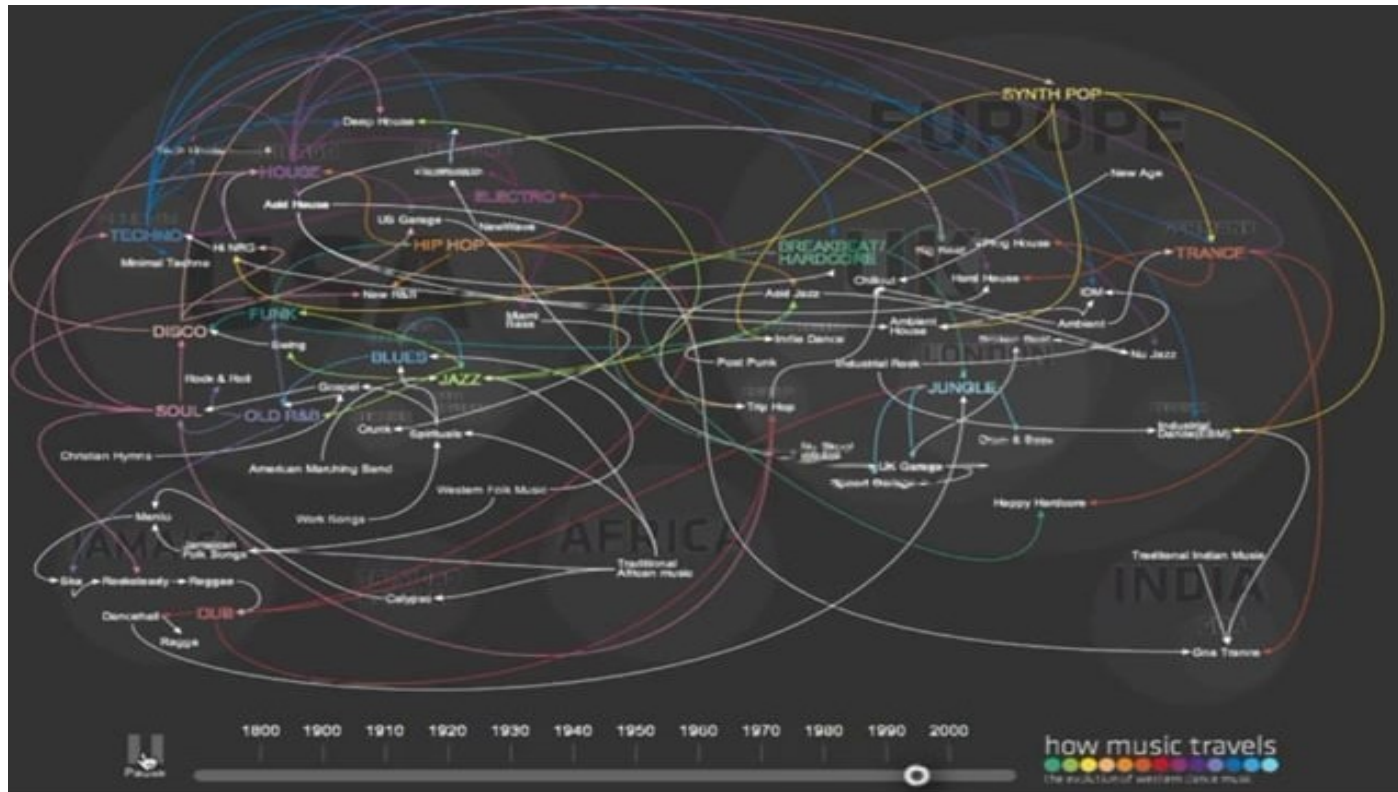


<https://machinelearningmastery.com/time-series-data-visualization-with-python/>

KOMPLEXNÍ VIZUALIZACE

- Více rozměrů než dva
 - Barva, velikost, tvar, vzdálenost,
- Zmenšení počtu dimenzí
 - Dimensionality reduction (t-SNE)
- Dynamické vizualizace
 - Interaktivní (e.g. Měřítko)
 - Automatické (e.g. Pohyblivá časová osa)

DYNAMICKÉ VIZUALIZACE



STORYTELLING NÁSTROJE A PROJEKTY

- Téma
 - Cíl
 - Cílová skupina
 - Scénář
 - Storyboard

 - Scalar
 - A Case of Hysteria
 - Narratives in Digital Libraries
 - Giant Squid
- Entity
 - Lidé
 - Události (časové údaje)
 - Místa
 - Koncepty
 - Vizualizace
 - Vztah mezi entitami
 - Klastrování
 - Časová osa
 - Multimédia
 - Obraz, video
 - Nahrávka, zvuk, hudba
 - Text



GENERATIVNÍ GRAFIKA A STORYTELLING

- Draw a researcher musing upon computer display full of data.
- Draw a visualization on a computer display, draw two friends watching the visualization. The two people are adults, looking fascinated and surprised by what they see on the display.

<https://openart.ai/>



DALŠÍ ČTENÍ/SLEDOVÁNÍ

- Edward Tufte: **THE WORK OF EDWARD TUFTE AND GRAPHICS PRESS**. <https://www.edwardtufte.com/tufte/>
- Interaction Design Foundation: **Guidelines for Good Visual Information Representations**. Interaction Design Foundation. 2019. <https://www.interaction-design.org/literature/article/guidelines-for-good-visual-information-representations>
- Vander P. J. 2016. **Python Data Science Handbook - Essential Tools for Working with Data**, O'Reilly. <https://jakevdp.github.io/PythonDataScienceHandbook/04.00-introduction-to-matplotlib.html>
- Jason Brownlee: **Time Series Data Visualization with Python**. Machine Learning Mastery. 2019. <https://machinelearningmastery.com/time-series-data-visualization-with-python/>
- **Pandas: Plotting with matplotlib**. <https://pandas.pydata.org/pandas-docs/version/0.13/visualization.html>

DALŠÍ ČTENÍ/SLEDOVÁNÍ

- Digital Humanities: Multimedia Storytelling. NYU Libraries.
<https://guides.nyu.edu/digital-humanities/tools-and-software/storytelling>
- Isabel Meirelles. (2013). Design for Information : An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations. Rockport Publishers.
- [The Beauty of Data Visualization, Ted Talk with David McCandless \(video 17 minutes, 40 seconds\)](https://guides.nyu.edu/digital-humanities/tools-and-software/visualization). <https://guides.nyu.edu/digital-humanities/tools-and-software/visualization>